Inteligencia Artificial 42 (2009) 3-11
doi: 10.4114/ia.v13i42.1041

INTELIGENCIA ARTIFICIAL

http://erevista.aepia.org/

A Trust and Reputation Model as Adaptive Mechanism for Multi-Agent Systems

Alberto Caballero, Teresa Garcia-Valverde, Juan A. Botia and Antonio Gomez-Skarmeta

Abstract Multi-agent systems (MAS) are complex software systems. Usually, this kind of software is situated in a constantly changing environment. In such conditions, the capability of adapting to the environment and to the rest of agents is crucial for each individual in the society. In this paper, we make a proposal to reach adaptiveness in MAS through using the notions of trust and reputation. We present the formal model and empirical evidence of its effectiveness.

1 Introduction

The theory of adaptivity has grown out of a variety of disciplines, including biology, psychology and artificial intelligence. Adaptation is usually understood as a means which generates better-performing solutions through interaction with the environment [7]. Thus, an adaptive system is a system that is able to adapt its behaviour according to changes in the environment.

Research on both adaptation and emergence of biological systems caused the new research discipline of Complex Adaptive System (CAS) to appear. A CAS consists of a network of agents interacting, where each agent makes decisions and acts individually, competing and cooperating among themselves. The overall behaviour of the system is the result of the behaviour of the individual agents [6]. In this sense, it is said that an individual agent is adaptive if it becomes better at achieving its goals with experience [9]. Thus, we can say that an adaptive MAS (AMAS) is compound by agents that are adaptive and seek to maximize some utility by evolving, increasing this measure over time [4]. Guessoum [5] defines an AMAS as an open system that continuously and dynamically self-modifies its structure.

It is possible to model simple agents which derive their emergent behaviour from the system as a whole. Notice that a system decomposed into a number of parts which interact in a common environment is a feature that allows the system adapt itself in a changing environment [5]. However, specifying complex adaptive systems with the multi-agent metaphore for modeling is not an easy task and it remains as an open issue. This is the reason why in the last years, the interest on adaptive MAS (AMAS) has increased. In this paper, we propose to use the notions of trust and reputation as the basis for AMAS. In this case, the feedback obtained comes from acquaintances in the society of agents, through interactions with them. For this, we define a model for Trust and Reputation based on SIMmilarity between tasks (TRSIM). The basic idea behind the model is that agents may change its behaviour as a consequence of changes in the environment. An agent may start improving its single utility if the conditions from the environment change towards more favourable ones. At the same time, an agent may degrade its single utility when the environment changes to unadequate conditions for it. If agents in the system maintain an adequate level of interaction among them, TRSIM is capable of detecting changes in the individuals of the society and accordingly, update trust and reputation to select, in any moment, whom to interact with. These changes are also perceived and managed for groups of agents.

The rest of the paper is structured as follows. Section 2 introduces structure, main characteristics and general functionality of the TRSIM model. Section 3 studies the adaptive capabilities of the model through empirical results. Finally, in section 4 some conclusions are given.

ISSN: 1988-3064(on-line)

2 The TRSIM model for management of trust and reputation

This section describes the structure and functionality of TRSIM, a trust and reputation model based on similarity criteria between tasks. According to the classification given by Ramchurn et al. [10], TRSIM is an adaptive model based on the learning and evolution of trust and reputation measures at individual level. In this classification, there are two complementary conceptual levels to conceptualise trust. The first one is individual-level trust, whereby an agent has some beliefs about the honesty or reciprocative nature of its interaction partners. The second one is system-level trust, whereby the actors in the system are forced to be trustworthiness by the rules of encounter (i.e. protocols and mechanisms) that regulate the system.

Protocols and mechanisms guarantee the trustworthy the agents into the system, but they cannot always achieve this objective without some loss in efficiency. In such cases, individual trust models are important in guiding an agent decision making. Similarly, where trust models at the individual level cannot cope with the uncertainty about the environment, system-level trust models aim to guide the interaction [10].

Generally, at individual-level, trust models consider that trust is a social phenomenon based on the interaction between agents, in order to select the most reliable partner and the strategy to adopt with it (i.e. who, when, how). Trust is a rating obtained from the interactions between agents: each of them stores information about the performance of others [11, 13].

