An Incentive Compatible Reputation Mechanism

Radu Jurca and Boi Faltings
Artificial Intelligence Laboratory (LIA),
Computer Science Department, Swiss Federal Institute of Technology (EPFL)
CH-1015 Ecublens, Switzerland
{radu.jurca, boi.faltings}@epfl.ch
http://liawww.epfl.ch/

Abstract

Traditional centralised approaches to security are difficult to apply to large, distributed marketplaces in which software agents operate. Developing a notion of trust that is based on the reputation of agents can provide a softer notion of security that is sufficient for many multi-agent applications.

In this paper, we address the issue of incentive-compatibility (i.e. how to make it optimal for agents to share reputation information truthfully), by introducing a side-payment scheme, organised through a set of broker agents, that makes it rational for software agents to truthfully share the reputation information they have acquired in their past experience. We also show how to use a cryptographic mechanism to protect the integrity of reputation information and to achieve a tight bounding between the identity and reputation of an agent.

1. Introduction

Software agents are a new and promising paradigm for open, distributed marketplaces. However, besides the many practical solutions this new paradigm provides, it also brings along a whole new set of unsolved questions. One of the issues that has attracted a lot of attention lately is security. Traditional, centralised approaches to security do no longer cope with the challenges arising from an open environment with distributed ownership in which agents interoperate. [16, 13, 14]

We focus in particular on the problem of trust, i.e. deciding whether another agent encountered in the network can be trusted, for example in a business transaction. In closed environments, trust is usually managed by authentication schemes that define what agents are to be trusted for a particular transaction. In an open environment, fixed classifications must be replaced by dynamic decisions. One important factor in such decisions is an agent's *reputation*, defined as information about its past behaviour.

The most reliable reputation information can be derived from an agent's own experience. However, much more data becomes available when reputation information is shared among an agent community. Such mechanisms have been proposed and also practically implemented. The various rating services on the Internet are examples of such mechanisms.

It is however not at all clear that it is in the best interest of an agent to truthfully report reputation information:

- by reporting any reputation information, it provides a competitive advantage to others, so it is not in its interest to report anything at all.

- by reporting positive ratings, the agent slightly decreases its own reputation with respect to the average of other agents, and also might create more demand for a scarce resource; so it is a disadvantage to report them truthfully.

- by reporting fake negative ratings, the agent can increase its own reputation with respect to others, and also can reduce the demand for a scarce resource; so it is an advantage to report them falsely.

Thus, it is interesting to consider how to make a reputation mechanism *incentive-compatible*, i.e. how to ensure that it is in the best interest of a rational agent to actually report reputation information truthfully. This is the problem we address in this research.

Section 2 describes the assumed model of the environment, Section 3 presents an example of a reputation mechanism that is incentive-compatible and some of its properties. Section 4 describes the implementation of the mechanism coupled with an identity mechanism that offers some guarantees of non-manipulation. Section 5 presents the simu-

	C	D
C	R,R	S,T
D	T,S	P,P

Figure 1. Payoff matrix for players in the Prisoners' Dilemma game.

lation results of our mechanism. Finally, related work is given, and a conclusion is drawn.

2. The Model

From the considerations given above, it is clear that an incentive-compatible mechanism should introduce side payments that make it rational for agents to share reputation information. Moreover, even if the action of reporting reputation information does not introduce negative consequences for an agent, without a side payment scheme a rational agent will be indifferent between reporting true or false information.

In our mechanism, these side payments are organised through a set of broker agents, called R-agents, that buy and sell reputation information.

The scenario is the following. We assume we have N rational agents: a_i for $i = 1 \ldots N$, that interact pairwise in an iterated Prisoner's Dilemma (IPD) environment [2]. The payoffs for an agent (Figure 1) are known as: temptation to defect (T), reward for cooperation (R), punishment for defection (P), and sucker's payoff (S). The following constraints on the payoffs hold: $T > R > P > S$ and $2R > T + S$. The game is an abstraction of social situations where an agent has to take one of the following two actions: *cooperate* (C) i.e. doing the socially responsible thing, or *defect* (D) i.e. maximise its immediate payoff regardless of how harmful this might be for the partner. The payoffs are such that each agent is better off defecting (or *cheating*) regardless of the opponent's choice, but the social outcome (the sum of the agents' payoffs) is maximised when both agents choose to cooperate. Thus the dilemma. We believe the IPD to be a representative model for many present electronic business interactions as traditional legal enforcement is impossible or very costly to implement (e.g. electronic auctions, peer-to-peer exchange systems).

