A Direct Reputation Model for VO Formation

Arturo Avila-Rosas[1] and Michael Luck[2]

[1] Instituto Mexicano del Petróleo,
Eje Central 152, México DF, CP 07730, México
aavilar@imp.mx

[2] School of Electronics and Computer Science,
University of Southampton,
Southampton, SO17 1BJ, United Kingdom
mml@ecs.soton.ac.uk

Abstract. We show that reputation is a basic ingredient in the Virtual Organisation (VO) formation process. Agents can use their experiences gained in direct past interactions to model other's reputation and deciding on either join a VO or determining who is the most suitable set of partners. Reputation values are computed using a reinforcement learning algorithm, so agents can learn and adapt their reputation models of their partners according to their recent behaviour. Our approach is especially powerful if the agent participates in a VO in which the members can change their behaviour to exploit their partners. The reputation model presented in this paper deals with the questions of deception and fraud that have been ignored in current models of VO formation.

1 Introduction

Recently, a large number of new collaborative, networked organisations have emerged, having as motivation the explosive progress in computer networks and communication systems, but also as a reaction to market pressures that demand customised, high quality products and services at lower costs and, at the same time, shorter production and marketing times. Promising greater flexibility, resource optimisation and responsiveness in *competitive open environments*, VOs are an example of this trend that has pervaded not only business domains but other areas such as e-science. The concept of a VO has been used to describe the aggregation of autonomous and independent organisations connected through a network and brought together to deliver a product or service in response to a customer need [6]. In this paper we take a VO to be *a temporary alliance composed of a number of autonomous entities (representing different individuals, departments and organisations) each of which has bounded problem solving capabilities and limited resources at their disposal, that come together to share skills or core competences and resources in order to better respond to customer needs or business opportunities, and whose cooperation is supported by computer networks* (adapted from [5]).

What distinguishes VOs from other forms of organisation is the full mutual dependence of their members to achieve their goal and therefore the need for

M. Pěchouček, P. Petta, and L.Z. Varga (Eds.): CEEMAS 2005, LNAI 3690, pp. 460–469, 2005.

cooperation. However, open environments in which VOs are embedded involve organisations and individuals that do not necessarily share the same objectives and interests that they might not know in advance, and where they might not trust each other, but should work together and help each other to achieve a common goal. One of the key omissions in the computational representation of VOs relates to the need to take into account more subjective facets like the *reputation* of the individuals, which helps to cope with heterogeneity, autonomy and diversity of interests among members. We observe that current solutions underestimate the possibility of swindle in VOs. A common flaw is assuming that the partners selected are fully competent and honest. Since partners represent organisations or individuals who want to maximise their utilities by joining a VO, they have a strong incentive to misrepresent the value of their contributions and enjoy more benefits of cooperative associations [1]. Further, partners are selected in relation to the abilities they claim to have, but it is possible that they do not have such abilities. However, due to lack of information about past interactions, it is difficult to detect and control these situations. This paper considers the introduction of reputation into VOs, by providing a reputation model based on the adaptive evaluation of direct experiences to identify trustworthy individuals to join VO.

The remainder of this paper is organised as follows. The requirement for reputation systems for VOs are discussed in Section 2. In Section 3 we present our reputation model for VOs which is based on reinforcement learning techniques. In Section 4, we describe the experiments undertaken to support the validity of our model, the results of which (and their comparison with two other models) are presented in Section 5. Section 6 reviews related work, and Section 7 present our conclusions.

2 Requirements

The objective of this section is to delineate the requirements for building a reputation system in order to serve as a decision-making variable in the selection of partners, promote cooperation, produce trust and induce *good* behaviour in the members of a VO.

1. *Distributed reputation management.* As they are distributed and dynamic, VOs do not depend on the presence of any centrally trusted authority. Moreover, individuals must maintain personalised models of the trustworthiness of others at a capability level so that they will be able to know which capability caused cooperation to fail and why [3]. By contrast, centralised management of reputation offers a biased perspective of reputation because it aggregates feedback, making no distinction between the preferences of partners submitting feedback that may not coincide with the interests of the VO.
2. *Dynamism.* VOs integrate multiple autonomous, diversely skilled partners under intense time pressures to create complex products or services [4]. Due to limitations in time and intense task pressures, VOs require that their

members develop mutual trust fast. Partners should be able to quickly use a reduced number of interactions to estimate the reputation of a partner and; at the same time, take partner selection decisions without having a significant impact, in terms of time consumption, on the formation of a VO.

