Trust Estimation Using Contextual Fitness

Joana Urbano, Ana Paula Rocha, Eugénio Oliveira

Faculdade de Engenharia da Universidade do Porto – DEI, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
LIACC – Laboratory of Artificial Intelligence and Computer Science at University of Porto
{joana.urbano, arocha, eco}@fe.up.pt

Abstract. Trust estimation is an essential process in several multi-agent systems domains. Although it is generally accepted that trust is situational, the majority of the Computational Trust and Reputation (CTR) systems existing today are not situation-aware. In this paper, we address the inclusion of the context in the trust management process. We first refer the benefits of considering context and make an overview of recently proposed situational-aware trust models. Then, we propose Contextual Fitness, a CTR component that brings context into the loop of trust management. We empirically show that this component optimizes the estimation of trustworthiness values in context-specific scenarios. Finally, we compare Contextual Fitness with another situation-aware trust approach proposed in the literature.

Keywords: Trust Systems; Situation-aware Trust; Multi-agent Systems

1 Introduction

Computational Trust and Reputation (CTR) systems are systems capable of collecting trust information about candidate partners and of computing trust scores for each one of these partners. In this document, we envision trust as the confidence that the trustier agent has on the capabilities and the willingness of a trustee partner in fulfilling its assigned tasks, in conformance to a given contract established between both parties.

Trust management is in these days a fundamental topic in agent and multi-agent systems, and its appliance concerns decision making processes in almost all electronic forms of commerce and social relationship. In this context, several trust and reputation models have been proposed. In the particular subfield that addresses the representation and the aggregation of social evaluations into trustworthiness scores, several models exist in the literature (e.g. [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]). However, these models do not bring context into the loop of trust estimation, even when the contextual property of trust has been referred in the literature for a long time. Indeed, the exploration of the context of the situation in evaluation to improve the decision making is a new trend of investigation on the trust management area.

1.1 Situation-aware Trust

It is commonly referred in the literature that trust holds the contextual property. In 1994, Marsh considered situational trust as "the amount of trust an agent has in another in a given situation", and gave a helpful example: "I may trust my brother to drive me to the airport, I most certainly would not trust him to fly the plane" [11]. In the same line of thought, Dimitrakos defines trust as a measurable belief that a given entity has on the competence of the trustee entity in behaving in a dependably way, in a given period of time, within a given context and relative to a specific task [12]. Both references relate to the *management of situation-aware trust*, where not all the past evidences are equally relevant to future interactions. However, the consideration of situation-aware trust extends its benefits over several other important dimensions:

- *Management of newcomers*, where the use of similarities between trustees and situations allows to infer trustworthiness during the *first encounter*;
- *Bootstrapping of unanticipated situations*, where missing information from a given target agent can be inferred from similar situations (e.g. “A person trusting Bob as a good car mechanic will not automatically trust him also in undertaking heart surgeries (…) [but] he probably could be quite good in repairing motorcycles” [13]);
- *Management of intuitive fear*, where additional environmental information is needed to support a trustier decision when it intuitively is not sure about a target agent (e.g. “A high tech company may fear to select a partner from a country of origin without high technology tradition, even though this partner has proved high quality work in the desired task in the recent past” [14]);
- *Reduction of the complexity of management of trust relationships* [15];
- *Allowance of the transitivity of trust,* by incorporating the situational dimension into the transitivity loop.[1]

The contribution of this paper is two-folded. First, we provide a state-of-the-art on situational trust. Second, we present a new trust technique based on an online process of stereotype extraction that can be used with any traditional trust aggregation engine. As we will show, this technique allows significant improvement in the trustworthiness estimation of an agent by dynamically detecting tendencies on the agent past behavior. Also, it distinguishes clearly from the recent proposals. As will be shown in the next section, those rely on hierarchical-based similarity measures, being much more complex to manage and also not adequate to manage the intuitive fear property.

This paper is organized as follows: Section 1 introduces the basic ideas about context and Section 2 provides a survey on related work. Section 3 proposes Contextual Fitness, a CTR component designed to manage the bootstrapping of unanticipated situation and to deal with the intuitive fear property. Section 4 presents the experiments done in order to evaluate the Contextual Fitness component and compare it with another situation-aware trust system. Section 5 concludes the paper.

