

The Impact of Competence and Benevolence in a Computational Model of Trust

Ameneh Deljoo, Tom Engers, Leon Gommans, Cees Laat

► **To cite this version:**

Ameneh Deljoo, Tom Engers, Leon Gommans, Cees Laat. The Impact of Competence and Benevolence in a Computational Model of Trust. 12th IFIP International Conference on Trust Management (TM), Jul 2018, Toronto, ON, Canada. pp.45-57, 10.1007/978-3-319-95276-5_4. hal-01855981

HAL Id: hal-01855981

<https://hal.inria.fr/hal-01855981>

Submitted on 9 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The impact of competence and benevolence in a computational model of Trust

Ameneh Deljoo¹, Tom van Engers², Leon Gommans^{1,3}, and Cees de Laat¹

¹ Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands
a.deljoo, delaat@uva.nl

² Leibniz Center for Law, University of Amsterdam, Amsterdam, The Netherlands
vanengers@uva.nl

³ Air France-KLM, Amsterdam, The Netherlands
Leon.Gommans@KLM.com

Abstract. Trust is a fundamental element of any social network. However, despite numerous studies on trust, few have conducted studies across disciplines to provide a complete picture of the different dimensions of trustworthiness, such as integrity, competence and benevolence. In this paper, we focus on two of these dimensions, competence and benevolence. We propose techniques to evaluate the competence of the trustee in specific situations and infer the benevolence of the trustee towards the trustor when the trust evaluation is made. Moreover, we evaluate both competence and benevolence on the perceived trustworthiness of the trustee, taking into consideration the development of the relationship between the trustor and the trustee over time. We identified different stages in this relationship development and use them to evaluate trustworthiness of trustee in the absence of evidence that could be used to evaluate trustworthiness. Finally, we set an experimental scenario implemented as an agent-based model to evaluate our approach. The results obtained from these experiments show that the proposed techniques can improve the reliability of the estimation of the trustworthiness of the agents.

Keywords: Computational trust, Social trust, Benevolence, Competence, Agent based modeling

1 Introduction

Trust is considered a fundamental basis in social societies. The social networks, however, evolve over time therefore, we need a sophisticated method that enables an agent to select the trusted peer to interact with. Trust has both a backward looking as well as a forward looking dimension, i.e. one expects certain behavior in the future based upon pastime experiences. Since one can never be sure that the conditions in the past are exactly the same as they are now (e.g, intentions of agents may change over time one may have to trust someone that had no experience with etc.), the estimation of trustworthiness is predominant to assess trust between each peer in such network. An agent-based model (ABM) of such

societies requires a computational model of trust.

Trust is a concept that will always need contextualization. It is something completely different to trust that someone will pay his debts, not commit a crime or is able to solve a complex mathematical calculation. In our pursuit for a computational model of trust that can be used in ABMs that allow us to study (non-)compliance in societies we focus on the trustworthiness of an individual in performing a certain task in a given situation, taking into account the agent’s competence in the matter, his overall integrity, and the stage of his relationships with the trustor. In order to better estimate the trustworthiness of trustees, it is important to consider these three dimensions individually, and to combine them in a dynamic way by taking into consideration the situation and the development of the relationship. However, the majority of the computational trust approaches presented in literature [1] estimates the trustworthiness of agents as one single factors rather than distinguishing these three trustees’ factors. In this paper, we present a computational trust approach grounded on the trust model presented by Marsh [2]. We will show that the resulting computational trust model is able to capture the competence and benevolence of the trustee. To estimate the trustee’s benevolence and competence, we use the evidence available to the trustee by considering different situations. Our approach combines the estimated competence of the trustee with the estimated benevolence into one single trustworthiness score. This score reflects the development of the relationship between trustor and trustee at the time of the assessment take into consideration. To estimate the competence of the trustee, we introduce three different stages of relationships between trustor and trustee. The results we obtain and present in this paper are very encouraging, as they show that there is a clear benefit in considering different stage of relationships between trustee and trustor in the described situations. The benevolence enhanced trust models allowed for a more accurate estimation of the trustees’ trustworthiness than the original computational trust models [3]. The remainder of the paper is organized as follows. Section 2 highlights the trust model and its antecedents. Then, Section 3 introduces our social computational model with its components. Section 4 deals with the simulation and results of the proposed solutions. Related work is covered by Section 5. Finally, Section 6 concludes the paper.

