In the Search of Better Deals Using Trust

Joana Urbano and Ana Paula Rocha and Eugénio Oliveira

Abstract. In current days, a great effort, on several scientific research areas, is being devoted to the automation of the procurement processes in the framework of business-to-business relations. However, several constraints limit the extension of the procurement process to truly open and global marketplaces. One such constraint is the lack of valid trust mechanisms that allow business agents to select partners outside the sphere of known relationships, in the hope of better deals. Following our firm belief that trust systems used in such hard environments shall be situation-aware and flexible, we present our approach for such a system. Also, we set up a multi-agent environment based on client-supplier transactions in order to evaluate two fundamental questions. First, whether the break of parochial relations allows for better deals without jeopardizing the overall utility of client agents by engaging on transactions with strangers. And second, how representative trust systems – including our own approach – support the exploration of new partners in a relative safe way.

1 INTRODUCTION

In current days, a strong I&D effort in different research areas is being put in the automation of the procurement processes of business-to-business relations. Ideally, the development of adequate agreement technologies would allow opening the space of business opportunities, where potential partners have the means and the confidence knowledge to search for good business opportunities outside their limited sphere of breeding relations.

However, there are actually constraints that limit this desired openness, and the fear of risking unknown partners is one of the biggest barriers to trade in a truly open market environment. In fact, economic exchanges with strangers can result in harm for the intervenient agents in both ways. Concerning a subcontracting scenario in the textile industry as an example, the client part of the relation can be deceived by the provider part in several different ways [1]:

- A delay in delivery, which affects all the supply chain;
- The quality received, as specified by affordability, safety and degree of uniqueness parameters;
- The quantity received (too much or too less);
- The violation of intellectual property rights;
- Ethical problems;
- Other problems, such as price rise and legislative changes.

Without some kind of a trust mechanism, it is reasonable to conclude that business partners would preferentially adopt parochial environments in detriment to more aggressive exploration of deals outside the already known partner relationships space. For instance, in the fashion retail industry, clients often rely on knowledge available through textile fairs and textile agents to make the bridge between brands and the trustable and reliable textile suppliers. However, even with this form of trust guarantees, the space of available suppliers is relatively small and strongly supported on the expected behavior of the partner, rather than on the real utility of the business transaction.

This paper addresses the need to develop computational trust systems that can be used in open and global markets, where the evidences that can be used to infer the trustworthiness of business agents are scarce, contextual and heterogeneous. It is our firm belief that these systems, in order to be acceptable, shall present the following characteristics:

- The aggregation of trust evidences into trustworthiness scores shall follow important properties of the dynamics of trust, as addressed in research areas related to social sciences and psychology;
- The trust inference process shall be fed from diversified and heterogeneous sources of information, such as past direct experiments, available reputation, specific recommendations, perceived roles, and credit ranking agencies;
- The trust system shall be able to detect different situations, and also to detect tendencies of agents’ behavior in the presence of such different situations;
- The trust system shall be able to infer the trustworthiness of agents even where the number of trust evidences is scarce.

In this paper, we set up a multi-agent simulation environment where textile client agents seek for optimal deals by selecting from a range of suppliers with different behaviors. Particularly, we address two main questions. First, we evaluate whether the break of the breeding parochial relations between business partners permits to increase the overall utility of clients or, on the contrary, jeopardize it as a consequence of the risk introduced by this strategy.

Then, we present our own approach to a trust system that was designed taking into consideration the hard requirements of open markets. Moreover, we evaluate how different trust methods, including our proposal, can support the exploration of new potential partners in such a way that the risk associated to trading with strangers is decreased by the method.

A third achievement of the work presented in this paper is the development of a simulation environment that can be used to support important studies about parochialism and trust that are being done in the social sciences area, such as the ones presented in [2] and in [3].

The rest of this paper is structured as follows. Section 2 presents our trust method, with particular emphasis on the Contextual Fitness component. Section 3 is devoted to the experiments we performed in order to evaluate the issues referred above. Finally, section 4 concludes the paper.

2 OUR CONTEXT-BASED TRUST METHOD

We developed a trust method envisioning its use in future automatic and open business-to-business markets, taking into special consideration the exigent requirements of such environments. Namely, we are concerned with the performance of the method when the trust evidences available about a given target agent are scarce and heterogeneous, and when the activity of the agents under evaluation can span through different situations and contexts.

