Using Trustworthy and Referee Agents to Secure Multi-Agent Systems

Jamal Bentahar and Babak Khosravifar
Concordia Institute for Information Systems Engineering, Concordia University, Canada
bentahar@ciise.concordia.ca, b_khosr@encs.concordia.ca

Abstract

Security is a crucial factor in multi-agent systems where agents dynamically leave and enter the system. In this paper, we continue our work on security for agent-based systems by proposing a new trust model allowing agents to decide about a target agent. The model combines two techniques: using reports provided by the trustworthy agents regarding to direct and indirect interaction histories, and using reports provided by the referee agents in the form of recommendations. In addition, after a periodic time, the actual behavior of the target agent is checked against the provided information by others to adjust their credibility. The paper considers different parameters for computing trust, which is represented as a quantitative value.

1 Introduction

During the past couple of years, agent communication languages and protocols have been of much interest in multi-agent systems. Agents are distributed in large scale network and mutually interact to share services with other agents. Therefore trust is essential in effective interactions within open multi-agent systems [4, 14]. In this paper we establish a framework allowing agents to evaluate how much trust they have in one another. Agents interact with each other by direct interaction and get the provided ratings. Inter-agent communication is regulated by protocols (shared amongst agents and thus public) and determined by strategies (internal to agents and thus private). Using this framework, agents are capable of evaluating the trust level of the agents which are not known (or not very well known) by consulting other agents who can provide suggestions about the trustworthiness level of other agents.

Generally in multi-agent systems agents reason using their current knowledge bases before making decisions, and can thus engage in flexible interactions [4]. In addition some of these existing models do not consider limitations in terms of false information provided or lack of information to perform evaluation process. They also do not provide trust propagation through the system. Trust models using direct experience need long term of interaction to reach a stage that agents can evaluate trust level of others. Thus it is important to be able to evaluate the credibility of the witnesses [3, 7, 9]. This is done either by direct experience used to estimate the trust level of these agents or by moving to the second level of evaluation process, asking other agents that are known to be trustworthy about the credibility of the witnesses. However, there is a problem if such *trustworthy agents* are not able to report on the testimony agents. This paper aims at overcoming these limitations by proposing a framework combining the use of *trustworthy agents* and *referee agents* proposed by the target agents. Also using the framework, the *requesting agents* (i.e. the agents requesting information about a target agent) perform maintenance after a period of direct interaction with a new agent in order to adjust the trustworthiness of the consulting agents who provided information regarding to the trust level of the new agent. In the maintenance process, the suggestions provided by other agents are compared with the actual behavior of the new agent in direct interaction. Some agent-based approaches to security exist in the literature, notably the one proposed in [12], but these focus on authentication and authorization issues whereas we focus on trust evaluation and propagation through a social network.

The remainder of this paper is organized as follows. In Section 2, we present the theoretical background, in particular *the trust function* and the notion of *direct trust*. In Section 3, we describe and discuss the details of computing the trust in the combined framework. Section 4 concludes the paper.

2 Background

In our framework, agents are equipped with Beliefs, Desires and Intentions (BDI). They use the BDI architecture when they interact with each other. They have reasoning capabilities allowing them to evaluate their interactions and to decide about the communicative acts to perform during interactions. These reasoning capabilities could be implemented using different theories such as argumentation

978-0-7695-3099-4/08 $25.00 © 2008 IEEE
DOI 10.1109/ITNG.2008.113

[6], game theory, decision theory [11], etc. The purpose of this paper is not to elaborate on these capabilities, but we simply assume their existence. To interact, agents use dialogue games, which are logical rules specifying the communication protocol [5]. Here We define an agent's trust in other agents as a probability function as follows:

Definition **(Trust Function):** Let $\mathcal{A}$ be a set of agents, and $\mathcal{D}$ be a set of domains or topics. The trust function Tr associates two agents from $\mathcal{A}$ and a domain from $\mathcal{D}$ with a trust value between 0 and 1:

$$Tr : \mathcal{A} \times \mathcal{A} \times \mathcal{D} \longrightarrow [0, 1]$$

Given some concrete agents Ag_a and Ag_b in $\mathcal{A}$ and some concrete domain D, $Tr(Ag_a, Ag_b, D)$ stands for "the trust value associated to the target agent Ag_b in domain D by the requesting agent Ag_a". The scale and dynamism of open environment make the participants able to rank each other's reputation level as they keep up interacting. Accordingly if a conflict happens between two agents, this will affect the confidence they have about each other. However, this is only related to the domain $\mathcal{D}$, and not generalized to other domains in which the two agents can trust each other. To simplify the notation, in the remainder we will omit the domain from all the formulas. Given agents Ag_a and Ag_b in $\mathcal{A}$, we will represent $Tr(Ag_a, Ag_b)$ in short as $Tr_{Ag_a}^{Ag_b}$.

