Version accepted for publication in IEEE Intelligent Systems.

editorial version: DOI: 10.1109/MIS.2015.98

Experimenting Certified Reputation in a Competitive Multi-Agent Scenario

Francesco Buccafurri, *Member, IEEE*, Antonello Comi, Gianluca Lax, *Member, IEEE*, and Domenico Rosaci[1]

Abstract

In hostile environments, there is a risk of low efficiency of reputation, due to possible attacks coming from malicious agents. Indeed, it is widely accepted the result obtained on the platform ART that, in competitive scenarios, the use of direct knowledge about the environment (i.e., reliability) is more effective than its combination with indirect knowledge (i.e., reputation). In the open MAS research field, the notion of certified reputation has been proposed to improve the reputation effectiveness. However, no experience about the introduction of certified reputation in competitive environments has been provided. Even though it is obvious that the effectiveness of the reputation mechanism is generally improved when the certification is adopted, it is not clear when this improvement allows us to use the reputation mechanism to increase the individual gain, thus resuming the role of reputation. In this paper, we deal with this problem by assessing the use of certified reputation in the context of competitive agents.

***Index Terms*-Competitive agents, security, trust.**

[1] F. Buccafurri, A. Comi, G. Lax and D. Rosaci are with the Department DIIES, University of Reggio Calabria, Italy. email:{bucca, antonello.comi, lax, domenico.rosaci}@unirc.it

1 Introduction

The introduction of trust-based approaches in multi-agent systems (MASs) has been widely recognized as a promising solution to improve the effectiveness of these systems [9, 3, 18, 20, 14]. In this context, an agent that does not have a sufficient direct knowledge about another agent can exploit reputation to estimate its trustworthiness. This decreases immunity of reputation to malicious behavior of agents, so that a number of attacks become possible, performed by agents that provide incorrect recommendations. For example, *self-promoting* attacks concern the possibility that an agent increases its reputation by fake positive feedback, while *slandering* attacks are carried out to decrease the reputation of other agents by providing fake negative feedback about them. If an agent that receives a recommendation cannot verify the recommendation reliability, it can be lead to follow a bad suggestion.

To contrast this weakness, a number of solutions can be applied [12, 16, 8]. Among these, our paper is focused on the mechanism of *certification* of reputation. The notion of certified reputation has been initially introduced in the context of open MASs [11, 10, 6, 15, 21] and has been theoretically faced in [12]. However, no experience about the introduction of the certified reputation in competitive environments has been provided. A competitive scenario is composed of self-interested agents that compete for winning a game. This is a particular sub-case of the general self-interested agents' scenario [12], because, during the competition, the agents are rewarded or penalized for their choices by the environment, and this a-posteriori feedback can be exploited by each agent for updating its internal reputation model. A well-known example of application of this scenario is represented by e-Commerce. Even though it is straightforward that the effectiveness of the reputation mechanism in a general environment of self-interested agents is generally improved when the certification is adopted, in a competitive environment, obtaining information about an agent reputation has a cost for the requester agent. Thus, it is an important issue to determine the conditions under which the usage of reputation is actually an advantage to increase the individual gain in the competition. Currently, in the literature, the role of reputation in the competitive MASs has been devalued, due to some results demonstrating that reputation is not necessary [16]. In this work, we analyse in depth such an issue, by showing that the conclusion above is limited to environments where the reputation measure is based on unreliable recommendations. Conversely, if recommendations are someway certified, the role of reputation can be profitably resumed.

To accomplish the above goal, we use the platform *Agent Reputation and Trust* (ART) [2], where a set of agents participate to a competition. Specifically, each agent gives appraisals on paintings presented by its clients, asking opinions to other agents and requiring recommendations about the reputation of other agents. The competitions run on the ART testbed showed that reputation has a marginal role w.r.t. reliability. However, this result considers only scenarios in which the number of agents is limited and the percentage of expert agents is forced to be appreciable. Nevertheless, what happens when the number of agents grows in such a way that finding