TRSIM uses a unique concept of trust, combining direct trust (obtained from direct experiences, basically) and reputation (obtained from information interchanged between agents). This concept of trust, considered by TRSIM, consists of two dimensions: (1) the trust in an agent to offer solutions to a given task, and (2) the trust in an agent giving information about the performance of others.

2.1 Model structure

The model is structured and operates following the structural diagram given in figure 1. The TRSIM model is composed by a set of information bases of experiences that each agent stores about the behaviour of the others and a set of functions to operate with these bases of experiences. Functions appearing in the figure, represented with rounded boxes, produce values to guide the interactions between agents, based on trust and reputation concepts.

From the point of view of an individual agent, that needs to solve a given task, the model has two bases of experiences to obtain trust and reputation values for this task: base of experiences for trust (IET) and for reputation (IER). In these bases of experiences, the model stores information related with the performance of other agents. IET manages a binnacle, storing historical information related to the capabilities of other agents to other solutions to a given task. In the other hand, IER updates a unique value, for each neighbour, to represent its skills giving information about the performance of others, for a specific task. The number of experiences of each type is limited, taking into account the bounded computational and storage resources in each agent. According to the relevance of each type of trust information, IER manages an unique value associated to reputation capabilities of each agent. Trust information is more relevant than reputation one.

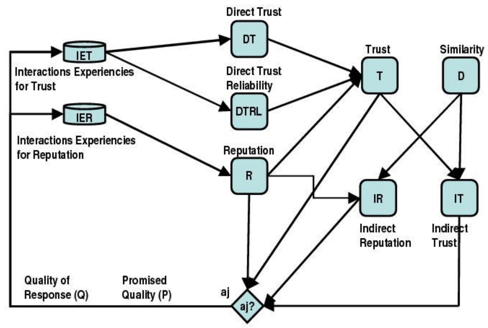

Figure 1: Relationship between different parts of the model

Using the bases of experiences, trust and reputation values are produced by means of the right combination of some functions, and information interchanged between agents. We are inspired in the effective and efficient way to combine informational sources, used by ReGreT [11]. Following its ideas, first, by introspection of the bases of experiences, the model calculates direct trust (DT), reputation (R) and reliability of DT (DT RL), and combines them to produce a unique value of global trust, using the function T. The value, aggregated from direct trust DT, its reliability DT RL and reputation R, is used to select the artners in the interaction, to ask about the solution or about others.

If the bases of experiences is empty for a given task, the model obtains the values of trust (by means of function DT) and reputation (by means of function R) for a similar well-known task and combines these values with the similarity degree between two tasks, given by function D. For that, the model uses functions IT and IR to select the partners in the interaction.

In this way, the model assists agent decisions making, recommending a given provider agent to solve its task, using responses offered by him. To evaluate the interaction with the partner, and to update its own bases of experiences, the model gives two functions: fulfillment of the promised satisfaction (P) and quality of the response (Q).

2.2 Adaptiveness of TRSIM

The main characteristics of the adaptive mechanisms of TRSIM are given by the way that the model updates the bases of experiences. The base of experience IET is represented by a set of tuples, where the agent stores information about others when they are contracted for to execute the task:

$$IET = \{(a_j , s_k , et_{j,k,l})|a_j \in A, s_k \in S, et_{j,k,l} \in [0, 1]\}$$

where A is the set of agents in the system, S is the set of possible specifications of tasks that the agent needs to contract, etj,k,l is the satisfaction degree of the agent (that maintains the base of experiences) when agent aj offers a solution to the task sk for the l-th time.

The base of experiences IER stores data about the reliability of other agents when they o□er reputation information:

$$IER = \{(a_j , s_k , er_{j,k})|a_j \in A, s_k \in S, er_{j,k} \in [0, 1]\}$$

where erj,k is the satisfaction degree of the agent when aj offers reputation values about the capabilities of ther agents to perform the task sk .

