1

Trust Evaluation for Reliable Electronic Transactions Between Business Partners

Joana Urbano, Ana Paula Rocha, and Eugénio Oliveira

LIACC / Departamento de Engenharia Informática
Faculdade de Engenharia, Universidade do Porto
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
{joana.urbano,arocha,eco}@fe.up.pt
http://paginas.fe.up.pt/~niadr/

Summary. In this era of digital economy, commercial relationships between business partners are increasing in flexibility, with new business binds being created whenever a business opportunity arises. Moreover, the instability in demand increases the need for enterprises to procure new partners as well as the associated risk of dealing with partners that may be unknown beforehand. Therefore, enterprises need mechanisms that allow to evaluate the confidence they have on their current and potential new, unknown, partners, and to monitor this confidence in a continuous and automatic way. This paper presents our computational trust model, which was inspired in the concept of the hysteresis of trust and betrayal and in the asymmetry principle of human psychology. Our model allows to estimate the trustworthiness of agents using different features of the dynamics of trust. Additionally, we present a study on the effect of preselecting partners based on their trustworthiness in automated negotiation processes. The study was conducted experimentally using our agent-based Electronic Institution framework for e-Contracting, which includes a normative environment and an automatic negotiation service, as well as the mentioned computational trust service. The results obtained show that, in identified conditions, business clients benefit from preselecting partners based on trust prior to the negotiation phase.

Key words: Computational Trust Systems; Dynamics of Trust; Multi-Agent Systems

1.1 Introduction

In the new era of digital economy, commercial relationships between business partners are increasing in flexibility, with business binds being created whenever a business opportunity arises. The emergent need for new products and services, with increased quality, shorter time to market and low prices, and the instability in product demand is forcing enterprises to risk new, sometimes unknown, suppliers, possibly from all over the world. This new reality brings new technological, social, ethical, and economical challenges and risks to the industry.

Moreover, the desired automation of business inter-organizational relationships encounters some barriers in key stages, such as the selection of partners, negotiation and contract elaboration, particularly when there is a large number of partners in the play (e.g. textile industry) that are unknown beforehand. The construction of reliable and more widely accepted mechanisms of trust and reputation will allow organizations to continuously update their confidence level on current and potentially new partners.

The potential benefits of such mechanisms are two-fold: i) they allow for a broader selection of partners, as it will be possible to infer confidence values for a larger number of partners (e.g. through the use of social evaluations or through reasoning on organizational characteristics); ii) they make it safer for an organization to increment the degree of tasks that can be automated, both in the partner selection process and in the automated negotiation of contracts. Trust and reputation mechanisms are, in fact, getting great attention from different research areas, from social science and psychology to economics and computer science, particularly in the multi-agent and the service oriented architecture communities.

Our current research work focuses on the automation of inter-organizational interactions in two different but yet complementary tasks: the selection of partners and the negotiation of contracts in dynamic environments, both supported by the use of confidence knowledge derived from trust and reputation. In our work, we follow the paradigm of Virtual Organizations and Electronic Institutions [1][2].

In this paper, we study the effect of filtering the space of candidate partners in automated negotiation processes taking into consideration the partners' estimated trustworthiness. In the first part of the article, we present our model of computational trust, which is used to estimate the trustworthiness of the partner agents. It is composed of an engine that aggregates the available trust evidences using heuristics derived from studies on social trust. The main motivation for the development of this trust model was the intuition that the journey along the way to gain trustworthiness must not necessarily be the same as the one that takes place when trustworthiness decreases; although this is a consensual concept, most of the computational trust approaches do not implement it in a practical way.

We proceed by proposing the use of contractual evidences as input to the trust model. We consider that contractual data should be used as an additional source of trust information, whenever it is available, as it can be considered more reliable than indirect social evaluations that are transmitted through peer agents, such as recommendations and reputation.

Finally, we describe the experiments we have run in our Electronic Institution framework in order to study the effects of trust-based preselecting of partners on the overall process of negotiation between business partners. In these experiments, we consider different alternative values for the impact of negative interactions for the selecting agent, as well as different variability in the expected utilities provided by the different candidates to the joint action. Also, we use a simple contract of sale and adapt our trust aggregation engine in order to accept as input contractual information managed by the Electronic Institution. The results we obtained show that prefiltering of proposals from candidate agents based on trust increased the utility that clients got from the negotiation, provided that the received proposals are not completely disparate.

The remaining of this article is structured as follows: in Section 1.1.1, we present related work. In Section 1.2, we describe the theoretical foundations that led to our model of computational trust, *Sinalpha*, which is presented in Section 1.3. Section 1.4 describes the simulated textile scenario that we built using the concept of an Electronic Institution and Section 1.5 presents the experiments we have run over this scenario in order to evaluate the benefits of using our trust model in the process of selecting partners. Finally, Section 1.6 summarizes the major conclusions of this paper.

1.1.1 Related Work

In the last years, several computational trust models have been proposed in the literature. Most of the approaches deal with the representation of trust evidences (i.e. structured information available on the target agent in evaluation) and the aggregation of these evidences in order to compute the trustworthiness of agents.

The earliest computational trust models use relatively simple aggregation algorithms. Some of them simply sum up integer values (e.g. the eBay reputation system). Others aggregate classifications using means and weighted means (e.g. [3][4][5][6]), Beta distributions ([7][8]) or Dirichlet distributions ([9]). Still other approaches focus on Bayesian techniques ([10][11][12]) and trust learning approaches ([13][14][15]). More complex models are implemented using complex beliefs, desires and intentions (BDI) architectures [6][16].

Despite the interesting characteristics exhibited by these approaches, they tend to ignore important knowledge on trust generated in the areas of Psychology, Sociology and Economics. One such piece of knowledge has motivated our research on trust, namely, the concept of the *asymmetry principle* of human psychology [17], which states that trust is hard to build and easy to loose.