an expert agent becomes difficult? This question has been addressed in [4], where the authors highlight that in large agent spaces, it can be difficult for an agent to obtain sufficient evidence about a potential partner's past behavior to build predictive models of trust capable of making evaluations about new, unknown agents. Moreover, the ART simulator assigns a static expertise value to each agent. However, in a general context, agents' expertise changes over time, introducing a new difficulty in evaluating the expertise of the other agents. We argue that the sole reliability does not give satisfactory results.
To experiment certified reputation in ART, we adopt a trust model, called *Certified Reputation In Trust* (CRIT), in which recommendations are provided along with a *level of assurance* that is a measure of the trustworthiness of the recommendation that the requested agent provides to the requester agent about a third agent.
We study this issue for several populations of agents characterized by: (i) different population size; (ii) different number of expert agents; and (iii) dynamic agents' expertise. We observe that in presence of large populations with few *expert* agents, the use of the sole reliability is not sufficient. In this situation, our model gives the best results and outperforms those methods that use both reliability and reputation, which are not effective in presence of a relevant number of unreliable recommendations.
The most recent results presented in the context of competitive multi agent systems are, to the best of our knowledge, those obtained in the ART competitions 2006-2008. Therefore, we have considered as competitors the two winner agents of the last editions of that competition, namely UNO 2008 and IAM, and we have also compared CRIT with RRAF, which has been proposed in 2011 for studying the role of reputation in competitive multi agent systems having large dimension. Another two interesting approaches, called CRM [13] and DTMAS [1], consider both direct and undirect trust, as in our paper. However, they are not designed to be applied in a competitive scenario, and particular concepts as *winner of the competition*, *prize*, *penalization* are not obviously considered when defining trust measures. As a consequence, a comparison with CRIT would not be fair.
The remaining of the paper is organized as follows. After having introduced the multi-agent scenario in Section 2, we present our trust model in Section 3. Then, in Section 4, we describe the experiments performed and, in Section 5, we draw our conclusions.

2 The multi-agent scenario

We deal with a multi-agent system, whose agents can provide a set of services to clients. When a client needs a service falling into a given category *cat*, it sends a request to the Agent System Manager ASM, which assigns the request to an agent. At each step, ASM examines all the service requests submitted by clients, and assigns them to the agents. The client pays a given price *sp* to the selected agent to obtain the service. The interactions among agents follow this protocol:

To provide a client with a service of a given category *cat*, an agent *a* may decide to require the collaboration of another agent *b*. Moreover, *a* can ask a third agent *c* for a recommendation about *b*. If *b* accepts the

request, then *a* must pay a given price (reputation price) *rp* to *c*. Moreover, *a* can also ask *b* itself for providing an auto-declaration of its expertise.

The interactions between agents are executed by following an *assurance protocol* that allows the interlocutors to mutually exchange a *proof* synthetically describing the interaction. In this paper, we do not deal in depth with the issue of the assurance protocol, because this aspect is orthogonal to the core of the proposal in the sense that, in principle, we could choose any assurance protocol able to produce, as a final state, a [0, 1]-real *level of assurance* (where 1 is the maximum level of assurance), representing a measure of the trustworthiness of the recommendation that a requested agent *c* provides to a requester agent *a* about a third agent *b*. Anyway, we give a very short sketch about how this level of assurance can be obtained in our protocol, in order to make plausible the overall proposal. The level of assurance is obtained by evaluating the proof that *c* is able to show to *a* in order to guarantee the level of assurance of the provided recommendation. The maximum level of assurance corresponds to the case in which all the transactions occurred between *b* and *c*, on which *c* produces its recommendation, are traced through messages whose authenticity and non-repudiation are based on a Third-Trusted-Party-granted certification. In contrast, the presence of some transaction traced through a weaker mechanism (for example, in our model the transactions can be traced just on the basis of some randomly chosen witnesses), reduces the level of assurance of the recommendation, until the minimum value corresponding to the case of all transactions with no proof. More in detail, given a transaction T between two agents *b* and *c*, we say that T is *witnessed* by *w* if *b* and *c* have notified to the (randomly) agent *w* all the data of T. We say that T is *signed* if T is provided with the digital signature of both *b* and *c*. Therefore, the *level of assurance* LoA(T) of T is computed as follows: LoA(T) = 1 if T is both witnessed and signed; LoA(T) = 0.5 if either T is witnessed but it is not signed or T is signed but it is not witnessed; LoA(T) = 0 if T is neither signed or witnessed. Finally, given a recommendation *r* provided by an agent *c* to a requester agent *a* about a third agent *b*, we compute the *level of assurance* $LoA(r)$ of *r*, as the maximum LoA(T) among all the transactions T occurred between *b* and *c*.