The players are selected randomly for each game, with uniform probability. Before the game begins, agents can choose whether or not to play the game with the chosen partner. Pregame contracts (contracts negotiated before the game begins) are allowed but they are not binding (there is no central authority to enforce the contract). No transfer of payoff is possible between agents.

The behaviour of an agent is influenced by an a priori type (defined by a fixed probability of cooperation) and a time variant component which depends on the last k actions of the agent. The a priori type can be interpreted as innate (genetic) information about the agent behaviour, while the time variant component is the present "mood" of an agent with "memory" of length k.

Definition 1 *An agent A behaves according to a dynamic type with memory k, if there is an a priori probability of cooperation p and a conditional probability of cooperation depending on the sequence $a_{t-1}a_{t-2} \ldots a_{t-k}$ of last k actions took by the agent, such that the probability of agent A cooperating at time t is defined as:*

$$P_t(C) = \Gamma(p, P(C|a_{t-1}a_{t-2} \ldots a_{t-k})); \quad (1)$$

where $\Gamma : [0,1] \times [0,1] \to [0,1]$ is some function.

Each agent can buy reputation information about another agent from an R-agent at a cost F, and later sell reputation information to the same R-agent at a price F'. R-agents scale F' and F such that the mechanism breaks even in the long run. Agents report either 0 for a defection (D) or 1 for cooperation (C).

Reputation information about an agent is represented by the M last reports submitted about that agent, where M is some integer. This choice of taking into consideration only the last M reports is motivated by both game theoretic results (in [8] it is proven that infinite monitoring cannot induce cooperation in games where mistakes or errors can occur) and empirical studies of the eBay[1] reputation system (a survey [12] of empirical studies on the eBay feedback mechanism draws the conclusion that only recent ratings influence buyer behaviour).

In our scenario, agents systematically buy reputation information before playing the game and they are only allowed to sell a reputation report for another agent when they have previously bought reputation information for that agent.

To conclude, the interaction protocol is the following: As two agents (A and B) are randomly selected to play the game, each of them will select an R-agent whom to ask reputation information. The algorithm for selecting the R-agent is discussed in detail in Section 4.2. Agents ask their selected R-agent the reputation of their partner, and pay for that information. The currency used for reputation payments is different from the one used for the game payoffs and there is no exchange possible between the two. An agent that loses all its reputation money can no longer use the reputation service.

After knowing the reputation of their partners, agents can decide whether or not to play the game. If both agents decide to play the game, they enter in a pregame contracting phase where a contract can be negotiated. If both agents are

---

[1] www.ebay.com

satisfied with the contract, they will actually play the game, and receive the corresponding payoffs. From the payoffs agents can exactly determine the behaviour of their partner, and can submit a report to the selected R-agent. Based on the outcome of the game agents also update their view on the utility of different R-agents. (Section 4.2).

3 Incentive Compatible Reputation Reporting

Let us look in more detail at payment functions that can elicit honest reputation information reporting from agents. As there is no central authority to provide irrefutable information, the payment function can be based only on the reports of other agents. Since we cannot assume that past information can be kept secret from the reporting agent, the payment function can only dependent on future, unknown reports.

Let $\pi : \{0,1\} \times \Psi \rightarrow \mathbb{R}$ be such a function. $\Psi = \{\{0,1\}^K | K \in \mathbb{N}\}$ is the set of all possible sets containing K reputation reports (0 or 1) and $\pi(s, S)$ is the payment an agent A obtains when submitting a report $s \in \{0,1\}$ about B and the set S of reputation reports will be observed about the same agent B. π is eliciting truthful reporting if it is maximised when the agents tells the truth, i.e. at time t, $\pi(C, S^C) > \pi(D, S^C)$ if a C is observed at t and S^C is the set of future reports submitted by other agents conditional on the fact that C occurred at t, and similarly $\pi(D, S^D) > \pi(C, S^D)$ if a D is observed at t and S^D is the set of future reports submitted by other agents conditional on the fact that D occurred at t.