2 Trust

Trust is an essential part of social interaction. Trust is a broad concept studied in areas such as sociology and psychology [4]. The concept of trust has received ample attention from various disciplines, and although prior research has put forth diverse interpretations of trust, a common core emerges.

Trust antecedents framework, illustrated in Figure 1, was proposed by Mayer *et al.* [4]. Following Mayer *et al.* [4] we use the following trust definition: “Trust is the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control other party [4]”.

Specifically, trust is a decision of the trustor regarding the perceived trustworthiness of a trustee. Four trust factors are generic and important during the trust evaluation, namely competence, benevolence and integrity of the trustee, and the trust propensity of the trustor. The decision to trust means that the trustor is willing to take any possible risk caused by the trustee or environment, no matter whether s/he has the competence to monitor or control the trustee or environment. Building on this definition, we define inter-organizational trust as the expectation held by one firm that another will not exploit its vulnerabilities when faced with the opportunity to do so [4–6]. This expectation is confirmed when parties:

- demonstrate competence relate to the potential ability of the evaluated entity to do a given task,
- act accordingly to fulfill the commitments even when acting on them is not in self-interest and accept the consequences, and
- do good and act out of kindness even if unforeseen contingencies arise.

Our definition thus bases inter-organizational trust on three related components: competence, integrity, and benevolence, which have been proposed by Mayer *et al.* [4]. Computational trust is considered as an enabler of technology in virtual societies, and the estimation of trustworthiness is paramount to assess the trust that a trustor agent has on a given trustee. An individual is more or less trustworthy in performing a task in a given situation depending on his competence in the matter, his overall integrity, and the stage of his relationships with the trustor. Therefore, in order to better estimate the trustworthiness of agents, it is important to consider these three dimensions individually, and to combine them in a dynamic way, by taking into consideration the situation and the development of the relationship. However, the majority of the computational trust approaches presented in literature estimate the trustworthiness of agents as a block and do not distinguish between these trustees' attributions.

Proposition 1. Close and long-term relationships have a direct impact on the competence and benevolence of partners.

In following, we present a computational trust approach grounded on multidisciplinary literature on trust that is able to capture the competence and benevolence of the agent under evaluation. This framework has been implemented with ABM.

3 Computational trust model

In this section, we introduce the social computational trust model. Our aim is to define a mechanism for estimating trustworthiness of a trustee that can be used by the trustor to evaluate trust and make decisions about the future relationship with the trustee. Extracting trustworthiness of trustee based on Mayer *et al.*'s model [4] has been only implemented by few scholars such as [3, 7]. Most of these computational trust approaches estimate the trustees' trustworthiness using individual items of evidence about these trustees' behavior in the past

interactions, either with the trustor or with third party agents [8–10]. However, none of these approaches is able to estimate the benevolence of the trustee. We claim that understanding the benevolence and competence of the trustee towards the trustor at the moment of the trust decision is fundamental for being able to accurately estimate the latter’s trustworthiness. With this in mind, we present the main hypothesis of this work as follows;

Hypothesis 1. The extraction of benevolence-competence based information from the set of evidence on the trustee under evaluation and its use in adequate stages of the relationship between trustor and trustee shows that trustee’s trustworthiness improves by increasing the number of interactions between trustor and trustee.