If *a* decides to obtain the collaboration of *b*, it needs to pay a price *cp* to *b*.

At the end of the step, *a* receives a feedback from ASM for each service provided by *a* in the current step.

The Agent UML of the protocol is reported in Fig. 1.

Figure 1: The protocol of agent interaction.

3 CRIT: Certified Reputation in Trust

This section aims at describing the CRIT framework. First, we present the formal definition of the trust model, and then we describe the process of the trust update.

3.1 Trust Model: Formal Definitions

We denote by A the list containing all the agents belonging to the multi-agent systems, and by a_i the i-th element of A. A set of four mappings, denoted by SR_i, R_i, β_i, and P_i is associated with each agent a_i, where each mapping receives an agent a_j and a category *cat* as input and yields as output a trust measure that a_i assigns to a_j, in relation to the category *cat*. Each measure is represented by a real number belonging to the interval [0, 1], where 0 (1, resp.) is the minimum (maximum, resp.) value of trust. In particular:
$SR_i(a_j, cat)$ represents the *service reliability* that a_i assigns to the services provided by a_j for the category *cat*.
$R_i(a_j, cat)$ represents the *reputation* that a_i assigns to a_j for the category *cat*. Reputation is a measure of trust that an agent assigns to another agent based on some recommendations coming from the agents of the community.
$\beta_i(a_j, cat)$, called *reliability preference*, represents the *preference* that a_i assigns to the usage of reliability with respect to reputation in evaluating a_j for the category *cat*. In other words, when a_i has to compute the overall trust score of an agent a_j in a category *cat*, it considers both the contribution of the service reliability $SR_i(a_j, cat)$ and the reputation $R_i(a_j, cat)$. The importance to give to reliability is represented by the value $\beta_i(a_j, cat)$. In our framework, β_i is computed by the agent a_i based on the *assurance information*, provided together with the recommendations by the contacted agents.
$P_i(a_j, cat)$ represents the overall preference that a_i assigns to a_j for

the category *cat*, based on both the reliability and reputation perceived by a_i.
We also define a mapping denoted by $RECC_i$, representing the *recommendations* obtained by agent a_i, where a *recommendation* is as a pair $r = \langle v, l \rangle$, such that v and l are two [0, 1]-real numbers called *recommendation value* and *recommendation level of assurance*, respectively.
Formally, $RECC_i$ is a mapping that receives two agents a_j and a_k and a category *cat* as input, and yields as output a recommendation $RECC_i(a_j, a_k, cat)$ representing the recommendation that the agent a_j provided to the agent a_i about the agent a_k for the category *cat*, together with a measure of the level of assurance that can be assigned to this recommendation.

3.2 The trust updating algorithm

The mappings are updated by the agent a_i at each step, using the following algorithm:

1. Reception of the recommendations from the other agents;
2. Computation of *SR* mapping using the feedback sent by ASM;
3. Computation of *R* and β mappings using the available certificated recommendations;
4. Computation of *P* mapping;
5. Selection of the best candidate agents to request collaboration.

Below, we describe into detail the different phases of the algorithm.