Theorem 1 *For agents of dynamic type (Definition 1) if the function Γ (Equation (1)) does not depend on the agent's previous actions, there is no payment function that can be used by a reputation system in order to elicit honest behaviour.*

PROOF. If Γ does not depend on the previous actions of the agent, the probability of cooperation of an agent A at time t is:

$$
\begin{aligned}
P_A^t(C) &= \Gamma(p, P(C|a_{t-1}a_{t-2}\ldots a_{t-k})) \\
&= \Gamma(p_A) = const.
\end{aligned}
$$

for all $t \in \mathbb{N}$, where p_A is the innate probability of cooperation of agent A. As $P_A^t(C)$ does not change over time, eventually the reputation information r_A about agent A will accurately predict $\Gamma(p_A)$. Because the choice of future actions does not depend on the action chosen in time t, if the reporting agent knows $\Gamma(p_A)$, it will be able to predict the set $S = S^C = S^D$ of future reports about A that is going to be considered for paying the present announcement. If

S is known, there is no π such that: $\pi(C, S) > \pi(D, S)$ (when a C is observed), and $\pi(D, S) > \pi(C, S)$ (when a D is observed). $\square$

The results of Theorem 1 are quite striking since many existing reputation systems assume an agent behaviour influenced only by an a priori type [3, 4, 17, 19, 22].

We propose a simple payment scheme that makes it rational for agents to truthfully share the reputation information. The basic idea is that : R-agents pay the reputation report of agent A about agent B only if it matches the next report submitted about B. The payment function which implements this rule is:

$$
\pi(s, S) = \begin{cases} F' & \text{if } S = \{s\} \\ 0 & \text{otherwise} \end{cases} \tag{2}
$$

where S is the set containing only the next report submitted about B.

Theorem 2 *If an agent is of dynamic type (Definition 1), when the function Γ is not known, the payment function presented in Equation (2) enforces truthful reputation reporting as a Nash equilibrium. (i.e. if the other agents tell the truth, it is in the best interest of any agent to also report the truth).*

PROOF. As agent A does not have information about the behaviour of the partner agent B (the Γ function is not known), for A the behaviour of B is independently identically distributed for the present and following game (i.e. B cooperates with some probability $\overline{p}$). The probability that agent B chooses the same action in two consecutive games is: $(1 - \overline{p})^2 + \overline{p}^2 = 1 - 2\overline{p} + 2\overline{p}^2$ which is bounded by $[0.5, 1]$. On the other hand, the probability that agent B will change its behaviour in two consecutive games is: $2\overline{p}(1-\overline{p})$, which is bounded by $[0,0.5]$. Assuming that the next agent to report about B will report the truth and that there is no collusion among the agents, the best strategy for agent A is to report the behaviour of B truthfully, since this means its report will be paid by the R-agent with probability of at least 0.5. $\square$

As a direct consequence of Theorem 2, the payment function in Equation (2) can be safely used from the very first interaction within the system. Its simplicity makes it easy to implement and be understood by the actors in the system.

However, for π (as defined by Equation (2)) to supply stable truth-telling incentives in the long run supplementary constraints have to be imposed on the parameters of the agents behavioural model.

Theorem 3 *If $Pr(C_{t+1}|C_t) = \Gamma(p, P(C|C)) > 0.5$ and $Pr(D_{t+1}|D_t) = (1 - \Gamma(p, P(C|D)) > 0.5$, where: $P(C|C) = \sum_{a_{t-2}\ldots a_{t-k}} P(C|Ca_{t-2}\ldots a_{t-k})$;*

$P(C|D) = \sum_{a_{t-2}...a_{t-k}} P(C|Da_{t-2}...a_{t-k})$; *the payment function defined by Equation (2) induces truthful reporting.*

PROOF. Having observed a C from agent B, A can report 0 (lie) or 1 (tell the truth). The expected payoff:

$$\begin{aligned}
E[\pi(1,S)] &= Pr_B(C_{t+1}|C_t)\pi(1,\{1\}) + \\
&\quad Pr_B(D_{t+1}|C_t)\pi(1,\{0\}) \\
&> Pr_B(D_{t+1}|C_t)\pi(0,\{0\}) + \\
&\quad Pr_B(C_{t+1}|C_t)\pi(0,\{1\}) \\
&= E[\pi(0,S)];
\end{aligned}$$

Therefore A is better off by reporting the truth. Similarly, having observed a D the agent A is better off by reporting the truth. □

There are results [17] that require less strong constraints on the agent behaviour model. However, the payment scheme in [17] relies on accurately knowing the parameters of the model which is not possible in the starting period of the system, or if the system is dynamic. Our method can accommodate both these situations at a cost of supplementary constraints on the parameters of the behaviour model.