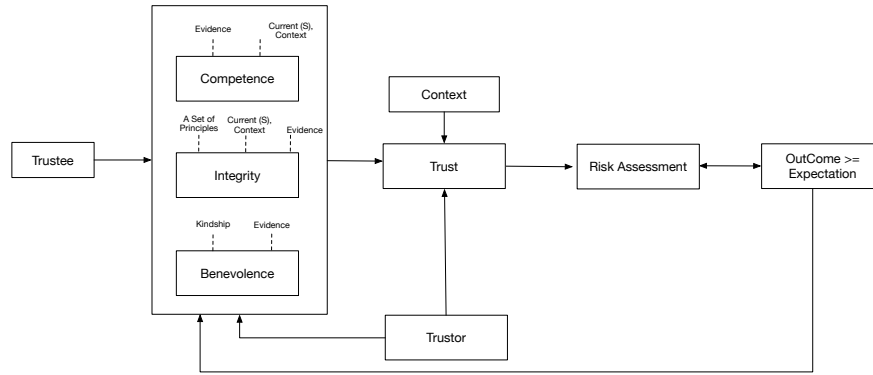


Fig. 1. Trust framework and its antecedents.

3.1 Basic Notation

Our generic computational trust model is applied to environments where trustor agents select the best trustees to interact with, having the posterior establishment of dyadic agreements between partners. We define the society of agents by A , where x and y represent as trustor and trustee respectively and $x, y \in A$. We represent $T_{(x,y)}$ as the amount of trust x has upon y based on the realization of a given task $t \in T$ with respect to the situation s_i , where $T = \{t_1, t_2, \dots, t_m\}$ is the set of all possible m tasks in the society and $s_i = \{s_1, s_2, \dots, s_i\}$ is the set of all possible situations in the society. These set of tasks bring duties for the trustee that need to be fulfilled. As we mentioned before, context plays an important role in our model. In order to characterize and describe the situation leading to an agreement, we consider the definition of context as including four main types of context: identity, time, location, and activity [11]. Urbano *et al.* [3] identified eight dimensions of context $\{d_1, d_2, \dots, d_8\}$ ⁴, where dimensions

⁴ Dimensions are defined based on the scenarios

d_1 and d_2 represent the agents trustor and the trustee, respectively; d_3 and d_4 represent the time and the location of agreement; and d_5 , d_6 , d_7 and d_8 identify and characterize the type of the task, its complexity, deadline, and outcome of its realization, respectively. We adapt the Urbano *et al.* [3] context definition in this paper. For the simplicity, we assume that all agreements performed in the society of agents refer to the same type of task t (d_5), although tasks have different degrees of complexity (d_6) and deadlines (d_7). In this paper, we only consider four dimensions of context, which are d_1 , d_2 , d_5 and d_8 . We exclude time, complexity and deadlines from our formalization. They will be addressed in a future publication.

Urbano *et al.* considered three different outcomes (d_8) is defined by $O = \{FD, FDD, V\}$, where FD (fulfill duty) means the trustor considers that the trustee performed whatever matter he had to perform on time, FDD (fulfill duty with delay) means the trustee was presented with an unexpected delay in the performance of the task (or duty), and V (violation) means the trustee did not perform the given task. Considering $O = \{FD, FDD, V\}$, possible values for this function are $val(FDD) = 1.0$, $val(FD) = 0.5$, and $val(V) = 0.5$.

Finally, the set of all the existed evidence on given trustee is represented by $E_{*,y} = \{e_i \in \varepsilon : v_2^{e_i} = y\}$, where ε represents all evidence available on trustee from the all the trustee's direct neighbor. Following, $E_{(x,y)}$ shows all the evidence about the direct interactions between trustor and trustee $E_{(x,y)} = \{e_i \in \varepsilon : v_1^{e_i} = x, v_2^{e_i} = y\}$.