Phase 1: Reception of the Recommendations. The agent a_i receives some recommendations from the other agents, in response to previous recommendation requests. Such recommendations are encoded by *RECC* mapping. In particular, each recommendation coming from an agent a_j and related to an agent a_k is contained in a *recommendation message m*, which is a tuple $\langle v, l \rangle$, whose elements are stored by a_i in the mapping $RECC_i(a_j, a_k, cat).v$ (the recommendation value) and $RECC_i(a_j, a_k, cat).l$ (the recommendation level of assurance), respectively.
Phase 2: Computation of SR mapping: ASM sends a_i the feedback for each service *s* provided in the past step, where the contributions given by other agents to a_i are evaluated. These feedback values are contained in a mapping *FEED*, where each feedback $FEED_i(s, a_j)$ is a real number belonging to [0, 1], representing the quality of the collaboration that the agent a_j provided to the agent a_i concerning the service *s*. A feedback equal to 0 (1, resp.) means minimum (maximum, resp.) quality of the service.
Based on this feedback, the agent a_i updates the mappings *SR* by computing the current reliability shown by an agent a_j by averaging all the feedback values concerning a_j. Therefore, denoting by $Services_i(a_j, cat)$ the set of services of the category *cat* provided by a_i with the collaboration of a_j, the current service reliability shown by a_j,

which we denote by $sr(j, cat)$, is computed as

$$sr_i(a_j, cat) = \frac{\sum_{s \in Services_i(a_j, cat)} FEED_i(s, a_j)}{|Services_i(a_j, cat)|}$$

At each new step, the current reliability is taken into account for updating the element SR_i by averaging the value of SR_i at the previous step $t-1$ and the current reliability computed at the new step t, denoted by SR^t. Thus:

$$SR_i^t(a_j, cat) = \alpha \cdot SR_i^{t-1}(a_j, cat) + (1-\alpha) \cdot sr_i^t(a_j, cat)$$

where α is a real value belonging to [0, 1], representing the importance that a_i gives to the past evaluations of reliability with respect to the current evaluation.

Phase 3: Computation of R and β. The recommendations contained in the mapping $RECC_i$ are used by the agent a_i to compute the reputations of the other agents of the community. In particular, a_i computes the reputation of another agent a_j as a weighted mean of all the recommendations received from the other agents of the community concerning a_j (let us denote by AS this set), where the weight of each recommendation value is the corresponding level of assurance. Thus, $R_i(a_j, cat)$ is equal to:

$$\frac{\sum_{k \in AS, k \neq i} RECC_i(a_k, a_j, cat).v \cdot RECC_i(a_k, a_j, cat).l}{\sum_{k \in AS, k \neq i} RECC_i(a_k, a_j, cat).l}$$

where, we recall, $RECC_i(a_k, a_j, cat).v$ (resp., $RECC_i(a_k, a_j, cat).l$) is the value (resp., the level of assurance) of the recommendation that the agent a_k provided to the agent a_i about the agent a_j.

The β coefficient associated with the agent a_i is recorded in the mapping β_i. The computation of the average level of assurance of the recommendations related to an agent a_j in the category cat, denoted by $\beta_i(a_j, cat)$, is obtained by averaging the level of assurance associated with all the recommendations related to a_j in the category cat:

$$\beta_i(a_j, cat) = \frac{\sum_{k \in AS, k \neq i} RECC_i(a_k, a_j, cat).l}{|AS| - 1}$$

Phase 4: Computation of P. The agent a_i finally computes the overall preference measure $P_i(a_j, cat)$ in agent a_j by taking into account both the service reliability $SR_i(a_j, cat)$ and the reputation $R_i(a_j, cat)$. In particular, the value of the mapping $\beta_i(a_j, cat)$ is used to weight the importance of the service reliability with respect to reputation:

$$P_i(a_j, cat) = \beta_i(a_j, cat) \cdot SR_i(a_j, cat) + (1 - \beta_i(a_j, cat) \cdot R_i(a_j, cat))$$

At each step, the agent a_i exploits the mapping P to select the most suitable candidates to require a collaboration.

4 Evaluation

In this section, we describe the experimental campaign aimed to evaluate the advantages and the limitations introduced by CRIT. In our experiment, we used the ART platform, widely used for computing fair comparisons of trust models [7, 5, 17].
The results of the ART competitions, together with the analysis made in [16], concluded that the use of reputation in agent spaces having small size is not useful, being sufficient to exploit the sole reliability measure. To the best of our knowledge, no further analysis has been proposed in the literature until 2012, when RRAF model [19] has been presented to test the possibility that reputation can introduce some advantage in large agent spaces. Therefore, in this paper we have compared our approach, which uses reliability and reputation with certificated recommendations, with the best algorithms presented in the past that do not use certificated recommendations, namely RRAF, which exploits both reliability and reputation, UNO 2008, which is the best algorithm using the sole reliability, and IAM [22], which is the unique winner of an ART competition using both reliability and reputation.
The game is supervised by a *simulator*, operating as follows:

The clients are simulated by ART and require opinions on paintings to the appraiser agents. Each painting belongs to an *era*, therefore the set of the categories *SC* is the set of all possible eras. For each appraisal, an agent earns a given money amount *sp* that is stored in its bank account *BA*.
Each agent has a specific expertise level in each era, assigned by the simulator. The error made by an agent while appraising a painting depends on this expertise and the price the appraiser decides to spend. The agent's expertise, defined as its ability to generate an opinion about the value of a painting, is described by a normal distribution of the error between the agent's opinion and the painting value. Agents know their levels of expertise for each era but the simulator does not inform them of other agents' expertise levels. The values of paintings presented by clients are chosen from a uniform distribution. Likewise, the eras which paintings belong to are also uniformly distributed among the set of eras.
Each agent can obtain recommendations about another agent by other agents. Each recommendation has a given price *rp*, which in our experiments has been set to the value 0.01 used in the ART competition. A recommendation is simulated by a pair $<v, l>$ of real values ranging in [0, 1], where v is the value of the recommendation and l is its level of assurance.
The simulator sends the feedback $FEED_i(a_j, cat)$ to each agent a_i for each agent a_j contacted by a_i in the previous step. By using this feedback, the agent a_i updates the mapping *SR*, which at the initial time is set to 0, thus assigning no reliability to unknown agents. This way, at the initial time, each agent considers all the other agents equally unknown. Therefore, initially, each agent has the same probability as other agents to be contacted and the performance shown at the first contact will be used to update *SR*.
The recommendations stored in the mapping *RECC* are used by the agent a_i to compute the reputations of the other agents.

Each agent a_i computes the overall preference measure $P_i(a_j, cat)$ in the agent a_j.
Although initially clients are distributed evenly among agents, agents whose final appraisals are the most accurate are rewarded with a larger share of the client base in subsequent steps.

We have set sp = 100 and cp = 10, as in the ART competition 2008. We have also used a value α = 0.5, as in the original RRAF model [19]. Our purpose was to compare the performances of agents exploiting the trust model here presented, and three other approaches, namely UNO2008 [16] (the winner of the ART 2008 competition - it uses only reliability in its trust model), RRAF [19] and IAM [22].
To simulate different types of agent populations, we have built, besides the agent CRIT, UNO2008, RRAF and IAM, which are the subjects of our comparison, three types of dummy agent:

CERTIFIED agent: it is an agent provided with a random quality of expertise $e \in [0, 1]$, generated by a uniform distribution. It responds to a recommendation request with a certified recommendation, whose level of assurance $l \in [0, 1]$ is generated by a uniform random distribution.
INEXPERT agent: it is an agent provided with a random quality of expertise $e \in [0, 0.2]$, generated by a uniform distribution. Like the CERTIFIED agent, it responds to a recommendation request with a certified recommendation having level of assurance generated by a uniform random distribution.
UNCERTIFIED agent: like the CERTIFIED agent, it is provided with a random quality of expertise $e \in [0, 1]$, generated by a uniform distribution. Differently from CERTIFIED and INEXPERT agent, it does not provide any level of assurance when it responds to a recommendation request. For our analysis, we just need to model an agent that makes fake recommendations, thus collapsing the different (slandering and self-promoting) attackers into this generic category UNCERTIFIED.

Figure 2: Average BA vs % of INEXPERT agents.

4.1 Experiments with static expertise

We have run two different sub-categories of experiments with static expertise. The experiments of the first sub-category deal with the variation of the number of INEXPERT agents, while those of the second subcategory analyse the variation of the usage of the certificates.