3.1. Robustness against lying agents

In real applications, not every agent will behave rationally and follow the equilibrium described in Theorem 2. It is reasonable to assume that a certain percentage of the agents present in the environment will not always tell the truth. We will further call this type of irrational agents that lie with a certain probability *lying agents*. Clearly, an increasing number of lying agents will affect the incentive-compatibility property of the mechanism, as it can no longer be guaranteed that truthful reports are paid with higher probability. In the remains of this section we would like to derive a theoretical threshold for the percentage of lying agents for which the property of incentive-compatibility still holds.

Let us consider that α percent of the agents within the system are lying agents that lie with probability β. The probability that the next report submitted about an agent A is a true report is: $1 - \alpha\beta$. Let g be the fixed point solution of the Γ function, i.e. $\Gamma(p,g) = g$. For an infinite run, the probability of cooperation for an agent of dynamic type is g.

A true reputation report filed by an agent A about agent B is paid by the R-agent only if the next report about agent B is also true and agent B adopted the same behaviour in the next game, or if the next report about B is false, but B changed its behaviour in the next game. The probability that a true report is paid is thus:

$$E_T = (1-\alpha\beta)(1-2g+2g^2) + 2\alpha\beta(g-g^2);$$

$E_T \cdot F'$ represents the expected payoff for a truthful reputation report.

On the other hand, a false report about agent B is being paid by the R-agents if the next report about B is true but B changed its behaviour in the next game, or if the next report about B is false but B displayed the same behaviour in the next game. The probability that a false report is paid is thus:

$$E_F = (1-\alpha\beta)2(g-g^2) + \alpha\beta(1-2g+2g^2);$$

$E_F \cdot F'$ represents the expected payoff for a false reputation report.

As long as $E_T > E_F$ rational agents will have the incentive to report the truth. Relation (3) gives an upper bound for the percentage of false reports (the product $\alpha\beta$) for which the property of incentive-compatibility is guaranteed.

$$E_{true} > E_{false} \Leftrightarrow \alpha\beta < 0.5 \qquad (3)$$

Because false reports are on the average paid less than true reports, agents that submit false reports will gradually loose their reputation money and will be prevented from using the reputation mechanism at all. As a consequence, the percentage of false reports submitted will converge to 0 for an infinite run. Evidence of that can be seen in Section 5.

4. Mechanism Implementation

We are implementing the above described mechanism as part of a *reputation service* deployed on the Agentcities network [1]. As opposed to other reputation mechanism implementations [3, 4, 21] we designed our reputation service independent of the application domain. Our aim is to provide a flexible service that can be deployed in a real world multi-agent environment (such as the Agentcities network) where different applications can use the service. In order to achieve this goal, the service has to provide some security guarantees that make it trustworthy enough to be used by real world applications. In the remains of this section we will address two security issues: (1) identity (i.e. agents cannot impersonate other agents and steal their reputation) and (2) integrity (i.e the reputation of one agent cannot be modified without the agent's consent). The interaction protocols and the local logic of the agents are also described in this section.

4.1. Security Issues

When designing the reputation service, we had the following security goals in mind:

1. Agents should not be able to tamper with agent A's reputation without A's consent (i.e. A played the game and agreed that its partner can file reputation reports at the end of the game);

2. An agent should be able to read the reputation r_A of agent A without being able to modify it;

3. Agents should be identifiable. (i.e. agent A should be able to prove that reputation r_A really refers to itself).

We propose the following solution. Each agent has only one pair of keys (sk, pk), sk being secret and pk being public. By $E(k, m)$ we denote the encryption of message m with the key k. By $D(k, n)$ we denote the decryption of the encrypted message n using the key k. The following relations hold: $D(sk, E(pk, m)) = m$ and $D(pk, E(sk, m)) = m$.

The reputation information r_A about an agent A is kept by an R-agent as a set of the last M reports submitted about agent A to that R-agent: $r_A = \{(report_i, t_i) | i = 1, \ldots, M\}$ where t_i is a time-stamp that uniquely identifies each report. By (r_A++) we denote the set of reports about agent A updated with one more positive report. Similarly, by (r_A--) we denote the set of reports about agent A updated with one more negative report.