---

[1] Although existing models of reputation are generally based on the transitivity of trust, some authors consider that trust is not transitive [16], unless some sort of situational dimension is incorporated into the model [17].

2 Related Work

Few computational trust models make use of the context in order to estimate trustworthiness values. [18] overview some of the first of such approaches, and in this section we complement this study referring to most recent proposals.

[13], [17], [18] propose the *Context Management Framework*, a model where trust relations in one domain are used to infer trust relations in similar domains. The model uses case-base reasoning techniques to estimate trust in unanticipated situations, by retrieving the most similar cases from a case base. In order to represent the similarity, the model uses a context-specific trust ontology; also, the authors propose to use relational similarity, based on the SimRank algorithm [19]. The major drawback of this approach resides in weak assumption made by this algorithm about the similarity between different objects. As the authors recognize, in more complicated cases, the similarity of two context models is itself context depend [13].

[20], [21], [22] propose the *Context Space and Reference Contexts*, a model that defines the context space as a Q-dimensional metric space with one dimension per each represented situation feature. In this model, trustworthiness values are updated relatively to a set of reference contexts, placed regularly over the context space or adaptively, for every new trust evidence. In the presence of a specific situation, the most similar reference contexts are used and the trust score is computed summing up their trustworthiness values, weighted by the similarity between the new situation and the reference contexts. The major drawback of this model is the consideration of multiple dimensions that can lead to an exponential number of reference contexts that each trustier needs to keep for every trustee. Also, this model can only complement traditional CTR systems that aggregate evidences using weighted means approaches.

[23] propose a model where the expected behaviour of an agent on a given situation is represented as a conditional probability distribution function (PDF) over the possible observations *given the possible agreements*. As in the previous model, a concrete experience about a commitment can be used to update the expectation of behaviour over semantically close commitments, allowing for faster bootstrapping of unanticipated situations. The model also uses an ontology for describing concepts (for instances, business orders). Although the authors state that the PDF is initialized using background knowledge on the other agents, the model suffers from the a priori information limitation of conditional probability distributions.

[24] propose a model of trust that uses as information sources the direct experience with the trustee and also information provided by organizational roles, where the role taxonomy is dynamically updated from trust information maintained by the agents using clustering techniques. When computing trust scores, the trustworthiness value of the trustee on different roles is weighted according to their similarity to the role of current situation.

The majority of the models above is based on ontologies or taxonomies, and all of them imply the definition of domain-based similarity measures of situations or roles. However, there are sometimes subtleties in situations that are not captured by hierarchical-based similarity measures. For instance, these models can reach the conclusion that the situation *delivery of one container of cotton from Asia to Europe* is quite similar to the situation *delivery of one container of chiffon from Asia to Europe*, but they would probably fail to discover that, although the trustee is *generally* considered good in providing the service, he tends to fail it when the contractualized delivery time is low (*management of intuitive fear* property). The model we propose here distinguishes from the above models by performing online evaluation of the trust evidences, as it is able to dynamically extract stereotypes of the (probably changing over time) behavior of the trustee, on a given specific situation. I.e., the similarity measure is not defined *a priori* but extracted from experience.

3 The Proposed Trust System

The computational trust and reputation system (CTR) we propose is illustrated in Fig. 1. It is composed of three core components: the *Aggregator*, an aggregation engine that computes trustworthiness values based on the existing trust evidences; the *Contextual Fitness* component, a module that measures how well the behaviour of a given trustee agent fits to the specificities of the current situation; and the *Similarity Analyser*, an inference engine that compares the characteristics of newcomers' agents with the characteristics of agents for which there exists trust information.

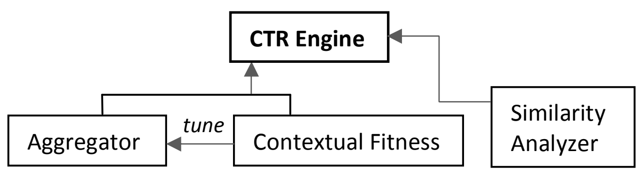

Fig. 1. Architecture of the proposed trust model

The *Aggregator* component of the model can be any aggregating engine that is able to compute trustworthiness values from trust information sources (e.g. evidences from direct experiences, witnesses and organizational roles). The *Similarity Analyzer* is a component that performs conceptual clustering on organizational characteristics in order to generate profiles of business entities. When a trust score for a newcomer agent is to be computed, the CTR engine asks the Similarity Analyzer to compare the characteristics of this agent to the profiles of business entities for which there is some trust information and an estimation of the newcomer trustworthiness is inferred. The way the organizational characteristics of agents are captured is not addressed in this paper, although several mechanisms may be considered (e.g. disseminated information on the Web, transmission of knowledge between communicating agents, public or private directory services). Finally, the *Contextual Fitness* component is the subject of this paper, and it would be thoroughly addressed in the next sections.

