

Journal of Theoretical and Applied Electronic  
Commerce Research

E-ISSN: 0718-1876

ncerpa@utalca.cl

Universidad de Talca  
Chile

Roosmand, Omid; Nematbakhsh, Mohammad Ali; Baraani, Ahmad  
An Electronic Marketplace Based on Reputation and Learning  
Journal of Theoretical and Applied Electronic Commerce Research, vol. 2, núm. 1, april, 2007, pp. 1-  
17  
Universidad de Talca  
Curicó, Chile

Available in: <http://www.redalyc.org/articulo.oa?id=96520102>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System  
Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal  
Non-profit academic project, developed under the open access initiative

## An Electronic Marketplace Based on Reputation and Learning

Omid Roozmand<sup>1</sup>, Mohammad Ali Nematbakhsh<sup>2</sup>, Ahmad Baraani<sup>3</sup>

<sup>1</sup> University of Isfahan, Iran, Department of Computer Engineering, [roozmand@eng.ui.ac.ir](mailto:roozmand@eng.ui.ac.ir)

<sup>2</sup> University of Isfahan, Iran, Department of Computer Engineering, [nematbakhsh@eng.ui.ac.ir](mailto:nematbakhsh@eng.ui.ac.ir)

<sup>3</sup> University of Isfahan, Iran, Department of Computer Engineering, [ahmadb@eng.ui.ac.ir](mailto:ahmadb@eng.ui.ac.ir)

Received 27 September 2006; received in revised form 22 October 2006; accepted 09 January 2007

### Abstract

In this paper, we propose a market model which is based on reputation and reinforcement learning algorithms for buying and selling agents. Three important factors: quality, price and delivery-time are considered in the model. We take into account the fact that buying agents can have different priorities on quality, price and delivery-time of their goods and selling agents adjust their bids according to buying agents preferences. Also we have assumed that multiple selling agents may offer the same goods with different qualities, prices and delivery-times. In our model, selling agents learn to maximize their expected profits by using reinforcement learning to adjust product quality, price and delivery-time. Also each selling agent models the reputation of buying agents based on their profits for that seller and uses this reputation to consider discount for reputable buying agents. Buying agents learn to model the reputation of selling agents based on different features of goods: reputation on quality, reputation on price and reputation on delivery-time to avoid interaction with disreputable selling agents. The model has been implemented with Aglet and tested in a large-sized marketplace. The results show that selling/buying agents that model the reputation of buying/selling agents obtain more satisfaction rather than selling/buying agents who only use the reinforcement learning.

**Key words:** Reputation, Reinforcement Learning, Electronic Commerce Agents

## 1 Introduction

With the advent of mobile and intelligent agent technology, e-commerce has been entered in a new era of its life [28]. Also agent architecture provides a flexible environment to model the other fields of research [8], [12], [20]. Agent-Based e-Marketplace is one of the most important results of using agent technology over e-Commerce. Electronic marketplace provides a single location for many buyers and sellers to congregate electronically and complete their own transactions. In the recent years, the extensive research is focused on designing agent-based e-Marketplaces [2], [6], [14], [15], [19]. Moreover, there are some research on personal intelligent agents for e-commerce applications [5], [7], [8], [10], [29]. But the most important problem that can be mentioned in these works is poor intelligence of trading agents.

In addition, reinforcement learning [17] has been studied for various multi-agent problems [4], [16], [21], [22]. However, these efforts are not directly modeled as economic agents and market environments. There are some research on reputation and trust modeling which do not use reinforcement learning [3], [9], [11], [18], [30]. A number of agent models for electronic market environments have been proposed. Jango [10] is a shopping agent that assists customers in getting product information. Given a specific product by a customer, Jango simultaneously queries multiple online merchants (from a list maintained by NetBot, Inc.) for the product availability, price, and important product features. Jango then displays the query results to the customer. Although Jango provides customers with useful information for merchant comparison, at least three shortcomings may be identified: (i) The task of analyzing the resultant information and selecting appropriate merchants is completely left for customers, (ii) The algorithm underlying its operation does not consider product quality which is of great importance for the merchant selection task, (iii) Jango is not equipped with any learning capability to help customers choose more and more appropriate merchants. Another interesting agent model is Kasbah [5], designed by the MIT Media Lab. Kasbah is a multi-agent electronic marketplace where selling and buying agents can negotiate with one another to find the "best possible deal" for their users. The main advantage of Kasbah is that its agents are autonomous in making decisions, thus freeing users from having to find and negotiate with buyers and sellers. However, as admitted in [5], Kasbah's agents are not very smart as they do not make use of any AI learning techniques.

Vidal and Durfee [27] address the problem of how buying and selling agents should behave in an information economy such as the University of Michigan Digital Library. They divide agents into classes corresponding to the agents' capabilities of modeling other agents: Zero-level agents are the agents that learn from the observations they make about their environment, and from any environmental rewards they receive. One-level agents are those agents that model agents as zero-level agents. Two-level agents are those that model agents as one-level agents. Higher level agents are recursively defined in the same manner. It should be intuitive that the agents with more complete models of others will always do better. However, because of the computational costs associated with maintaining deeper (i.e., more complex) models, there should be a level at which the gains and the costs of having deeper models balance out for each agent. The main problem addressed in this model is to answer the question of when an agent benefits from having deeper models of others. Also reinforcement learning has been applied in market environments for buying and selling agents, but reputation has not been used as a means to protect buyers from purchasing low quality goods. Moreover, selling agents do not consider altering the quality of their products while learning to maximize their profits.

Tran and Cohen in [23]-[26] exploit reinforcement learning for buying agents to model the reputation of selling agents to protect buyers from communicating with non-reputable sellers. Nevertheless, buyers in this model should have fixed priorities on quality and price of their desired goods. In this way, they can not change their preferences to buy a good in a sequence of purchases. That is, a buying agent can not purchase a good in an auction with priority on quality and willing to buy the same good in another auction with priority on price. In addition, selling agents do not model the reputation of buyers to consider discount and just only focuses on two factors of quality and price.

In our proposed learning algorithms, each selling agent models the reputation of buyers and dedicates them discounts based on their reputation. This model focuses on three important factors in market: quality, price and delivery. Because of the existence of buying agents with different preferences and priorities on their desired goods, the buying agents model the reputation of selling agents based on quality, price and delivery separately. For example, a buyer may need a good with high quality now, but with low price later. The proposed model has been implemented with Aglet [1], [13].

The paper is organized as follows: section 2 introduces our proposed market and learning algorithms. Section 3 discusses current experimental results and outlines proposed future experimentations with the model. Finally, Section 4 provides conclusion and some future research directions.

## 2 The Proposed Algorithm

In this section we propose our marketplace model and learning algorithm for buying and selling agents based on reinforcement learning and reputation modeling.

## 2.1 General Architecture for Agent-Based e-Marketplace

The proposed architecture of e-Marketplace is shown in Figure 1. There are three types of server in the proposed architecture for e-Marketplace, they are: (1) Marketplace (2) Buying Agent Server, and (3) Selling Agent Server. Each server includes several stationary agents and mobile agents and some important transactions between different agents in the marketplace. They are described as follows:

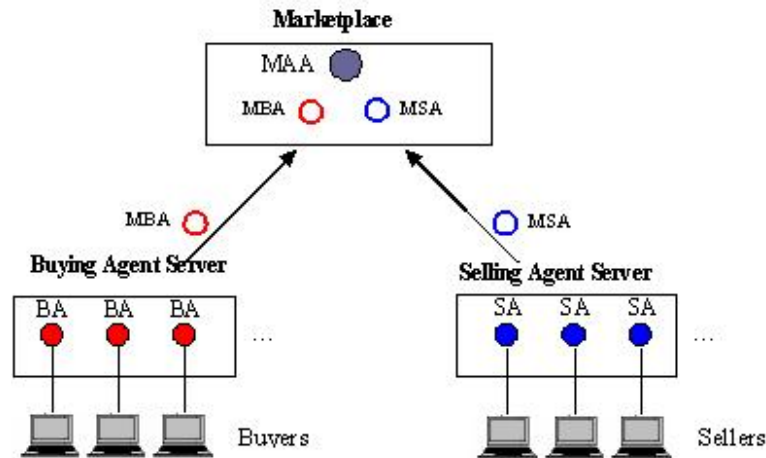


Figure 1: The architecture of electronic commerce environments

### 2.1.1 Marketplace

Marketplace is a platform that supports the transaction facilities for mobile agent of sellers and buyers. There is a static Agent (MAA: Market Assistant Agent) and two kinds of Mobile Agent in the Marketplace:

