Which Adequate Trust Model for Trust Networks?

Dimitri Melaye[1], Yves Demazeau[1], and Thierry Bouron[2]

[1] IMAG - Laboratoire Leibniz,
46, avenue Felix Viallet, 38031 Grenoble Cedex, France,
dimitri.melaye@imag.fr, yves.demazeau@imag.fr
[2] FTR&D,
38-40 rue du Général Leclerc 92794,
Issy les Moulineaux cedex 9 PARIS, France,
thierry.bouron@francetelecom.com

Abstract. This article deals with the choice of individual trust models adapted to networks. We consider trust as a social and effective multi-agent process. We introduce the notion of trust networks viewed as a set of one-to-one trust relationships, we wonder which trust model should be chosen to build and exploit it. We extract five criteria for comparison of trust models. We then evaluate two trust models and discuss what could be a relevant trust model in a multi-agent setting.

1 Introduction

The expansion of the distributed systems such as electronic trade or services for citizens highlights new problematics where trust plays a crucial role [1] . Trust has been recently identified as an essential notion in the business to business applications where the relationship are supported electronically in an open environment. It is viewed as "the subjective probability by which an agent expects that another agent performs a given action on which its welfare depends" [2].

Following Castelfranchi and Falcone's work [3] and Demazeau's VOWELS methodology [4], a trust process can be based on three of the four vowels of the VOWELS methodology: A (Agent), I (Interaction), E (Environment) and O (Organisation). We distinguish the trust model and the decision-taking. The trust model computes a trust level between two agents, and is function of I and E. The decision-taking is agent-centered (A) and is grounded on the *trust level* and on the *risk* taken by relying (typically, a decision-taking with a high risk requires a high trust level).

The aim of our work is to consider trust as a social process and to pass from individual trust to social networks. Thus, an agent will be able to take into account its social and organizational resources. It necessitates to consider the fourth vowel of [4], the organization (O). For this purpose, we introduce reasoning about the other agents' trust in the previous trust process. Reasoning is based on a *trust network* that can be viewed as a set of trust one-to-one models. To build and exploit the *trust networks*, the choice of the one-to-one trust model is crucial for the implementation of the trust networks. In [5], we have exhibited several characteristics of a trust model, among which the number of the considered relationships (one-to-one, n-to-one and n-to-p). In

Please use the following format when citing this chapter:
Melaye, Dimitri, Demazeau, Yves, Bouron, Thierry, 2006, in IFIP International Federation for Information Processing, Volume 204, Artificial Intelligence Applications and Innovations, eds. Maglogiannis, I., Karpouzis, K., Bramer, M., (Boston: Springer), pp. 236–244

this paper, we propose a focus on the comparison of one-to-one models for the trust networks. We concentrate on the relation between the trust model, the trust network, and the decision-taking.

First we introduce the notion of trust networks inspired by Sichman's networks [6]. We assume a trust network is built from one-to-one models. We then derive and formalize five criteria for its building and its exploitation: observability, understandability, handlability, social exploitability, and context-scalability. We finally evaluate two trust models regarding these criteria, and discuss what could be a relevant trust model.

1.1 Motivation and principles

A multi-agent system cannot be reduced to the simple sum of its agents: social reasoning and environment have a crucial role in the coordination and the interactions between cognitive agents. Indeed, it is necessary to evaluate the agents' social resources. Thus, trust should be considered as a network of relationships rather than independent one-to-one relationships. But usually, MAS contribution only consider one-to-one trust models.

In order to pass from individual trust to social networks, we have taken an inspiration from Sichman's dependence networks [6], where we have replaced dependence with trust. Each agent can compute the supposed trust relationship between the other agents. The set of these relationships makes up *its trust network*. We assume these trust relationships are. As an agent's trust network is its *subjective* representation of the trust relationships, it generally different of the "objective" trust network that emerges from the effective trust relationships between the agents of the system.

1.2 Formalization

Formally, an agent is characterized by an external description: each agent has structured information about the other agents' trust. These information are private and own.