In order to achieve the security goals mentioned above, R-agents store $E(r_A, sk_A)$, r_A encrypted with sk_A, the secret key of agent A. As part of the pregame contract agents ask from their partners signed versions of $(r++)$ and $(r--)$. If agents A and B are going to play the game, as part of the pregame contract A will have $E((r_B++), sk_B)$ and $E((r_B--), sk_B)$ while agent B will have $E((r_A++), sk_A)$ and $E((r_A--), sk_A)$. At the end of the game, agents are free to file the positive or the negative report, by sending one of $E((r++), sk)$ or $E((r--), sk)$ to the R-agent. Before accepting the report, the R-agent checks the validity of the report (verifies the signature) and the integrity of the time-stamps (the oldest report has been replaced by a report with a newer time-stamp). If both checks are successful, the R-agent replaces the old reputation information by the new one. R-agents accept reports only from agents who have previously bought reputation information from them, and do not accept reports submitted by agents about themselves.

Assuming there is no collusion among the agents in the system (either between any two agents or between agents and R-agents), the above described mechanism has the following properties:

Lemma 1 *The reputation of an agent is uniquely tied to its identity. (No agent A can "steal" the reputation of agent B)*

PROOF. Let us suppose that agent A tries to impersonate agent B in order to take advantage of the higher reputation of the latter while playing the game with agent C. When interrogating an R-agent about agent's A reputation, agent C will get $E(r_B, sk_B)$. Agent A's fraud can easily be observed by asking a signature sample (for example the signature of a randomly generated number n) from A. As A does

not have access to the secret key sk_B of agent B (which is a reasonable assumption), C will be able to detect that $E(r_b, sk_B)$ is forged. $\square$

Lemma 2 *No agent can by its own will falsely increase or decrease its reputation.*

PROOF. The reputation of agent A can only be modified when an R-agent accepts a value $E(r_A^{new}, sk_A)$. Any R-agent will accept such a report only if it is filed by another agent B who has previously bought information about A and probably played the game with A. Since there is no collusion between A and B (our assumption) there is no way agent A can increase its reputation falsely. $\square$

Lemma 3 *No agent can tamper with agent A's reputation without A's consent*

PROOF. Since agent A is the only agent who can produce $E(r_A, sk_A)$ (assumption of secrecy of private key sk_A) no modification of A's reputation can be made without the explicit consent of A (manifested through the delivery of $E(r_A, sk_A)$)

NOTE. The R-agents can still manipulate the reputation mechanism by delivering an old value for the reputation of an agent. However, they cannot report just any value, but only a value which was true at some moment in time. Besides, malicious R-agents can be detected by the agents who get asked to sign older versions of their reputation. $\square$

The three lemmas above provide some minimum non-manipulation guarantees that make the reputation service trustworthy.

4.2. Selection of R-agents

In the mechanism we have described in Section 3, a set of broker agents (called R-agents) buy and sell reputation information. Typically there will be more R-agents present in the system, and each of them will have its own view on the reputation of the agents that play the game. R-agents centralise reputation information (which speeds up the process of building the information in a system where a large number of agents play the game) but there are no synchronisation requirements among different R-agents. Some R-agents may possess more accurate information than others. It is therefore important that agents learn to recognise the R-agents that have qualitative information and develop a trust model for the R-agents themselves. This is also a measure of robustness, since it will protect agents (to some extent) against malicious R-agents.

We implemented a non-stationary reinforcement algorithm [20] that allows agents to learn which R-agents provide useful information. The algorithm works by estimating the values of different actions an agent might perform. $Q(a)$ is defined as the expected value of choosing action a. Once

the values $Q(a)$ have been learned, the optimal action to be chosen is the one with the highest Q-value. In our case, action a corresponds to choosing R-agent R_a as reputation information provider.