1. MAA (Market Assistant Agent): The MAA is responsible for registering mobile buying and selling agents in the buyer and seller database of marketplace. The buyer database of marketplace contains: owner of mobile buying agent, buying agent server, a unique identifier, and proxy address of agent provided by aglet context and time of registration. The seller database of marketplace contains: owner of mobile selling agent, selling agent server, a unique identifier, address proxy of selling agent provided by aglet context, goods which is available for mobile selling agent to sell and time of registration. Agent A can communicate with agent B through the proxy address of agent B and vice versa. Also the MAA answers to the mobile buying agent request by retrieving proxy address of sellers, from seller database, who have good g to sell and send the list to the mobile buying agent.
2. MBA (Mobile Buying Agent): stands for the buyer, moves to the Marketplace and trades with mobile selling agents and learns, based on reinforcement learning, that which sellers can satisfy its preferences. Also the MBA measures the reputation of each mobile selling agent on different factors: quality, price and delivery and focuses its business on reputable sellers and prevent to interact with non reputable ones.
3. MSA (Mobile Selling Agent): stands for the seller, moves to the Marketplace and trades with mobile buying agents and learns how to adjust its bids according to the preferences of the buying agents while trying to maximize its expected profit. Also models the reputation of mobile buying agents to dedicate discount for them based on their reputation.

### 2.1.2 Buying Agent Server

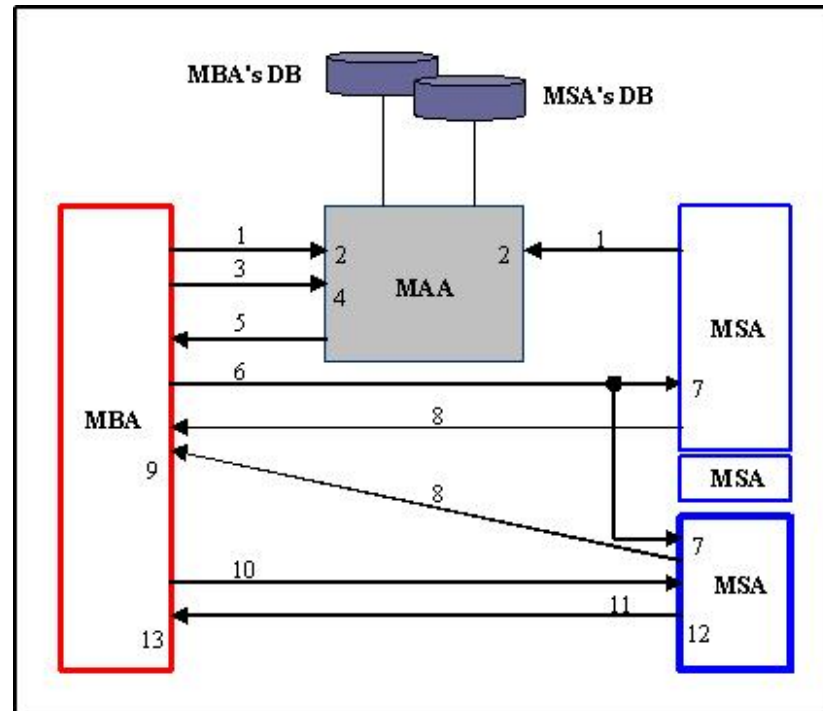
The Buying Agent Server provides the interface of Buying Agent (BA) that lets users to initialize and control their buying agent to carry the e-commerce activation out. Buying Agent Server stores the information of buyer in the database and will produce Mobile Buying Agent (MBA) according to the requirements of the user. It stands for the user to go to the marketplace to make bargains.

### 2.1.3 Selling Agent Server

Each seller, which wants to join this e-marketplace, should build a Seller Server. There are two main Agents in a Seller Server, include: (1) Selling Agent (SA): which is provided by Selling Agent Server that lets the seller to initialize its selling agent and specify the goods which is available to sell, and (2) Mobile Selling Agent (MSA) is created by Selling Agent Server and migrates to the Marketplace and try to sell goods with maximum profit for its owner.

### 2.1.4 Transactions in Marketplace

Figure 2 shows the process of trading using thirteen transactions:



MBA: Mobile Buying Agent  
MBA's DB: Mobile Buying Agent's Database  
MAA: Market Assistant Agent  
MSA: Mobile Selling Agent  
MSA's DB: Mobile Selling Agent's Database

Figure 2: Transactions in Agent-Based e-Marketplace

1. BA submits registration request to MAA. Also SA submits registration request to MAA.
2. MAA stores BA's and SA's registration information in B's and S's Databases.
3. BA requests from MAA for list of relevant sellers who sell specified product.
4. MAA retrieves relevant sellers for requested product.
5. MAA sends list of relevant sellers to BA.
6. BA multicasts its requests to relevant sellers for specified product.
7. Each of those SA's prepares bid for BA based on his reputation and purchases.
8. Each of those SA's sends bid to BA.
9. BA receives all bids, evaluates their value and selects the best bid.
10. BA announces the chosen bid owner and pays it.
11. Chosen SA delivers the product to BA.
12. Chosen SA updates the reputation of BA.
13. BA estimates the real value of good and updates the trust and reputation of this SA.

By considering some assumptions, we make the market more realistic and simpler. Therefore, we assume that:

1. Quality, price and delivery offered by different selling agents can be variable.
2. Each selling agent considers discounts for buying agents based on their reputation.
3. There may be some dishonest selling agents in the market who lie on quality and delivery.
4. Buying agents in the market are not dishonest.
5. A buyer can purchase a good in different conditions with variant priorities on quality, price and delivery instead of fixed priorities.
6. Each buyer has his own preferences and priorities on quality, price and delivery.
7. Product delivery is done by transferring message between seller and buying agents.
8. Maximum quality of a good presented in the market is definite so that all selling and buying agents know that.
9. If a seller wants to deliver his product later, buyer expects more reduction in price from that seller based on late time units
10. Buyer can estimate the quality of the good he purchases only after receiving the good from the selected seller.

In the following sections 2.2 and 2.3, we present the seller and buyer algorithms, respectively. Also in section 2.4 an example is described to show how the buyer and seller algorithms work.

## 2.2 Seller Algorithm

Let  $S$  be the set of sellers,  $G$  be the set of goods,  $B$  be the set of buyers,  $Q$  be the set of qualities,  $P$  be the set of prices and  $D$  be the set of deliveries, and  $S$ ,  $G$ ,  $B$ ,  $Q$ ,  $P$  and  $D$  are finite sets (It means that  $q_{\min} \in Q$  and  $q_{\max} \in Q$  represent minimum and maximum quality of goods that can be available in the market and all sellers and buyers know this). Assume that seller  $s \in S$  has received a request from buyer  $b \in B$  on good  $g \in G$ . Seller  $s$  has to decide on the quality, price and delivery of good  $g$  to be delivered to buyer  $b$ . Assume that  $R$  be the set of real numbers. Let function  $e^s : G \times Q \times P \times D \times B \rightarrow R$  estimate the expected profit for seller  $s$  if it sells good  $g$  with quality  $q$  at price  $p$  and delivery  $d$  to buyer  $b$ . Let  $c^s(g, q, b)$  be the cost that seller  $s$  incurs to produce good  $g$  with quality  $q$  for buyer  $b$ . Seller  $s$  produce different versions of good  $g$  based on buyers requirements. The price that seller  $s$  chooses to sell good  $g$  to buyer  $b$  is greater than or even equal to  $c^s(g, q, b)$ . Function  $e^s$  chooses a bid that has the maximum profit for seller  $s$ . If seller  $s$  produces good  $g$  with the cost of  $c^s(g, q, b)$ , the maximum price for seller  $s$  is calculated as follows:

$$p_{\max} = c^s(g, q, b) + c^s(g, q, b) * \kappa \quad (1)$$

In which,  $\kappa$  is the maximum percent of profit for seller  $s$ . Moreover, seller  $s$  models the reputation of all buyers in the market using function  $r^s : B \rightarrow (0,1)$  that is called the reputation function of  $s$ . Initially seller  $s$  sets the reputation rating  $r^s(b) = 0$  for each buyer  $b \in B$ . We do not use the negative reputation for buyers, because we assumed all buyers are honest and no sellers are interested to lose their customers. The sellers want to satisfy the buyers' requirements so that they compete with each other to increase the number of their own customers.