3.2 Our social computational trust Model

The benevolence-competence based computational model of trust that we present in this paper is a part of a larger framework of social trust based on Mayer's trust model. It integrates three distinct functions: the competence evaluation function ($Com_{(x,y)} : s_i \times Ex_{(*,y)} \in [0, 1]$), the benevolence evaluation function ($Ben_{(x,y)} : E_{(x,y)} \in [0, 1]$), and the trustworthiness evaluation function ($TW_{(x,y)} \in [0, 1]$) in the set of situations s_i . We describe each of these functions in the following subsections. We illustrated the computational model in Figure 2.

Benevolence function Benevolence is considered as a key element of trust and an antecedent of trustworthiness by several scholars (e.g. [12,13]). The estimated value of the benevolence of trustor x toward trustee y , $Ben_{(x,y)}$, is derived from the direct interactions (i.e. $E_{(x,y)}$) between trustee and trustor in the situations s_i . The output of the benevolence evaluation function $Ben_{(x,y)}$, defined in $[0, 1]$, is

$$Ben_{(x,y)} = \frac{1}{|S|} \sum_{E_{(x,y)}} (val(E_{(x,y)})). \quad (1)$$

Where S is the set of situations, in which x has interactions with y .

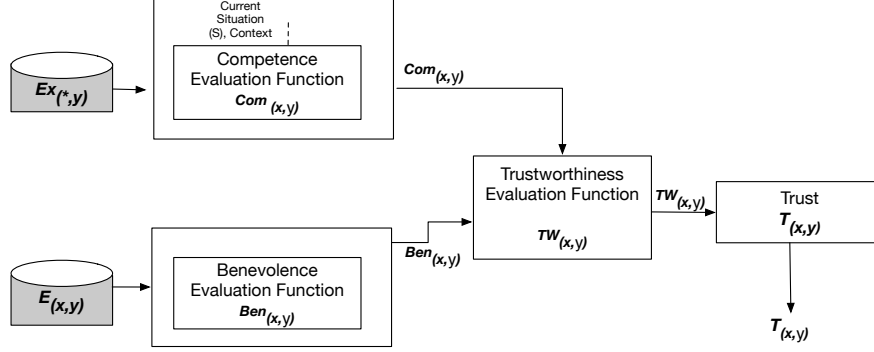


Fig. 2. Our benevolence-competence computational model of trust.

Competence function The competence evaluation function $Com_{(x,y)}$ estimates the general ability of the trustee under evaluation in performing a given task t in a specific situation s . This function takes all the evidence available on the trustee under evaluation $Ex_{(*,y)}$ as input. The output of the competence evaluation function is the estimated competence of the agent, $Com_{(x,y)}$, defined in $[0, 1]$. Competence, as risk, involves an agent making a judgment about the trustee's ability to perform the given task. We consider three different possible situations to evaluate trustee's ability.

1. There is no evidence available from the trustee. To judge the trustee's competence, the trustor will calculate the risk of trusting a stranger and decide based on the risk as

$$Risk = \frac{Cost \times (1 - Pr)}{benefit \times Pr}, \quad (2)$$

where Pr is the probability of performing the task by the given trustee.

2. Situation β : there are some evidence but not for the considered context. In this situation, the trustor collects all the evidence from other agents and evaluates the competence of trustee based on them as

$$Com = \frac{1}{|N|} \sum_{\beta \in N} (val(E_{*,y}) \times \hat{T}_x(y, \beta)), \quad (3)$$

where $\hat{T}_x(y, \beta)$ denotes the basic trust that x has on y and β is the set of all situations in which x has interactions with y . The basic trust y calculated as $1/|N| \sum_{\beta \in N} T_{(x,y)}$. N denotes as the set of situations similar to the present situation (s_i) in which x has interactions with y .

3. Situation α : there is related evidence about the agent in this context or similar situation.

$$Com = \frac{1}{|N|} \sum_{\alpha \in N} (val(E_{*,y})), \quad (4)$$

where α is the set of all situations in which x has interactions with y . These three situations (no evidence available, β and α) are assumed to help an individual to make the decision. In this paper, we assume that trustor and trustee collaborate in the similar context (i.e. situation α).