The current implementation of our system that encompasses the proposed method is composed of two different modules, as depicted in Figure 1:

---

1 LIACC – Laboratory for Artificial Intelligence and Computer Science, Faculdade de Engenharia da Universidade do Porto – DEI, Rua Dr. Roberto Frias, 4200-465, Porto, Portugal, email: joana.urbano@fe.up.pt, arocha@fe.up.pt, ecor@fe.up.pt
The aggregator component, which is responsible for aggregating the available trust evidences of an agent into a trustworthiness score for this agent. Several trust engines that are defined in the literature can be used (some of the more representative models are reviewed in [4]), although we are particularly interested on engines that model the dynamics of trust, as those described in [5] and [6], as they appear to perform better than the traditional statistical approaches;

- The Contextual Fitness component, which tunes the outcome of the aggregating step by taking into consideration the specificities of the current business opportunity and the adequacy of the target agent to the specific situation under assessment.

![Diagram of the current implementation of our trust system](image)

The idea beyond this extension is that if the trust system detects that a target agent has some kind of handicap related to the current necessity, the overall trustworthiness of the agent for this necessity will be zero; otherwise, the trustworthiness score is the value computed by the simple trust aggregator for the same situation. One good characteristic of this approach is that it can be used with any conventional trust aggregation engine, being it based on statistical, probabilistic or heuristic models, as it is the case of those reviewed in [4].

Before we describe the Contextual Fitness component in more detail, we first introduce the notation and the scenario used all over this paper.

### 2.1 Scenario and Notation

In the context of this paper, we define $\text{trust}_{ac}(As) \in [0, 1]$ as the trustworthiness value of a trustee agent $As$, in the eye of a trustee agent $Ac$, as computed by a traditional trust aggregator engine; and $\text{adequacy trust } ad(As, at) \in [0, 1]$ as a binary operator for situation-awareness purposes, where:

- $Ac \in C$ is an agent from the set $C$ of client agents;
- $As \in S$ is an agent from the set $S$ of supplier agents;
- $at \in AT$ describes the need, i.e. an instance of the space $AT$ of all possible combinations of attribute-value pairs that describe the need (good, product or service).

In the textile scenario mentioned in the introductory section, a need is announced through a call for proposals issued by a client, and concerns the delivery of some quantity of a fabric due in some delivery time. Thus, an example of an instance of the $AT$ space is the following: $\{\text{fabric='cotton', quantity='900000', delivery time='15'}\}$. It is worth to note that, for the sake of scalability, all quantitative values are previously quantified using fuzzy techniques.

Therefore, the trustworthiness value of agent $As$ as seen by agent $Ac$ in the specific context $at$ is given by the following equation:

$$\text{trust}_{ac}(As, at) = \text{trust}_{ac}(As) \cdot ad_{ac}(As, at) \quad (1)$$

This is the same as to say that, in a given moment, an agent may be qualified as trustworthy in some situation and as untrustworthy in a (maybe slightly) different situation.

Finally, a contractual evidence represents a transaction that took place between a client agent $Ac$ and the selected partner agent $As$, for which an outcome $o \in \{\text{true, false}\}$ is generated. I.e., agent $As$ either succeed to provide the good in the contractual terms or violate the contract. Each supplier agent $As$ will, therefore, have an history of its past contractual evidences, each one represented by the tuple $<Ac, As, at, t, o, \sigma>, \text{ where } t$ is the timestamp of the transaction, needed when in use with aggregation systems that weight evidences by their recency.

#### 2.2 The Contextual Fitness Component

The Contextual Fitness component is based on an online, incremental and flexible technique of behavior tendencies extraction that we have developed. Although we have been testing different methods for extracting these tendencies from the historical set of agents’ evidences, our current version uses the information gain metric. This is a metric used in the machine learning area (such as in the simple decision tree learning algorithm ID3 [7]) for classification purposes. It is typically used as an offline process, implying that the training and testing phases occur before the actual classification of new instances is performed.

The information gain metric is based on the entropy concept of information theory, and is defined as following:

$$Gain(S, A) \equiv \text{Entropy}(S) - \sum_{s \in \text{Values}(A)} \frac{|s|}{|S|} \text{Entropy}(s) \quad (2)$$

where $Gain(S, A)$ is the information gain of attribute $A$ relative to a collection of samples $S$, $Values(A)$ is the set of all possible values for attribute $A$, and $s$ is the subset of $S$ for which attribute $A$ has value $v$ ($\{7\}$).

In our approach, we use this metric to dynamically learn a decision tree from the history of evidences of a given agent $As$, every time it is necessary to verify the adequacy of the agent proposal to the current need announced by the client agent. In fact, we use all the evidences available about the supplier to build a decision tree, which normally consists of a dataset with some dozens of evidences, if that much. This means that no training or testing phases are performed. After that, the failure tendencies of the agent in evaluation are extracted from the rules pointing to false outcomes. Figure 2 depicts a decision tree that was learned for a given supplier in a specific experiment we have run (we use the Weka API [8] in our simulations).