When an interaction is finalized, in order to update the base of experiences IET , for each agent a_j , that gives the solution w_j to the task s_k, the agent adds a new experience, $edt_j,k = (a_j, s_k, et_{j,k})$, where the experience value $et_{j,k}$ is a measure obtained from the real quality of the solution (Q) and the fulfillment (P) of the promised satisfaction ($ec_{j,k}$) is $et_{j,k} = Q(w_j, s_k)\cdot P (ec_{j,k} , Q(w_j, s_k))$. The definitions for functions Q and P are given in previous works [1, 2]. The base of experiences IER has an unique value of reputation $er_{j,k}$ to indicate the reliability of agent a_j to give reputation information about other agents performing task s_k . To update this value, the model takes into account the variation (produced during the interaction) on trust about agents recommended by a_j. Agents recommended by a_j for the task s_k is denoted by $M_j(s_k)$ (for more information, please, see [1]).

For each recommended agent $a_r \in Mj (s_k)$, the model obtains the trust value that agent had about ar (denoted by $f^{(0)}_{r,k}$) at the beginning of the interaction and the new value (denoted by $f_{r,k}$) at the end. The trust value to give reputation information at the beginning of the interaction $er^{(0)}_{j,k}$ is modified combining the mean of all differences between final and previous trust values for each agent a_r, about agent a_j:

$$er_{j,k} = \begin{cases} er_{j,k}^{(0)} + \left(1 - er_{j,k}^{(0)}\right) * \Delta; \ \Delta \geq 0 \\ er_{j,k}^{(0)} * (1 + \Delta); \Delta \leq 0 \end{cases}$$

where Δ is the mean of all differences between final and previous trust values:

$$\Delta = \frac{\sum_{a_r \in M_j(s_k)} f_{r,k} - f_{r,k}^{(0)}}{|M_j(s_k)|}$$

This function exposes a suitable behaviour in updating the value of confidence on an agent given information about the performance of others. The value of the reputation erj,k will be better than er(0)j,k when the trust on recommended agents from Mj(sk) is improved during the interaction. The confidence of the agent, giving reputation information, is increased when the mean of the differences of trust values is positive. If the trust of recommended agents is decreased, the confidence of the agent as recommender is decreased too.

3 Assessing the adaptive capabilities

We develop a set of experiments devoted to asses the ability of the model to adapt to changes on the behaviour of provider agents. For that, after the stabilization of the system, we induce some changes in the quality of the response of the provider agents and analyze how the model reacts to these changes.

The main results are related with the experimental situations where the behaviour of the agents changes, individually or in a group. Sometimes, in a multi-agent system, agents are forming coalitions. Coalitions are established in the way that, inside each group, agents coordinate their activities in a different way that they carry out them with agents outside the group. For instance, agents inside the same coalition can coordinate their tasks but not with agents of other groups [14]. A member of a given coalition can execute any action that causes, as prejudice, the degradation of the quality of all members in he coalition that it belongs to. Given the importance of this type of situation, where a group level change in the behaviour of agents takes place, the experiments analyze the ability of TRSIM to adapt to them.

3.1 Experiments Scenario

Experiments have been developed considering a single simulation scenario, defined from a unified set of experimental conditions. The scenario is based on interactions among agents. Agents play two roles: provider or consumer of resources. All agents play the provider role, but, for each round, only one agent acts as consumer, according to the requirements of the user randomly assigned at the beginning of the round.

These sets consist of 6 tasks and 10 responses, respectively. The 1/3 of the tasks belong to High-demanding tasks, the 1/3 to Medium-demanding tasks, and 1/3 to Low-demanding tasks. The degree of demand of a given task indicates the level of diffculty to satisfy this task using the set of the solutions that agents may offer. We consider that the demand of a task is given by the values of its non-functional attributes. The degree of demand of a task is greater than other when it has higher values for its non-functional attributes.

One solution (or response), from the set of solutions W , is assigned to each agent when it acts as provider, defining three types of agents. The distribution of agents having High-, Medium- and Low-quality responses varies depending on the type of experiment. At the beginning of each replica of the simulation, only one consumer agent is assigned to represent the requirements of the users, randomly obtained from the set of task specifications S.

A round is a minimal unit of time considered in the simulation. It begins when the requirements of the initiator agent are established, by means of the task specification selected, and finalizes when the bases of experiences are updated. A number of 400 rounds are carried out in each simulation. The experimental evidences are the result from the simulation using our own implementation in Java. The values, shown in each figure, are the mean of the values obtained from 20 replicas of the experiments.