When seller  $s$  sends his bid to buyer  $b$ , there are the two following possibilities:

Seller  $s$  succeeds to sell good  $g$  with quality  $q$  at price  $p$  and delivery  $d$  to buyer  $b$ . It means that seller  $s$  has presented a bid better than the other sellers' bids to buyer  $b$ . Therefore, seller  $s$  may be re-selected by buyer  $b$  if seller  $s$  repeats this bid again for buyer  $b$  for specified good  $g$ . Seller  $s$  delivers product to buyer  $b$  and updates the reputation of buyer  $b$  using reinforcement learning:

$$r^s(b) = r^s(b) + \mu(1 - r^s(b)) \quad (2)$$

Where,  $\mu$  is a positive factor called cooperative factor and is equal to:

$$\mu = \frac{p - c^s(g, q, b)}{p_{\max} - c^s(g, q, b)} \quad (3)$$

In which,  $p_{\max} - c^s(g, q, b)$  is the maximum profit for seller  $s$  if it could sell good  $g$  to  $b$ .

So the new bid for buyer  $b$  based on its new reputation is calculated by seller  $s$  as follows:

$$p_{\text{new}} = p_s - (r_{\text{new}}^s(b) - r_{\text{before}}^s(b)) * p_s \quad (4)$$

Seller  $s$  does not succeed to sell good  $g$  with quality  $q$  at price  $p$  and delivery  $d$  to buyer  $b$ . It means that, the bid of seller  $s$  has not satisfied the buyer  $b$ . If seller  $s$  repeats the previous bid to buyer  $b$ , the possibility of success in selling good  $g$  to buyer  $b$  is low. Therefore, it is needed to alter the price, delivery and may be quality of the good to be offered to buyer  $b$ . In a real market, buyers expect that if a seller want to deliver his good later than his offered time, the seller should reduce the price according to a formula based on price and delivery. Let  $rp$  be a variable that specifies the reduction percent of price for seller who want to delivers his product late. That is, he should reduce the price of his product according to this value. In addition, for preparing a new bid for buyer  $b$ , the reputation of the buyer is also used to determine the new price. The quality remains as before but new price is updated with reinforcement learning as follows:

$$p_{\text{new}} = p - (rp * p) - \beta * r^s(b) * p \quad (5)$$

In which,  $\beta$  ( $0 < \beta < 1$ ) is a variable that denotes discount. It means that seller who wants to consider more discount for his customer, sets  $\beta$  with greater value, and vice versa, instead he increments the product delivery value:

$$dt_{new} = dt + dt_{inc}$$

(6)

Which  $dt_{inc}$  is increasing rate for delivery corresponding to price reduction which has been assumed by seller  $s$ . According to the fact that a seller does not sell his goods with a price lower than the production cost of the good, therefore if  $p_{new} < c^s(g, q, b)$ , then seller  $s$  does not suggest the same good with previous quality. So that, he may optionally raise the value of quality by increasing its production cost as follows:

$$c^s(g, q, b) = (1 + inc)c^s(g, q, b) \quad (7)$$

Where,  $inc$  is a specific constant called seller  $s$ 's quality increasing factor.

### 2.3 Buyer Algorithm

Assume that buyer  $b$  wants to buy good  $g$ . Buyer  $b$  broadcasts his request to all sellers which they have good  $g$  to sell (According to what discussed earlier in Figure 1, list of these sellers has been already retrieved from MAA.). Sellers answer the request by sending bids to buyer  $b$ . Buyer  $b$  receives all bids and selects the suitable bid. Buyer  $b$  models the reputation of all sellers and selects the suitable bid from a reputable seller. Buyer  $b$  models the reputation of each seller based on three factors of quality, price and delivery, separately. To model the reputation of each seller, buyer  $b$  uses functions  $r_q^b: S \rightarrow (-1, 1)$ ,  $r_p^b: S \rightarrow (-1, 1)$  and  $r_d^b: S \rightarrow (-1, 1)$  that are called reputation function of  $b$

based on factors quality ( $q$ ), price ( $p$ ) and delivery ( $d$ ), respectively. For example  $r_q^b(s)$  represents the reputation of seller  $s$  on quality computed by buyer  $b$ . Initially, buyer  $b$  sets the reputation ratings  $r_q^b(s) = 0$ ,  $r_p^b(s) = 0$  and  $r_d^b(s) = 0$  for every seller  $s \in S$ . Seller  $s$  is reputable for buyer  $b$  on quality iff  $r_q^b(s) \geq \Theta_q^b$ , where  $\Theta_q^b$  is buyer  $b$ 's reputable threshold on quality ( $0 < \Theta_q^b < 1$ ). A seller  $s$  is considered as disreputable for buyer  $b$  on quality iff  $r_q^b(s) \leq \theta_q^b$ , where  $\theta_q^b$  is buyer  $b$ 's disreputable threshold on quality ( $-1 < \theta_q^b < 0$ ). Similarly, we define buyer  $b$ 's reputable and disreputable thresholds based on price and delivery by replacing  $q$  with  $p$  and  $d$  in the above inequalities, respectively.

Let  $S_{r-q}^b$  be the set of sellers with good reputation on quality to buyer  $b$ ; that is;  $S_{r-q}^b$  contains the sellers that have served  $b$  with expected quality of  $b$  in the past and are therefore reputable on quality by  $b$ . Hence,  $S_{r-q}^b \subseteq S$  and is initially empty, i.e.,

$$S_{r-q}^b = \left\{ s \in S \mid r_q^b(s) \geq \Theta_q^b \right\} \subseteq S \quad (8)$$

Also Let  $S_{r-p}^b$  and  $S_{r-d}^b$  be the set of sellers with good reputation on price and delivery, respectively. ( $S_{r-p}^b$  and  $S_{r-d}^b$ )  $\subseteq S$ ,  $S_{r-p}^b$  and  $S_{r-d}^b$  are initially empty too, i.e.,

$$S_{r-p}^b = \left\{ s \in S \mid r_p^b(s) \geq \Theta_p^b \right\} \subseteq S \quad (9)$$

$$S_{r-d}^b = \left\{ s \in S \mid r_d^b(s) \geq \Theta_d^b \right\} \subseteq S \quad (10)$$

Assume that  $S_{nr-q}^b$  be the set of sellers with bad reputation on quality to buyer  $b$ ; that is;  $S_{nr-q}^b$  contains the sellers that have served  $b$  with not expected quality of  $b$  and are known as non reputable sellers on quality by  $b$ .  $S_{nr-q}^b \subseteq S$  and is initially empty, i.e.,

$$S_{nr-q}^b = \left\{ s \in S \mid r_q^b(s) \leq \theta_q^b \right\} \subseteq S \quad (11)$$

Also Let  $S_{nr-p}^b$  and  $S_{nr-d}^b$  be the set of sellers with bad reputation on price and delivery to buyer  $b$ , respectively. ( $S_{nr-p}^b$  and  $S_{nr-d}^b$ )  $\subseteq S$ ,  $S_{nr-p}^b$  and  $S_{nr-d}^b$  are initially empty, i.e.,

$$S_{nr-p}^b = \left\{ s \in S \mid r_p^b(s) \leq \theta_p^b \right\} \subseteq S \quad (12)$$



$$S_{nr\_d}^b = \left\{ s \in S \mid r_{nr\_d}^b(s) \leq \theta_d^b \right\} \subseteq S \quad (13)$$

Let  $w_q$ ,  $w_p$  and  $w_d$  be the weight of values quality, price and delivery for buyer b so that  $w_q + w_p + w_d = 1$ . We define the buyer b's general reputable threshold as follows:

$$\Theta^b = w_q * \Theta_q^b + w_p * \Theta_p^b + w_d * \Theta_d^b \quad (14)$$

while buyer b's general disreputable threshold is:

$$\theta^b = w_q * \theta_q^b + w_p * \theta_p^b + w_d * \theta_d^b \quad (15)$$

In the same way, we calculate the general reputation of seller s as follows:

$$r^b(s) = w_q * r_q^b(s) + w_p * r_p^b(s) + w_d * r_d^b(s) \quad (16)$$

Let  $S_r^b$  and  $S_{dr}^b$  be the sets of reputable and disreputable sellers to buyer b respectively, i.e.,

$$S_r^b = \left\{ s \in S \mid r^b(s) \geq \Theta^b \right\} \subseteq S \quad (17)$$

And

$$S_{nr}^b = \left\{ s \in S \mid r^b(s) \leq \theta^b \right\} \subseteq S. \quad (18)$$

Which  $\Theta^b$  and  $\theta^b$  are general reputable threshold and general disreputable threshold respectively. Buyer b will focus his business on the reputable sellers and stays away from disreputable ones.