3. *Adaptability.* VOs operate under high levels of demand uncertainty generated by unknown and rapid shifts in consumer preferences [4]. Demand uncertainty creates changes in the structure of the VO, which is forced to adapt itself by reallocating tasks or redefining them. In these circumstances, organisations feeds into periodic evaluations of the VO which, in turn, leads partners to make adjustments to their relationships and identify when changes in the efficiency of partners is due to the adaptation process or due to abusive behaviour [4]. This suggests that the updating process of reputation values should be a *learning* process about another's true abilities, that captures the observed performance through the reputation of the partner.
4. *Predictability.* The behaviour of each partner in a VO usually offers clues to the others about its capabilities and *intentions*, so it is possible to make predictions about its future behaviour. The main objective of *predictions* is to detect any misconduct of the partner early enough, so that the VO can take necessary steps to protect itself from adverse effects of partner misbehaviour. Reputation must provide information to predict the future performance of a partner and eventually the risk involved of interacting with it.

3 Direct Reputation Model

In this section we introduce our model of reputation, which meets some of the requirements discussed in the previous section. We start by defining mathematically the concepts of reputation and impressions. Next we describe the methods used in our model for updating reputation.

3.1 Reputation

We define the reputation of an agent as *a perception regarding its intention and competences, which is held by other agents through the formation and dissemination of subjective evaluations based on experiences and observations of past actions.* Here, these evaluations are called *impressions*. From the definition, the observed behaviour of others is collected through: (i) direct experiences, with interaction histories serving as a strong evidence for estimating someone reputation or (ii) via the testimony of others, known as recommenders. On the basis of the source of reputation, two concepts of reputation may be derived, namely *direct reputation* and *social reputation.* Although important, the concept of social reputation lies beyond the scope of our research and is not defined; we only make reference to it as another source of reputation different from direct reputation.

3.2 Direct Reputation

Direct reputation (DR) is defined as the weighted average evaluation that an agent makes of another's competence, and gives the extent to which the target is *good* or *bad* with respect to a given behaviour or action. Direct reputation is context-dependent so that an agent is reputed according to the service provided. For example, an agent may be well reputed as a printing service provider but poorly reputed as a file storing service provider. VOs provide an environment in which agents may offer the same service with different qualities for different reasons such as demand uncertainty or as a result of dishonest behaviour. In this sense, we adopt the ideas of Shapiro [8] expressed in his analysis of the economic effects of reputation in such environments. Shapiro proved that the most efficient way to estimate a seller's reputation (i.e., the way that induces the seller to produce at the highest quality level) is a time-discounted average of the recent ratings evaluating its reputation. Hence, direct reputation is computed as the average of *impressions* received within the most recent time window,

$$W = [t - \epsilon, t], \tag{1}$$

where ϵ defines a time interval that limits the set of interactions and in which impressions are used to compute a direct reputation value. Impressions are weighted from 0 to 1 to indicate the notion of importance of an impression in relation to others for calculating reputation. Taking only the most recent impressions is equivalent to using an average calculation where weights are non-zero for impressions received within the time window and 0 otherwise. The direct reputation values vary in the range of [0,1] and are used only to represent comparative values in this continuous space from bad reputation (values near 0) to good reputation (values near 1). The direct reputation of i in the perspective of j in context k is represented as:

$$DR_{ij}^{k} \in [0, 1].$$

3.3 Impression

We define an impression as an evaluative opinion that is formed by any entity (individual, organisation, etc.) based on a discrete experience with another partner, coupled with the partner's performance. Computationally, an impression is the value assigned to a service that indicates the proximity of the service provided by an agent i to the expectations of agent j requesting the service. An impression is related with a dimension that describes just one of the qualities of the service as required by agent j. For example, a partner can get different impressions for its efficiency or the quality of its services. The group of dimensions needed for evaluating the whole performance of a service provider is denoted by the set of enabling qualities Q and it is context-dependent. For example, to evaluate an agent in the context of a printing service, two dimensions may be taken into account: printing quality and rapidity. Mathematically, the impression appear as follows,

$$imp_{ij}^{d} \in [0, 1],$$

$$Q_{ij} = \{d \in k | k \textit{ is a context}\}, \tag{2}$$

where i is the service provider whose interaction with the service consumer j left in it the strong impression *imp* in relation to dimension d, and Q_{ij} is the set of dimensions for evaluating a service provider in context k. The numbers used for impressions are merely reference values for making comparisons, each consumer establishes a personal threshold of *acceptable* values for the dimension d evaluated. This personal threshold may be based on:

- the agreed values of a contract, when interactions are fixed by contractual terms; or
- the values that constitute a standard for delivering a service, when standards are available to indicate the permissible values of a particular dimension; or
- the values obtained empirically, when the consumer has previous experience of consuming a particular service and can estimate optimal values for the dimensions involved.

Once a personal threshold of acceptable values is established, it is compared with the actual values of each dimension after providing a service.

3.4 Updating Direct Reputation

Each agent updates its reputation value of a service provider every time it receives impressions from either direct (immediate or observed interactions) or indirect experiences. Our first proposal to update the reputation values (after receiving t rated experiences or impressions) consists in the use of the following reinforcement learning based action update rules:

$$DR_t = DR_{t-1} + \alpha \cdot [imp_t - DR_{t-1}]. \tag{3}$$

Reputation, in Eq.(3), can be interpreted as the aggregation of the previous value of reputation plus a factor that strengthens or weakens that value. This factor indicates the proximity of the recent impression to the past reputation, and shows of how well the previous reputation predicts the latest given impression. Note that although we omit the indices k, i and j to make the expression more readable, DR_t makes reference to the reputation of an agent i in the opinion of agent j for the context k. The update rule in Eq.(3) is a linear function which is required in an open environment where the number of prior interactions may be reduced, and reputation cannot be updated in the long term through a non-linear function because an agent could cheat on many occasions before the reputation is updated. Instead, reputation must be updated immediately after any interaction. If α is near 1 then all the previous history will be forgotten, otherwise, if α is near 0 then the previous history will be preserved.

The factor α is also known as a learning rate, and is an indicator of how long past experiences will last in the memory of the system. For example, while low values of α mean that early experiences will have more influence in the system than recent ratings, high values of α indicate that early experiences will soon be forgotten. For our purposes, we consider α as a function $\alpha(DR_{t-1}, imp_t)$ with the following properties that are based on the ideas of Carbo et al. [2]:

- The function $\alpha(DR_{t-1}, imp_t)$ determines how fast the reputation value changes after an experience and how this affects the memory of the system.

This depends on the accuracy of the predictions suggested by the *impressions* received; that is, how much similarity exists between the expectation formed by the previous reputation values and the last rating. We consider the initial value for the function $\alpha(DR_{t0}, imp_{t0}))$ to be 0.5. That is, as the agent starts to learn, it will be careful with the first impressions until it learns how to better estimate its predictions.

– Similarity will be estimated through a similarity function $\beta(DR_{t-1}, imp_t) \in (0, 1)$:

$$\beta(DR_{t-1}, imp_t) = 1 - e^{-10 \cdot ABS(E - imp)}, \tag{4}$$

where E is the estimated rating based on the past reputation and rating:

$$E = \frac{DR_{t-1} + imp_{t-1}}{2}. \tag{5}$$

– Finally, the function $\alpha(DR_{t-1}, imp)$ is updated as follows:

$$\alpha(DR_t, imp) = \frac{\alpha(DR_{t-1}, imp) + \beta(DR_{t-1}, imp)}{2}. \tag{6}$$

4 Experiments

We performed two sets of experiments to evaluate DIRECT (our algorithm for computing reputation based on direct interactions) and show its feasibility and effectiveness. For comparison purpose, we use two existing models of reputation, SPORAS [9] and REGRET [7]. These models were chosen because reputation systems for VOs should consider the time when the interactions take place in order to update their reputation values, and both SPORAS and REGRET meet this requirement. In order to compare similar values of reputation, SPORAS and REGRET were modified to produce reputation values normalised in the range [0,1]. Additionally, in the case of REGRET, just the *individual dimension* of reputation is considered because it is the only one associated with direct interactions.