Each figure shows in a separate way the evolution for High- and Medium-demanding tasks. We do not consider Low-demanding tasks because High- and Medium-demanding tasks are the unique cases in which the model is really useful. Low-demanding tasks can be satisfied efficiently for any type of task. For Low-demanding tasks, similar high satisfaction values can be obtained by randomly selected responses.

3.2 Simulation of changes in the behaviour of provider agents

Behaviour of agents depends on the values of attributes which describe responses to requests of tasks. Hence, simulating changes in the behaviour of provider agents implies modifying these values. Behavioural changes are produced gradually, during 20 rounds of the simulation, and they take place once the model reaches a stable condition. Changes begin at round $t = 200$. The curves show the metrics related to the model performance, from the round $t = 190$ to the round $t = 250$.

Figure 2 defines the variations produced in the attributes of the responses, given by agents that modify its behaviour during several rounds. Figure 2.a shows the worsening and recovery tendency of attributes of a given response: the value of each attribute is gradually decreased during 10 rounds and then, with the same regime, recovers to its original value. In the other hand, figure 2.b shows the improvement and worsening of attributes of a given response: the value of each attribute is gradually improved during 10 rounds and then decreased, in the same gradual way, to its initial value.

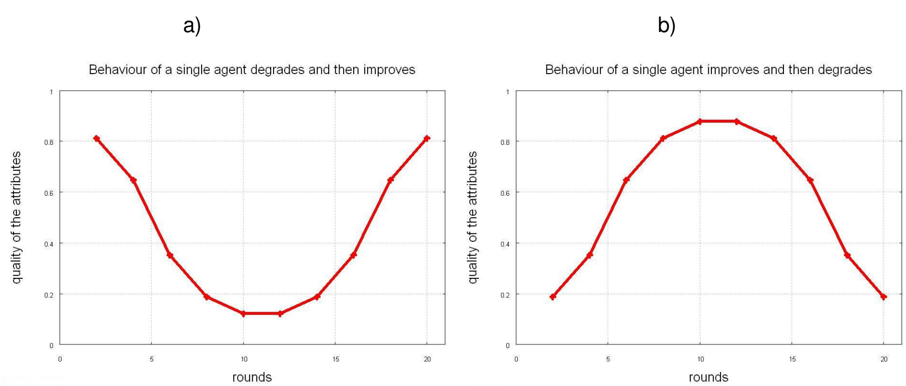

Figure 2: Types of changes in the behaviour of agents: a) behaviour of a single agent degrades and and then improves, b) behaviour of a single agent improves and then degrades.

3.3 Behavioural changes of *High-quality* agents

Experiments considered in this section take into account that the distribution of population of provider agents in the system consists of 30% of High-quality agents, 30% of Low-quality and 40% of Medium- quality.

The experiments suppose that all High-quality agents are part of a coalition, and, for a specific reason, the quality of solutions that they offer are negatively affected. We study the situation given by the behavioral changes, where all High-quality agents simultaneously degrade their behaviour and again, simultaneously return to a good shape.

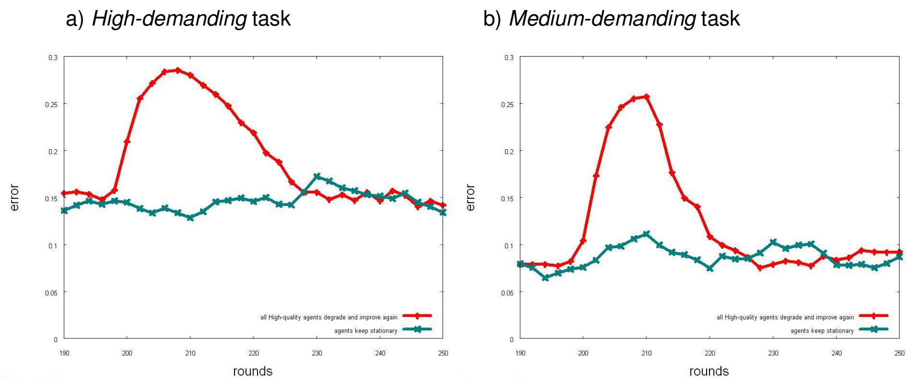

Figure 3: Comparison of error in recommending the suitable response, according to the type of the task to solve in each round (a) *High-demanding*, b) *Medium-demanding)*, when the agents do not change during the simulation and when all *High-quality* agents degrade their behaviour and then improve it.