In this section we consider the case where agents in the system know each other because they had a prior interaction history and can thus compute the trust value of all agents (and thus the Tr function) *directly*. In general, agents can evaluate the outcomes of their interactions using more flexible values such as "*very good*", "*good*", "*fair*", "*bad*", and "*very bad*". In the general case, they can evaluate their interactions according to a scale of n types numbered from 1 (the most successful interaction) to n (the less successful interaction), such that the first m interaction types ($m < n$) are successful (for example of type "*very good*", "*good*", and "*fair*"). Let $NI_i{}_{Ag_a}^{Ag_b}$ be the number of interactions of type i that Ag_a had with Ag_b. Then Tr can be computed by Equation 1 below as the ratio of the "*number of successful outcomes*" to the "*total number of possible outcomes*".:

$$Tr_{Ag_a}^{Ag_b} = \frac{\sum_{i=1}^{m} w_i NI_i{}_{Ag_a}^{Ag_b}}{\sum_{i=1}^{n} w_i NI_i{}_{Ag_a}^{Ag_b}} \quad (1)$$

where w_i is the weight associated to the interaction type i.

Agents can use several strategies when weighting the interaction types. For example, to minimize the risk of dealing with untrustworthy agents, the weight of "*very bad*" interactions could be higher than the one of "*very good*" interactions. Therefore unsuccessful interactions are more valuable when assessing the agents' trust, and agents should perform well and avoid bad behavior in order to get a better trust value. However, less demanding agents could give the same weight to all interaction types, or give more weight to the "*very good*" and "*good*" interactions.

In open multi-agent systems, agents are known to be autonomous and not know everything about their dynamic environment and there is no central manager for control. Therefore trust is set between two agents which are supposed to interact. Trustworthiness is a dynamic characteristic that changes according to the interactions taking place between two agentsHowever, if the number of interactions with some agent is low (e.g. because the agents has only recently joined the system), agents are not able to compute their trust value directly, but may need to rely upon information provided to them by other agents. Different protocols have been emerged capable of consulting other agents in order to get a better idea about a particular agent's trust level. As proposed in [1, 2, 7], each agent uses two kinds of beliefs: local beliefs and total beliefs. Local beliefs are based on the direct experience of interaction agents. Total beliefs are based on the combination of the different testimonies of other agents that we call *witnesses*. In our framework, local beliefs are given by Equation 1. Total beliefs require studying how different probability measures offered by witnesses can be combined. We use two kinds of witnesses: trustworthy agents that the requesting agent trusts, and referee agents that the target agent suggests to report on his trust. In our framework the suggestions gathered from different types of consulting agents are counted based on the credibility of the agent providing the information.

3 Trust Evaluation in The Combined Framework

Suppose that an agent Ag_a wants to evaluate the trustworthiness of a target agent Ag_b with who he never (or not enough) interacted before. Ag_a may want to consult some other agents to get better and more accurate information about Ag_b's reputation as illustrated in figure 1. The interfering agents are either known by Ag_a to be trustworthy (we call these agents trustworthy agents) or known by Ag_b and have been introduced by him to report on his trust level based on their past experience (we call these agents referee agents). Consequently, we distinguish the community of trustworthy agents from the community of referee agents.

In this model, agents actively interact and find such referees in order to gain the trust of their potential partners. To determine whether an agent is trustworthy, a trustworthiness threshold W must be fixed. Thus, Ag_b will be considered trustworthy by Ag_a if $Tr_{Ag_a}^{Ag_b}$ is higher or equal to W. This threshold can differ depending on type of agent.