4.1 Accuracy

The objective of this experiment is to evaluate the accuracy of the reputation model. The value of reputation must provide a measure of the true capabilities of a service provider (SP) for providing a service. We generate 10 series of data representing the quality perceived of a service (QP) during 60 interactions. This data varies randomly in the interval *[-z, +z]* from a mean value q of the actual quality (QoS). We want to model the fact that although SP delivers its services with the same quality, that is $q = 0.5$, the consumer (CA) may perceive such a quality in distinct ways. In our experiment, CA's perceptions vary around the actual quality q with a standard deviation $\sigma = 0.03$, and according to a normal distribution.

4.2 Abuse

The experiment here described is similar to that described in [9] where a SP who joins a VO behaves reliably until it reaches a high reputation value and then starts committing fraud. Thus, in this experiment we aim to show quantitatively which model of reputation offers a mechanism for dealing with deceit. Deceit in a VO is found when a partner deteriorates the quality of its services once it has reached a certain level of reputation, in order to exploit others. We measure the rapidity with which agents *learn* the *new* behaviour of their partners in terms of the minimum number of interactions to adjust the reputation of a partner towards true quality of a service. We generated two sets of 10 data series representing the quality of a service during 120 interactions, both data varying using a normal distribution in the interval *[-z, +z]* from a mean value *q1* and *q2* of quality. In the first set of data the agent provides the highest quality during the first 25% of the interactions and, after that it decreases the quality in 6.25%, 12.5%, 18.75%, 25%, 31.25%, 37.5%, 43.75% and 50% for the rest of the interactions. We want to model the fact that after delivering its services with the same quality, $q1 = 0.8$, during the first 30 interactions, this may be perceived by CA in distinct way. SP then reduces the quality of its services to milk the reputation already built. As in the previous experiment, CA's perceptions vary around the actual qualities *q1* and *q2* with a standard deviation $\sigma = 0.03$, and according to a normal distribution.

5 Results

5.1 Results of Reputation Evaluation Accuracy

In Figure 1 the reputation values computed with the three algorithms are shown. As can be seen, our proposal obtains similar results as REGRET in a similar number of interactions. DIRECT and REGRET establish the reputation value faster than SPORAS. Although REGRET and our proposal DIRECT use different aggregation algorithms for computing reputation, both obtain accurate results when a SP maintains the provision of its service without change. In our simulations, we compute the number of interactions before the reputation curves generated by each of the algorithms converge towards the actual QoS. The convergence is considered when the calculated values of reputation are in the interval $[q - \sigma, q + \sigma]$.

5.2 Results of Abuse of Prior Performance

In Figure 2 we can see that REGRET requires in general more interactions to adapt its reputation values to the change of behaviour of service providers. In contrast, DIRECT and SPORAS show a more adaptive behaviour and require fewer interactions. On the other hand, REGRET updates reputation values very slowly, and opportunistic providers might take advantage of this by getting high values of reputation to be considered as well reputed and then start to cheat.

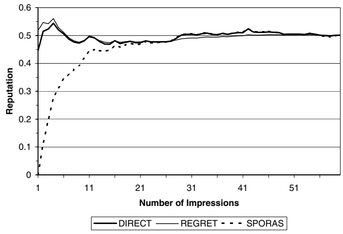

Fig. 1. Building Reputation

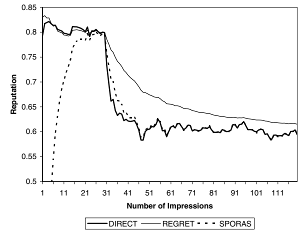

Fig. 2. Reputation is decreased in 25.0%

In Figure 3 the minimum number of interactions is shown to adjust the decrement in the reputation value of 6.25%, 12.5%, 18.75%, 25%, 31.25%, 37.5%, 43.75% and 50%. As can be seen SPORAS and DIRECT require less than 20 interactions to adjust their reputation values to the change in quality of the service, regardless of the percentage in which the quality is reduced. In REGRET, due to the accumulation of experiences, the effect of past experiences on the computation of reputation provokes accelerated increment in number of interactions required to adjust its reputation value.