Figure 3 shows the evolution of satisfaction degree with recommended solution when *High-quality* agents degrade their behaviour and then improve it. We contrast the situation when the behaviour of *High-quality* agents changes with the situation when agents keep their behaviour stationary. We analyze the cases when the initiator agent tries to solve *High-demanding* and *Medium-demanding* tasks.

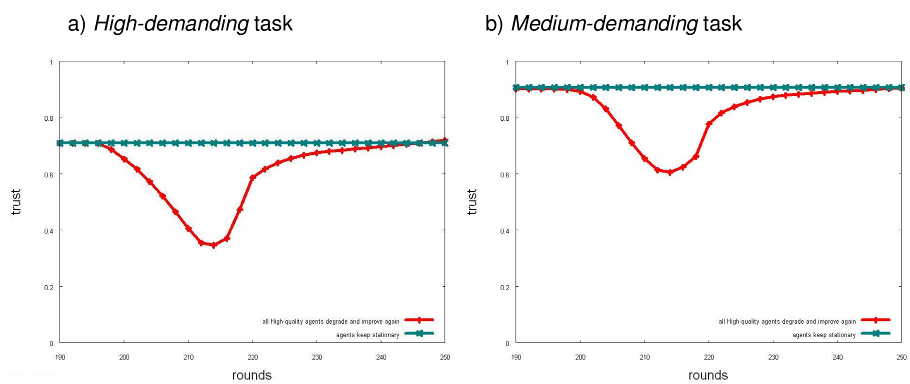

Figure 4: Comparison of trust evolution in the *High-quality* agent degrades its behaviour and then improves it, according to the type of the task to solve in each round (a) *High-demanding*, b) *Medium-demanding*).

For both types of tasks, figure 3 shows that the model is capable to adapt to this type of changes in solutions given by *High-quality* agents. During the change, the error in recommending the suitable solution increases and decreases, according to the way that the quality of the agent solution changes. The error is increased after the qualities become worse, and then it is gradually decreased when qualities are recovered. This behaviour is similar to other situations, in which only a *High-quality* agent degrades the quality and improves afterwards.

Figure 4 shows the variations in the mean of trust values maintained by the contracting agent in the *High-quality* agents that degrade their behaviour and then improve it, according to the type of the task to solve in each round. It compares the cases when agents behaviour keep stationary during the simulation and when all *High-quality* agents degrade their behaviour and then improve it.

This group of curves evidences that the model is capable of representing, using trust values, this type of behavioural changes of agents. In this case, the model decreases the trust when agents degrade their behaviour (during the first part of the change). But, when agents recover their qualities, trust is increased reaching similar values that it has at the beginning of the change.

Both figures show that model is capable to adapt to this type of behavioural change of *High-quality* agents.

3.4 Behavioural changes of *Low-quality* agents

Here, we offer experimental results related with the performance of the model in situations where the behaviour of Low-quality agents changes. There is not variation in the stability of the model, when a Low-quality agent changes individually. There are not significant variations in the main metrics. These measures do not change when there are a lot of High-quality agents in the system, with high trust and reputation values prior to the changes. These High-quality agents are selected during the change, offering

the same satisfaction degree, error and trust in the recommended solution. In these cases, the variations in the Low-quality agents do not a□ect the system performance.