Ag_a attributes a trust measure $Tr_{Ag_a}^{Ag_i}$ to each of the agents $Ag_i (i = 1 \dots k)$ she considers *trustworthy*, and a trust measure $Tr_{Ag_a}^{Rf_j}$ to the referee agents $Rf_j (j = 1 \dots k')$

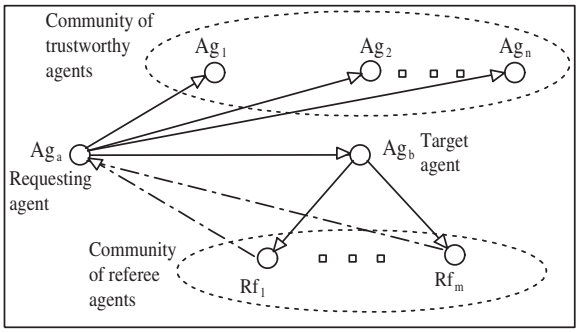

Figure 1. Overall trustworthy and referee agents protocol topology.

she knows. In general, when an (*evaluator*) agent assesses the trustworthiness of another (*evaluated*) agent, the former may consider the latter either trustworthy or untrustworthy depending on the trust measure he assigns to this evaluated agent and some *threshold* fixed by the evaluator. We will define Ag_i (resp. Rf_j) trustworthy by Ag_a when the trust measure $Tr_{Ag_a}^{Ag_i}$ (resp. $Tr_{Ag_a}^{Rf_j}$), given by Equation 1, is greater than a threshold W_a (resp. W_b) fixed by Ag_a. Respectively the referee agent will be considered trustworthy when the trust measure $Tr_{Ag_a}^{Rf_j}$, given by equation 1, is greater than the threshold W_b associated with the recommenders by Ag_a. There is a possibility that the referee agent is not that reliable since the referee agent can collude with agents that intends to support by providing some falsely references for them. On the other hand, Ag_b chooses its referee agents to put forward and a rational agent only presents its best ones. Therefore Ag_a can expect some exaggerated information regarding to Ag_b's creditability.

We assume that consulting agents Ag_i and Rf_j also use equation 1 to assess the trust value of the agents they know, and in particular Ag_b. Thus, the problem consists in Ag_a evaluating Ag_b's trust measure combining the trust values transmitted by trustworthy agents and referee agents to Ag_a. Once this value is computed, Ag_a decides to consider Ag_b trustworthy or not depending again on the threshold W_a.

3.1 Evaluation 1

We notice that this problem cannot be formulated as a problem of conditional probability. Consequently, it is not possible to use *Bayes' theorem or total probability theorem either*. The reason is that events in our problem are not mutually exclusive, whereas this condition is necessary for these two theorems. To solve this problem, we must investigate the distribution of the random variable X representing the trustworthiness of Ag_b. Since X takes only two values: 0 (the agent is not trustworthy) or 1 (the agent is trustworthy), variable X follows a *Bernoulli distribution* $\beta(1, p)$. Accordingly $E(X) = p$ where $E(X)$ is the expectation of the variable X and p is the probability that the agent is trustworthy. Here, p is the probability we seek. Therefore it is enough to evaluate the expectation $E(X)$ to find $Tr_{Ag_a}^{Ag_b}$. However, this expectation is a theoretical mean that we must estimate. To this end, we can use the *Central Limit Theorem (CLT) and the law of large numbers*. The CLT states that whenever a sample of size n $(X_1, \ldots, X_n)$ is taken from any distribution with mean μ, then the sample mean $(X_1 + \cdots + X_n)/n$ will be approximately normally distributed with mean μ. As an application of this theorem, the arithmetic mean (average) $(X_1 + \cdots + X_n)/n$ approaches a normal distribution of mean μ, the expectation and standard deviation $\sigma/\sqrt{n}$. Generally, and according to *the law of large numbers*, the expectation can be estimated by the weighted arithmetic mean.