Definition 1. *The agent's trust network is the set of the trust model considered by it between the other agents. Formally, let N_i be the number of agents considered by an agent i. Let $\mathcal{T}_i$ be its trust network. We have $\mathcal{T}_i = \{M_{ijk} \mid 1 \leqslant j \leqslant N_i,\ 1 \leqslant k \leqslant N_i,\ j \neq k\}$, where M_{ijk} is the trust model considered by the agent i between the agent j and k.*

$\mathcal{T}_i$ can be viewed as the oriented graph $G_i(X_i, E_i)$ where X_i is the set of the N_i nodes (agents known by i), and E_i is the set of the edges (relationships between these agents). Each edge e_{ijk} of this graph is valuated by the trust model considered by the agent i between the agent j and the agent k. The update of the trust models and the update of the graph are independent: the latter is updated only at the time of its use.

Such a net synthesizes the supposed trust relationships between the agents. In the framework of a social multi-agent system, the agents do not always interact each other directly. So some intermediary agents are often necessary to satisfy their goals. For instance, concerning the bottom-up coalition formation, even if an agent must consider

the trusted agents, it should also consider the agents that trust in itself and that can rely on other agents. Thus, it is not sufficient to take into account trust only at an agent-centered level, and it is necessary to consider the others' trust.

1.3 Which one-to-one trust model do we need?

We have assumed that an agent's trust network can be viewed as a set of one-to-one models of trust connecting every agents. Therefore, technically, which trust model to choose to derive the agent's trust and to take some decisions? In the remainder of this article, we precise the necessary properties of each M_{ijk} trust model. We propose five criteria to compare one-to-one trust models: observability, understandability, handlability, social exploitability, and context-scalability. These criteria are based on three of the four vowels of the VOWELS methodology [4]: A (Agent), I (Interaction), and E (Environment). The I vowel corresponds to the building of the trust network from the trust models. The A vowel corresponds to the social exploitability of the trust models by the agents. The E vowel corresponds to the problem of scalability of our model: it is linked up with the property of context-sensitivity of trust. The fourth vowel, the organization (O), is supported by the trust networks themselves, and is not used actually yet.

Interaction: observability, understandability and handlability As trust is built from interactions and observations [7], the interaction vowel is used for to the building of the trust networks. For this purpose, we distinguish three criteria: observability, understandability and handlability.

Observability In order to build its trust network, an agent should be able to derive the trust relationships that exist between the other agents. On the contrary of Fullam and Barber's evaluation [8], the important thing is not that an agent i is able to determine the target agent k's objective trustworthiness, but that it is able to determine the agent j's trust in k. For this purpose, either the agent must be able to use the observations of the interactions between the agent j and the agent k (direct observation or indirect observation if it does not take part in ones), or the agent must be able to use other agents' opinion. These both sources of trust correspond to the Observer role and to the Evaluator role in Mulller's work [9]. Let us notice the observations should respect the privacy of the interactions: they can be only some "clues".

Proposition 1. *Let $\mathcal{A}_i$ the set of the agents knwon by the agent i. Let M_{ijk} be the trust model built by the agent i between the agent j and the agent k. The criterion of observability is fulfilled iff M_{ijk} is equipped with a function $observability_{ijk}$:*

$$observability_{ijk} : \{O_{njk}\}_{n,j,k \in \mathcal{A}_i} \times \{E_{njk}\}_{n,j,k \in \mathcal{A}_i} \rightarrow T$$

where

- *T is the set of the final trust levels.*

- $\{O_{njk}\}_{n,j,k\in\mathcal{A}_i}$ *is the set of the observations reported by the agents* n *about the relationship between* j *and* k. *When* $n = j$, *it corresponds to a direct observation. When* $n \neq i$, *it corresponds to an indirect observation.*
- $\{E_{njk}\}_{n,j,k\in\mathcal{A}_i}$ *is the set of the final trust values between* j *and* k *according to* n. *In other words, it is* n*'s evaluation about agent* j*'s trust in the agent* k. *If* $n = j$, *the agent* n *corresponds to a trust third party.*

Understandability and handlability The model should not be reduced to a final trust level obtained by an obscure mechanism. Since trust consists in *beliefs* [3], trust should be broken up into beliefs corresponding to the different components of trust (e.g. personal, reputation, generalization or competence dimension). The aim of this breaking up is to be able to have different computable components to reason about them at the network level. Indeed, these components correspond to orthogonal networks: for instance, let us consider the framework of online auctions of specialized objects. Non-specialized occasional buyer's trust is based on the generalization and/or reputation dimension. What is more, these agents trust in specialists' opinions easily. However, it is not the case of the specialized agents: they belong to the same dense network of specialists, and do not trust in other specialists blindly. In fact, their trust is supported by the personal dimension. Thus, according to the agents, the basis of trust is not the same and is supported by different dimensions. Moreover, if a non-specialized agent specializes, trust between it and the other specialized agents switches from generalized trust to personal trust.

