Trustworthiness Tendency Incremental Extraction Using Information Gain

Joana Urbano, Ana Paula Rocha, Eugénio Oliveira
LIACC – Laboratory for Artificial Intelligence and Computer Science, Faculdade de Engenharia da Universidade do Porto – DEI, Rua Dr. Roberto Frias, 4200-465, Porto, Portugal
joana.urbano@fe.up.pt, arocha@fe.up.pt, eco@fe.up.pt

Abstract

Computational trust systems are getting popular in several domains such as social networks, grid computing and business-to-business systems. However, the estimation of the trustworthiness of agents is not trivial in scenarios where the existing trust evidences are scarce. We propose an online, situation-aware trust model that uses the information gain metric to dynamically extract tendencies of failure of target agents, improving the process of selection of partners in a relevant way. Experimental results presented in this paper show that our proposal outperforms other trust approaches in contextual scenarios.

1. Introduction

The trust research field is generating interesting models of trustworthiness representation and estimation that can be used in application fields such as online auctions, recommenders systems, and social networks (see [1] for an overview). In these domains, the existence of trust evidences for the agents in evaluation is normally large. However, there are other application domains where the trust evidences available on target agents might be scarce. It is our conviction that in these cases the trust models should be equipped with capabilities that are not present in the traditional approaches referred above, namely, the capacity of dynamically discovering tendencies of behavior of agents in *specific situations*.

We set up a social simulation scenario where trading client agents attempt to place orders of textile fabrics to the best available partners, taking into account the evaluated trustworthiness of these partners (other conditions, such as price or payment conditions, are not considered). Typical buying leads discriminate fabric, quantity and the delivery time, and we use these to represent contracts between clients and suppliers. Finally, this is an open-market scenario, where client agents risk for better opportunities outside their sphere of embedded relationships, which might include transacting with unknown partners. In such a scenario, it is our conviction that the clients need an *online*, *situation-aware* computational trust system that is able to detect changes of behavior of agents and their failure tendencies in specific conditions, in situations where the available trust evidences can be scarce.

Although the need for situation-aware trust has been identified in the literature for some years, only few, recent trust models implement this functionality (see [3] for an overview). Also, the majority of these proposals relies on ontologies and implies the definition of domain-based similarity measures for situation comparison, which limits the scalability and the extensibility of the process. We propose a different approach for a situation-aware trust model. It is a contextual extension to traditional trust systems that dynamically extracts failure tendencies of target agents taking into account the specific situation in assessment. This is a flexible and online process that does not need previous knowledge of the agents in evaluation.

In previous work [3], we used the *increase in frequency* metric [4] for the extraction of tendencies. Although the results obtained were promising, we detected that the accuracy of the tendencies extracted could be improved. Therefore, we propose in this paper an enhanced version of our trust model that performs incremental characterization of negative evidences using the *information gain* metric proposed in [5]. Despite the inherent simplicity of this metric, its innovative use as an online, incremental process instead of the traditional historic based classification process revealed to be an effective alternative to tendencies extraction when the number of available evidences is scarce, as will be presented in the experimentation section.

The remaining of the paper is organized as follows: section 2 describes our proposed trust mechanism. Section 3 describes the experiments performed and. presents concluding remarks and future work.

2. Online extraction of behavior tendencies

Our proposal extends traditional trust systems by generating a value of the *adequacy* of a given target agent to the specific situation in assessment. If the trust system detects that an agent has some kind of *handicap* related to the current necessity, the overall trustworthiness of the agent for this necessity will be zero; otherwise, it is the value computed by the trust aggregator for the same situation.