In fact, only a few more recent computational trust models are incorporating this feature in their algorithms. Our model of trust presents a sigmoid-like curve for building trust where a specific parameter (λ) is configured with positive values (λ^+) for the aggregation of positive outcomes and with negative values (λ^-) for the aggregation of negative outcomes ($|\lambda^+| \leq |\lambda^-|$).

The model proposed in [18] presents an asymmetric trust update function, using two different equations for the aggregation of positive and negative experiences, respectively. In [19], we compared this model with our trust approach and with a weighted means by recency model (such as the one used in

FIRE [5]), by running different experiments in a simple scenario; the results have shown similar performances for both models that implement the asymmetry property, both achieving better results than the weighted means by recency approach. Although the model in [18] performed quite well, we think that this model scales worse to multivalued outcomes than our model, due to the fact that it uses separate functions to aggregate positive and negative outcomes.

Another property of trust that we explore in our model is related to the notion given in [20] that an entity must be trustworthy for a long time before others realize it is worth trusting. In our model, we call it the *maturity* property. All the approaches referred above that are based on sums, means and distributions lack the maturity property; instead, some of them use an additional *credibility* value that represents the confidence that the truster has on the estimated trustworthiness value based on the number of evidences that were used in the aggregation process.

Moreover, approaches that lack the maturity property normally present the *indistinguishable past* property, which, as pointed in [21], 'expresses that only the experiences themselves count and not the point in time at which they were experienced'. As considered in [21], this is not a natural property of trust. In our trust model, as well as in [18], different patterns of past outcomes lead to different trustworthiness estimations, which means that both models implement the inverse *distinguishable past* property.

A recent trend of investigation in trust is the exploration of context to improve the decision making, raising significantly the number and type of information that the evaluator has in order to compute trust. This means that, along with social evaluations given by direct experience or through witnesses, a plethora of new information related to the context of the business and of the organizations involved can improve the prediction of behavior of partners in a very significant way. However, few proposals have been made on this specific area (e.g. [22][23][24][25]), leaving an enormous world of research. The trust model that we present in this paper does not directly consider the situation in which the evidences are produced. Instead, we have developed a situation-aware component (cf. [26][27]) that can be used along with any trust aggregator in order to tune the estimated trustworthiness to the current situation.

Finally, we consider in this paper the prefiltering of candidate partners as a prior step to the negotiation phase in electronic business transactions. The work in [28] also considers the possibility of selection of candidate partners based on the reputation of agents, prior to the negotiation phase. However, the empirical evaluation of their trust model is focused on testing its resistance to attacks, and they do not model negotiation in their experiments.

1.2 A Model for the Hysteresis of Trust and Betrayal

As we have mentioned in the previous section, most of the existing computational trust models aggregate evidences doing simple additions, simple means, weighted means or even sum ups using distributions (e.g. Beta and Dirichlet). These models, although computationally efficient, do not allow for an easy integration of features normally associated with the process of creating trust. Therefore, we propose in this paper a model based on heuristics that implements important features of the dynamics of trust.[1]

Our first motivation was to translate into a model the intuition that the journey along the way to gain trustworthiness must not necessarily be the same as the one that happens when trustworthiness decreases, but, instead, must describe a route similar to a hysteresis curve. Interesting enough, we came across the work of Straker, where he describes the hysteresis of trust and betrayal [20]. In his model, Straker assumes that betrayal often happens in the balance between the trustworthiness of a self and the trust placed in this self, and he defines four different regions that arise out of this balance: *creating trust*, *trust is given*, *taking advantage* and *betrayal*. Figure 1.1 illustrates this model.

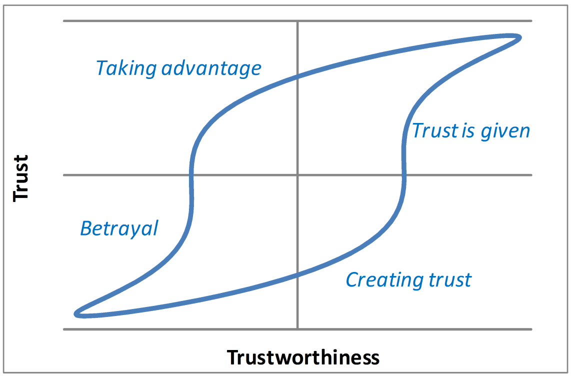

Fig. 1.1. The hysteresis of trust and betrayal

The model depicted in Figure 1.1 assumes that it is hard to trust a target agent when it is not very trustworthy, and that it takes time and effort for the agent to increase its trustworthiness in order to be trusted by others (the *creating trust* phase). Eventually, after proving its trustworthiness for a long time, the agent can be considered worth trusting (the *trust is given* phase). From now on, the target agent is blindly believed by other agents even if

[1] Some parts of the work contained in this paper must be understood as companion to the work in [29].

it starts behaving less trustworthy. The target agent is now in the *taking advantage* phase. However, if the agent keeps behaving in an untrustworthy way, the other agents will eventually realize it and feel betrayed and hurt by the actions of the target agent, and the trust on it decays abruptly. This phase, *betrayal*, can be highly damaging.

The non-linear behavior just described could be easily incorporated into our trust aggregation engine using, for example, the Lapshin formula [30] depicted in Equation 1.1.

$$x(\alpha) = a\cos^m \alpha + b_x \sin^n \alpha, \qquad y(\alpha) = b_y \sin \alpha. \tag{1.1}$$

In the formula, a represents the coersitivity parameter, m and n are integers used to fit the curve, and b_x and b_y are the saturation parameters. Figure 1.1 was obtained with the following configuration for these parameters: $a = 0.1$, $b_x = 0.5$, $b_y = 0.5$, $n = 3$ and $m = 1$.