Our random variable X is the weighted average of n independent variables X_i that correspond to Ag_b's trust level according to the point of view of trustworthy agents Ag_i and referee agents Rf_j. These variables follow the same *Bernoulli distribution*. Consequently, the variable X follows a normal distribution whose average is the weighted average of the expectations of the independent variables X_i. The mathematical estimation of expectation $E(X)$ is given by the following equation:

$$M = \frac{\sum_{i=1}^{n} \left(Tr_{Ag_a}^{Ag_i}.Tr_{Ag_i}^{Ag_b}\right) + \sum_{j=1}^{m} \left(Tr_{Ag_a}^{Rf_j}.Tr_{Rf_j}^{Ag_b}\right)}{\sum_{i=1}^{n} Tr_{Ag_a}^{Ag_i} + \sum_{j=1}^{m} Tr_{Ag_a}^{Rf_j}} \quad (2)$$

The number of requested references (m) is defined by Ag_a and is related to the number of trustworthy agents Ag_a has (n) in order to ensure that enough number of third parties have been involved to participate in evaluation. Although Ag_b may cannot (or refuse to) provide the requested number of referees. That would be a bit affective in his trust evaluation process.

In order to introduce its referee agents, Ag_b forwards Ag_a's information to each one of referees he wants to introduce. Referee agents reply directly to Ag_a the trust level of Ag_b regarding to their past experience of direct interaction with Ag_b. Categorizing the referee agent, there are three possibilities: (1)if the referee agent is also a trustworthy agent of Ag_a, which gets more priority; (2)if Ag_a knows the referee agent, then it considers the referee agent's suggestion by including his trustworthiness value based on previous reputation that she has made and (3)if Ag_a does not know the referee agent; here Ag_a can adopt different policies, she can just accept the referee agents to whom she had direct experience and thus the corresponding assigned trustworthiness value is used. In this policy, using direct

trust estimate in equation 1, such referee agents of no interaction history with Ag_a will automatically removed. However Ag_a may take the policy of assigning a default value η gained by overall reputation of such referee agents and start over. The value η is specific for each referee agent. Therefore we can advance equation 1 in order to make the consulting agents' trust estimation more flexible:

$$Tr_{Ag_a}^{Rf_j} = \begin{cases} \frac{\sum_{i=1}^{m} w_i NI_i{}_{Ag_a}^{Rf_j}}{\sum_{i=1}^{n} w_i NI_i{}_{Ag_a}^{Rf_j}}, & \text{n>0;} \\ \eta, & \text{n=0.} \end{cases} \quad (3)$$

In this equation if the assigned value η is still 0, that means Ag_a does not know the introduced referee agent at all. In this case Ag_a does not consider his suggestion about Ag_b, but it saves the referee's suggestion anyway in order to compare by the real behavior Ag_b performs after starting interaction with Ag_a.

3.2 Evaluation 2

The estimation M in equation 2, however, does not take into account the *number of interactions* between the trustworthy/referee agents and Ag_b. These numbers are important factors because they promote information coming from agents knowing more about Ag_b. The agents who had high number of interactions by Ag_b are considered as good sources of information about his trustworthiness. In addition, another factor might be used to reflect the *timely relevance* of transmitted information. This is because the agent's environment is dynamic and may change quickly. The idea is to promote recent information and to deal with out-of-date information with less emphasis. The timely relevance could be represented as a coefficient when computing the agent's trust. In our model, we assess the factor $TR(\Delta t_{Ag_i}^{Ag_b})$ by using the function defined in equation 4.

$$TR(\Delta t_{Ag_i}^{Ag_b}) = e^{-\lambda\ ln(\Delta t_{Ag_i}^{Ag_b})} \quad (4)$$

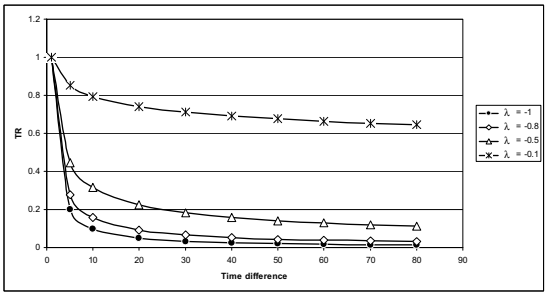

Figure 2. Timely relevance behavior.

Δt is the time difference between the current time and the time at which Ag_i updates her information about Ag_b's trust. λ is an application-independent coefficient. The intuition behind this formula is to use a function decreasing with the time difference like what is shown in figure 2. Consequently the more recent the information is, the higher is the timely relevance coefficient. The function is used for the computational reasons when dealing with large numbers. In fact, this function is similar to the well known *reliability function* for systems engineering ($R(t) = e^{-\lambda t}$).