Assume that each seller send its bid in triple  $bid(q_s, p_s, d_s)$  to buyer b. Then buyer b guesses the value of bids offered by each seller by using this function:

$$G^b(q_s, p_s, d_s, s) = w_q * \frac{q_s}{q_{\max}} - w_p * \frac{p_s}{p_{\max}} - w_d * \frac{d_s}{d_{\max}} \quad (19)$$

Where  $q_{\max}$  is the maximum quality of good g in the market,  $p_{\max}$  is the maximum price for good with quality  $q_{\max}$  and  $d_{\max}$  is the maximum time for seller s to deliver good g late. Then buyer b selects the seller  $\hat{s}$  who belongs to the set of reputable sellers for buyer b whose bid value for buyer b is more than the other sellers', i.e.,

$$\hat{s} = \arg \max_{s \in S_r^b} G^b(q_s, p_s, d_s, s) \quad (20)$$

Where,  $\arg$  is an operator such that  $\arg G^b(s)$  returns s. In addition, if no sellers in  $S_r^b$  submit bids for delivering g (i.e.,  $S_r^b = \emptyset$ ), then buyer b has to choose a seller  $\hat{s}$  among sellers who are not reputable nor disreputable:

$$\hat{s} = \arg \max_{s \notin (S_r^b \cup S_{dr}^b)} G^b(q_s, p_s, d_s, s) \quad (21)$$

After paying seller  $\hat{s}$  and receiving good g, buyer b examine the quality  $q \in Q$  of good g. Assume that buyer b find the quality  $\hat{q}$  that has been delivered at  $\hat{d}$ . Let the expected quality, price and delivery for the buyer be  $q_b$ ,  $p_b$  and  $d_b$ , respectively. Updating reputation on quality, price and delivery of seller  $\hat{s}$  is illustrated in next parts respectively. In addition, with a probability  $\rho$  buyer b chooses to explore (rather than exploit) the marketplace by randomly selecting a seller  $\hat{s}$  from the set of all sellers. Initially, the value of  $\rho$  should be set to 1, and then decreased over time to some fixed minimum value determined by the buyer.

Updating Reputation on Quality:

If  $\hat{q} \geq q_b$  then the reputation of seller  $\hat{s}$  on quality is updated using reinforcement learning as follows:

$$r_q^b(s) = \begin{cases} r_q^b(s) + \mu_q (1 - r_q^b(s)) & \text{if } r_q^b(s) \geq 0 \\ r_q^b(s) + \mu_q (1 + r_q^b(s)) & \text{if } r_q^b(s) < 0 \end{cases} \quad (22)$$

Where,  $\mu_q$  is a positive factor called the cooperation factor.  $\mu_q$  is calculated as follows:



$$\mu_q = \begin{cases} \frac{\hat{q} - q_b}{q_{\max}} & \text{if } \frac{\hat{q} - q_b}{q_{\max}} > \mu_{\min\_q} \\ \mu_{\min\_q} & \text{otherwise} \end{cases} \quad (23)$$

That is, seller  $\hat{s}$  offers good  $g$  with a quality greater than or equal to the value that buyer  $b$  demanded for quality of good  $g$  and therefore the reputation of seller  $\hat{s}$  on quality is increased by equation (22) accordingly.  $\mu_{\min\_q}$  is a positive factor called minimum cooperation factor for quality.

If  $\hat{q} < q_b$  then the reputation of seller  $\hat{s}$  on quality is updated as follows:

$$r_q^b(s) = \begin{cases} r_q^b(s) + v_q(1 - r_q^b(s)) & \text{if } r_q^b(s) \geq 0 \\ r_q^b(s) + v_q(1 + r_q^b(s)) & \text{if } r_q^b(s) < 0 \end{cases} \quad (24)$$

Where,  $v_q$  is a negative factor called the non-cooperation factor.  $v_q$  is calculated as follows:

$$v_q = \lambda_q \frac{\hat{q} - q_b}{q_{\max}} \quad (25)$$

In which,  $\lambda_q$  ( $\lambda_q > 1$ ) is called the penalty factor so that  $|v_q| > |\mu_q|$  to implement the traditional assumption that reputation should be difficult to build up, but easy to tear down.

Updating Reputation on Price:

1) If  $p_b \geq p_s$  then the reputation of seller  $\hat{s}$  on price is updated using reinforcement learning as follows:

$$r_p^b(s) = \begin{cases} r_p^b(s) + \mu_p(1 - r_p^b(s)) & \text{if } r_p^b(s) \geq 0 \\ r_p^b(s) + \mu_p(1 + r_p^b(s)) & \text{if } r_p^b(s) < 0 \end{cases} \quad (26)$$

Where,  $\mu_p$  is a positive factor called the cooperation factor.  $\mu_p$  is calculated as follows:

$$\mu_p = \begin{cases} \frac{p_b - p_s}{p_{\max}} & \text{if } \frac{p_b - p_s}{p_{\max}} > \mu_{\min\_p} \\ \mu_{\min\_p} & \text{otherwise} \end{cases} \quad (27)$$

That is, seller  $\hat{s}$  offers good  $g$  with a price lower than or equal to the value that buyer  $b$  demanded for price of good  $g$  and therefore the reputation of seller  $\hat{s}$  on price is increased by equation (26) accordingly. It implements this fact that buyer  $b$  expects to buy goods with low price, therefore sellers who offer goods with lower price than the other, set more reputation on price for themselves to buyer  $b$  and those sellers have positive reputation on price that their price is lower than expected price of buyer  $b$ .  $\mu_{\min\_p}$  is a positive factor called minimum cooperation factor for price.

2) If  $p_b < p_s$  then the reputation of seller  $\hat{s}$  on price is updated as follows:

$$r_p^b(s) = \begin{cases} r_p^b(s) + v_p(1 - r_p^b(s)) & \text{if } r_p^b(s) \geq 0 \\ r_p^b(s) + v_p(1 + r_p^b(s)) & \text{if } r_p^b(s) < 0 \end{cases} \quad (28)$$

Where,  $v_p$  is a negative factor called the non-cooperation factor.  $v_p$  is calculated as follows:

$$v_p = \lambda_p \frac{p_b - p_s}{p_{\max}} \quad (29)$$

In which,  $\lambda_p$  ( $\lambda_p > 1$ ) is called the penalty factor so that  $|v_p| > |\mu_p|$ .