Concerning the tree below, our algorithm was able to identify that, at the time of this particular assessment, the supplier showed a tendency to fail contracts that match the tendencies $\{\text{good = cotton, \_dtime = low}\}$ and $\{\text{good = voile,\_\_}\}$. Therefore, the trustworthiness value $\text{trust}_{ac}(As, at)$ of agent $As$, as given by Equation 1, would be zero if situation $at$ matched any of the tendencies derived from the learned decision tree. Otherwise, the trustworthiness value of the target agent for the considered situation would be given by the $\text{trust}_{ac}(As)$ component of equation 1.

Several issues may arise from the use of the information gain criteria in our technique, such as the need to prune the generated trees or the need to use similar metrics that permit heterogeneous evidences (e.g. the gain ratio metric presented in [9]). We address the first question in [10], and leave the second one to future work. On the other hand, we are interested, in this paper, in evaluating the adequacy of our technique when applied to open market environments, where clients risk trading with suppliers that reside outside the space of the clients’ breeding environment.
In the experiments, we generate a population of clients that have different perspectives concerning the selection of suppliers, depending if it is within the space of their embedded relationships or outside it. Moreover, all suppliers have different handicaps on performing some particular aspect of a business transaction. For instance, some suppliers tend to fail to deliver fabric in short delivery dates, while others might fail to deliver high quantities of any fabric type. The aim of these experiments is to evaluate whether clients that better explore the space of available suppliers would achieve, in the end, higher utility than the clients that adopt a parochial, conservative, strategy, and how a particular trust aggregation technique can better assist this decision.

In this paper, we run two different sets of experiments, as described in the following sections. In the next section, we describe the generic configuration common to both sets of experiments.

3 EXPERIMENTS

We set up a multi-agent simulation scenario where business clients in the textile industry try to select the best suppliers of textile fabric, i.e. the ones that would maximize the utility of the clients. In the experiments, we generate a population of clients that have different perspectives concerning the selection of suppliers, depending if it is within the space of their embedded relationships or outside it. Moreover, all suppliers have different handicaps on performing some particular aspect of a business transaction. For instance, some suppliers tend to fail to deliver fabric in short delivery dates, while others might fail to deliver high quantities of any fabric type. The aim of these experiments is to evaluate whether clients that better explore the space of available suppliers would achieve, in the end, higher utility than the clients that adopt a parochial, conservative, strategy, and how a particular trust aggregation technique can better assist this decision.

In this paper, we run two different sets of experiments, as described in the following sections. In the next section, we describe the generic configuration common to both sets of experiments.

3.1 Generic Configuration

Table 1 presents the configuration parameters that are common to both sets of experiments.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabrics</td>
<td>{Chiffon, Cotton, Voile}</td>
</tr>
<tr>
<td>Quantities</td>
<td>{Low, Medium, High}</td>
</tr>
<tr>
<td>Delivery Time</td>
<td>{Low, Medium, Big}</td>
</tr>
<tr>
<td># buyers</td>
<td>20</td>
</tr>
<tr>
<td># of sellers</td>
<td>50</td>
</tr>
<tr>
<td>Seller stocks</td>
<td>4 contracts per round</td>
</tr>
<tr>
<td>Types of sellers</td>
<td>Uniform distribution over the</td>
</tr>
<tr>
<td></td>
<td>types considered in population</td>
</tr>
<tr>
<td># rounds</td>
<td>60</td>
</tr>
<tr>
<td># runs/ experiment</td>
<td>15</td>
</tr>
</tbody>
</table>

In all experiments, every client selects a supplier that is expected to maximize the utility of the client, where the notion of utility is specific of each set of experiments, as described in the next two sections. It is assumed that every supplier is able to provide any different type of fabric. Also, every client is initialized with a specific business need, represented in a form of a call for proposals (cfp) that is an instance $e AT$, as described in section 2.1. The clients keep their need constant in all the experiment rounds. Moreover, all values of possible quantities and delivery times are fuzzified in the categorical values depicted in Table 1.