Equation 5 gives us an estimation of $Tr_{Ag_a}^{Ag_b}$ if we take into account the explained factors. This equation is composed of values got from two different consulting communities involved in trust evaluation. The function $\Omega(S,k)$ is defined as the summation of the trust values estimated by the k agents type S *(trustworthy/referee)* together with their related self-trustworthiness (from Ag_a's point of view), timely relevance and the number of interactions done by the target agent Ag_b; This number can be identified by the total number of Ag_b's commitments and arguments. Respectively $\Omega'(S,k)$ represents the similar function without the Ag_b's credibility from agent S point of view.

$$Tr_{Ag_a}^{Ag_b}(\Delta t) = \frac{\Omega(Ag,n) + \Omega(Rf,m)}{\Omega'(Ag,n) + \Omega'(Rf,m)} \quad (5)$$

$$\Omega(S,k) = \sum_{i=1}^{n} Tr_{Ag_a}^{S} \times Tr_{S}^{Ag_b} \times TR(\Delta t_{S}^{Ag_b}) \times N_{Ag_b}^{S} \quad (6)$$

$$\Omega'(S,k) = \sum_{i=1}^{n} \left(Tr_{Ag_a}^{S} \times N_{Ag_b}^{S} \times TR(\Delta t_{S}^{Ag_b}) \right) \quad (7)$$

There is a possibility that the referee agent is untrustworthy from Ag_a's point of view. That means Ag_a would rate a negative value to Rf_j as a result of bad past direct experience Ag_a had with Rf_j. Consequently Rf_j's suggestion could not be helpful in the evaluation process of Ag_b and thus Ag_a may decide not to consider it at all (set flag F to 0). However it would be reasonable if Ag_a also considers Rf_j's suggestion about Ag_b and consequently decreases the Ag_b's credibility as he introduced a bad recommender (set flag F to 1). Therefore we advance equation 1 by replacing the $Tr_{Ag_a}^{Rf_j}$ by a new overall trust evaluation value $Tr''^{Rf_j}_{Ag_a}$. Equations 8 and 9 identify the value of $Tr''^{Rf_j}_{Ag_a}$ as follows:

$$Tr''^{Rf_j}_{Ag_a} = \begin{cases} Tr_{Ag_a}^{Rf_j}, & Tr_{Ag_a}^{Rf_j} > 0; \\ Tr'^{Rf_j}_{Ag_a}, & Tr_{Ag_a}^{Rf_j} < 0. \end{cases} \quad (8)$$

$$Tr'^{Rf_j}_{Ag_a} = \begin{cases} 0, & F\text{=0;} \\ Tr_{Ag_a}^{Rf_j} \times F, & F\text{=1 and } |Tr_{Ag_a}^{Rf_j}| > \lambda; \\ -\lambda \times F, & F\text{=1 and } |Tr_{Ag_a}^{Rf_j}| < \lambda. \end{cases} \quad (9)$$

Here the $Tr''^{Rf_j}_{Ag_a}$ would be $Tr_{Ag_a}^{Rf_j}$ itself if the referee agent has a positive credibility from Ag_a's point of view, but we would give $Tr'^{Rf_j}_{Ag_a}$ value to the overall referee's trust level

if it has a negative credibility in Ag_a's side. Consequently $Tr'^{Rf_j}_{Ag_a}$ would be evaluated by three cases. If the flag F is set to 0, that means the negative creditable referee agents would not be considered, but if the flag is 1, Ag_a will consider the referee agent's suggestion and respectively overall it would make a negative value. That will affect the trust evaluation of the Ag_b while it would be decreased. Ag_b will get this penalty because he introduced a bad referee agent. But we define a limit here that guarantees a least penalty an agent would get in such cases. That means if the absolute value was less than a threshold λ then overall it would not affect that much and agents may not care to introduce bad agents. Then in such cases the value λ would be considered to provide the appropriate negative trust measurement.