The proposed trust model is being developed in the context of the selection of partners in business-to-business (B2B) processes, where trading partners of the textile industry attempt to place orders to the best available partners. However, the model can be also used in other scenarios, such as virtual organizations and social networks. The trust evidences used in our model are represented by the tuple $<A_c, A_p, At_1..At_n, o>$, where:

- $A_c \subset C$ is an agent from the set C of clients' agents (i.e., the *trustier* agent);

- $A_p \subset P$ is an agent from the set P of providers' agents (i.e., the *trustee* agent);
- $At_i \subset AT$ is an attribute from the set AT of n contract attributes, described as attribute-value pairs (e.g. *fabric=cotton*, *quantity=360000*, *deliveryTime*=7);
- $o \subset \{T, F\}$ is the outcome of the contract, either representing successful (*True*) or violated (*False*) contracts by the provider partner.[2]

3.1 The Contextual Fitness Component

The *Contextual Fitness* (CF) is the CTR component responsible for tuning the trustworthiness values computed by the Aggregation component. It is intended to manage *intuitive fear* situations, i.e. to capture a certain degree of intuition that trustier agents (e.g. executive managers) have by the time they make a decision (e.g. selecting partners in the context of a given business opportunity). As an example, let us consider that a given selector agent knows that a candidate partner is trustworthy or well reputed in fulfilling agreements when selling blue cotton zippers to European countries, but that it ignores the ability of the candidate agent in providing high quantities of the material in a short period of time. We name this situation the *contextual ignorance*. In this scenario, the trustier agent already knows the trustworthiness score of the trustee in the specific situation and role in evaluation; however, it fears (let us assume the trustier is very sensitive to risk) that the trustee is not able to succeed in a rather subtle, but relevant different type of contract.

3.2 The Contextual Fitness Algorithm

In the scenario used in this document, the trustier describes its business needs in the form of a call for proposals (CFP). When an agent issues a CFP and receives several proposals from candidate partners, the CF component checks the adequacy of each proposal to the CFP requirements using the following steps:

1. For each candidate partner that proposed, the agent performs conceptual clustering over its contractual past evidences;
2. For each created cluster, a stereotype is extracted;
3. The stereotypes are compared to the current CFP using a similarity analysis approach, and a contextual fitness value $cf \subset [0,1]$ is derived;
4. The values computed by the Aggregator and the CF components are combined and a global trustworthiness value is derived for the agent in evaluation.

3.3 Comments about Current Implementation

Current implementation of the CF component is a simplification of the algorithm presented above. Therefore, at step 1 classification is done over the *outcome* $o \subset \{T, F\}$ attribute of the evidence. This means that the contractual evidences of each trustee agent are classified into two different classes: one with all the evidences related to successful contracts, and the other containing evidences related to violated contracts.

[2] A more diversified representation of outcome information is left for future work.

Then, at step 2, a stereotype is extracted for each generated class, using the metric illustrated in Equation 1 that measures the increase in the frequency of a category *c* within a community [25].

$$\propto = \left(\frac{\#InstAttCluster}{\#InstCluster}\right)^2 - \left(\frac{\#InstAttTotal}{\#InstTotal}\right)^2 . \quad \textbf{(1)}$$

In the equation above, *#InstAttCluster* is the number of times (trustee evidences) 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. The purpose of these two first steps is, therefore, to detect *tendencies* of success and failure for each particular trustee agent.

At step 3, the stereotypes extracted for each trustee are compared to the current CFP and a contextual fitness value is derived. In the current implementation, the similarity analysis is quite simple, and a non-correlated comparison for each one of the attributes is performed. As a result, a binary value is derived: a zero (0) value for full match with a *false* stereotype, or a one (1) value in all remaining cases. Fig. 2 shows an example of a match between a trustee stereotype and the current CFP.