4.1.1 Performance versus number of INEXPERT agents

This campaign was composed of 11 experiments, each associated with a different agent population A_i, with $i = 1, \ldots, 11$. Each population A_i is composed of one CRIT agent, one UNO2008 agent, one RRAF agent and other 200 agents. Among these 200 agents, $20 \cdot (i - 1)$ are INEXPERT agents and the remaining ones are CERTIFIED agents. In other words, the agent population A_i has a percentage of INEXPERT agents increasing with i. We have run 10 games for each

experiment. In Fig. 2, we have plotted the average bank amount *BA* of each agent involved in the comparison for each experiment corresponding to a different percentage of INEXPERT agents, where the average is computed on all the experiments. The results clearly show that CRIT is always the best performing agent (except the case of 100% of INEXPERT agents, where obviously the performances of all the agents drastically degrade) and that its advantage over the other agents generally increases as the percentage of INEXPERT agents increases. The advantage over the second performer (i.e., UNO2008) is equal to 19.84% for a population with no INEXPERT agents, and the maximum advantage, equal to 31%, is reached for a population containing 80% of INEXPERT agents.
To study the role of the size of the agent population, we have repeated the experiments above for different sizes. In particular, we have considered three percentages of INEXPERT agents, namely a *low* percentage (20%), a medium percentage (50%) and a high percentage (80%). For each of these percentages, we have compared the average bank amounts of the four competitors for different population sizes. The results we have obtained point out that the variation of the population size does not significantly influence the performances of the competitors, except the case of both a very small size (i.e., 50 agents) and a high percentage of INEXPERT agents, where all the competitors show the worst results. Only in this situation CRIT obtained an average bank amount slightly smaller than UNO2008, while in all the other cases it is always the winner.

Figure 3: Advantage of CRIT w.r.t UNO2008 vs % of CERTIFIED agents.

4.1.2 Performance versus percentage of CERTIFIED agents

To analyse the role of certified reputation, we have performed a second campaign of 11 experiments, each composed of 10 games run on a population of 200 agents, containing $20 \cdot (i - 1)$ CERTIFIED agents, with $i = 1, \ldots, 11$, while the rest of the population is composed of UNCERTIFIED agents. In Fig. 3, for each percentage of CERTIFIED agents, we have plotted the average percentage of advantage obtained by CRIT over UNO2008, which was the second performer in almost all games.
Fig. 3 highlights that, if the number of agents using certified reputation is less than 20% of CERTIFIED agents), our approach performs significantly worse than UNO2008. Instead, if certified reputation is used enough (almost 20% of CERTIFIED agents), CRIT agent drastically wins, with an advantage that is about 20%, confirming the result obtained in the previous experiment (corresponding to the initial point of Fig. 2, with 0 INEXPERT agents).

4.2 Experiments with dynamic expertise

We have performed a third experiment to study how performances change when agents cannot rely on the a-priori knowledge of a static expertise. To do this, we enable the random variation of the expertise of CERTIFIED agents during the game. In detail, we have

performed a campaign of 11 experiments, where each experiment is composed of 10 games run on a population containing the 4 competitors and other 200 agents. In the i-th experiment (with $i = 1, \ldots, 11$), among these 200 agents, are CERTIFIED agents to which the simulator assigns a new random quality of expertise at each step of the simulation. The remaining $200 - 20 \cdot (i-1)$ agents are UNCERTIFIED agents. In Fig. 4, for each percentage of CERTIFIED agents corresponding to each experiment, we have plotted the performances of the competitor agents. The result shows that CRIT is the best performer with a significant advantage over its competitors. It is worth highlighting that, when the percentage of CERTIFIED agents increases, the performances of all the competitors generally tend to make worse, due to the effect of "confusion" introduced by the variability of the expertise. However, among all the competitors, CRIT is the agent with the best capability to adapt to the variability of expertise.
The experiments described above can be synthetically summarized in the following results:
When the number of expert agents is sufficiently small, it is generally difficult for all the examined techniques to find them to obtain the best collaboration. In this situation, CRIT almost always exhibits the best performances and the gap to the other techniques increases with the size of the agent population. Only if the population is small, all the techniques are capable of finding the rare expert agents, and CRIT does not show any advantage (in particular, UNO2008 achieves comparable results using the sole reliability). Otherwise (i.e., for non-small populations), the introduction of certified reputation allows CRIT to clearly overcome the other competitors.
Instead, if the number of expert agents is sufficiently high, all the techniques show comparable results, independently of the population size. The introduction of certified reputation produces only limited advantages in this case.