Function $TW_{(x,y)}$ The trustworthiness evaluation function $TW_{(x,y)}$ takes into consideration the perception of the competence and benevolence of the trustee (Algorithm 1). We assume the same weight to the competence and benevolence dimension, this weight when both trustor and trustee are practically strangers ($W_{(x,y)} = 0$), and progressively increasing the weight of these to dimension as the partners get to know each other better, this weight becomes one when partners are considered to be closed (Proposition 2). In the algorithm 1 $N_{(x,y)}$

Algorithm 1 Calculate $TW_{(x,y)}$

Require: $E_{(*,y)}$: the set of all evidence about trustee y .

Require: $N_{ben} = 2$: minimum interaction between trustee and trustor.

Ensure: $E_{(x,y)} \notin E_{(*,y)}$

- 1: $N_{(x,y)} \leftarrow |E_{(x,y)}|$
 - 2: $Ben_{(x,y)} \leftarrow F(Ben_{(x,y)})$
 - 3: $Com_{(x,y)} \leftarrow F(Com_{(x,y)})$
 - 4: $TW_{(x,y)} = Com_{(x,y)} + Ben_{(x,y)}$
 - 5: **return** $TW_{(x,y)}$
-

represents the number of interactions between x and y , (line 1), and defined a minimum number of interactions ($N_{ben} = 2$) between trustor x and trustee y . $N_{(x,y)}$ used to combine the estimated value of the trustee's benevolence as returned by $F(Ben_{(x,y)})$ (line 2) with the estimated value of its competence as returned by $F(Com_{(x,y)})$ (line 3). Finally, the estimated value of the trustee's trustworthiness $TW_{(x,y)}$ is computed as a sum of $Com_{(x,y)}$ and $Ben_{(x,y)}$ (line 4).

4 Simulation setup

In this set of experiments, we want to test **Hypothesis 1**, which we reformulated as follows: The extraction of benevolence-competence based on the information from the set of evidence on the trustee under evaluation and its use in adequate stages of the relationship between trustor and trustee shows that trustee's trustworthiness improves by increasing the number of interactions between trustor and trustee.

The experiments were conducted in Jadex [14] environment. We set up a social network, shown in Figure 3 to represent our network. We setup a collaborative network of organizations. This social network represents a collaborative network of organizations like the ones we study in our SARNET research project where

service providers collaborate and act on behalf of partners, acts that may harm the individual interests, all in order to protect the collaborative network against cyber attacks. Each node represents an autonomous organization that needs to trust other parties and share sensitive information with them. For the simplicity, we assumed that there is only one task being negotiated by all nodes which mitigates an attack and shares the attack information with other parties⁵.

We define four different situations: s_1 provides a specific number of samples within 24 hours, s_2 provides a specific type of resource (e.g. allocating resources), s_3 blocks a link, and s_4 monitors a specific traffic.

This model starts after the establishment of an agreement between the trustor and the selected trustee, thus excluding the selection process itself. It focuses on both types of agents' decision concerning the fulfillment of the established agreement: the trustees may opt to fulfill the agreement (the trustors will report the outcome FDD) or to delay its realization. Accordingly, the trustors may respond to a delay by either retaliating, denouncing the breach (reporting outcome V) or forgiving the contingency (reporting outcome FD).

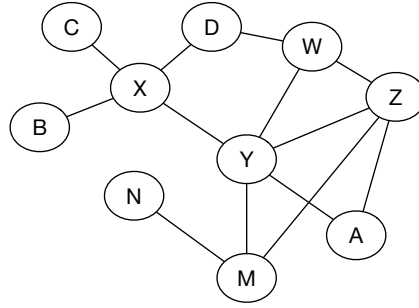


Fig. 3. Social network schema.