An interesting characteristic of Straker's model is the introduction of the four phases in trust building. Also, it reflects the difference between trust and trustworthiness, where the former is exhibited by the one that trusts, in relation to the one that is trusted, and the later is a property of the one who is trusted, in relation to the truster ([31][32]). The model also assumes Hardin's idea about the existence of a causal connection between trustworthiness and trust ([33]).

Although Straker's model depicts the balance between trust and trustworthiness in an attractive way, it presents some drawbacks at the conceptual level. In fact, the model assumes that trust depends exclusively on trustworthiness. However, as stated by Castelfranchi and Falcone in [32], although trustworthiness is a component of trust, trust cannot be reduced to trustworthiness. Other dimensions, such as the personality and disposition of the truster, or even the characteristics of the one trusted, must also be considered.

In the same way, a direct implementation of Straker's model would allow for an agent located at the top end of the *trust is given* phase or at the *taking advantage* phase to behave in an intermittent way without being greatly penalized by this undesirable behavior. Moreover, we intuitively feel that the *taking advantage* phase is too long and allows for severe deceptive behaviors from agents that have proved to be trustworthy in the past.

Given these considerations, we reconsidered our trust approach, aiming at developing a model that retains the good characteristics of the hysteresis-based model of trust and betrayal and, at the same time, addresses the limitations mentioned above.

We named this new model *Sinalpha*, and we present it in the next section. *Sinalpha* is an engine that aggregates evidences on a target in evaluation in order to estimate its trustworthiness. Therefore, it is just one component of a broader trust system that may include other components, such as the situation-aware tuner mentioned in Section 1.1.1 and/or any dispositional-based component that may be developed in the future. This way, we attenuate the strong connection between trust and trustworthiness that exists in Straker's model.

1.3 Sinalpha

Figure 1.2 shows *Sinalpha*, the sinusoidal curve that serves as the basis for our new aggregation engine. The mathematics underlying *Sinalpha* are presented in Equation 1.2.

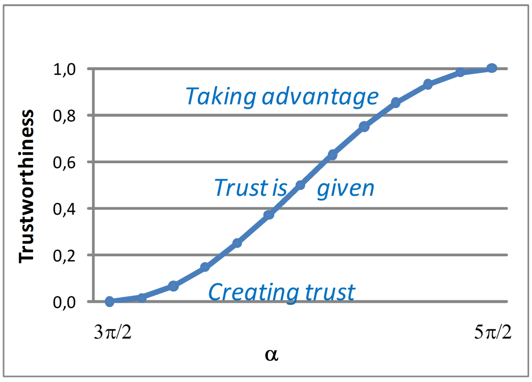

Fig. 1.2. *Sinalpha*, the simplified version of the hysteresis-based aggregation curve

$$\begin{aligned} y(\alpha) &= \delta \cdot sin\alpha + \delta, \\ \alpha_0 = \alpha_{min} = \frac{3\pi}{2}, \alpha_{max} &= \frac{5\pi}{2}, \\ \alpha &= \alpha + \lambda \times \omega \end{aligned} \tag{1.2}$$

As can be observed in Figures 1.1 and 1.2, the new model is a simplified version of the hysteresis-based trust aggregation engine, in that the ascending and descending paths on trust creation are now the same.[2]

Equation 1.2 shows the update function used for estimating the trustworthiness of the agent in evaluation. In the equation, parameter δ has a fixed

[2] In the new trustworthiness formula, we adapted $y(\alpha)$ from Lapshin's model. Other formulas, probably more intuitive, could be used, such as the one proposed in [34], which models hysteresis as a linear function of x. We leave the experimentation of alternative formulas for future work.

value of +0.5 and it is used to normalize the trustworthiness values into the range $[0, 1]$.

When there are no evidences on an agent's past behavior, α assumes the value $\frac{3\pi}{2}$, which corresponds to a trustworthiness value of zero. The agent is then at the *creating trust* phase. If there is a new evidence about the agent in evaluation, the value of α is updated by a value of $\lambda \times \omega$, as shown in Equation 1.2. The value of α increases with positive outcomes and decreases with negative outcomes, being limited to the range $[\alpha_{min}, \alpha_{max}]$.

Parameters λ and ω allow for the introduction of interesting properties of the dynamics of trustworthiness building, which we have identified in Section 1.1.1. Parameter ω permits to set how fast or how slow an agent can be perceived as highly trustworthy. In Figure 1.2, ω was set to have value $\frac{\pi}{12}$, meaning that a supplier that always presents positive outcomes reaches the maximum estimated trustworthiness in twelve steps. A good characteristic of ω is that it can be configured differently according to the personality of the trusting agent (e.g. more or less cautious) or even to the specificity of the situation in assessment. We leave the study of ω possibilities for future work.

Parameter λ is used to adapt the asymmetry principle of human psychology – which refers that trust is hard to build and easy to loose – to the estimation of the trustworthiness of agents. In Figure 1.2, this parameter assumes the value $\lambda^+ = +1.0$ when aggregating positive outcomes and $\lambda^- = -1.5$ when aggregating negative outcomes. By setting $|\lambda^+| < |\lambda^-|$, the model is able to capture sudden changes in the behavior of the agents in evaluation, effectively penalizing the undesirable patterns of intermittent behavior that could happen with a direct implementation of Straker's model.

A new observation of Figure 1.2 allows to conclude that *Sinalpha* preserves the *maturity* property of the hysteresis-based approach. In fact, by assuming three phases of growth/decay (*creating trust*, *trust is given* and *taking advantage*), the model avoids that the agent in evaluation is considered trustworthy before it provides enough evidences about its trustworthiness. It also signifies that the impact of negative results is weaker when the trustworthiness value of the agent is low (the *creating trust* phase) or high (the *taking advantage* phase), and stronger when the agent is in the process of acquiring trustworthiness (the *taking advantage* phase).