Stereotype: Agent X, null, null, low, false
CFP: chiffon, 1080000, 7

Fig. 2. Example of a CFP and a stereotype for trustee agent X

In the figure above, a delivery time of 7 days was previously quantified to the 'low' category. The extracted stereotype means that the trustee agent has a tendency to fail any kind of contract that involves low delivery times, independently of the fabric and the quantity provided by the agent in the past. Therefore, the contextual fitness value for the trustee proposal is set to zero, which means that the overall trust score for the trustee is also zero. Finally, it is worthy to note that all quantitative values of CFP parameters and target contractual attributes are fuzzyfied prior to their usage in the classification process. For example, numeric quantity values are translated into labels such as 'low', 'medium' and 'high'.

4 Experiments

In the experiments performed, we evaluated the Contextual Fitness (CF) component, i.e. the ability of the component in tuning the trustworthiness estimation of trustee agents using business contextual information. For this, we used a social simulation where trading client agents attempt to place orders to the best available partners. Every client in the simulation has a specific business need translated as a call for proposals (CFP), specifying values for a fabric, a quantity and a delivery time (cf. Fig. 2). At every round of the simulation, each client issues its specific CFP and waits for proposals from providers' agents. The selection of the best proposal at every round is done using a given *selection approach*.

In the experiments, we generated three different types of provider agents, as resumed in Table 1. The type was assigned to each agent at initialization, following a uniform distribution over the three possible values.

Table 1. Characterization of target agents

Supplier Type	Description
S_{HQT}	Probabilistically succeeds 95% of the established contracts, except the ones that involve the delivery of *high quantities*, which probabilistic fails 95% of the time
S_{HDT}	Probabilistically succeeds 95% of the established contracts, except the ones where the *delivery time is low*, which probabilistic fails 95% of the time
S_{HFB}	Probabilistically succeeds 95% of the established contracts, except the ones that refer to a *specific fabric*, which probabilistic fails 95% of the time

This way, each provider agent can transact with different clients in several distinct business situations, and its contractual history includes contextual situations where it was successful and others where it actually suffers from some kind of handicap (e.g. the provider may have a tendency to fail contracts with short delivery times). Table 2 resumes the remaining configuration parameters of the experiments.

Table 2. Configuration of the experiments

Parameter	Value
Fabrics	{Chiffon, Cotton, Voile}
Quantities	{Low, Medium, High}
Delivery Time	{Low, Medium, Big}
# buyers	20
# of sellers	50
Types of sellers	Uniform dist. {"S_{HQT}", "S_{HDT}", "S_{HFB}"}
# rounds	100
# runs per experiment	20
TradCTR parameters	As described in [18] (no bootstrapping)

Selection Approaches. In order to evaluate the benefits of the contextual fitness module, we tested three different approaches: the *TradCTR* approach is an aggregating engine proposed in [14] that represents a traditional, non situation-aware aggregation engine; the *CF* approach is the *TradCTR* engine complemented by the Contextual Fitness module; and the *ContSpace* is the Context Spaces and Reference Contexts model described in section 2, with reference contexts being placed regularly over the combinations of possible values of CFP attributes.

Performance Metrics. In every experiment, we measured the number of successful contracts per type of target agents and per approach, and averaged this number over the total number of rounds. In the best case, each client is able to identify the handicap of every provider and to select the best proposal, leading to an average of 95% of successful contracts (cf. Table 1).

4.1 Results

Fig. 3 presents the results obtained in the experiments.

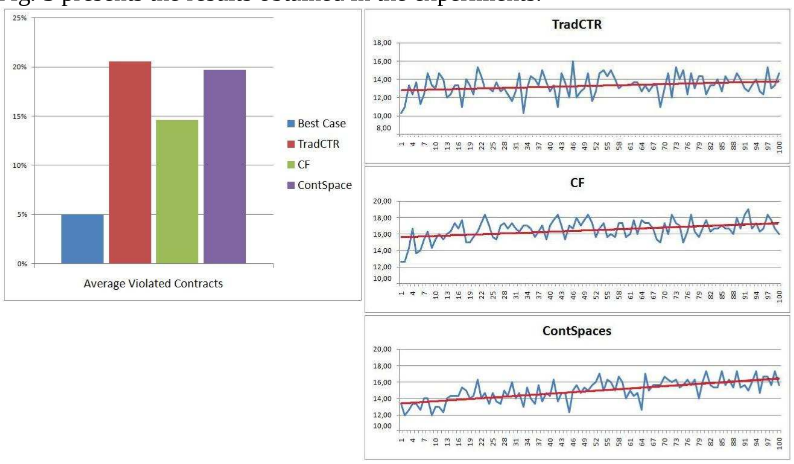

Fig. 3. Average number of violated contracts (*left*) and linear trend lines (*right*) per approach