Updating Reputation on Delivery:

If  $d_b \geq \hat{d}$  then the reputation of seller  $\hat{s}$  on delivery is updated using reinforcement learning as follows:

$$r_d^b(s) = \begin{cases} r_d^b(s) + \mu_d(1 - r_d^b(s)) & \text{if } r_d^b(s) \geq 0 \\ r_d^b(s) + \mu_d(1 + r_d^b(s)) & \text{if } r_d^b(s) < 0 \end{cases} \quad (30)$$

Where,  $\mu_d$  is a positive factor called the cooperation factor.  $\mu_d$  is calculated as follows:

$$\mu_d = \begin{cases} \frac{d_b - \hat{d}}{d_{\max}} & \text{if } \frac{d_b - \hat{d}}{d_{\max}} > \mu_{\min\_d} \\ \mu_{\min\_d} & \text{otherwise} \end{cases} \quad (31)$$

That is, seller  $\hat{s}$  offers good g with a delivery lower than or equal to the value that buyer b demanded for delivery of good g and therefore the reputation of seller  $\hat{s}$  on delivery is increased by equation (30) accordingly. It means that sellers who deliver their product more quickly, set more reputation on delivery for themselves and those sellers have positive reputation on delivery that their delivery is lower than expected delivery of buyer b.  $\mu_{\min\_d}$  is a positive factor called minimum cooperation factor for delivery.

If  $d_b < \hat{d}$  then the reputation of seller  $\hat{s}$  on delivery is updated as follows:

$$r_d^b(s) = \begin{cases} r_d^b(s) + v_d(1 - r_d^b(s)) & \text{if } r_d^b(s) \geq 0 \\ r_d^b(s) + v_d(1 + r_d^b(s)) & \text{if } r_d^b(s) < 0 \end{cases} \quad (32)$$

Where,  $v_d$  is a negative factor called the non-cooperation factor.  $v_d$  is calculated as follows:

$$v_d = \lambda_d \frac{d_b - \hat{d}}{d_{\max}} \quad (33)$$

In which,  $\lambda_d$  ( $\lambda_d > 1$ ) is called the penalty factor so that  $|v_d| > |\mu_d|$ .

## 2.4 An Example

This subsection provides a numerical example illustrating the proposed algorithm for buyers and sellers, respectively.

### 2.4.1 Buyer Situation

Consider a simple buyer situation where a buyer b announces its need of some good g to all sellers which they have good g to sell (According to what discussed earlier in Figure 1, list of these sellers has been already retrieved from MAA.). Suppose that there are 5 sellers in the marketplace, i.e.,

$$S = \{s_i \mid i = 1 \dots 5\}$$

Furthermore, suppose that:

1. A seller can produce goods at the maximum quality of 50.
2. Maximum and delivery of goods are 60 and 20 respectively.
3. In addition, some parameters are applied for buyers as follows:
4. For all buyers, reputable thresholds for quality, price and delivery are equal to 0.5, 0.4, 0.4, while their corresponding disreputable thresholds are -0.8, -0.5 and -0.5, respectively.

Expected values for buyer b on quality, price and delivery are 40, 45 and 5, respectively while weights  $w_q$ ,  $w_p$  and  $w_d$  are 0.65, 0.25 and 0.1, respectively.

If  $\hat{q} \geq q$ , we define  $\mu_{\min\_q}$  in equation (23) equals to 0.05. Also we suppose  $\mu_{\min\_p} = \mu_{\min\_d} = 0.05$ .

If  $\hat{q} < q$ , we get  $\lambda_q = 1.5$  in equation (25). We also define  $\lambda_p$  and  $\lambda_d$  equal to 1.5.

Assume that after some interactions between buyer b and sellers, the reputation rating on quality  $r_q^b(s_i)$  for each seller by buyer are as follows (Table 1):

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r_q^b(s_i)$	0.7	0.75	0.25	0.62	0.3

Table 1: Reputation ratings on quality of different sellers to buyer b

Also the reputation ratings on price and delivery-time of different sellers to buyer b are shown in Table 2 and Table 3 respectively.

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r_p^b(s_i)$	0.3	0.35	0.4	0.1	0.32

Table 2: Reputation ratings on price of different sellers to buyer b

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r_d^b(s_i)$	0.3	0.5	0.3	0.65	0.7

Table 3: Reputation ratings on delivery of different sellers to buyer b

General reputation threshold and general reputation of sellers are computed based on (14) and (16) respectively. General reputation of each seller is shown in Table 4,

$$\Theta^b = 0.65 * 0.5 + 0.25 * 0.4 + 0.1 * 0.4 = 0.465$$

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r^b(s_i)$	0.56	0.625	0.2925	0.493	0.345

Table 4: Reputation ratings on delivery of different sellers to buyer b

So, sellers with general reputation equal or greater than  $\Theta^b = 0.465$ , ( $r^b(s_i) \geq 0.465$ ) are reputable to buyer b, Hence sellers  $s_1$ ,  $s_2$  and  $s_4$  have the chance to be chosen by buyer b in current auction. Also set of reputable sellers updated by buyer b based on (17) as follows:

$$S_r^b = \{s_1, s_2, s_4\} \subseteq S$$

After b's announcement of its request for good g to all sellers which they have good g to sell, the sellers bid with the following specification to deliver g to buyer b have been shown in Table 5. Let triplet bid(quality, price, delivery) be the structure of a bid's specification.

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$bid(q, p, d)$	(47,48.44,7)	(46.2,54,5)	(45.5,52,1)	(48,50,3)	(49,52,2)

Table 5: Bid's offered by different sellers for good g to buyer b

Now buyer b guesses the value of each bid offered by sellers based on equation (19). Results are shown in Table 6.

$s_i$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$G^b(q_{s_i}, p_{s_i}, d_{s_i}, s)$	0.37417	0.3506	0.36983	0.4007	0.4104

Table 6: Bid's offered by different sellers for good g to buyer b

Then buyer b selects the seller  $\hat{s}$  who belongs to the set of reputable sellers for buyer b ( $\{s_1, s_2, s_4\}$ ) whose bid value for buyer b is more than the other sellers' by equation (20). So b buys good g from  $s_4$  with guessed value  $G^b(q_{s_4}, p_{s_4}, d_{s_4}, s_4) = 0.4007$  for its good by b. guessed value for good offered by  $s_5$  is 0.4104, but because

this seller has not served b well in the past auctions and has been known as non reputable seller by b, therefore buyer b does not interact with seller  $s_5$  and selects  $s_4$  as winner of this auction and buys good g from  $s_4$ .

Suppose that after paying, seller  $s_4$  deliver good g to b and b examines the quality of good g and finds  $\hat{q} = 48$  that has been delivered at  $\hat{d} = 4$  (seller  $s_4$  offered delivery time equal to  $d=3$  and deliver its product one unit of delivery late). Buyer b now updates the reputation of seller  $s_4$  on quality by equation (22) and (23) as follows:

$$r_q^b(s) = 0.62 + \frac{48 - 40}{50} * 0.62 = 0.7192$$

Reputation on price and delivery is updated similarly as follows:

$$r_p^b(s) = 0.1 + \frac{43 - 50}{60} * 0.1 = 0.0884$$

$$r_d^b(s) = 0.65 + \frac{5 - 3}{20} * 0.65 = 0.715$$

we see that reputation on price of seller  $s_4$  updated with smaller value than before. It is because of this fact that maximum price expected of b is smaller than price of the good it has purchased. But high quality and good delivery of good g delivered by  $s_4$  ratio to what b expected, increases the value of good g offered by  $s_4$ .

Thus by providing good g with high value, seller  $s_4$  has improved its general reputation to buyer b and increases its chance to be selected again by buyer b in the next auctions and remain in set  $S_r^b$  of reputable sellers to b.

#### 2.4.2 Seller Situation

Consider how a seller in the above-said marketplace, behaves according to the proposed seller algorithms. In this example we investigate behavior of seller  $s_4$  and  $s_1$  in the marketplace. Assume these assumptions:

We define the maximum percent of profit  $\kappa = 0.2$ . Therefore, according to equation (1) if a good costs 47, then the maximum price that seller s can dedicate is equal to 56.4.

We assume reduction percent of price (rp) and discount variable ( $\beta$ ) in equation (5) are equal to 0.015 and 0.05, respectively.

Sellers increase cost and quality of goods in equation (7) with the *inc* rate of 0.04.

Increasing rate of delivery is  $d_{inc} = 1$ .

Reputation of buyer b near seller  $s_1$  and  $s_4$  are 0.3 and 0.32 respectively.