Finally, *Sinalpha* also preserves the *distinguishable past* from the hysteresis-based approach, meaning that the model distinguishes between the possible sequence of past outcomes of the target agent in the evaluation.

The *Sinalpha* model was extensively tested and compared to a weighted-means by recency trust approach based on the work in [5]. The results obtained and published in [29][19] show that the three properties of the dynamics of trust described above, namely, *asymmetry*, *maturity* and *distinguishable past*, do allow for a better understanding of the characteristics of the target agent in evaluation, increasing the accuracy of the trustworthiness estimations and, consequently, providing a better support to the process of selection of partners.

1.4 The Simulated Textile Scenario

In this section, we describe the simulated textile scenario that we used in order to evaluate the benefits of using trust to preselect the candidate partners in a negotiation process. This scenario is based on the concept of our *Electronic Institution* [2], an agent-based framework implemented in Jade [35] that we have developed for business-to-business e-contracting, which provides supporting services such as *automatic negotiation*, *contract monitoring and enforcement* and *computational trust*.

In our simulated scenario, textile client agents select the best suppliers of textile fabrics through a multi-round, multi-attribute negotiation phase. Every client has a business need randomly assigned at setup, which consists of a given component (fabric) and the associated preferential values for unit price, quantity and delivery time. Then, at every simulation *episode*, each one of the client agents announces its business need by issuing a call for proposal (CFP) to candidate suppliers.

In current experiments, every supplier registered in the Electronic Institution is able to provide any type of fabric. When a supplier agent receives a CFP concerning the provision of a given component, it generates a proposal based on its own preferential values for this specific component in terms of unit price, quantity and delivery times. This means that the client that issued the CFP has the expectation of receiving proposals from the candidate suppliers with different *utility* values.[3]

At the end of the negotiation phase, which involves the exchange of proposals and counter-proposals between the business partners (more concretely, between their *negotiation mediator agents*), the client agent finally selects the supplier whose proposal yields him the maximum utility, and a contract between both business agents is drafted by their respective negotiation mediators. From this moment on, the *normative environment* is responsible for monitoring the contract and to act upon offenders, if necessary.

1.4.1 Preselection of Trust

In addition to the basic negotiation process described above, every client agent registered in the Electronic Institution that desires to initiate a new negotiation process in order to satisfy a new business opportunity has now the option to preselect the suppliers that are able to proceed to the negotiation phase, based on the trust put on these suppliers. In the work presented in this paper, the trust assessment is done using *contractual information* derived from past business transactions performed on the Electronic Institution.

Figure 1.3 illustrates the main activities associated with the process of selecting partners, including the *Trust Assessment* activity performed by the

[3] The utility that a client derives from a received proposal is inversely proportional to the *deviation* between the values in the proposal and the preferential values of the client concerning the issued CFP, as we will see in Section 1.5.

Computational Trust agent. As we can observe in this figure, there is a connection between the output of the *Contract Monitoring* activity, performed by the *Normative Environment*, and the input to the *Trust Assessment* activity, performed by the *Computational Trust* service, which is related to the generation and management of the repository of contractual trust evidences of the Electronic Institution. We address this issue next.

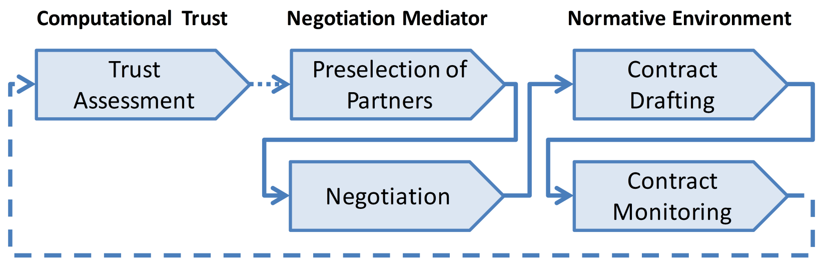

Fig. 1.3. Main activities associated with the process of partners selection and corresponding services

1.4.2 Generation of Contractual Trust Evidences

As a result of its monitoring activity, the *Normative Environment* agent asynchronously broadcasts to the *Computational Trust* agent contractual events concerning the execution of contracts by the entities involved in the contracts.

In the current simulated scenario, we use a single contract of sale, which stipulates that the supplier must deliver the component to the client before a specified deadline. If the component is delivered by this deadline, the client now has an obligation to pay the amount due to the supplier. However, if the supplier is late in providing the component, the deadline is violated and automatically the client denounces the violation of the delivery obligation. This simple contract is depicted in Figure 1.4.

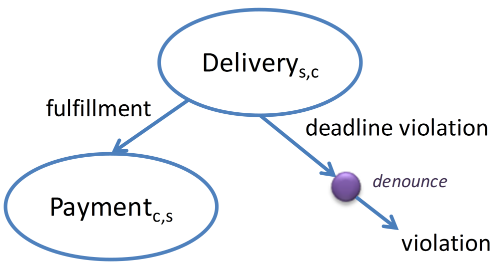

Fig. 1.4. The simple Contract of Sale

In this very simple example, two evidences must be generated by the *Computational Trust* service. One concerns the delivery of the component, where the supplier s acts as the bearer and the client c acts as the counterparty ($evd_{c,s}$), and the other concerns the payment from the client to the supplier ($evd_{s,c}$). For the sake of simplicity, we only consider here the generation of evidence $evd_{c,s}$.