We define $trust_{Ac}\ (As) \in [0,1]$ as the confidence of agent *Ac* in agent *As*, as computed by a traditional trust aggregator engine; *adequacy trust* $ad_{Ac}(As, at\ ) \in \{0,1\}$ is a situation-awareness binary operator, 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, and $at \in AT$ describes the *necessity*, i.e. an instance of the space *AT* of all possible combinations of attribute-value pairs that describe the necessity. In our textile scenario, a necessity is given by a buying lead issued by a client concerning the delivery of some quantity of a fabric due in some delivery time. The trustworthiness value of agent *As* as seen by agent *Ac* in the specific context *at* is given by the following equation:

$$trust_{Ac}\ (As, at) = trust_{Ac}\ (As) \cdot ad_{Ac}(As, at\ ) \quad (1)$$

Finally, a *contractual evidence* represents a transaction at time t between a client agent *Ac* and an agent *As*, for which an outcome $o \in \{true, false\}$ 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 $< Ac, As, at, t, o >$.

2.1. Extraction using frequency increase

In [3] we proposed the use of the increase in frequency metric (2) for characterizing tendencies of behavior of an agent. The negative trust evidences (e.g. violated contracts) are put in the false class and the stereotype of this class is extracted through characterization of concept, following the equation:

$$\propto = \left(\frac{|InstAttCluster|}{|InstCluster|}\right)^2 - \left(\frac{|InstAttTotal|}{|InstTotal|}\right)^2, \quad (2)$$

where |*InstAttCluster*| is the number of times that a given attribute-value pair appears in the class; |*InstCluster*| is the total number of evidences in the class; |*InstAttTotal*| is the number of times that the same appears in all classes, and |*InstTotal*| is the total number of evidences kept for the trustee.

Preliminary experiments ([3]) showed that this approach is effective when compared to non situation-aware approaches; however, analysis and experimentation showed us that the used metric sometimes fails to capture existing correlations between these attributes, as can be seen in the example presented in Table 1.

Table 1. Past contractual history of a supplier (fabric, quantity, delivery time, outcome)

cotton,medium,big,true	*voile,medium,medium,true*
cotton,high,medium,true	*voile,medium,big,true*
cotton,medium,big,true	*voile,medium,medium,true*
voile,medium,medium,true	*cotton,low,big,true*
*voile,high,low,**false***	*cotton,high,medium,**false***
cotton,high,medium,true	*voile,medium,low,true*
voile,medium,big,true	*voile,medium,big,true*
voile,high,big,true	*voile,medium,low,true*
cotton,low,big,true	

Considering a threshold for $\propto$ of 0.25 in (2), the approach identified the failure tendency (*, *high*, *) for the false class of Table 1, which means that *As* has a tendency to fail contracts that imply the provision of high quantities of fabric. Using the same example and the information gain approach described in the next section, the extracted tendency was (*, *high*, *low*), which corresponds to the profile set for the target agent in evaluation. In other cases, the metric used was not able to extract any tendency, even when we analytically observed the presence of behavior patterns that ideally would be captured as tendencies.

2.2. Extraction using information gain

The information gain is a metric used in machine learning for offline classification purposes ([5]), with the phases of training and testing occurring before the actual classification of new instances. It is based on the entropy concept of information theory:

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \epsilon Values(A)} \frac{|s_v|}{|S|} Entropy(s_v) \quad (3)$$

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_v is the subset of *S* for which attribute *A* has value *v* ([5]).

We use this metric to dynamically learn a decision tree from the available (normally scarce) history of evidences of an agent, *every time* it is necessary to verify the adequacy of the agent proposal to the current necessity of the client agent. No training or testing phases are performed. After that, the failure tendency of the agent in evaluation is extracted from the rules for false outcomes. Figure 1 depicts a decision tree that

was learned for a given supplier in a particular experiment we run.

```
good = cotton
|   dtime = low: false
|   dtime = medium: true
|   dtime = big: null
good = chiffon: null
good = voile: false
```

Figure 1: decision tree generated in our simulations

The tree above indicates that at the time of the trust assessment the supplier showed a tendency to fail contracts that match the tendencies *(cotton,*,low)* and *(voile,*,*)*. Therefore, the trustworthiness value $trust_{Ac}\,(As, at)$ 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, it would be given by the $trust_{Ac}\,(As)$ component of equation 1. We generate the trees without post-prunning (i.e. data might be overfitted) since this is an online process that is dynamically updated at every trust assessment of the target agent and the dataset is generally small.