Proposition 2. *Let be* $\mathcal{B}_i$ *the set of the possible states of agent* i*'s beliefs. The criterion of understandability is fulfilled iff the function* $observability_{ijk}$ *is the composition of two function* g_{ijk} *and* $understand_{ijk} : \{O_{njk}\}_{n,j,k\in\mathcal{A}_i} \times \{E_{njk}\}_{n,j,k\in\mathcal{A}_i} \rightarrow \mathcal{B}_i$ *such that:*

$$observability_{ijk} = g_{ijk} \circ understand_{ijk}$$

Proposition 3. *Let be* $\mathcal{B}_i$ *the set of the possible states of agent* i*'s beliefs. The criterion of handlability is fulfilled iff the function* $observability_{ijk}$ *is the composition of two function* $handle_{ijk} : \mathcal{B}_i \rightarrow T$ *and* h_{ijk} *such that:*

$$observability_{ijk} = handle_{ijk} \circ h_{ijk}$$

We can notice the criteria of understandability and handlability are satisfied both iff $observability_{ijk}$ is the composition of $understand_{ijk}$ and $handle_{ijk}$.

Agent: social exploitability The final aim of a trust network is to take a decision by taking into account the other agents, and not only one. This trust decision, called "reliance", should end in a social action [6] based on several other agents.

Proposition 4. *Let* T_i^n *be the set of the trust levels of the* n *agents known by the agent* i *and* A_i *be the set of the agent* i*'s possible actions. The criterion of social exploitability is fulfilled iff it exists the function* $social\_exploitability_i$ *such that:*

$$social\_exploitability_i : T_i^n \rightarrow A_i \text{ with } |n > 1|.$$

Environment: context-scalability Trust is a contextual notion. It refers to a given context. Thus, in a naïve approach, an agent should have one trust network per context. However, it can lead to a too heavy system, since the potential number of models used by an agent's trust network increases quadratically with the number of the considered agents. The contexts should be handled in a way not to have one model per context. In this purpose, we should have a mechanism to derive a model for a given context from a model for another context . It permits to reduce the number of model and to define their interdependences.

Proposition 5. *Let $\mathcal{W}$ be the set of the possible contexts, and M^{ω}_{ijk} be the trust model of the agent i concerning j's trust in k for a context $\omega \in \mathcal{W}$. The criterion of context-scalability is fulfilled iff it exists the function $context\_scalability_i$ permitting to pass from some known n contexts to another context:*

$$context\_scalability_i : \mathcal{M}_i^n \rightarrow \mathcal{M}_i$$
$$context\_scalability_i(M^{\omega_1}_{ijk}, M^{\omega_2}_{ijk}, ..., M^{\omega_n}_{ijk}) = M^{\omega_{n+1}}_{ijk}$$

We have described the five criteria permitting to evaluate one-to-one trust models adapted to social networks. They must account for their relevance in an effective multi-agent context. In the remainder of this article, we put them into practice by examining and discussing three trust models.

2 Model examination

We have chosen two trust models for evaluation: Sabater and Sierra's ReGreT system [10], and Castelfranchi and Falcone's model [3]. They permit an agent to determine trust in another agent, so that they are good candidates for the building of a trust network. According to our social view of trust, we have deliberately avoided the game theory-oriented systems to concentrate on more socio-cognitive and reputation trust models. The choice has been fixed on a set of models that are representative of different views of trust and that are completed enough to be evaluated and implemented actually. Due to lack of space, we evaluate only two models and we do not detail them to concentrate on the relevant points for our evaluation. In the following sections, we evaluate these models according to the previous criteria. We summarize te result of each evaluation in a table.

2.1 The model by Sabater and Sierra [10]

The ReGreT system may be one of the most complete systems. The model is based on three dimensions of trust that are integrated numerically by weighting to obtain a reputation measure: the individual dimension, the social dimension (witness, neighbourhood, and system reputation) and the ontological dimension. The individual dimension deals with direct interactions. The social dimension uses other agents' information. The ontological dimension supports the context notion.

Observability. Although nothing is evoked about the exterior observability of the interactions, the function $observability_{ijk}$ is implementable (observations and evaluations). However, the strong restriction is that the agents can exchange information only coming from their own direct observations or interactions: formally, the sets $\{O_{njk}\}_{n,j,k \in A_i}$ and $\{E_{njk}\}_{n,j,k \in A_i}$ are defined with the restriction $n = j$.