The results illustrated in the above figure (left) show that the *TradCTR* approach is in average 15.56% worse than the best case, and that this number is reduced to 9.62% when the contextual dimension is added (*CF* approach). The *ContSpace* approach is in average 14.72% worse than the best case, performing a little better than *TradCTR* but significantly worse than *CF*, for the evaluated scenario.

The figure also illustrates the trendlines of each approach over the rounds (plots on the right). We can observe from them that the *CF* approach starts to learn earlier in time (i.e. from a few number of the trustee evidences), while the *ContSpace* approach has a more accentuated learning curve without, however, ever reaching *CF* performance.

4.2 Interpretation of Results

The results obtained in this set of experiments showed that, even oversimplified, the current implementation of the Contextual Fitness significantly improves the ability of the trust system in estimating trustworthiness scores of handicap-like trustee agents. By tracing each one of the experiments, we could also observe that, whenever the *CF* approach was able to extract stereotypes, it efficiently picked up candidate partners in a selective way, taking into account the characteristics of the current CFP. The traces also showed that stereotypes are generally correctly extracted with a very little number of past contractual evidences, for the population used in these experiments. Observing the trendlines at Fig. 3, we realize that the traditional CTR approach performs poorly at round 100 than the situation-aware version of it at the first rounds.

The comparison between *CF* and *ContSpace* has shown that the dynamic extraction of stereotypes allows for a faster bootstrapping than the consideration of reference contexts, for the evaluated scenario. This is an interesting result, as the *CF* approach promises to scale better than *ContSpace* in bigger context spaces.

5 Conclusions and Future Work

This paper focused on situation-aware trust management and overviewed recent approaches that complement traditional CTR systems with the consideration of context. These approaches are scarce, and are generally based on hierarchical-like structures for measuring similarity between situations or roles. They are mainly intended for the managing of unanticipated situations. However, the granularity of the approaches seems to respond poorly to another important contextual situation, the management of intuitive fear. In this context, we present Contextual Fitness (CF), a situation-aware component that complements traditional CTR systems. CF component is able to dynamically detect behavioural tendencies of trustee agents by following an online process of stereotypes extraction.

Experimental evaluation of the CF component showed that the use of machine learning techniques in the estimation process (namely, the classification and consequent extraction of stereotypes from the historical contractual evidences of candidate agents) allows the selective choice of these agents taking into account the current business situation, significantly increasing the trustier utility.

We are currently working on the full implementation of the Contextual Fitness component, namely, on its steps 3 and 4, as described in section 3.2. We believe that it would allow further improving of the trustworthiness estimation process. We leave as future work the specification and the implementation of the Similarity Analyzer, a machine learning-based module intended to infer contractual characteristics for newcomers (i.e. agents for which there is no contractual or other trust/reputation information), based on their organizational properties.

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

6 References

1. Ramchurn, S., Sierra, C., Godo, L., Jennings, N.R.: Devising a trust model for multi-agent interactions using confidence and reputation. In: Int. J. Applied Artificial Intelligence (18), pp. 833–852 (2004)
2. Sabater, J.: Trust and Reputation for Agent Societies. Number 20 in Monografies de l'institut d'investigacio en intelligència artificial. IIIA-CSIC (2003)
3. Carbo, J., Molina, J., Davila, J.: A BDI Agent Architecture for Reasoning about Reputation. In: IEEE International Conference on Systems, Man, and Cybernetics (2001)
4. Jøsang, A., Ismail, R.: The Beta Reputation System. In: Proceedings of the 15th Bled Electronic Commerce Conference, Sloven (2002)
5. Reece, S., Rogers, A., Roberts, S., Jennings, N.R.: A Multi-Dimensional Trust Model for Heterogeneous Contract Observations. In: 22nd AAAI Conf. on Artificial Intelligence (2007)