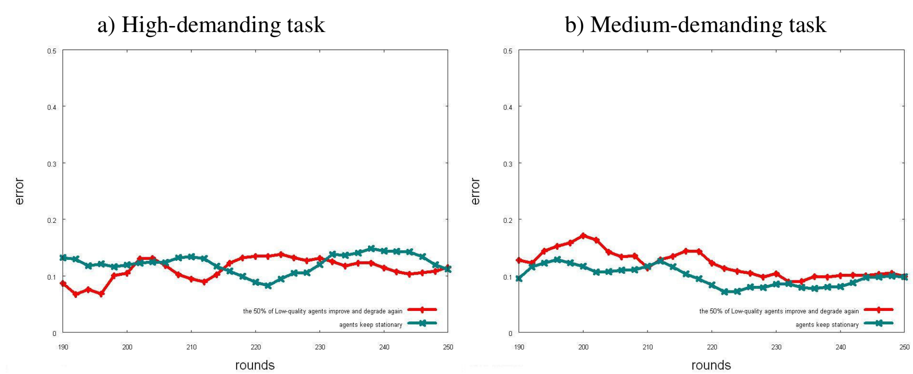

Figure 5: Comparison of error in recommending the suitable response, according to the type of the task to solve in each round (a) *High-demanding*, b) *Medium-demanding*), when the agents do not change during the simulation and when the 50% of *Low-quality* improve their behaviour and then degrade again.

In the following experiment, we consider that there is not any *High-quality* agent. This way, we study the performance of the model when several *Low-quality* agents change their behaviour in absence of *High-quality* ones. This experiment take into account that the 50% of *Low-quality* agents, which are part of a coalition, improve their behaviour and degrade it again, when the population of agents in the simulation consists of 60% of *Low-quality* agents and 40% *Medium-quality* ones. The rest of experimental conditions are the same that the previous experiments shown in section 3.3.

Similarly to previous experiments, we show separately evidences related with *High-demanding* and *Medium-demanding* tasks. In this type of situations, the model does not show significant variations in the evolution of error recommending the suitable response (Figure 5). The evolution of error values points out the ability of the model to recommend the most suitable solution using trust information, in every round. The small variation in error evidences that the model is capable to identify the agent offers the best solution in every moment. In other words, the model keeps a stable status when the 50% of *Low-quality* agents improve its behaviour and degrade again.

Figure 6 shows the evolution of the satisfaction degree with the recommended solution when some *Low-quality* agents change their behaviour, comparing with the case when agents keep stationary.

For both types of tasks, figure 6 shows that the satisfaction degree is increased and, then, returned to degrade in the same way that the quality of the agents changes. The satisfaction is increased during the first part of the change (when quality is increased) because the quality of these agents produce better satisfaction degrees that Medium-quality. (Before the change, the Medium-quality agents produced the best satisfaction values.) After that, satisfaction values are decreased until similar values at the beginning of the change. The differences of this metric between both situations are higher for High-demanding tasks than Medium-demanding one.

Also, similarly to the situation in which High-quality agents change (presented in the previous section), the model is capable to represent, using trust values, this type of behavioural changes of agents. In this case, during the first part of the change (when agents improve), the model increases the trust in these agents. After that, trust values are decreased until similar values at the beginning of the change.

a) *High-demanding* task b) *Medium-demanding* task

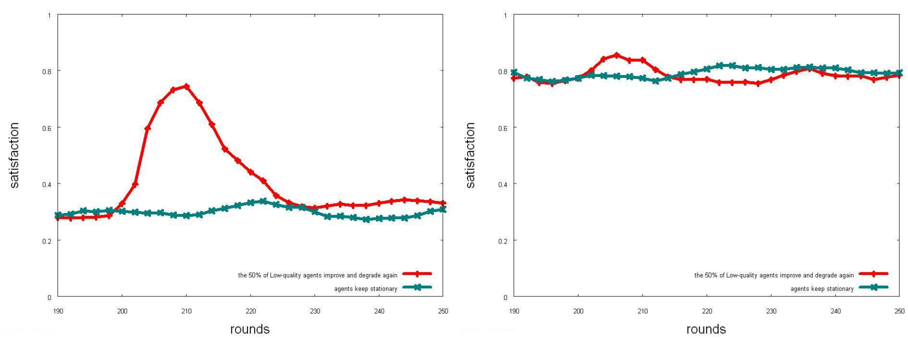

Figure 6: Comparison of satisfaction degree with the recommended solution, according to the type of the task to solve in each round (a) *High-demanding*, b) *Medium-demanding*), when the agents do not change during the simulation and when the 50% of *Low-quality* improve their behaviour and then degrade again.