After that buyer b selected  $s_4$  as winner of auction, it sends its announcement to all sellers which they had sent bid to buyer b. Behaviors of sellers  $s_1$  and  $s_4$  after receiving this announcement are as follows:

Seller  $s_4$  calculates the achieved profit from this auction and updates the reputation of buyer b. The cost of production of good g is 48 and we have:

$$\text{Profit} = 50 - 48 = 2$$

Then  $\mu$  is updated based on equation (3) as follows:

$$\mu = \frac{2}{57.6 - 48} = 0.2084$$

And reputation of buyer b is updated by equation (2):

$$r^s(b) = 0.32 + 0.2084 * (1 - 0.32) = 0.461712$$

So in the next auction, seller  $s_4$  consider discount for buyer b based on new reputation of b as follows:

$$p_{new} = 50 - 0.05 * (0.461712 - 0.32) * 50 = 49.661712$$

Therefore, new bid by seller  $s_4$  to buyer b for good g is bid (48, 49.661712, and 3).

Seller  $s_1$  should alter its bid to increase the chance to be selected by buyer b in the next auction. Seller  $s_1$  decreases the price of good g by equation (5):

$$p_{new} = 48.44 - (0.015 * 48.44) - 0.05 * 0.3 * 48.44 = 46.9868$$

As we said in seller algorithm, the price should be offered by a seller can not be smaller than cost of production of the good. Now ( $p_{new} = 46.9868 < c^{s_1}(g, q, b) = 48$ ), therefore seller  $s_1$  does not propose the new price and try to alter the quality of good g. Seller  $s_1$  produce good g with higher quality than before by equation (7) as follows:

$$c^{s_1}(g, q, b) = (1 + 0.04) * 48 = 49.92$$

It means that quality of new good g is 50.4; therefore maximum price for good g with quality 50.4 is calculated by equation (1) as follows:

$$p_{max} = 49.92 + 49.92 * 0.2 = 59.904$$

Seller  $s_1$  calculates the discount for buyer b by considering the reputation of b, i.e.,

$$discount = 0.05 * 0.3 * 59.904 = 0.89856$$

So, price  $p_s$  that should be offered by seller  $s_1$  is:

$$p_s = 59.904 - 0.89856 = 59.00544$$

Therefore, new bid by seller  $s_1$  to buyer b for good g is bid (49.92, 59.00544, and 1). Seller  $s_1$  start bid with maximum expected profit and delivery 1, and then if does not succeed to sell good g reduce price and increase delivery based on equation (5) and (6) respectively.

### 3 Experimental Results

We have implemented the proposed model with Aglets that are java based stationary and mobile agents built in the aglet environment. Our results show that when seller agent models the reputation of buyer agents and dedicates discount to those that are reputable, obtains greater satisfaction compare to the situation when he only alters the quality, price and delivery of his goods. Also buyer agents that follow proposed algorithms are more flexible in different conditions for selecting goods. We have tested our proposed model, both for buyer and seller agents, in extensive experimentation. In parts 3.1 and 3.2 the seller agents satisfaction and buyer agents satisfaction are presented.

#### 3.1 Seller Satisfaction

In the test for evaluating seller algorithm, there are 25 seller and 20 buyer agents in our simulated marketplace. Assume that buyers arrange totally 2000 auctions. Let triplet g(quality, price, delivery) be the structure of a good's specification. All buyer agents use the proposed algorithm in this paper for buyer and seller agents which are divided into five groups:

4. Group A consists of five sellers  $s_0, s_1, \dots, \text{and } s_4$ . These are dishonest sellers on quality who try to attract buyers with high quality goods and then cheat them with really low quality ones. They offer g(48,50, and 2) and then deliver its good as g(38,50, and 2).
5. Group B consists of five sellers  $s_5, s_6, \dots, \text{and } s_9$ . These are dishonest sellers on delivery who try to attract buyers by offering the best delivery along with suitable quality but then cheat them by delivering goods so late. They offer g(48,50, and 2) but deliver their good as g(48,50, and 13).
6. Group C consists of five sellers  $s_{10}, s_{11}, \dots, \text{and } s_{14}$  that do not cheat buyers and use fixed bid for any buyer. They offer and deliver goods as g (40, 44, and 7).
7. Group D consists of five sellers  $s_{15}, s_{16}, \dots, \text{and } s_{19}$  which alter quality, price and delivery of their goods but do not model the reputation of buyers. Moreover, they do not consider discount for buyers. They start their bids as g (38, 45.6, and 1) and then alter their offers based on buyers' requirements.
8. Group E consists of five sellers  $s_{20}, s_{21}, \dots, \text{and } s_{24}$  that in addition to altering the quality, price and delivery of their goods, model the reputation of buyers and also consider discount for them based on their reputation. They start their bids as g(38,45.6, and 1) and then use the proposed algorithms to alter their bids.

In addition, there are other parameters are considered for sellers:

1. Quality is chosen equal to cost to support the common assumption that it costs more to produce high quality goods. That is, a good in quality of 38 costs just 38.
2. We define the maximum percent of profit  $\kappa = 0.2$ . Therefore, according to equation (1) if a good costs 38, then the maximum price that seller s can dedicate is equal to 45.6.

3. We assume reduction percent of price (rp) and discount variable ( $\beta$ ) in equation (4) are equal to 0.015 and 0.05, respectively.
4. Sellers increase cost and quality of goods in equation (5) with the *inc* rate of 0.02.
5. A seller can produce goods at the maximum quality of 50.

All buyers use the buyer agents algorithm proposed in this paper and the parameters that are applied are as the following:

1. For all buyers, reputable thresholds for quality, price and delivery are equal to 0.4, while their corresponding disreputable thresholds are -0.8, -0.5 and -0.5, respectively.
2. Expected values for buyer b on quality, price and delivery are 40, 43 and 8, respectively while weights  $w_q$ ,  $w_p$  and  $w_d$  are 0.65, 0.25 and 0.1, respectively.
3. If  $\hat{q} \geq q$ , we define  $\mu_{\min\_q}$  in equation (23) equals to 0.05. Also we suppose  $\mu_{\min\_p} = \mu_{\min\_d} = 0.05$ .
4. If  $\hat{q} < q$ , we get  $\lambda_q = 1.5$  in equation (25). We also define  $\lambda_p$  and  $\lambda_d$  equal to 1.5.

The results of this experiment confirm that sellers who exploit the proposed algorithms (i.e., group E), achieve better satisfaction than the other sellers. In addition, buyers learn to focus their business on sellers who have reached enough reputation and prevent to interact with disreputable ones. Average and total number of sales made by each of these five groups of sellers is shown in Table 7.

Group	A	B	C	D	E
Total # of sales	100	100	262	427	1111
Average # of sales	20	20	52.4	85.4	222.2

Table 7: Total and average number of sales made by five groups of seller agent

Sellers of groups A and B are dishonest sellers that lie on quality and delivery, respectively. In real markets, it is expected that when buyers purchase from a seller who tries to cheat them, they will not deal with him for their future purchases. Table 7 confirms this matter so that each buyer purchases from dishonest sellers no more than once. There are 20 buyers in the market and each of them was cheated by a dishonest seller once. Therefore each dishonest seller can cheat each buyer one time and totally wins in 20 auctions. Buyers model the reputation of dishonest seller and consider the reputation for the seller lower than disreputable threshold,  $\theta^b$ , as described in equation 15. Actually buyers learn to stay away from disreputable sellers.

Sellers of group C, offer goods in fixed quality, price and delivery. Although they may sell some of their goods in their first deals, but because of the existence of sellers of the other groups who alter their bids to offer goods in high quality, buyers will no longer purchase from sellers of this group, since they can not visit the buyers' requirements. Sellers of group D alter their bids based on buyer requirements and they achieve further sales in comparison to sellers of groups A, B and C.