As for the suppliers, as in order to simulate realistic behaviors, they are initialized with a specific behavior that denotes a specific handicap in specified contexts. For example, a supplier initialized with the $HQT$ behavior (standing for 'Handicap in Quantity') has a handicap in providing high quantities of any fabric; this way, if it is selected to a business transaction that involves the delivery of high quantities of fabric, it has a probability of 95% of failing the contract. In any other transaction that the supplier is involved, it will have a probability of only 5% of failing the contract. In our experiments, we used five other types of behavior that represent five other different types of handicap: on a given fabric ($HFAB$), on low delivery times ($HDT$), on high quantities of a given fabric ($HFABQ$), on low delivery times for a given fabric ($HFABDT$), and on high quantities to be delivered in low delivery times ($HQTDT$). As happen with the example before, a supplier has a probability of 95% of violating a contract if the current cfp matches the supplier’s handicap, and 95% of probability of succeeding the contract otherwise.

3.2 First set of experiments

3.2.1 Testbed and methodology

The first set of experiments was designed in order to evaluate the tendency of different trust models in exploring new business opportunities and how this capacity translates in terms of succeeded transactions. We used in this set of experiments three different models:

- **$SA$**: this model represents $SinAlpha$, a traditional trust aggregation system that uses properties of the dynamics of trust. The model was developed by us and is described in [6]. In there, we claim that this model gets better results than traditional statistical aggregation engines using weighted means;
- **$CS$**: this model goes one step further traditional trust models by considering contextual aspects of the business in assessment. It represents the model described in [11], a situation-aware technique that defines a context space as a n-dimensional metric space with one dimension per each represented situation feature. It is able to estimate trustworthiness values in unanticipated situations using the similarity between both situations. In the current experiments, we placed the reference contexts regularly over the combinations of possible values of the contractual attributes. This approach in a way represents situation-aware proposals that use domain specific, predefined similarity metrics to predict unanticipated situations ([12], [13], [14]);
- **$CF$**: this is the Contextual Fitness technique described in section 2.2 that is used in these experiments complimentary to the $SA$ approach. As with the approach $CS$ defined in the previous point, it is a situation-aware trust model. It was designed to fit well to non parochial open market scenarios, where the number of available trust evidences for a particular partner agent might be scarce.

3.2.2 Evaluation metrics

In this stage of experiments, we want to evaluate how client agents tend to behave in terms of selection of partners – and how good it is their decision on that – when using each one of the approaches defined above, representing respectively the traditional trust approach, more recent situation-aware trust models, and our own proposal for situation-aware trust assessment in scenarios that might involve scarcity of trust evidences.

Therefore, we use here two different metrics: the average utility got by all clients at every negotiation round, measured by the ratio given by the number of succeed contracts over the number of all contracts negotiated in the round, and the number of different suppliers that were selected by all the negotiating clients at every round.
3.2.3 Results

Figure 3 shows the results obtained in the first set of experiments. As can be observed from the graphic (bottom), both the SA and the CS approaches are relatively conservative concerning the selection of partners, where the 20 clients of the experiments choose in average between 9 and 10.5 different suppliers at each round.

![Figure 3. Average utility (top) obtained versus the average number of suppliers selected (bottom) at every negotiation round.](image)

On the other hand, the CF approach starts, since the first rounds, exploring a larger number of different suppliers and keeps showing this behavior all over the rounds. The described behavior of the approaches seems to be related with the utility that is achieved by them, as can be observed from the top plots of the graphic. In fact, the approach that is able to select from a greater number of different suppliers (the CF approach) also gets in average significant better utility (90.46%) than the other two approaches (83.30% for SA, and 85.87% for CS).

Another important result obtained with the CF approach is that, after some quick learning at the first rounds, the number of succeeded contracts with this method is very close to the maximum of 19 contracts that can succeed per round, i.e. to the 95% probabilistic limit imposed in the population of suppliers generated for our experiments.

3.2.4 Interpretation of results

At a first sight, we could expect that an approach that explores more business partners in the scenario described would have a smaller number of succeed contracts, at least in the first rounds of suppliers’ exploration. However, the results show that the CF approach does not perform worse than the remaining representative approaches at this first exploration phase and performs significantly better than the others in the remaining steps of the experiments. This is due to the fact that the CF approach is able to extract tendencies of behavior with a reduced number of trust evidences.

In fact, when aggregating the trust evidences in order to compute the trustworthiness score of the agent in assessment, the CS model weights each trust evidence with the relative similarity between the evidence and the current cfp situation. When the number of trust evidences is scarce, it is not possible to populate all reference contexts that are sampled in the multi-dimensional evidence space, and the differences between different situations remain tenuous. In these conditions, the approach has a tendency to select, from the set of the more fitted suppliers, the ones that have already been involved in more contracts. Related to the SA approach, the CS approach has the benefit of contextualizing the decision, therefore achieving higher utility. But its relatively embedded parochial behavior explains why the results obtained are worse when compared to the CF approach.