Therefore, when an instance of the contract of sale is enacted, several contractual events are asynchronously sent from the *Normative Environment* agent to the *Computational Trust* agent: those concerning the establishment of a new contract and the end of the contract, which start and finish the process to generate trust evidences for this contract, respectively; the one that announces that an obligation (e.g. the delivery obligation) is triggered; and those concerning either the fulfillment of an obligation or its deadline violation followed by the event concerning the violation of the obligation. Contracts including more obligations may generate complex webs of related events. With the simple contract of sale considered above, the *Computational Trust* agent maps the generated events related to the delivery obligation into evidence $evd_{c,s} = (c, s, F)$ for the cases where the supplier s delivers the component on time to client c (therefore, s fulfilled its part of the contract – F) or into evidence $evd_{c,s} = (c, s, V)$, for the cases where s violates (V) its part of the contract.

1.5 Experiments

1.5.1 The Testbed

All experiments described in this document were performed using the Electronic Institution framework and the scenario described in the previous section.

In order to study the impact of filtering the candidate suppliers prior to the negotiation process, we configured three different experiments: in experiment $T5$ (top 5), client agents only allow for 25% of the most trusted agents to enter the negotiation phase; in experiment $T10$ (top 10), this number is raised to 50%; and in experiment NF (no filtering), no preselection is performed, meaning that all registered suppliers are allowed to enter the negotiation phase and, thus, to be selected to trade with the client. In the first two cases, the trustworthiness assessment is made using the *Sinalpha* trust model.

In all the experiments, we consider 10 clients and 20 suppliers. In every experiment, there are 15 *episodes* and, at every episode, a new negotiation cycle is started for each one of the 10 clients, with the announcing of a new business need contained in a CFP. The negotiation protocol used in all episodes is a 3-round protocol. At the first episode of each experiment, the repository of trust evidences is cleaned, which means that the estimated trustworthiness of all suppliers is zero. Finally, we run every experiment 30 times; at every new run, clients and suppliers change their preferential values for components, by randomly picking up from the values presented in Table 1.1.

Table 1.1. Configuration values for components, price, quantity (in *meters*) and delivery time (in *days*)

component	cotton, chiffon, voile
price	$1 - 10$
quantity	180, 270, 360, 450, 540, 630, 720, 810, 900, 1080
delivery time	7, 10, 13, 19, 19, 22, 25, 28, 31, 34

All suppliers are able to trade any component from the ones presented in Table 1.1, and they face no stock limitations. Each one of them has an intrinsic degree of *performance* reflecting the fact that it has some *handicap* in providing the required component in certain *circumstances*. Therefore, at setup time, each supplier is randomly assigned a handicap following a uniform distribution over the set of available handicaps, which are presented in Table 1.2.

Table 1.2. Different types of suppliers used in the experiments.

Handicap	Description
HFab	Handicap in providing a specific fabric
HQt	Handicap in providing high quantities
HDt	Handicap in providing low delivery times
HFabQt	Handicap in providing a specific fabric in high quantities
HFabDt	Handicap in providing a specific fabric in low delivery times
HQtDt	Handicap in providing high quantities in low delivery times

When a supplier is selected to entail business with a client, it will fulfill the established contract with a probability of 95% if it does not present a handicap on the agreed contractualized terms. Otherwise, it has a probability of 95% of violating the deadline of the contract and failing it. [4]

1.5.2 Evaluation Metrics

At every episode, the selection of the best proposal is done taking into consideration the *utility* of each one of the proposals received by the client in the

[4] In these experiments, the behavior of the agents is defined by given probabilities, meaning that they are not able to control their behavior. This is a restrictive approach that, however, is adopted in most of the works on computational trust (see, as examples, [8][7][5]). In fact, most of the real data available on agents' trustworthy behavior is derived from trust games and it is mostly restricted to correlations between measures of generalized trust collected over time (cf. [33]).

last round of negotiation. At every negotiation round, the *utility of a proposal* is given by the inverse of the *deviation* between the client preferences specified in the CFP, for all the negotiable items price, quantity and delivery time, and what is offered in the proposal in evaluation. Equation 1.3 shows the formula used to evaluate the proposals; if a received proposal meets exactly the client requirements, the formula is not applied and this proposal is considered the winner proposal in current round.

$$\text{utility of proposal} = \left(\sum_{i}^{k} \frac{v_{pref_i} - v_i}{max_i - min_i}\right)^{-1}. \tag{1.3}$$

In Equation 1.3, which is an adapted version of the one presented in [36], v_{pref_i} is the client's preferred value for negotiation item i, v_i is the value of the negotiation item i of the current proposal in evaluation, and max_i and min_i are the maximum and minimum values that the client accepts for item i, respectively. The metric *average utility of proposals* gives the average utility of the proposals (as measured in Equation 1.3) selected by all clients in every episode and every run of the simulation.

In these experiments, we consider that the *effective utility* gained by a client after a business transaction is the same as the utility of the proposal selected by the client in this transaction, if the associated contract is fulfilled by the supplier. However, if the contract is violated, the effective utility depends of the effect that a contractual breach has on the client. Equation 1.4) shows the formula used in our experiments to derive the effective utility of a transaction.

$$\text{effective utility} = \begin{cases} \text{potential utility, if fulfilled,} \\ \frac{-k}{100} \times \text{potential utility, otherwise.} \end{cases} \tag{1.4}$$

In the equation above, k is a parameter that reflects the impact that violated contracts have on the effective utility gained by the clients in the interactions with the selected suppliers. As an example, if a given client selects a proposal with potential utility of 2.40 units and posteriorly the associated contract is violated by the supplier, the effective utility achieved by the client is -0.24 for $k = 10$ and -0.6 for $k = 25$. In the experiments, we calculate the effective utility for distinct values of k, such that $k \in \{0, 10, 15, 20, 25, 30\}$. The metric *average effective utility* gives the average utility *effectively gained* by all clients in an episode, averaged over all episodes and all runs of the simulation.