Understandability and handlability. With this model, the function $observability_{ijk}$ is decomposable: $understand_{ijk}$ breaks up trust into three dimensions (individual, social, ontological) and $handle_{ijk}$ is implemented by a simple numerical combination of the values of these dimensions.

Social exploitability. The model is exploited with a negotiation model in a one-to-one framework: the decision-taking is based on a function $rely$ that takes into account only one agent ($social\_exploitability_i : T_i^1 \rightarrow A_i$)

Context-scalability. With its ontological dimension, this model proposes a naïve implementation of the function $context\_scalability_i$: the reputation in a given context is obtained by the numerical combination of the reputation of the less general context, according to a ontological structure.

Table 1. Summary of the examination of Sabater and Sierra's model [10].

	Observ.	Underst.	Handl.	Social expl.	Context-scal.
Evaluation	+	++	+	-	+

The examination is summarized in the table 2. This model is relevant for three criteria: understandability, handlability, and context-scalability. It is interesting for the observability criterion too.

2.2 The model by Castelfranchi and Falcone [3]

This model contrasts with the previous models, because it assumes the agents can be irrational, that is to say they had not to maximize their utility. Moreover, on the contrary of the previous models based on quantitative approaches, this cognitive model is based on the BDI approach, and highlights the importance of the intention in the trust process.

Observability. In this model, the observability is not evoked.

Understandability and handlability. Based on a cognitive structure of trust, [3] implementes this model with cognitive fuzzy map. Indeed, the function $understand_{ijk}$ is implemented: the model computes a set of belief values (ability, willingness, dependence) derived from the sources of trust. The function $handle_{ijk}$ is implemented too by computing a trust level from these beliefs.

Social exploitability. The decision-taking is based on a function $social\_exploitability_i$ that takes into account only one one agent ($social\_exploitability_i : T_i^1 \rightarrow A_i$)

Context-scalability. The problem is not asked.

The examination is summarized in the table 3. This model is relevant for two criteria: understandability and handlability.

Table 2. Summary of the examination of Castelfranchi and Falcone's model [3].

	Observ.	Underst.	Handl.	Social expl.	Context-scal.
Evaluation	-	++	++	-	-

3 Discussion

As we have seen in the previous section, the recent research in trust for MAS lead to many interesting and different models. The common characteristic of the examined models is their decentralized approach. It seems it is established that the centralized systems of trust are not relevant in a multi-agent context: the experience showed these systems such as the Ebay system where the opinions are centralized are perfectible. According to our criteria, the results of our examination are quite disappointing. No model is really adapted to pass from individual trust to social networks. In fact, all these trust models are not thought out as a multi-agent mechanism. They correspond to an agent-centered mechanism above all, even if some multi-agent aspects (like reputation) are plugged.

3.1 About the measures

Observability. The models do not fulfil the criteria of observability. It is not surprising since their purpose is only to provide an agent with a way to determine the true trustworthiness of other agents (on the contrary of our need: a model that can determine the subjective agent's trust in a target agent). It is a drawback, because observability is necessary to build the trust networks. Sabater and Sierra proposes a similar mechanism to communicate the evaluations. But they are limited to the agents' own experiences too: it is a drawback for our trust networks, because it limits the sources of trust (it would be as if there would be no reputation mechanism in an agent-centered approach).

Understandability and handlability. These both criteria are fulfilled by the ReGreT system[10] and Castelfranchi and Falcone's model [3]. However, there is a difference in their implementation: the computed beliefs in Sabater and Sierra's model correspond to the sources of trust in Falcone and Castelfranchi's. Indeed, the computed beliefs in [3] correspond to a more abstract concept. We show the main difference between these both models: Falcone and Castelfranchi's model is based on a cognitive approach, whereas Sabater and Sierra is not. In the next section, we will precise our interest for such a model.

Social exploitability. The criteria is the less fulfilled one. The models exploit trust only for one target agent. One more time, they are not thought out as a multi-agent mechanism.

Context-scalability. Only Sabater and Sierra's model proposes a relevant mechanism for the management of the context. Generally, the context aspects are not studied in the one-to-one trust model, since they do not correspond to a need of scalability: the scalability is easier in an agent-centered framework.