3.3 Second set of experiments

3.3.1 Testbed and methodology

In the second set of experiments we generated two different populations of client agents in order to evaluate the potential benefits of choosing open market strategies, in terms of the global utility achieved by the clients. In these experiments, we used the CF approach, based on the results obtained in the first set of experiments that show that this approach is the more adequate to selecting partners in open market scenarios.

Therefore, we used the following client type of populations:

- **Parochial**: this population includes client agents that favor known trustworthy suppliers instead of risking new, probably better suppliers. Their decision process of selecting partners is exclusively based on the trustworthiness scores of each supplier under assessment;

- **Non parochial**: this population includes client agents that seek to maximize their utility in the transacted operations. The decision process of these agents is based not only on the trustworthiness values of each supplier in assessment, but also in the expected value that come arise from the transaction with the supplier. In reality, these clients select the partner with whom they will trade based on their calculated utility, i.e. the product of the trustworthiness score of agents and their internal value, as explained below.

In this set of experiments, we introduced the notion of value of a contract. This value can be expressed through several characteristics of real world supplier connections, such as the price promised by the suppliers, the convenience of transacting with a given supplier, or even ethical and ecological concerns related to that particular partner, etc. In order to perform our experiments, each supplier has now an internal value that is assigned at the initialization time, following a uniform distribution on the set \{0.50, 0.60, 0.70, 0.80, 0.90\}.

Also, it is assumed that the value of an unknown supplier is 1.0, and the true value of the supplier is only presented to a given client after this client has already traded with the supplier. Through these settings, we expect that non parochial-based clients are motivated to explore new partners, as the potential calculated utility that arises from exploring outside the embedded space of relationships is high.

3.3.2 Evaluation metrics

In order to evaluate whether the exploration of new partners can increase the utility achieved by the clients, we compare both parochial and non parochial strategies based on the following different metrics: the number of successful contracts achieved by
all clients at every negotiation round and respective average number of contracts over all rounds; the number of different suppliers selected at every round; the average utility achieved by the clients at every round and its average score over all rounds.

3.3.3 Results

After performing the experiments, we observed that both approaches got similar results concerning the average number of successful contracts per client (parochial: 90.89%; non-parochial: 90.67%). However, the non-parochial strategy leads to a significant higher utility (75.25%) than the parochial strategy (68.59%).

Figure 4 plots the number of successful contracts, the number of different suppliers chosen and the utility observed per round, for each one of the two strategies. As expected, due to the use of the above introduced internal values (< 1.0), the achieved utility is always lower than the number of successful contracts per round.

3.3.4 Interpretation of results

In this second set of experiments, we used the same trust method to support the process of partners’ selection, which explains the similar results obtained by both strategies concerning the number of successful contracts and the number of different suppliers selected per round. This is due to the CF capability in distinguishing between different handicapped partners and to reason based on their adequacy to the current business situation.

However, the most interesting conclusion extracted from this set of experiments is tied to the results obtained when the selection of partners is done taking into consideration not only the trustworthiness estimation of each supplier, but also the estimated value of each supplier, which mirrors a much more realistic situation. The results obtained have shown that the flexibility of the online tendency extraction of CF allows to safely exploring a larger space of opportunities, permitting the identification of different characteristics of suppliers. For instance, using the CF approach, a client agent feels safe to explore other potential trustable partners outside its previously known group of trustable suppliers, which in turn can bring extra utility (e.g. better prices) to the client business.

4 CONCLUSIONS

In this paper, we strongly support our previous firm belief that true open business-to-business markets need robust, computational assisted trust mechanisms that deal with heterogeneous, contextual and probably scarce evidences in order to effectively compute trustworthiness scores for business agents. However, current computational trust model proposals do not address these issues in a practical way. Therefore, we introduced in this paper a trust technique that is able to effectively extract tendencies of agents’ behavior in different scenarios and contexts, even when the number of trust evidences is reduced. This technique can be used with any conventional trust aggregation engine.

Then, we experimentally evaluated how different trust model approaches, including our proposal, behave in environments where agents can seek business partners outside their breeding trading acquaintances. We verified that the focus on the online and incremental extraction of behavior features proposed by our technique effectively supports the exploration of new potential partners and, consequently, of new business opportunities, without jeopardizing the overall utility of business agents.

ACKNOWLEDGMENTS

The first author enjoys a PhD grant with reference SFRH/BD/39070/2007 from the Portuguese Fundação para a Ciência e a Tecnologia.

REFERENCES