After being initialised to arbitrary numbers, Q-values are updated on the basis of experience as follows:

$$Q(a) \leftarrow Q(a) + \delta[reward_a - Q(a)]$$

where $\delta, 0 < \delta \leq 1$, is a constant called *learning rate* and $reward_a$ is the payoff obtained for choosing action a. In our case, we define the following payoffs:

$$reward = \begin{cases} 1 & \text{if partner cooperated and was recommended as trustworthy by the R-agent;} \\ 0 & \text{if partner defected and was recommended as trustworthy by the R-agent;} \\ 0.5 & \text{if partner was recommended as not trustworthy and consequently the game was not played;} \end{cases}$$

For selecting the action to be taken in the present step, we used an ϵ-*greedy* selection rule. The action with the highest Q-value is selected with $(1 - \epsilon)$ probability, while in $\epsilon\%$ of the cases, another action is chosen at random. This ensures *exploration* of all the action space, while still *exploiting* the best action. The set of Q-values each agent possesses about the R-agents in the system is also a *direct, interaction-derived, reputation mechanism* [18] that takes into account only the previous experience of the agents.

5. Experimental Results

We built a simulation in order to test the performance of the reputation mechanism. We deployed five thousand player agents and ten R-agents into the system. All agents were of dynamic type with memory of length one characterised by a linear function:

$$\Gamma(p, P(C|a)) = w_1 p + (1 - w_1)P(C|a)$$

with the following parameters: $P(C|C) = 0.9, P(C|D) = 0.3, w_1 = 0.2$ and $p = 0.8$.

The payoffs for the game (Figure 1) had the following values: $T = 5, R = 3, P = 1, S = 0$. The game was played on average one thousand times by each agent. For each game, the two agents were selected randomly with equal probability.

Agents chose their R-agents according to the procedure described in Section 4.2. The learning rate, δ of the reinforcement algorithm was set to 0.9 and the exploration factor ϵ of the ϵ-*greedy* action selection rule was set to 5%.
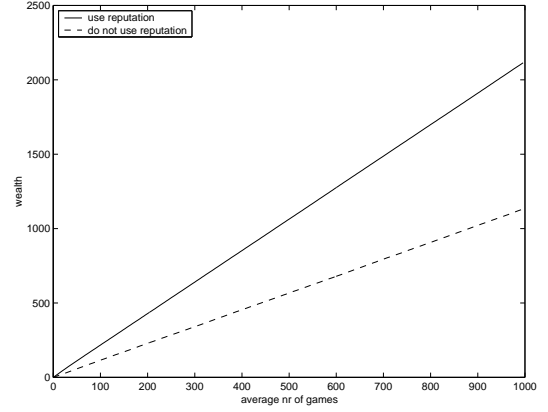

Figure 2. Average fitness of cooperative agents that use and of cooperative agents that do not use the reputation service

Experiments show that approximately 40% of the non-cooperative interactions (interaction in which one of the agents defects) are eliminated as a consequence of the use of the reputation mechanism. This is quite strong evidence of the utility of the reputation mechanism.

Figure 2 plots comparatively the average wealth of the agents that use the reputation mechanism and of those that do not use the reputation mechanism. Clearly, the agents that use reputation information are better off then agents that do not use reputation information.

The incentive compatibility property of the mechanism can be seen in Figure 3. We introduced in the environment 10% of lying agents (agents that do not act rationally and do not always tell the truth) that lie all the time ($\beta = 1$), and another 10% of lying agents that lie with probability $\beta = 50\%$. Figure 3 plots the average reputation money of lying and truthful agents. The points marked on the diagram corresponds to the moment when lying agents lost their reputation money. From that point on, lying agents would be prevented from using the reputation service.

6. Related Work

In [16] the authors present a definition of trust by identifying its constructs: *trust-related behaviour, trusting intentions, trusting beliefs, institution-based trust* and *disposition to trust*. In the present paper, we present a simple trust model that uses only the *trusting beliefs* construct (i.e. the extent to which one believes that the other person has characteristics beneficial to one) under the name of *reputation*. For simplicity, the four different aspects of reputation (competence, benevolence, integrity and predictability) were combined into one number.
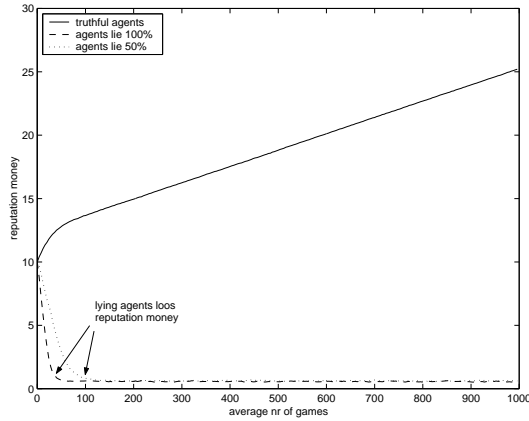
Figure 3. Average reputation money for lying and truthful agents