3.2 Choosing individual trust model for networks

According to our examination, there is no trust model which we can use as it is for our purpose. However, we can extract two avenues of research in order to have a model fulfilled the criteria.

First the model should be based on a symbolic approach. Indeed, these latter are more relevant for our trust networks, since they support reasoning and explicit handling of trust components. Thus, they are more adapted to the fulfilment of the criteria of understandability and context-scalability. Liau's model [11] have proposed a trust model based on modal logic that formalizes the internal components of trust. Although this model does not include social aspects and does not seem have been implemented, the modal logic approach is an interesting avenue to fulfil the criteria of understandability and context-scalability.

Secondly we are interested in a Castelfranchi-inspired model [3]. The breaking up of trust and the social approach of this model is adapted to our real multi-agent view. We think the fulfilment of the observability and social exploitability criteria necessitates such an approach, since the "others" are taken into account. The main critic about this socio-cognitive model deals with the problem of the rationality of the agents [12]: Castelfranchi's model would be motivated from humans and would not adapted to a rational approach. Although this critic is pertinent in a pure artificial system, it is not relevant in a multi-agent system that could include the user. Now, as we evoked in introduction, we are interested in such systems and we need a model that is generic enough to support the potential irrationality of the user.

4 Conclusion

This article has presented a study of trust models for *trust networks*. By regarding trust as a social and effective multi-agent process, we have introduced the notion of trust networks. An agent's trust network is its subjective representation of the trust relationships between the agents of the system. It can be considered as a set of one-to-one trust relationships, so we have wondered which trust model to choose. We have derived from the notion of trust networks five criteria for comparison of trust models: observability, understandability and handlability, social exploitability, and context-scalability. We have then evaluated three trust models regarding these criteria.

The results of our examination show that no model is really adapted to our need. However, we think a more symbolic approach like modal logic is relevant. Thus, in the future, our work will deal with the building of a trust model inspired by Castelfranchi and Falcone's work and formalized by such a tool. It will address as much as possible the criteria that are not fulfilled.

Acknowledgements We thank France Telecom R&D for supporting the research reported in this paper (CRE 46128855 France Telecom - CNRS).

References

1. Urban, G.: Building an internet trust generator: Adaptive experimentation on the internet to improve site trust and sales. Technical report, MIT (2003)

2. Gambetta, D.: Can we trust trust? In Gambetta, D., ed.: Trust: Making and Breaking Cooperative Relations. Department of Sociology, University of Oxford (2000) 213–237
3. Castelfranchi, C., Falcone, R., Pezzulo, G.: Trust in information sources as a source for trust: a fuzzy approach. In: Proceedings of AAMAS'03, ACM Press (2003) 89–96
4. Demazeau, Y.: Steps towards multi-agent oriented programming. In: IWMAS'97. (1997)
5. Melaye, D., Demazeau, Y.: Modles et rseaux de confiance, analyse bibliographique. Cahiers du Leibniz **142** (2005)
6. Sichman, J., Conte, R., Castelfranchi, C., Demazeau, Y.: A social reasoning mechanism based on dependence networks. In Cohn, A.G., ed.: Proceedings of ECAI'94, Chichester, John Wiley & Sons (1994) 188–192
7. Jonker, C.M., Treur, J.: Formal analysis of models for the dynamics of trust based on experiences. In Garijo, F.J., Boman, M., eds.: Proceedings of MAAMAW'99. Volume 1647., Berlin, Springer-Verlag: Heidelberg, Germany (1999) 221–231
8. Fullam, K.K., Barber, K.S.: Evaluating approaches for trust and reputation research: Exploring a competition testbed. In: Proceedings of The Workshop on Reputation in Agent Societies at Intelligent Agent Technology, Beijing (2004)
9. Muller, G., Vercouter, L., Boissier, O.: Towards a general definition of trust and its application to openness in mas. In: Sixth International Workshop on Trust, Privacy, Deception, and Fraud in Agent Societies, Melbourne (2003)
10. Sabater, J., Sierra, C.: Social regret, a reputation model based on social relations. SIGecom Exch. **3** (2002) 44–56
11. Liau, C.: Belief, information acquisition, and trust in multi-agent systems- a modal logic formulation. Artificial Intelligence **149** (2003) 31–60
12. Ramchurn, S.D., Huynh, D., Jennings, N.R.: Trust in multi-agent systems. The Knowledge Engineering Review **19** (2004)