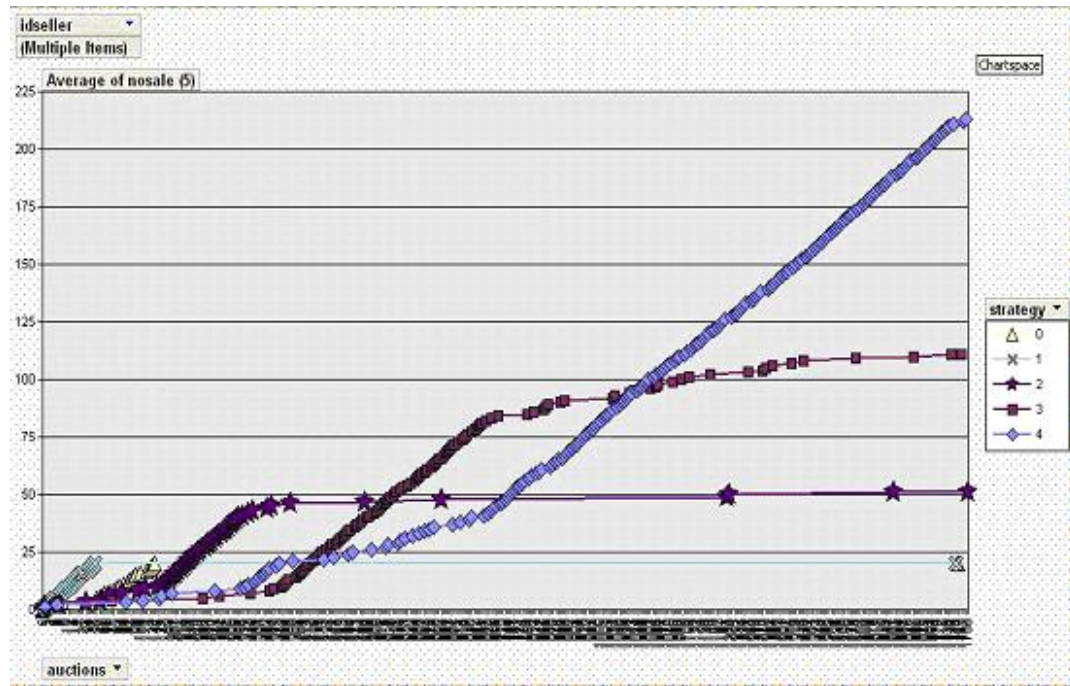


Figure 3: Comparison of sales made by sellers  $s_0$ ,  $s_5$ ,  $s_{10}$ ,  $s_{15}$  and  $s_{20}$  with strategies 0 (Group A), 1 (Group B), 2 (Group C), 3 (Group D), and 4 (Group E), respectively.

In real markets, sellers pay tribute to buyers in order to attract and keep them as their own customers for long time. Discount is one of the important factors that sellers can promote for their own reputable buyers. Sellers of group E, apply this marketing strategy to increase the number of their customers. The results shown in Table 7 confirm this hypothesis. Buyers gradually learn to purchase their required goods from sellers who offer goods in high quality and best delivery while dedicate discounts. In order to investigate the hypothesis mentioned above, in Figure 3 we show the number of sales made by sellers  $s_0$  (from group A),  $s_5$  (from group B),  $s_{10}$  (from group C),  $s_{15}$  (from group D) and  $s_{20}$  (from group E) during 2000 auctions in market. Curves numbered 0 through 4 belong to groups A, B, C, D, and E, respectively.

The results obtained from these five groups show the superiority of our presented model, so that sellers who exploit this model (group E), made sales in an average number equals to 222.2 while the other groups (A, B, C and D) did 20, 20, 52.4 and 85.4, and the average profit of sellers who exploit the proposed model (group E) is equal to 588.2 while the other groups (A, B, C and D) achieved the average profit 200, 180, 209.6 256.7, respectively. Figure 3 shows that dishonest sellers  $s_0$  and  $s_5$ , initially do have good sales in the market. However, number of sales of honest sellers  $s_{10}$  and  $s_{15}$  are smoothly increased over the time. In the long time, seller  $s_{20}$  ultimately outdistances from the other sellers in the market. Because buyers have learned to buy from sellers who honestly offers goods in high quality in addition to discount dedication based on their reputation.

### 3.2 Buyer Satisfaction

In the test for validation buyer algorithm, there are 25 seller and 20 buyer agents in our simulated marketplace assuming that buyers arrange totally 2000 auctions. Seller agents are divided into five groups as described in part 3.1. In this test we have simulated buyer agents into four groups:

1. Group I consist of five buyers  $b_0, b_1, \dots, b_4$ : These buyers do not model the reputation of sellers, at all.
2. Group II consists of five buyers  $b_5, b_6, \dots, b_9$ : These buyers just model the reputation of sellers on quality and avoid interacting with sellers which are not reputable on quality, but they may interact with sellers that lie on delivery-time or sellers who sell their goods so expensive. They do not model the reputation of sellers on delivery-time and price.
3. Group III consists of five buyers  $b_{10}, b_{11}, \dots, b_{14}$ : These buyers just model the reputation of sellers on delivery-time and avoid interacting with non-reputable sellers on delivery-time, but they may interact with sellers that lie on quality or sellers who sell their goods very expensive because they do not model the reputation of sellers on quality and price.



4. Group V consists of five buyers  $b_{15}, b_{16}, \dots, b_{19}$ : These buyers model the reputation of buyers on quality and delivery-time and avoid interacting with non-reputable sellers on both delivery-time and quality.

The other parameters for buyers and sellers are similar to the parameters considered in prior section. The results of this experiment show that buyers who apply the proposed algorithm (i.e., group V) achieve more satisfaction than the other buyers. Table 8 shows that each group of buyers has focused on which group of sellers for doing their trade.

Group of Buyer/Seller	A	B	C	D	E
$b_0$ (Group I)	51	40	2	2	5
$b_5$ (Group II)	5	78	5	4	8
$b_{10}$ (group III)	81	5	5	3	6
$b_{15}$ (Group V)	5	5	8	28	54

Table 8: Number of purchases made from each groups of seller by buyers:  
 $b_0$  (Group I),  $b_5$  (Group II),  $b_{10}$  (Group III) and  $b_{15}$  (Group V)

Table 8 shows that the buyer following the proposed algorithm,  $b_{15}$ , makes more purchases from sellers in group E but fewer purchases from the dishonest sellers, in comparison with the buyer not using a reputation mechanism  $b_0$ , and buyers who just model the reputation of sellers on quality and does not model the reputation of sellers on delivery-time  $b_5$ , and buyers that just model the reputation of sellers on delivery-time and does not model their reputation on quality  $b_{10}$ . We know that sellers of group E make best offers for buyers and are more honest in comparison with the other groups of sellers. So we expect that buyers focus their trades on sellers of group E and then D, in order to obtain more satisfaction. Table 8 shows that buyer  $b_0$  makes 51% of its purchases from sellers of group A, which are dishonest on quality and 40% from sellers of group B that are dishonest on delivery-time. Also  $b_0$  makes just 2%, 2% and 5% of its purchases from groups C, D and E respectively. Purchases from group C, D, and E are done in random. Remember that buyer  $b$  with a probability  $\rho$  chooses to explore (rather than exploit) the marketplace by randomly selecting a seller  $\hat{s}$  from the set of all sellers as described in buyer algorithm. Other sellers of groups D and E alter their bids but in comparison with group A's bid obtain less value because sellers of group A bid to buyer with very high quality and cheat buyers. So if buyer does not model the reputation of seller, it considers very high value for sellers' bid and selects them as winner in auctions much more than once. Buyer  $b_5$  just models the reputation of sellers on quality and avoids interacting with disreputable sellers on quality. Table 8 shows that buyer  $b_5$  makes 5% of its purchases from sellers of group A who are dishonest on quality, it means that  $b_5$  by modeling the reputation of sellers on quality; learn to avoid interacting with sellers who lie about the quality of their goods. But  $b_5$  has made 78% of its purchases from group B. Sellers of group B attract buyers with high quality and very soon delivery-time but they deliver their goods so late. It is expectable that buyers do not interact with disreputable sellers but because  $b_5$  do not model the reputation of sellers on delivery-time, so he was cheated much more than once by sellers of group B. Behavior of  $b_{10}$  is like the behavior of  $b_5$ , but  $b_{10}$  just models the reputation of sellers on delivery-time. Table 8 shows that  $b_{10}$  do 81% of its purchases from sellers of group A, because it does not model the reputation of sellers on quality and was cheated by sellers of group A who lie on quality of their goods. As we said before it is reasonable that buyers focus their trades on sellers of group E and then D, to achieve more satisfaction.  $b_{15}$  models the reputation of sellers on quality and delivery-time and avoids interacting with disreputable sellers.  $b_{15}$  has done 5% of its purchases with sellers of group A and 5% with group B. It means that  $b_{15}$  evaluates the reputation of sellers and avoids interacting with disreputable ones at all.  $b_{15}$  that applies the proposed algorithm for buyers presented in this paper has obtained more satisfaction in comparison with buyers  $b_0$ ,  $b_5$  and  $b_{10}$ .  $b_{15}$  learns to focus its trades on sellers who alter their bids and increase the quality of their goods (group D and E) and in long time learns to focus on sellers which in addition to altering bids and increasing the quality of goods consider

discount for buyers(Group E). Sellers of Group E, as described before, model the reputation of buyers and then dedicate discount for them based on their reputation. So it is expectable that buyers make more trades with sellers of group E.  $b_{15}$  which use the proposed buyer algorithm do 28% and 54% of its purchases with sellers of groups D and E, respectively and make totally 10% of its purchases with disreputable sellers. It is clear that  $b_{15}$  obtains more satisfaction than the other groups of buyers in the market.