4.1 Result

Our result consists of two parts. First, we calculate benevolence of agent y by considering all the evidence (i.e direct interactions) that x has on y . Second, evaluate the competence of agent given trustees a , z and y in four mentioned situations from the trustor's x view. Hence, we perform four different types of situations simultaneously, each with six agents. We assume that agents are honest and there is no conflict on the evidence and message are encrypted (the interminable agents cannot manipulate the message). In order to compare all

⁵ The technical details and code of this research can be found in <http://delaat.net/sarnet/index.html>

approaches, we measure and average the number of agreements with outcomes FD , FDD and V . We are able to calculate the benevolence and competence of each trustor by equation 1 and 4. To calculate the $Ben_{(x,y)}$, we extract all the evidence that x has on y . We perform the simulation for 20, 50 and 100 rounds of interactions. We also consider that each agent can freely fulfill duty, fulfill duty with delay and violate the agreement. We have summarized the result in Table 1 (including mean M and standard deviation SD).

Table 1. $Ben_{(x,y)}$ evaluation for the number of rounds

No. of rounds \rightarrow	20	50	100
$Ben_{(x,y)}$	<i>0.22</i>	0.31	0.86
SD	<i>0.113</i>	0.105	0.081
M	<i>0.762</i>	0.777	0.810

To evaluate the competence function, we select three agents a , z and y from the set of agents and calculate the competence of these three agents from trustor' x view. Agent x will collect (collecting the evidence is done by sending a query to each agent and asking its opinion) all the available evidence from the neighbour's of a , z and y which are reported in Table 2. The simulation has been performed for four different situations and three different rounds. As result, we compared

Table 2. Agent x asks different Agents' opinion about agent z in the four situations

Situation \rightarrow	s_1	s_2	s_3	s_4
Y	FDD,FD	FD,FD	FDD,FD	FDD,FDD
M	FDD,FD	FD,FD	FDD,FD	FDD,FDD
W	FDD,FD	FD,FD	FDD,FD	FDD,FDD
D	FDD,FD	FD,FD	FDD,FD	FDD,FDD

mean (M) and standard deviation(SD) and competence of each trustee for different rounds and we detailed them in Table 3 and Table 1. In Table 3 and Table 1, the highest values are in bold and the lowest values are in italics. We verified that in Table 1 the benevolence of trustees are increased as the number of interaction increases. For instance, with only 20 rounds, when the number of interactions between any two partners is not large the benevolence is small. By increasing the number of rounds, the benevolence increases significantly. Indeed, this confirms that the number of interactions is, in fact, impact the benevolence existing between any pair of trustor-trustee (We have proved the main hypothesis of this work3). In the case of competence (see Table 3), we also observed the

Table 3. Competence evaluation for agents a , z and y and the number of rounds

No. of rounds \rightarrow	20	50	100
$Com_{(x,a)}$	0.21	0.40	0.65
$Com_{(x,z)}$	0.28	0.43	0.88
$Com_{(x,y)}$	0.18	0.33	0.54
SD_a	0.059	0.051	0.048
SD_z	0.068	0.058	0.042
SD_y	0.081	0.066	0.041
M_a	0.927	0.938	0.941
M_z	0.910	0.927	0.947
M_y	0.917	0.936	0.962

same behavior from the simulation. The competence of agent is influence directly by the number of interactions. As we mentioned before, the simulation has been repeated for 50 and 100 rounds. The benevolence value reaches the maximum value of one by increasing the number of interactions between partners and the partners that are considered to be in a close relationship (Proposition 2). Indeed, each trustor can conclude the trustworthiness of trustee in each round and make a decision. Overall, we are able to confirm the truthfulness of Hypothesis 1.

5 Related work

Many computational trust models have been presented by different scholars, nevertheless, only a few models are actually social computational models. Adalie *et al.* [15] presented a conceptual model of social trust based on Kelton *et al.*'s model [16]. Kelton *et al.*'s consider ability, positive intentions, ethics, and predictability as the trustworthiness components. Adalie *et al.* used a probabilistic approach to implement the model, but by recognizing the limits of such approach in the treatment of the social concepts, their model was not implemented.