Figure 4: Average BA vs % of CERTIFIED agents.

Certified reputation produces very good results when a high number of agents uses certification. Concerning this aspect, we have also highlighted that, keeping constant the percentage of agents using certification, the advantage of CRIT increases with the population size, because the total number of certified recommendations generally increases too.

In case of dynamic expertise, the use of certified reputation is advantageous. A plausible interpretation of this result is that the presence of a percentage of agents with dynamic expertise introduces a sort of "confusion", which favors certified reputation, especially when the number of the agents with dynamic expertise increases.

The first and the second results, considered on the whole, give an answer to the question posed in the introduction, that is, "what happens when the number of agents grows in such a way that finding an expert agent becomes really difficult?". The third result provides an answer to the other issue stated in the introduction, concerning the effects produced on the agent performances when expertise changes over time.

5 Conclusion and Future Work

The use of reputation seems necessary in competitive MASs when a large size of the agent space makes difficult for an agent to obtain a complete knowledge about the expertise of other agents. However, the most recent studies in the trust-based agent community highlighted that reputation is not effective enough, due to the impossibility of verifying trustworthiness of recommendation providers. Certifying reputation is a way to fortify reputation mechanisms. This claim has been fully confirmed in our paper, but we added some new important results, analysing in detail what happens in a competitive MAS when certified reputation is enabled. We have compared our proposal with other trust-based approaches running on the standard ART platform. While the results of the last ART competition rewarded UNO2008, which does not use reputation measures at all, our experiments, carried out on a set of agents larger than that of the ART competition and including a significant percentage of deceiving agents, clearly show the importance of using reputation, but only if it is combined with certification. In our experiments, when a sufficient percentage of agents uses our assurance information, CRIT outperforms the other approaches, showing its best results in presence of an agent population characterized by very low expertise values. It is worth remarking that our study is currently limited to a sensitivity analysis of the trust algorithms against randomly generated behaviour, and that an extension of such an analysis to the case of smart opponents is subject of our ongoing and future research. Moreover, although the overhead due to the introduction of additional message passing does not appear relevant, we are planning to further analyse the aspect of computational complexity, for evaluating the impact of adopting different assurance models.

References

[1] A.M. Aref and T.T. Tran. A decentralized trustworthiness estimation model for open, multiagent systems (dtmas). *Journal of Trust Management*, 2(1):1-20, 2015.
[2] ART url, 2012. http://megatron.iiia.csic.es/art-testbed.
[3] F. Buccafurri, A. Comi, G. Lax, and D. Rosaci. The roles of reliability and reputation in competitive multi agent systems. *Lecture Notes in*

Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7566 LNCS (PART 2). Proceedings of On the Move to Meaningful Internet Systems: OTM 2012, pages 326-339, 2012.
[4] C. Burnett, T.J. Norman, and K. Sycara. Sources of stereotypical trust in multi-agent systems. In *Proceedings of the 14th International Workshop on Trust in Agent Societies*, page 25, 2011.
[5] J. Carbo and J.M. Molina-Lopez. An extension of a fuzzy reputation agent trust model (afras) in the art testbed. *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, 14(8):821-831, 2010.
[6] Reputation-based systems: a security analysis, 2012. European Network and Information Security Agency. Available online at http://www.enisa.europa.eu/publications/archive/reputation-based-systems-a-security-analysis.
[7] E. Erriquez, W. Van der Hoek, and M. Wooldridge. Building and using social structures: A case study using the agent art testbed. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(2):25, 2013.
[8] M. Gomez, J. Sabater-Mir, J. Carbo, and G. Muller. Improving the art-testbed, thoughts and reflections. In *Proceedings of the Workshop on Competitive Agents in the Agent Reputation and Trust Testbed at CAEPIA-2007, Salamanca, Spain*, pages 1-15, 2007.
[9] F. Hendrikx, K. Bubendorfer, and R. Chard. Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing*, 75:184-197, 2015.
[10] T.D. Huynh, N.R. Jennings, and N.R. Shadbolt. Certified reputation: how an agent can trust a stranger. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1217-1224. ACM, 2006.
[11] T.D. Huynh, N.R. Jennings, and N.R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2):119-154, 2006.
[12] A. Jøsang and J. Golbeck. Challenges for robust trust and reputation systems. In *Proceedings of the 5th International Workshop on Security and Trust Management (SMT 2009), Saint Malo, France*, 2009.
[13] B. Khosravifar, J. Bentahar, M. Gomrokchi, and R. Alam. CRM: An efficient trust and reputation model for agent computing. *Knowledge-Based Systems*, 30:1-16, 2012.
[14] P. Massa. A survey of trust use and modeling in real online systems. *Trust in E-services: Technologies, Practices and Challenges*, 2007.
[15] M.F. Mohammed, C.P. Lim, and A. Quteishat. A novel trust measurement method based on certified belief in strength for a multi-agent classifier system. *Neural Computing and Applications*, 24(2):421-429, 2014.
[16] J. Murillo and V. Munoz. Agent uno: Winner in the 2007 spanish art testbed competition. In *Workshop on Competitive agents in Agent Reputation and Trust Testbed, The Conference of the Spanish Association for Artificial Intelligence (CAEPIA)*, 2007.
[17] J. Murillo, V. Muñoz, B. López, and D. Busquets. Developing strategies for the art domain. *Current Topics in Artificial Intelligence*, pages 171-180, 2010.
[18] I. Pinyol and J. Sabater-Mir. Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review*, 40(1):1-25, 2013.
[19] D. Rosaci. Trust measures for competitive agents. *Knowledge-Based*

Systems, 28:38–46, 2012.
[20] D. Rosaci, G.M.L. Sarné, and S. Garruzzo. Integrating trust measures in multiagent systems. *International Journal of Intelligent Systems*, 27(1):1–15, 2012.
[21] M. Srikanth and K.B. Madhuri. Secure and effective p2p reputation system using trust management and self certified cryptographic exchanges. *International Journal of Computer Science Issues (IJCSI)*, 10(2), 2013.
[22] W.T. Teacy, T.D. Huynh, R. Dash, N. Jennings, J. Patel, and M. Luck. The art of iam: The winning strategy for the 2006 competition. *The Workshop on Trust in Agent Societies at The Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2007), Honolulu, Hawaii, USA*, 2007.

Francesco Buccafurri is a full professor of computer science at the University Mediterranea of Reggio Calabria, Italy. In 1995 he took the PhD degree in computer science at the University of Calabria. In 1996 he was visiting researcher at the database and knowledge representation group of Vienna University of Technology. His research interests include information security and privacy, social networks, deductive-databases, knowledge-representation and non-monotonic reasoning, model checking, data compression, data streams, agents, P2P systems. He has published more than 140 papers in top-level international journals and conference proceedings. He serves as a referee for international journals and he is a member of a number of conference PCs. Francesco Buccafurri is Associate Editor of Information Sciences (Elsevier), he is included in the editorial board of a number of other international journals, and played the role of PC chair and PC member in many international conferences. He is member of the IEEE computer society. Contact him at bucca@unirc.it.
Antonello Comi is a post-doc of Computer Science at the University Mediterranea of Reggio Calabria, Italy. In 2014, he took the PhD in Information Engineering. His research interests include multi-agent systems, trust and reputation in social communities. Contact him at antonello.comi@unirc.it.
Gianluca Lax is an assistant professor of Computer Science at the University Mediterranea of Reggio Calabria, Italy. In 2005, he took the PhD in Computer Science at the University of Calabria. His research interests include social networks, information security, data streams, user modelling, P2P systems, and e-commerce. Contact him at lax@unirc.it.
Domenico Rosaci is an assistant professor of Computer Science at the University Mediterranea of Reggio Calabria, Italy. In 1999, he took the PhD in Electronic Engineering. His research interests include distributed artificial intelligence, multi-agent systems, trust and reputation in social communities. He has published more than 120 papers in outstanding international journals and conference proceedings. He is a member of a number of conference PCs and he is Associate Editor of Journal of Universal Computer Science (Springer). Contact him at domenico.rosaci@unirc.it.