3. Experiments

We evaluated the performance of our trust mechanism in situation-aware scenarios with small datasets and compared it with other trust approaches. All simulations were done using our textile scenario (fabric: cotton, chiffon and voile; quantity: low, medium, high; delivery time: low, medium, big). At every round, each client broadcasts a specific buying lead and the suppliers that have stock on this fabric make a proposal. The selection of the best supplier is based on trust. We used 20 clients, 50 suppliers, 100 rounds, 40 runs, and seller stock of 1 to 10 contracts per round.

The trust approaches in evaluation are *TradCTR* ([6]), a non situation-aware aggregation engine enhanced with dynamics of trust; *ContSpace* ([2]), a situation-aware model representing situation-aware proposals that use pre-defined similarity metrics; *IncrFreq* ([3]), representing the technique of tendency extraction using the frequency increase metric; and *InfGain*, our proposal to extract tendencies of behavior using the information gain criteria.

Table 2 summarizes the different type of suppliers' populations used. Population A is composed of agents with a handicap in one dimension of the contract (*HFab* for fabric, *HQt* for high quantity and *HDt* for low delivery time). Population B has additional suppliers that have composite handicaps (95% of failing a contract in the presence of correlated handicap, 10% one half handicap and 95% of success otherwise (the remaining 5% is considered noise). Population C is constituted by suppliers that are not sensitive to context. We use it to evaluate the performance of situation-aware approaches when dealing with pure probabilistic data.

Table 2: Characterization of populations

Supplier	Prob. Success	PopA	PopB	PopC
Good,Fair	0.95,0.70			√
Bad	0.50			
HFab,HQt	0.05 (hndcap)	√	√	
HDt	0.95 (othrwse)			
HFabQt	0.05 (hndcap)			
HFabDt	0.90 (1/2 hnd)		√	
HQtDt	0.95 (othrwse)			

Utility measures the percentage of successful contracts per approach averaged over all clients and all runs of the experiments (in the best possible case, as seen in Table 2, each client is able to identify the handicap of every provider and gets an average of 95% utility). *Average number of successful contracts* averages the number of succeeded contracts per round (for 20 clients) over the full set of runs of the experiments. Both metrics measure the performance of each approach in selecting the more adequate partners taking into account the context of the actual necessity.

In order to further compare *InfGain* and *IncrFreq*, we use the *number of unfitting choices per round* metric. It averages the number of suppliers that *were selected being identified as not adequate* to the specific necessity, and is used in order to evaluate if the overfitting feature of the *InfGain* approach prevents the choice of adequate suppliers, at any round.

3.1. Results

Figure 2 compares the utility achieved by each approach, for populations A, B and C.

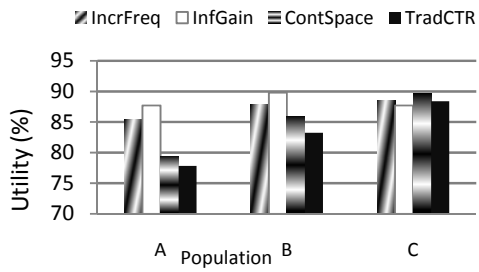

Figure 2. Utility of populations A, B and C

Figure 3 illustrates the performance of each

approach per round (i.e. the number of evidences used to estimate the agents' trustworthiness) for population A. We see that *InfGain* and *IncrFreq* are more able to understand the handicaps of suppliers when the number of evidences is scarce than the other two approaches, as they succeed to do better choices of suppliers. Also, the *InfGain* approach is the fastest learner.