Mui et al. [18] present an extensive reputation typology classified by the means of collecting the reputation information. Our trust model employs two of the categories: the *direct interaction-derived reputation* and the *propagated (from other agents) indirect reputation.*

There are a number of systems that implement trust mechanisms based only on direct interaction-derived reputation: [3, 4, 15, 21, 5]. All these systems work in an environment with a relatively small number of agents where direct reputation can be build. They are however not appropriate for a very large environment because the time necessary to build direct reputation would be too large. All these reputation mechanisms are also application dependent and cannot be deployed in an open environment such as a multi-agent system.

[19, 22] propose solutions that take into account the reputation information reported by other agents. However, these solutions we believe not to be realistic as they do not provide any incentive for the agents to report the reputation information. Besides, each agent has to implement a rather complicated mechanism for judging the information it has received from its peers.

[9] and [10] describe methods for protecting the reputation mechanism from unfair ratings. The approach the author takes is centralised, and relies on a central authority to ensure the safety properties of the mechanism. The same author [11] studies theoretical properties of eBay-like online reputation mechanism without taking into account though the problem of incentive-compatibility.

[7] and [6] present very interesting results for incentive compatible trading mechanisms for which the seller is not completely trustworthy. By scaling the amount of the traded product the authors prove that it is possible to make it rational for sellers to truthfully declare their trustworthiness.

Truthful declaration of one's trustworthiness eliminates the need of reputation mechanisms and significantly reduces the cost of trust management. However, the assumptions the authors make about the trading environment (i.e. the form of the cost function and the selling price which is supposed to be smaller than the marginal cost) are not common in most electronic markets. The solution might apply to specific domains, such as CPU or bandwidth allocation, and points out a very interesting possibility.

A significant contribution towards eliciting honest behaviour reporting is made in [17]. The authors propose scoring rules as payment functions which induce rational honest reporting. The scoring rules however, cannot be implemented without accurately knowing the parameters of the agents' behaviour model, which can be a real problem in real-world systems. Our mechanism provides an easily implementable payment function that achieves the same incentive-compatible properties.

7. Conclusions and Future Work

In our work, we built a successful incentive compatible reputation mechanism that works in an environment where a big number of agents play an IPD game. The mechanism was implemented for an open multi-agent environment (such as the Agentcities platform) and provides security guarantees that make it usable in a real life application. A cryptographic mechanism was used to ensure the integrity of reputation information and a tight bounding between the reputation of an agent and its identity. We used both direct interaction-derived reputation and propagated indirect reputation in order to speed up the information building phase.

There are still, however, a number of issues that need to be addressed in our future work. One of them is the startup of the system, as there is no incentive to perform the first purchase of reputation when it is null. This can be solved if R-agents will accept (and pay) the first few reports about each agent without requiring that information be bought in advance by the reporter. This involves an initial investment on behalf of the R-agents, which will later be recuperated through the price setting mechanism.

As regarding the implementation of the mechanism, in our current version, agents can be involved in just one game at a particular moment in time. This shortcoming is enforced by the cryptographic mechanism used for ensuring the coupling between the agent's identity and its reputation.

Let us suppose that both agents C and B would like to play the game with agent A, and let us suppose that both C and B choose R-agent R_A as their reputation information provider. Both C and B will get $E(r_A, sk_A)$ from R-agent R_A as the reputation of A. As part of the pregame contract, both C and B will get from A the same $E(r_A + +, sk_A)$ and $E(r_A - -, sk_A)$. When C and B

will try to submit their reports, only one (the first submitted) will be accepted by the R-agent, while the second will be ignored. This makes R-agents lose information, but also perturbs the reputation payment scheme (the second report is not considered for payment). As a possible solution, we will consider making agents submit not only $\{E(r++,sk),E(r--,sk)\}$ as part of the pregame contract, but $\{E(r\pm\pm,sk),E(r\pm2,sk),\dots,(E(r\pm p),sk)\}$, where p is the maxim number of simultaneous games an agent is allowed to play.