In addition to the utility-based metrics, we used two other metrics: the *percentage of successful contracts*, averaged over all clients, all episodes and all runs; and the *percentage of distinct suppliers* selected at every episode by the 10 clients, averaged over all runs of the experiments.

1.5.3 Results

Table 1.3 presents the results obtained in experiments $T5$, $T10$ and NF related to the metrics based on utility. In the second row of the table, *proposal utility deviation* gives the standard deviation of the utilities of all proposals received by a client in the last negotiation round of one episode, averaged over all episodes, clients, and runs. The last six rows show the results obtained for the effective utility measure for the different values of k.

Table 1.3. The results obtained for utility-based metrics

	$T5$	$T10$	NF
average utility of proposals	2.60	3.25	4.63
proposal utility deviation	0.75	0.84	1.05
effective utility (k=0)	1.84	2.12	2.39
effective utility (k=10)	1.76	2.01	2.09
effective utility (k=15)	1.73	1.96	1.97
effective utility (k=20)	1.69	1.90	1.86
effective utility (k=25)	1.65	1.84	1.74
effective utility (k=30)	1.61	1.79	1.63

Table 1.4 presents the average percentage of successful contracts and the average percentage of different suppliers selected by episode, for experiments $T5$, $T10$ and NF.

Table 1.4. The results obtained for the percentage of successful contracts and for the percentage of different suppliers metrics

	$T5$	$T10$	NF
successful contracts (%)	68.77	64.38	49.91
different selected suppliers (%)	54.43	72.48	82.40

1.5.4 Interpretation of the Results

The first conclusion we can derive from the results presented above is that the preselection of candidate partners decreases the potential utility that can be achieved in a negotiation, whenever the utility of a proposal is calculated as a trade-off between several negotiation attributes (e.g. price, quantity and delivery time). This happens because the N most trustworthy suppliers selected for the negotiation phase are not necessarily the ones that present the proposals with higher utility to the client. In Figure 1.5, we plot the relative potential utility and the relative proposal utility deviation obtained by $T5$

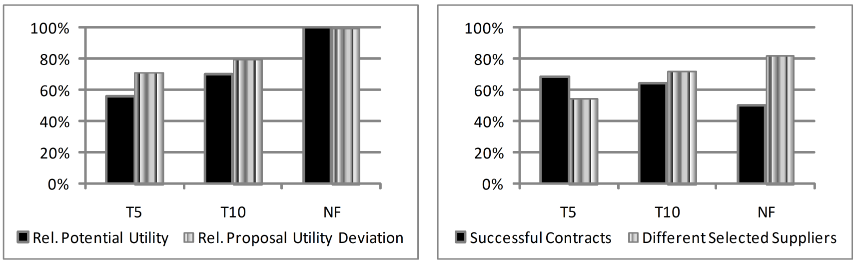

Fig. 1.5. Relative values of potential utility and proposal utility deviation for $T5$ and $T10$, when compared to those of NF (*left*); and percentage of successful contracts and percentage of different suppliers selected per episode (*right*)

and $T10$, when compared to the same values obtained by NF, in the absence of trust-based preselection of suppliers.

However, this apparent drawback of using trust-based preselection can result in improved utility for the clients in case the population of candidate suppliers has different attitudes towards the enforcement of contracts. For instance, both trust-based experiment types obtained significantly less violated contracts than experiment NF(cf. Figure 1.5 (*right*)).

The exact trade-off between getting more potential utility and less violated contracts strongly depends on the population of candidate partners and on the risk the clients are willing to take in the selection of these partners. If the population is relatively homogeneous in terms of contractual behavior, the contribution of trust in the selection decision can be relaxed. However, in open and global marketplaces, it is not expected that service providers share the same will or possibilities to fulfill contracts. In the same way, this trade-off depends on the meaning that clients give to breached contracts.

As can be observed in Figure 1.6, for the initial setup of the experiments with 3 rounds of negotiation (*down left graph*), not prefiltering the candidate partners is beneficial when k is less than 15, i.e. when the effective utility of unsuccessful interactions is higher than -15% of the potential utility. For higher values of k, the clients can get more value using approach $T10$. However, in the conditions set for these experiments, the finer filtering of candidates using the $T5$ approach is always worse in terms of utility for the clients than the other two approaches.

In order to better understand the trade-off between increasing the potential utility of clients, by allowing more candidate partners to get into the negotiation phase, and preventing partners to be too deceptive, by filtering these partners, we reran experiments $T5$, $T10$ and NF with two and four

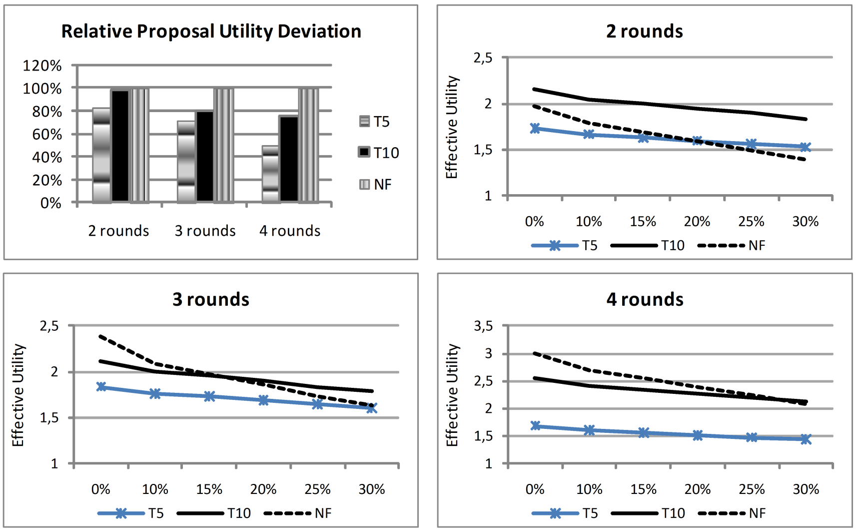

Fig. 1.6. Utility based results for experiments $T5$, $T10$ and NF with 2, 3 and 4 rounds of negotiation

rounds of negotiation.[5] Figure 1.6 graphically depicts the results obtained for the three types of experiments concerning the utility-based metrics.