Among all the presented computational trust models [17], the only computational approach that includes a comprehensive set of features grounded on the theory of trust is the socio-cognitive model of trust by Castelfranchi and Falcone [18]. This model considers that the trustor has a goal that can be achieved by the action of the trustee. In their view, trust is formed by considering the different beliefs that the trustor has about the trustee, either internal (beliefs on competence, disposition, and harmfulness) or external (opportunities and dangers). The values of these beliefs are further modulated by meta-beliefs about the relative strength of each belief. The richness of this model makes it hard to implement in practice. In fact, the current implementation of the model (e.g., [18]) requires extensive manual configuration by domain experts for each trustee and task under assessment and oversimplifies the theoretical model. Moreover, it requires explicit information about the competence and disposition (or similar

beliefs) of the agent under evaluation, which may be hard to get in dynamic agent-based environments.

Another social trust model was presented by Urbano *et al.* [3] called situation-aware and social computational trust Model (SOLUM). Their computational model consists of two parts: the first part is a general framework of computational trust, which is based on two fundamental characteristics of trust, the trustor’s disposition and emotional state. They adapted also Mayer’s trustworthiness dimension that includes the ability, integrity, and benevolence to determine trust. For the second part, they proposed a set of distinct techniques to extract information about the individual dimensions of the agent’s trustworthiness from the set of structured evidence available to the agent. The main difference between our model and Urbano’s model is that we consider different stages of relationships for the competence function. We slightly adapted and modified the Marsh [2] competence formulization by considering three different situations for trustor to make a decision about the (future) collaboration with the trustee.

Finally, Herzig *et al.* [19] formalized the model of Castelfranchi and Falcone, in multi-modal logic, adding the notions of occurring trust and dispositional trust (i.e., trust in a general disposition of the trustee to perform a similar task some point in the future). Skopik *et al.* [20] purposed a semi-distributed information sharing platform where different organizations can share the incidents information with their trusted peers. Skopik *et al.*, proposed a fuzzy method to evaluate trust among members. The major aim of social trust in their model is to personalize online interactions and prioritize collaboration with trustworthy individuals. The author claimed that trusted relations can be defined manually by users, e.g., by declaring “friend-relations” or can be determined automatically through mining of interactions. However, their social model is based on personal experience of each member and suffers from the scalability issue.

6 Conclusion and Discussion

Computational trust is crucial for decision making regarding possible agents’ future joint activities such as the alliances. It heavily relies on the estimation of trustworthiness to assess the trust on particular trustees. To better estimate this trustworthiness, it is important to estimate, besides other relevant features, their competence and benevolence separately, and to combine them taking into consideration the particular situation and relationship. In this paper, we described a part of our trust computational model. We evaluated our approach in a simulated experimental environment. We have proved that the trustworthiness estimation grew with the increasing number of interactions between any trustor-trustee pair. Besides, we went beyond the traditional evaluation of computational trust models (such as Fuzzy logic evaluation) and used a model of agents’ behavior where both trustors and trustees evolve their behaviors based on different stages of existing relationships (i.e. presented in the competence formula) between the agents. Concerning future work, we intend to further identify

the particular circumstances in which the use of this sophisticated trust model is more relevant. Also, we intend to explore integrity as another dimension of trustworthiness, as well as exploring other ways of combining the trustworthiness dimensions, and to use other antecedents of trust, such as the trustors' own propensity to trust. A lot of future work remains to be done to ensure that the framework is functional in practice. We plan to conduct experiments to evaluate the robustness of the proposed trust mechanism against the badmouthing attacks and non-compliant members. Furthermore, evidential reasoning in the case of conflicting on evidence and evaluate trust based on this method will need to be elaborated and evaluated. Another interesting direction for future work is to develop guidelines in consultation with policy makers to define sounds policy and standard for determining the contract importance based on the risk associated with the proposed methods and the members' preferences. Many different adaptations, tests, and experiments for our model have been left for the future due to lack of time. Future work concerns the a deeper analysis of our proposed model and validates the model with different computational models that proposed by other scholars such as Urbano *et al.*

7 Acknowledgment

This work is funded by the Dutch Science Foundation project SARNET (grant no: CYBSEC.14.003/618.001.016) and the Dutch project COMMIT (WP20.11). Special thanks go to our research partner KLM. The authors would also like to thank anonymous reviewers for their comments.