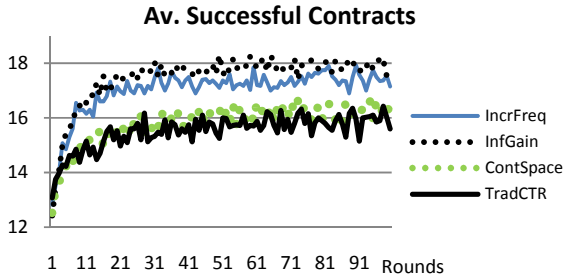

Figure 3: trendlines for each approach (pop. A)

Figure 4 shows the unfitting choices per round of both *IncrFreq* and *InfGain*.

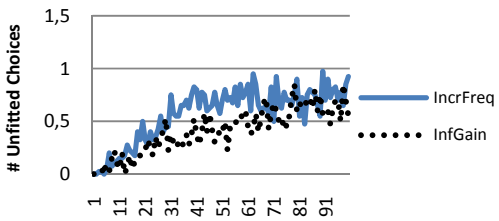

Figure 4: unfitting choices per round

3.2. Concluding remarks and future work

From the results obtained, we observe that all situation-aware proposals outperform the non contextual trust approach. A further analysis of the traces of the experiments shows that *TradCTR* has a tendency to select suppliers with longer contractual history, even when they often fail contracts, privileging a parochial strategy. This is due to the fact that this approach does not avoid partners with handicaps on the present necessity, relying only on the $trust_{Ac}\ (As)$ component of equation 1.

We also conclude that both proposals that dynamically extract tendencies of behavior of the agents in assessment generally outperform the *ContSpace* approach in the first rounds of the experiments, when the trust evidences available on partners are scarce. The *ContSpace* approach has a relative poor performance in population A, significantly increases in population B, when the number of failure patterns doubles, and gets the best result in the rather artificial population C, where the number of possible combinations (reference contexts) increases significantly. This approach based on similarity metrics is more adequate to the exploration of unanticipated situations when the phenomenon of small world occurs.

We also verify that the more accurate extracted tendencies allowed by the *InfGain* approach leads to a better selection of suppliers. In population B, the *InfGain* approach is only 5% worse than the (artificial due to stock limitations) optimal solution. Also, there is no penalty associated to the overfitting issue identified in section 2.2. In fact, by being able to better detect suppliers not fitted to the current necessity at any time, the *InfGain* approach allows a wider exploration of the proponent suppliers, and is able to do a more informed decision at any round. Finally, the results obtained for population C show that there is no evident penalty in using situation-aware processes in the estimation of trustworthiness scores in populations that do not exhibit context-based behaviors.

As future work, we propose to extract the best characteristics of information gain criteria and to capture them in a computationally lighter version of our approach, more adequate to online unsupervised processes, and to work with missing attributes.

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

[1] J. Sabater, M. Paolucci, On Representation and Aggregation of Social Evaluations. Computational Trust and Reputation Models, Int. J. Approx. Reasoning, 2007

[2] M. Rehak, M. Gregor, M. Pechoucek, Multidimensional context representations for situational trust. IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications, pp. 315-320, 2006

[3] J. Urbano, A. P. Rocha, E. Oliveira, Refining the Trustworthiness Assessment of Suppliers through Extraction of Stereotypes. In 12th International Conference on Enterprise Information Systems, Madeira – Portugal, 2010

[4] G. Paliouras, V. Karkaletsis, C. Papatheodorou, C. D. Pyropoulos, Exploiting Learning Techniques for the Acquisition of User Stereotypes and Communities. In: UM99 User Modeling: Proceedings of the 7th Int. Conference, 1999

[5] J. R. Quinlan, "Induction of Decision Trees", *Machine Learning*, 1(1), 1986, 81-106

[6] J. Urbano, A. P. Rocha, E. Oliveira, Computing Confidence Values: Does Trust Dynamics Matter?, in Proceedings of EPIA'09, Portugal, 2009