Observing this figure, it becomes evident that the preselection of candidate partners is beneficial when these do not diverge in a relevant way in the proposals they make to the clients – which is a reasonable assumption to do in real open and wide marketplaces –, and that the excessive prefiltering ($T5$) gives worse results than a more moderate selection of partners ($T10$). However, this advantage of the trust-based preselection of partners is lost when the clients receive proposals with significantly different utilities and when they do not attach great importance to violated contracts.

The results obtained with these experiments showed us the enormous potential of using our Electronic Institution framework to deepen our studies on computational trust models. In fact, this framework allows us to go further than the traditional trust assessment approaches that measure utility exclusively in terms of the success/failure of a transaction and that do not take into account other negotiation dimensions besides trust.

The conclusions we draw here are preliminary, in the sense that several other configurations must be tested, such as the introduction of different types of populations of suppliers. In the same way, we intend to use other types of contracts, with an increased number of obligations. Finally, we address as

[5] We must bear in mind that, with the specific negotiation protocol we are using, the different proposals received by a client in response to the launched CFP will become increasingly disparate as the number of rounds increase (cf. Figure 1.6).

future work the inclusion of trust as another dimension of the negotiation phase, in addition, or instead, of its usage prior to the negotiation phase.

1.6 Conclusions

In the new economy, companies increasingly feel the need to seek new business partners who can bring them better quality, lower prices, or supply a temporary need. However, current organizations still do not dare to break their parochial relations due to the lack of reliable trust technologies. In the particular field of computational trust systems, a lot of work still needs to be done. Particularly, the computational trust models need to implement in their core modules coherent theories of trust from the social sciences.

In this paper, we presented *Sinalpha*, a trust aggregation model used to estimate the trustworthiness of agents that was inspired by the concept of the hysteresis of trust and betrayal and by the asymmetry principle of human psychology. By retaining the good characteristics of these conceptual models and by simultaneously addressing some of the limitations that preclude the computational implementation of the hysteresis-based approach, we succeeded in embedding in *Sinalpha* with the *asymmetry*, *maturity* and *distinguishable past* properties of the dynamics of trust.

Next, we presented a simulation scenario based on the trading of textile fabrics that we have run in an agent-based Electronic Institution framework, composed of several services, such as automatic negotiation, normative environment and computational trust. This scenario allowed us to run several experiments that evaluated the effects of preselecting partners based on their trustworthiness as a prior stage to the negotiation process. The results we have obtained showed us that the preselection of partners can yield more utility to clients when the impact of failed contracts is highly negative to the latter, and when the proposals received by them are not extremely disparate in terms of potential utility.

As future work, we intend to study further the phenomenon of preselection in order to derive the optimal number/percentage of candidates to be filtered in this process. Also, we are currently working on the configuration of *Sinalpha*'s parameters in order to make the model more adaptive to the different personalities of trusting agents and to the business circumstances.

Acknowledgments

The authors would like to thank Henrique Lopes Cardoso, Pedro Brandão and Sérgio Moura for their work on the Electronic Institution platform. This research is supported by project PTDC/EIA-EIA/104420/2008 of the Fundação para a Ciência e a Tecnologia (FCT). The first author is supported by FCT under grant SFRH/BD/39070/2007.

References

1. Urbano, J, Lopes Cardoso, H., Oliveira, E.: Making Electronic Contracting Operational and Trustworthy, in Advances in Artificial Intelligence Proceedings of the 12th Ibero-American Conference on Artificial Intelligence (IBERAMIA 2010), LNAI V. 6433, Springer, ISBN 978-3-642-16951-9, pp. 264–273, Bahía Blanca, Argentina, November 1-5 (2010)
2. Urbano, J., Lopes Cardoso, H., Oliveira, E., Rocha, A. P.: Normative and Trust-based Systems as Enabler Technologies for Automated Negotiation, F. Lopes and H. Coelho (Eds.), *Negotiation and Argumentation in MAS*, Bentham Science Publishers, To appear, 2011
3. Ramchurn, S., C. Sierra, L. Godo, N.R. Jennings: Devising a trust model for multi-agent interactions using confidence and reputation, International Journal of Applied Artificial Intelligence (18) pp. 833–852, (2004)
4. Sabater, J., C. Sierra: REGRET: A Reputation Model for Gregarious Societies. In Fourth workshop on deception fraud and trust in agents societies, 61–70 (2001)
5. Huynh, T. D., N. R. Jennings, N. R. Shadbolt: An integrated trust and reputation model for open multi-agent systems, in Autonomous Agents and Multi-Agent Systems, Vol. 13, N. 2, pp. 119–154 (2006)
6. Carbo J., J. Molina, J. Davila: A BDI Agent Architecture for Reasoning About Reputation, in IEEE International Conference on Systems, Man, and Cybernetics, v2, 817–822 (2001)
7. Jøsang A., R. Ismail: The Beta Reputation System, in Proceedings of the 15th Bled Electronic Commerce Conference, Slovenia (2002)
8. Patel, J.: A Trust and Reputation Model for Agent-Based Virtual Organisations. PhD thesis, University of Southampton (2006)
9. Reece, S., A. Rogers, S. Roberts, and N. R. Jennings: A Multi-Dimensional Trust Model for Heterogeneous Contract Observations. In: Twenty-Second AAAI Conference on Artificial Intelligence, Vancouver, Canada (2007)
10. Zacharia, G., P. Maes: Trust management through reputation mechanisms. In Applied Artificial Intelligence, 14(9), 881–908 (2000)
11. Haller, J.: A Bayesian Reputation System for Virtual Organizations, In Dagstuhl Seminar Proceedings 06461, Negotiation and Market Engineering (2007)
12. Melaye, D., Y. Demazeau: Modles et Rseaux de Confiance, in Les Cahiers Leibniz, n. 142, ISSN: 1298-020X (2005)
13. Hang, C., Wang, Y., and Singh, M. P.: An adaptive probabilistic trust model and its evaluation. In Proceedings of the 7th international Joint Conference on Autonomous Agents and Multiagent Systems -V3 International Conference on Autonomous Agents. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1485–1488 (2008)
14. Erete, I., E. Ferguson, and S. Sen: Learning task-specific trust decisions. In Proceedings of the 7th international Joint Conference on Autonomous Agents and Multiagent Systems - Volume 3 International Conference on Autonomous Agents. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1477–1480 (2008)
15. Smith M., M. desJardins: Learning to trust in the Competence and Commitment of Agents, in Autonomous Agent Multi-Agent Systems, 18:36–82 (2009)
16. Sabater-Mir, J., I. Pinyol, D. Villatoro, G. Cun: Towards Hybrid Experiments on reputation mechanisms: BDI Agents and Humans in Electronic Institutions,