References

- [1] Sabater, J., Sierra, C.: Review on computational trust and reputation models. *Artificial intelligence review* **24**(1) (2005) 33–60
- [2] Marsh, S.P.: Formalising trust as a computational concept. (1994)
- [3] Urbano, J., Rocha, A.P., Oliveira, E.: The impact of benevolence in computational trust. In: *Agreement Technologies*. Springer (2013) 210–224
- [4] Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. *Academy of management review* **20**(3) (1995) 709–734
- [5] Barney, J.B., Hansen, M.H.: Trustworthiness as a source of competitive advantage. *Strategic management journal* **15**(S1) (1994) 175–190
- [6] Krishnan, R., Martin, X., Noorderhaven, N.G.: When does trust matter to alliance performance? *Academy of Management journal* **49**(5) (2006) 894–917
- [7] Guo, G., Zhang, J., Thalmann, D., Yorke-Smith, N.: Etaf: An extended trust antecedents framework for trust prediction. In: *Advances in Social Networks Analysis and Mining (ASONAM)*, 2014 IEEE/ACM International Conference on, IEEE (2014) 540–547
- [8] Sabater, J., Sierra, C.: Regret: reputation in gregarious societies. In: *Proceedings of the fifth international conference on Autonomous agents*, ACM (2001) 194–195
- [9] Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: *System Sciences*, 2000. Proceedings of the 33rd Annual Hawaii International Conference on, IEEE (2000) 9–pp

- [10] Mui, L., Mohtashemi, M., Halberstadt, A.: A computational model of trust and reputation. In: System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on, IEEE (2002) 2431–2439
- [11] Abowd, G., Dey, A., Brown, P., Davies, N., Smith, M., Steggles, P.: Towards a better understanding of context and context-awareness. In: Handheld and ubiquitous computing, Springer (1999) 304–307
- [12] Levin, D.Z., Cross, R., Abrams, L.C., Lesser, E.L.: Trust and knowledge sharing: A critical combination. IBM Institute for Knowledge-Based Organizations **19** (2002)
- [13] Kosciak, T.R., Tranel, D.: The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia* **49**(4) (2011) 602–611
- [14] Braubach, L., Lamersdorf, W., Pokahr, A.: Jadex: Implementing a bdi-infrastructure for jade agents. (2003)
- [15] Adali, S., Wallace, W., Qian, Y., Vijayakumar, P., Singh, M.: A unified framework for trust in composite networks. Proc. 14th AAMAS W. Trust in Agent Societies, Taipei (2011) 1–12
- [16] Kelton, K., Fleischmann, K.R., Wallace, W.A.: Trust in digital information. *Journal of the American Society for Information Science and Technology* **59**(3) (2008) 363–374
- [17] Pinyol, I., Sabater-Mir, J.: Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review* **40**(1) (2013) 1–25
- [18] Castelfranchi, C., Falcone, R.: Trust theory: A socio-cognitive and computational model. Volume 18. John Wiley & Sons (2010)
- [19] Herzig, A., Lorini, E., Hübner, J.F., Vercouter, L.: A logic of trust and reputation. *Logic Journal of IGPL* **18**(1) (2009) 214–244
- [20] Skopik, F., Schall, D., Dustdar, S.: Modeling and mining of dynamic trust in complex service-oriented systems. In: Socially Enhanced Services Computing, Springer (2011) 29–75