We denote here that removing $TR(\Delta t^{Ag_b}_{Ag_i})$ from the Equation 5, results in the classical probability equation used to calculate the expectation $E(X)$. Equation 5 takes into account the four most important factor: (1) the trustworthiness of trustworthy/referee agents according to the point of view of Ag_a ($Tr^{Ag_i}_{Ag_a}$ and $Tr^{Rf_j}_{Ag_a}$); (2) the Ag_b's trustworthiness according to the point of view of trustworthy/referee agents ($Tr^{Ag_b}_{Ag_i}$ and $Tr^{Ag_b}_{Rf_j}$); (3) the number of interactions between these trustworthy/referee agents and Ag_b ($TN^{Ag_i}_{Ag_b}$ and $TN^{Rf_j}_{Ag_b}$) and (4) the timely relevance of information transmitted by trustworthy/referee agents ($TR(\Delta t^{Ag_b}_{Ag_i})$ and $TR(\Delta t^{Ag_b}_{Rf_j})$), as communicated by Ag_i to Ag_a following the strategies previously indicated. This equation shows how trust can be obtained by merging the trust values transmitted by consulting agents. This merging method takes into account the proportional relevance of each trust value, rather than treating them equally. Generally in trust evaluation we try to minimize the adverse affects the consulting agents may produce. For instance, two agents who have a strong relationship can support each other and overestimate their trust level when they have been introduced as referee agents. That implicitly means Ag_a can expect a more accurate suggestion from a referee agent who had a large number of interactions with Ag_b comparing to one who had less. Thus we should try to give more emphasis to such agents that previously had large number of interactions by Ag_b in terms of accepting their idea. Respectively these agents should affect more when the opposite of their suggestion turned out to be true. Therefore Ag_a needs to perform a maintenance to adjust the consulting agents credibility. In the following we describe the maintenance procedure done by Ag_a after a certain period of time the interaction has been initialized by Ag_b.

3.3 Maintenance

Generally Ag_a is more confident about its trustworthy agents as they have shown an acceptable trustworthiness so far, but the referee agents are chosen by Ag_b, so we should always consider the possibilities like the cooperating partners may vote in favor of each other or competing agents may underrate their opponents. Therefore Ag_a after a period of interacting with Ag_b performs maintenance in order to evaluate the witness reputation to assess the consulting agents' trust level. In trust evaluation process done by Ag_a maybe there were some referee agents involved in but their suggestion were not took into account as they were not known by Ag_a and consequently not eligible to interfere. But Ag_a did not discard their suggestion; after the maintenance Ag_a would be able to estimate such referee agents' credibility as long as they are known (because of their referee history) by Ag_a from now on. The rational behind this maintenance is to compare the actual behavior Ag_b performed after starting interaction with Ag_a with the suggestions provided about Ag_b's credibility by others.

Now if Ag_a received a reference from referee agent Rf_j (this accuracy check is not just specified to referee agents; trustworthy agents are also checked), Ag_a then adjusts Rf_j's trust level by comparing the actual performance of Ag_b, as a result of a period of direct interaction experience, and what Rf_j provided as the suggested trust level for Ag_b. Therefore thresholds ν_T and ν_R are associated as inaccuracy tolerance thresholds for the trustworthy and referee agents. We assign two different thresholds because the referee agents were supposed to deliver a more accurate information about Ag_b comparing to the trustworthy agents because they have been introduced by Ag_b. By doing so, if the difference is greater than the associated threshold for the agent, the consulting agent's trust level should be dropped to some extent, otherwise it will be enhanced regarding to the importance of suggestion provided, this value ϕ is defined by Ag_a. The important thing here is the ratio of dropping down the trustworthiness value of the consulting agent after the comparison. Let us assume the trust level assigned by Ag_a to Ag_b is $Tr^{Ag_b}_{Ag_a}$ and the value provided by the referee agent is $Tr^{Ag_b}_{Rf_j}$. Therefore Ag_a adjusts the trustworthiness level of the agent Rf_j by the following equation:

$$Tr^{Rf_j}_{Ag_a} = \begin{cases} Tr^{Rf_j}_{Ag_a} - N^{Rf_j}_{Ag_b} \times D_R, & \text{if } D_R > \nu_R; \\ Tr^{Rf_j}_{Ag_a} + \phi, & \text{if } D_R < \nu_R. \end{cases}$$

$$D_R = |Tr^{Ag_b}_{Rf_j} - Tr^{Ag_b}_{Ag_a}|; \tag{10}$$

The value D_R defines the inaccuracy of the trust level regarding to the specific consulting agent. The inaccuracy is checked by the predefined threshold (here for referee agents ν_R) to recognize whether the referee agent's suggestion was apart from the real value. If Ag_a, after comparison, considers the referee agent trustworthy, it increases the current trust level by the value ϕ, otherwise it decreases the trust level by the ratio related to the corresponding number of interactions done by the referee agent Rf_j and Ag_b. The