4 Conclusions

We study the effectiveness of TRSIM model as adaptive mechanism in multi-agent systems. Taking into account an experimental consumer-provider scenario, we carry out some empirical evaluations about the performance of the model related with situations in which the behaviour of some agents change.

Two types of experiments are presented, each of them related to the type of agents that change their behaviour. TRSIM model shows desires behaviour in these experiments, for all considered experimental situations. In all cases, the model is capable of representing the behavioural changes of agents using trust and reputation values. Variations in trust, associated to agents that change their behaviour, capture the variations in the quality of their solutions. In this way, the model can be useful to guarantee adaptive mechanisms to withstand behavioural changes in provider agents.

Also, for each experiment, we offer evidences related to the stability of the model when agents change their behaviour. When High-quality agents degrade their behaviour and improve it again, the model loses the stability, reaching it again after a few rounds. In other hand, when Low-quality agents change, the model keeps a stable status during the variation of qualities of responses.

Future works include more experimentation, in order to analyse what is the senstivity of the system for concrete parameters which govern the behaviour of the system, for real application domains. We are particulary interested on characterising the number of interactions needed by the model to adapt to changes. We are also interested on characterising change. Our inmediate goal now is to relate types of changes with the necessary number of rounds to adapt to them.

References

[1] A. Caballero, J. Botía and A. Skarmeta. A New Model for Trust and Reputation Management with an Ontology Based Approach for Similarity Between Tasks. In E. A. K. Fischer, I. J. Timm and N. Zhong, editors, Multiagent System Technologies, LNCS 4196, pages 172–183. Springer, 2006.

[2] A. Caballero, J. Botía and A. Skarmeta. On the behaviour of TRSIM model for trust and reputation. In M. K. P. Petta, J. Muller and M. Georgeff, editors, Multiagent System Technologies, LNCS 4687, pages 182–193. Springer, 2007.

[3] A. Cardon and Z. Guessoum. Systèmes multi-agents adaptatifs. JFIADSMA (Journées Françaises d'Intelligence Artificielle Distribuèe et Systèmes Multi-Agents), pages 100–116, 2000.

[4] K. Dooley. A Complex Adaptive Systems Model of Organization Change. Nonlinear Dynamics, Psychology, and Life Sciences, 1(1):69–97, 1997. DOI: http://dx.doi.org/10.1023/A:1022375910940

[5] Z. Guessoum. Adaptive Agents and Multiagent Systems. IEEE Distributed Systems Online, 5(7):4, 2004. DOI: http://dx.doi.org/10.1109/MDSO.2004.10

[6] J. H. Holland. Hidden Order: How Adaptation Builds Complexity. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1995.

[7] K. A. D. Jong. An analysis of the behavior of a class of genetic adaptive systems. PhD thesis, University of Michigan, Ann Arbor, MI, USA, 1975.

[8] S. A. Levin. Ecosystems and the biosphere as complex adaptive systems. Ecosystems, 1(5):431–436, 1998. DOI: http://dx.doi.org/10.1007/s100219900037

[9] P. Maes. Modeling adaptive autonomous agents. Artif. Life, 1(1-2):135–162, 1994.

[10] S. Ramchurn, D. Huynh, and N. Jennings. Trust in Multi-Agent Systems. Knowledge Engineering Review, 1(19):1–25, 2004.

[11] J. Sabater and C. Sierra. Social ReGreT, a reputation model based on social relations. ACM SIGecom Exchanges, 3(1):44–56, 2002. DOI: http://dx.doi.org/10.1145/844331.844337

[12] R. M. Smith and M. A. Bedau. Is Echo a Complex Adaptive System? Evolutionay Computation, 8(4):419–442, 2000. DOI: 10.1162/106365600568248

[13] W. T. Teacy, J. Patel, N. R. Jennings, and M. Luck. "TRAVOS": Trust and Reputation in the Context of Inaccurate Information Sources. Autonomous Agents and Multi-Agent Systems, 12(2):183–198, 2006. DOI: http://dx.doi.org/10.1007/s10458-006-5952-x

[14] G. Weiss, editor. Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence. The MIT Press, 2001.