Another direction for future research is the problem of collusion. As most of the existing reputation mechanisms, our implementation is vulnerable to the collusion of even two agents. Any agent can falsely increase its own reputation by colluding with just one partner, and making it submit fake reports.

Finally, a future version of our mechanism will have to scale the update of the reputation with the total value of the transaction. Nothing stops at this moment an agent from building a very high reputation by cooperating in many small transactions and then cheat in a very large transaction. For the moment, our environment has uniform transactions (the payoff matrix remains the same for all agents and for every game); however, a real life reputation mechanism will have to also work with agents being involved in transactions of varying values.

References

[1] http://www.lausanne.agentcities.net/reputationservice.

[2] R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[3] A. Birk. Boosting Cooperation by Evolving Trust. *Applied Artificial Intelligence*, 14:769–784, 2000.

[4] A. Birk. Learning to Trust. In M. S. R. Falcone and Y.-H. Tan, editors, *Trust in Cyber-societies*, volume LNAI 2246, pages 133–144. Springer-Verlag, Berlin Heidelberg, 2001.

[5] A. Biswas, S. Sen, and S. Debnath. Limiting Deception in a Group of Social Agents. *Applied Artificial Intelligence*, 14:785–797, 2000.

[6] S. Braynov. Incentive Compatible Trading Mechanism for Trust Revelation. In *Proceeding of the IJCAI'01 Workshop Economic Agents, Models and Mechanisms*, Seattle, 2001.

[7] S. Braynov and T. Sandholm. Incentive Compatible Mechanism for Trust Revelation. In *Proceedings of the AAMAS*, Bologna, Italy, 2002.

[8] M. Cripps, G. Mailath, and L. Samuelson. Reputation in Perturbed Repeated Games. Working Paper, 2002.

[9] C. Dellarocas. Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behaviour. In *Proceedings of the 2nd ACM conference on Electronic Commerce*, Minneapolis, MN, 2000.

[10] C. Dellarocas. The Design of Reliable Trust Management Systems for Electronic Trading Communities. Working Paper, MIT, 2001.

[11] C. Dellarocas. Efficiency and Robustness of eBay-like Online Reputation Mechanisms in Environments with Moral Hazard. Center for eBusiness Working Paper #170, MIT, 2002.

[12] C. Dellarocas. The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms. MIT Working Paper, 2002.

[13] L. Kagal, T. Finin, and J. Anupam. Moving from Security to Distributed Trust in Ubiquitos Computing Environment. *IEEE Computer*, December 2001.

[14] L. Kagal, T. Finin, and J. Anupam. A Delegation-based Distributed Model for Multi Agent Systems System. http://www.csee.umbc.edu/~finin/papers/aa02, 2002.

[15] R. Kramer. Trust Rules for Trust Dilemmas: How Decision Makers Think and Act in the Shadow of Doubt. In M. S. R. Falcone and Y.-H. Tan, editors, *Trust in Cyber-societies*, volume LNAI 2246, pages 9–26. Springer-Verlag, Berlin Heidelberg, 2001.

[16] H. McKnight and N. Chervany. Trust and Distrust: One Bite at a Time. In M. S. R. Falcone and Y.-H. Tan, editors, *Trust in Cyber-societies*, volume LNAI 2246, pages 27–54. Springer-Verlag, Berlin Heidelberg, 2001.

[17] N. Miller, P. Resnick, and R. Zeckhauser. Eliciting Honest Feedback in Electronic Markets. Working Paper, 2002.

[18] L. Mui, A. Halberstadt, and M. Mohtashemi. Notions of Reputation in Multi-Agents Systems:A Review. In *Proceedings of the AAMAS*, Bologna, Italy, 2002.

[19] M. Schillo, P. Funk, and M. Rovatsos. Using Trust for Detecting Deceitful Agents in Artificial Societies. *Applied Artificial Intelligence*, 14:825–848, 2000.

[20] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, Massachusetts, 1998.

[21] M. Witkowski, A. Artikis, and J. Pitt. Experiments in building Experiential Trust in a Society of Objective-Trust Based Agents. In M. S. R. Falcone and Y.-H. Tan, editors, *Trust in Cyber-societies*, volume LNAI 2246, pages 111–132. Springer-Verlag, Berlin Heidelberg, 2001.

[22] B. Yu and M. Singh. An Evidential Model of Distributed Reputation Management. In *Proceedings of the AAMAS*, Bologna, Italy, 2002.