in XII Conference of the Spanish Association for Artificial Intelligence (CAEPIA-07), V2, pp. 299–308 (2007)
17. Slovic P.: Perceived risk, trust, and democracy, Risk Analysis, vol. 13, no. 6, pp. 675682 (1993)
18. Bosse T., Jonker, C. M., Treur, J., Tykhonov, D.: Formal Analysis of Trust Dynamics in Human and Software Agent Experiments. In Proceedings of the 11th international workshop on Cooperative Information Agents XI (CIA '07), Matthias Klusch, Koen V. Hindriks, Mike P. Papazoglou, and Leon Sterling (Eds.). Springer-Verlag, Berlin, Heidelberg, pp. 343–359 (2007)
19. Danek A., Urbano J., Rocha A. P., Oliveira E.: Engaging the dynamics of trust in computational trust and reputation systems, in KES-AMSTA (1), LNCS, P. Jedrzejowicz, N. T. Nguyen, R. J. Howlett, and L. C. Jain, Eds., vol. 6070. Springer, pp. 2231 (2010)
20. Straker D.: Changing Minds: in Detail. Syque Press (2008)
21. Jonker, C. M., Treur, J.: Formal analysis of models for the dynamics of trust based on experiences, pp. 221-231 (1999)
22. Rettinger, A., M. Nickles, and V. Tresp: A statistical relational model for trust learning. In Proceedings of the 7th international Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2, Richland, SC, pp. 763–770 (2008)
23. Tavakolifard M.: Situation-aware trust management, pp. 413-416, (2009)
24. Rehak, M., Gregor, M., Pechoucek, M.: Multidimensional context representations for situational trust, DIS 2006: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications, Proceedings, pp. 315-320 388 (2006)
25. Nguyen C. T., Camp O.: Using context information to improve computation of trust in ad hoc networks, in Networking and Communications, WIMOB08. IEEE International Conference on Wireless and Mobile Computing, pp. 619-624 (2008)
26. Urbano, J., Rocha, A. P., Oliveira E.: Trustworthiness tendency incremental extraction using information gain, Web Intelligence and Intelligent Agent Technology, IEEE/WIC/ACM International Conference on, vol. 2, pp. 411-414 (2010)
27. Urbano, J., Rocha, A. P., Oliveira E.: A Situation-aware Computational Trust Model for Selecting Partners, N.T. Nguyen and R. Kowalczyk (Eds.): Transactions on Computational Collective Intelligence III, LNCS 6560, To appear, 2011
28. Kerschbaum, F., Haller, J., Karabulut, Y., Robinson, P.: Pathtrust: A trust-based reputation service for virtual organization formation. In Trust Management, LNCS 3986, pp. 193-205 (2006)
29. Urbano, J., Rocha, A. P., Oliveira E.: Computing Confidence Values: Does Trust Dynamics Matter? in 14th Portuguese Conference on Artificial Intelligence (EPIA 2009), L. Seabra Lopes, N. Lau, P. Mariano, and L. M. Rocha, Eds., vol. LNAI. Aveiro, Portugal: Springer, pp. 520-531 (2009)
30. Lapshin, R. V.: Analytical model for the approximation of hysteresis loop and its application to the scanning tunneling microscope. In Review of Scientific Instruments, volume 66, number 9, pages 4718–4730 (1995)
31. Kiyonari, T., Yamagishi, T., Cook, K., Cheshire, C.: Does Trust Beget Trustworthiness? Trust and Trustworthiness in Two Games and Two Cultures: A Research Note. In Social Psychology Quarterly, Vol. 69, No. 3, pp. 270–283 (2006)
32. Castelfranchi, C., Falcone, R.: Trust Theory: A Socio-Cognitive and Computational Model, John Wiley and Sons (2010)
33. Hardin, R.: Conceptions and Explanations of Trust. Trust in Society (edited by Karen S. Cook) New York: Russell Sage Foundation (2001)

34. Takács, J.: Mathematics of Hysteretic Phenomena: The T(x) Model for the Description of Hysteresis. Wiley-VCH, Weinheim (2003)
35. Bellifemine, F., Caire, G., Poggi, A., Rimassa, G.: Jade: A software framework for developing multi-agent applications. lessons learned, Information and Software Technology, vol. 50, no. 1-2, pp. 10–21 (2008)
36. Rocha, A. P., Oliveira, E.: An Electronic Market Architecture for the Formation of Virtual Enterprises. PRO-VE 1999: 421–432 (1999)