## 4 Conclusion and Future Work

In this paper we proposed a marketplace based on reputation and reinforcement learning algorithms for buying and selling agents. Selling agents learn to maximize their expected profits by adjusting product prices, delivery-time and altering the quality of their products and more important considering discount for reputable buyers based on their reputation. We showed that sellers who exploit the proposed algorithms obtain better satisfaction compared to the others. Buyers also learn to purchase from sellers who tribute them by dedicating discounts. We have investigated this fact that marketing and consumer relationship management are two important factors in business, so that sellers who obey this fact construct better reputation for themselves among buyers and get greater profit in comparison to the others. This model is very flexible to develop marketing purposes and modeling a real market completely. However, proposed model and algorithms can be improved so that both sellers and buyers who exploit the improved model can obtain best results as fast as possible. For example, if buyers share their knowledge in cooperation with each other, they will quickly know honest sellers who present best promotion and accordingly will stay away altogether from dishonest sellers. Therefore the profits of those buyers and sellers will quickly increase. On the other hand, seller can learn to offer suitable bid to the new buyers based on the similarity of their preferences compared to the preferences and trading behaviors of previous buyers who have already purchased goods or services from the seller. Our future research aims to provide a set of feasible learning algorithms together with a clear characterization of different situations under which a particular algorithm is preferable. Also for making the effective economic agents and desirable market environments it is attractive to model the reputation of buyers and sellers based fuzzy logic.

## References

- [1] Aglet Research Group of NCKU [Online], Available: <http://turtle.ee.ncku.edu.tw/>.
- [2] J. Ajenstat, and P. Jones, Virtual decision maker for stock market trading as a network of cooperating autonomous intelligent agents, in Proceedings of the 37th Annual Hawaii International Conference on System Sciences, 5-8 Jan. 2004, pp. 34-41.
- [3] O. Bamasak, and N. Zhang, A Distributed Reputation Management Scheme for Mobile Agent-Based E-Commerce Applications, e-Technology, e-Commerce and e-Service, 2005. IEEE '05. Proceedings. The 2005 IEEE International Conference on 29-01 March 2005, pp. 270-275.
- [4] S. Braga, A.P.Araujo, and A.F.R. Wyatt, Incremental topological reinforcement learning agent in non-structured environments, International Conference on Systems, Man and Cybernetics, 10-13 Oct 2004, pp. 5567-5572.
- [5] A. Chavez, and P. Maes, Kasbah: An Agent Marketplace for Buying and Selling Goods. In Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology, 1996.
- [6] T. W. Cheng, W. L. Wang, and A. Chen, E-marketplace using artificial immune system as matchmaker. Proceedings of e-Commerce Technology, 6-9 July 2004, pp. 358-361.
- [7] P. Dasgupta, N. Narasimhan, E. Louise, and P. M. Melliar-Smith, MagNet: Mobile Agents for Networked Electronic Trading. IEEE transaction on Knowledge and Data Engineering, vol. 11, No. 4, pp. 509-525, July/August 1999.
- [8] J.K. Debenham and B. Henderson-Sellers, Designing Agent-Based Process Systems: Extending the OPEN Process Framework, in Intelligent Agent Software Engineering, V. Plekhanova (Ed), (Idea Group), 2003, pp. 160-190.
- [9] G. Derbas, A. Kayssi, H. Artail, and A. Chehab, TRUMMAR - a trust model for mobile agent systems based on reputation. Pervasive Services, 2004. ICPS 2004. Proceedings. The IEEE/ACS International Conference on 19-23 July 2004, pp. 113-120.
- [10] R. B. Doorenbos, O. Etzioni, and D. Weld, A Scalable Comparison-Shopping Agent for the World Wide Web. In Proceedings of the First International Conference on Autonomous Agents, February 1997, pp. 39-48.
- [11] 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, 2001, pp. 27-34.
- [12] I.T. Hawryszkiewicz, Agent Architectures to Support Collaborative Processes, Proceedings of the Fourth International Conference on Web Information Systems Engineering (WISE'03), 2003.
- [13] IBM Aglet Home Page [Online], Available: <http://www.tri.ibm.co.jp/aglets>.
- [14] Y.K. Jong, and J. Lee, Evolution of intelligent agent in auction market, IEEE International Conference on Fuzzy Systems, Vol. 1, 25-29 July 2004, pp. 331-336.
- [15] P.W. Keller, and F. Duguay, RedAgent-2003: an autonomous, market-based supply-chain management agent. Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), 19-23 July 2004, pp. 1182-1189.

- [16] M. L. Littman, Markov Games As Framework for Multi-Agent Reinforcement Learning. In Proceedings of the Eleventh International Conference on Machine Learning, 1994, pp. 157-163.
- [17] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997.
- [18] L. Mui, M. Mohtashemi, and A. Halberstadt, A computational model of trust and reputation. System Sciences, Proceedings of the 35th Annual Hawaii International Conference on 7-10 Jan. 2002, pp. 2431-2439.
- [19] F. Neri, Agent based simulation of information diffusion in a virtual market place. Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 20-24 Sept. 2004, pp. 333-336.
- [20] E. Sanchez, M. Lama, R. Amorim, A. Riera, J. Vila, and S. Barro, A multi-tiered agent-based architecture for a cooperative learning environment, Proceedings of the Eleventh Euromicro Conference on Parallel, Distributed and Network-Based Processing (Euro-PDP'03), 2003.
- [21] T. W. Sandholm, and R. H. Crites, Multi-Agent Reinforcement in the Iterated Prisoner's Dilemma, In Biosystems, 1995, pp. 147-166.
- [22] S. Sen, M. Sekaran, and J. Hale. Learning to Coordinate without Sharing Information. In Proceedings of the Twelfth National Conference on Artificial Intelligence, 1994, pp. 426-431.
- [23] T. Tran, A Reliability Modeling Based Strategy to Avoid Infinite Harm from Dishonest Sellers in Electronic Marketplaces, Published in Journal of Business and Technology (JBT), Special Issue on Business Agents and the Semantic Web, Vol. 1, No.1, pp. 69-76, 2005.
- [24] T. Tran, and R. Cohen, Improving user satisfaction in agent-based electronic marketplaces by reputation modeling and adjustable product quality, Proceedings of Autonomous Agents and Multiagent Systems, 19-23 July 2004, pp. 828-835.
- [25] T. Tran and R. Cohen, The Stability of Electronic Marketplaces Where Buying Agents Use Reliability Modeling to Select Selling Agents, Published in Journal of Business and Technology (JBT), Special Issue on Business Agents and the Semantic Web, Vol. 1, No. 1, pp. 62-68, 2005.
- [26] T. Tran, and R. Cohen, Reputation-Oriented Reinforcement Learning Strategies for Economically-Motivated Agents in Electronic Market Environments. PhD Dissertation, School of Computer Science, University of Waterloo, 2003.
- [27] 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, 1996, pp. 377-384.
- [28] G. Weiss, A Modern Approach to Distributed Artificial Intelligence. The MIT Press Cambridge, 1999.
- [29] 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, 1998, pp. 301-308.
- [30] 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, 2000, pp. 154-165.