number of interaction is used as a measure here as we assume the higher number of interactions, the more accurate information supposed to perform, consequently the more decrease when wrong information is provided. Therefore having recorded the ratings provided by agent R about other agents, Ag_a can evaluate or adjust R's credibility after performing maintenance and checking the differences. Obviously the ratio of adjustment is not very high and affective to the trustworthiness agents as they were not supposed to know Ag_b and thus provide an accurate information about him. However there may be a good increase in the trust level provided by Ag_a to the consulting agents regarding because of their accuracy in providing information about Ag_b.

4 Conclusion

The contribution of this paper is the proposition of a new probabilistic and statistic-based model to secure communication-based multi-agent systems. A trustworthiness and referee agent framework have been presented, as well as several models, of increasing sophistication, for agents to make use of the information communicated to them by other agents they consider trustworthy to determine the trust of further target agents. Our model has the advantage of being computationally efficient and of taking into account four important factors: (1) the trust (from the viewpoint of the evaluator agents) of the trustworthy agents; (2) the trust value assigned to target agents according to the point of view of trustworthy agents; (3) the number of interactions between trustworthy agents and the target agents; and (4) the timely relevance of information transmitted by trustworthy agents. Another contribution of this paper is maintenance activities that agents perform in order to evaluate the consulting agents' trust level by comparing the provided information regarding to the target agent's trust level and the actual behavior of the target agent since it has started interaction. The resulting model allows us to produce a comprehensive assessment of the agents' credibility in a software system.

References

[1] A. Abdul-Rahman, and S. Hailes. Supporting trust in virtual communities. In Proc. of the 33rd Hawaii Int. Conf. on System Sciences, IEEE Computer Society Press. 2000.

[2] F. Azzedin and M. Maheswaran. A trust brokering system and its application to resource management in piblic-resource grids. In Proc. of the 18th International Parralel and Sistributed Processing Symposium (IPDPS'04) pp. 22-31, 2004.

[3] J. Bentahar, F. Toni, J-J. Ch. Meyer and J. Labban. A security framework for agent-based systems. In the International Journal of Web Information Systems 3(4):341-362, Emerald, 2007.

[4] J. Bentahar and J-J. Ch. Meyer. A new quantitative trust model for negotiating agents using argumentation. In the International Journal of Computer Science and Applications,4(2):1-21, 2007.

[5] J. Bentahar, Z. Maamar, D. Benslimane, and Ph. Thiran. An argumentation framework for communities of web services, IEEE Intelligent Systems, 22(6), 2007.

[6] P. M. Dung, P. Mancarella and F. Toni. Computing ideal sceptical argumentation. Artificial Intelligence, Special Issue on Argumentation in Artificial Intelligence, 2007.

[7] T. Dong-Huynh, N. R. Jennings and N.R. Shadbolt. Certified reputation: How an agent can trust a stranger. In Proc. of The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 1217-1224, Hakodate, Japan, 2006.

[8] T. Dong-Huynh, N. R. Jennings and N.R. Shadbolt. Fire: An integrated trust and reputation model for open multi-agent systems. Journal of Autonomous Agents and Multi-Agent Systems 13(2) pp. 119-154, 2006.

[9] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. Journal of Autonomous Agents and Multi-Agent Systems AAMAS, 2006, 119-154.

[10] E. M. Maximilien, and M. P. Singh. Reputation and endorsement for web services. ACM SIGEcom Exchanges, 3(1):24-31, 2002.

[11] S. Parsons, P. J. Gmytrasiewicz, and M. J. Wooldridge (Eds.). Game Theory and Decision Theory in Agent-Based Systems. Springer, 2002.

[12] E. Shakshuki, L. Zhonghai, and G. Jing. An agent-based approach to security service. International Journal of Network and Computer Applications. Elsevier, 28(3): 183-208, 2005.

[13] Y. Wang, and M. P. Singh. Formal trust model for multiagent ststems. Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI), pp. 1551-1556, 2007.

[14] P. Yolum and M. P. Singh. Engineering self-organizing referral networks for trustworthy service selection. IEEE Transaction on systems, man, and cybernetics, 35(3):396-407, 2005.