6. Zacharia, G., Maes, P.: Trust management through reputation mechanisms. In: Applied Artificial Intelligence, 14(9), pp. 881–908 (2000)
7. Sabater, J., Paolucci, M.: On Representation and Aggregation of Social Evaluations. In: Computational Trust and Reputation Models, Int. J. Approx. Reasoning (2007)
8. Hang, C., Wang, Y., 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 – Vol. 3, Estoril, Portugal, May 12 – 16 (2008)
9. Huynh, T.D., Jennings, N.R., Shadbolt, N.R.: An integrated trust and reputation model for open multi-agent systems. In: Autonomous Agents and Multi-Agent Systems, Vol. 13, N. 2, pp. 119–15 (2006)
10. Jonker, C.M., Treur, J.: Formal Analysis of Models for the Dynamics of Trust Based on Experiences. In: Procs. of the 9th European Workshop on Modeling Autonomous Agents. LNCS, vol. 1647, pp. 221-231 (1999)
11. Marsh, S.P.: Formalising Trust as a Computational Concept. PhD Thesis, University of Stirling (1994)
12. Dimitrakos T.: System Models, e-Risk and e-Trust. Towards Bridging the Gap? In: Towards the E-Society: E-Business, E-Commerce, and E-Government. Kluwer (2001)
13. Tavakolifard, M.: Situation-aware trust management. In: RecSys'09: Proceedings of the third ACM conference on Recommender systems, pp. 413-416 (2009)
14. Urbano J., Rocha A.P., Oliveira, E.: A Trust Aggregation Engine that Uses Contextual Information. In: 7th European Workshop on Multi-Agent Systems, Cyprus (2009)
15. Neisse, R., Wegdam, M., Sinderen, M., Lenzini, G.: Trust management model and architecture for context-aware service platforms. In: On the Move to Meaningful Internet Systems. LNCS, pp. 1803-1820 (2009)
16. Christianson, B., Harbison, W.S.: Why isn’t trust transitive? In: Proceedings of the International Workshop on Security Protocols, pp. 171-176, London, UK (1997)
17. Tavakolifard, M., Herrmann, P., Ozturk, P.: Analogical trust reasoning. In: Trust Management III, chapter 10, pp. 149-163 (2009)
18. Tavakolifard, M., Knapskog, S.J., Herrmann, P.: Trust transferability among similar contexts. In: Q2SWinet’08: Proceedings of the 4th ACM symposium on QoS and security for wireless and mobile networks, pp. 91-97, New York, USA (2008)
19. Jeh, G., Widom, J.: SimRank: a measure of structural-context similarity. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 538-543. ACM Press New York, NY, USA (2002)
20. Rehak, M., Gregor, M., Pechoucek, M.: Multidimensional context representations for situational trust. IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications, pp. 315-320 (2006)
21. Rehak, M., Pechoucek, M.: Trust modeling with context representation and generalized identities. In: Matthias Klusch, Koen V. Hindriks, Mike P. Papazoglou, and Leon Sterling (eds.), CIA. LNCS, vol. 4676, pp. 298-312 (2007)
22. Rehak, M., Pechoucek, M., Grill, M., Bartos, K.: Trust-based classifier combination for network anomaly detection. In: Matthias Klusch, Michal Pechoucek, and Axel Polleres, editors, Cooperative Information Agents XII. LNCS, vol. 5180, pp. 116-130 (2008)
23. Fabregues, A., Madrenas-Ciurana, J.: SRM: a tool for supplier performance. AAMAS’2009, pp. 1375-1376 (2009)
24. Hermoso, R., Billhardt, H., Ossowski, S.: Dynamic evolution of role taxonomies through multidimensional clustering in multiagent organizations. In: Principles of Practice in Multi-Agent Systems, vol. 5925, chapter 45, pp. 587–594. Springer Berlin Heidelberg (2009)
25. Paliouras, G., Karkaletsis V., Papatheodorou, C., Pyropoulos, C.D.: Exploiting Learning Techniques for the Acquisition of User Stereotypes and Communities. In: UM99 User Modeling: Proceedings of the Seventh International Conference (1999)