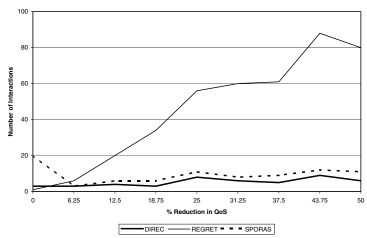

Fig. 3. Minimum number of interaction to adjust reputation values

The key difference between SPORAS and DIRECT is the ability to distinguish changes in behaviour based on the accuracy of the predictions. That is, DIRECT adjusts its values of reputation faster than SPORAS when the expectations created by the reputation values are closer to latest impressions. This can be seen in the slope of both curves. While the rapidity to detect the change in the quality of a service is alike for both curves, DIRECT presents values of reputation closer to the true quality of service.

6 Related Work

Zacharia and Maes in [9] present SPORAS, which is a *centralised* reputation system that establishes reputation for users in an on-line community, based on the aggregation of *rates* given by users after each transaction. Reputation in SPORAS aims to predict future performance of the users. In order to make accurate predictions using a small computational space, a recursive and adaptive algorithm for updating reputation is used. Reputation is calculated continuously using the previous value of reputation; and the previous value of reputation is reinforced or weakened depending on the rates obtained. This aggregation method then allows newer rates to count more than older ones. Because SPORAS is a centralised reputation system, it is not viable for VOs where partners need personalised reputation values calculated from assembled rates of those they trust already rather than those they do not know. Although the assumption made in SPORAS to make reputation values dependent on the reputation of the entity who is providing a feedback is correct, it mixes two different dimensions of reputation. While a user can be reputed as completely unable to cheat on deals, nonetheless that same user may be a bad evaluator of other users. That is, being an excellent service provider does not mean being an honest evaluator.

REGRET is a reputation system developed by Sabater and Sierra [7] that adopts a sociological approach for computing reputation in societies of agents trading well defined products inside an e-commerce environment. Although REGRET provides a very simple method for aggregating rates (or *impressions* that are the result of evaluating direct interactions) based on the weighted sum of the impressions (more relevance is given to the recent ones), its major contribution is the vision of reputation through of three dimensions. These dimensions are called the *individual dimension*, *social dimension* and *ontological dimension*. REGRET emphasises both individual and social components of social evaluations. That is, whereas the individual dimension is the effect of past experience with a given agent, the social dimension refers to reputation inherited by individuals from the *groups* they belong to. However, as discussed earlier, VOs require to a certain extent that the reputation of a partner is assessed in a *reactive* form to detect possible opportunistic behaviour. However, REGRET's main idea consists of emphasising the freshness of information. Computations in REGRET give a *fixed* high relevance to recent rates over older ones according to a time dependent function, and, moreover the rates are aggregated in a way that can be sensitive to noise since they are simply summed. Furthermore, VOs require that reputation be assessed swiftly in order to detect misbehaviour. REGRET, on its part, requires a minimum number of interactions to make correct evaluations of reputation but it is likely that partners will not interact the minimum number of times to provide a reliable reputation value. Finally, REGRET does not handle the problem of lying (strategically) among agents. Rates are obtained in a cooperative manner rather than in a competitive environment.

7 Conclusions and Future Work

We have provided a critical overview of the state of the art in the field of VOs and reputation. We argue that subjective aspects of partners such as their *competences* and *trustworthiness* should be taken into account in partner selection decisions, since these aspects ultimately influence cooperation between partners.

Moreover, we assert that reputation plays an important role in VOs when members decide who to interact with and when to interact, by providing information about the past behaviour of potential partners, their abilities and reliability. Additionally, we discussed the requirements for building reputation systems that pursue three basic objectives in the formation and operation of VOs: (1) they provide useful information about potential partners for selecting the most *appropriate*, and eventually enable the formation of VOs; (2) they foster trust among the partners of the VO by revealing each partner's capabilities and predicting its future behaviour; and (3) they offer a means for enhancing cooperation by detecting and deterring deceptive behaviour through imposing *collective sanctions* on defectors. Finally, we have provided experimental evidence to demonstrate the validity of the model developed and the fulfilment of the requirements mentioned above, including a comparative analysis of the model proposed in this thesis and two other models. In particular, two aspects were analysed regarding the accuracy of the values calculated and the ability to detect abuses.

Although this paper has answered how reputation is relevant to recognise cooperative partners through direct interactions, it opens up more research opportunities and questions that are unanswered. Moreover, there are other issues that were not faced in this paper, due to the bounds imposed on the research, and still need to be addressed.

References

1. S. Braynov and T. Sandholm. Trust revelation in multiagent interaction. In *Proceedings of CHI'02 Workshop on Philosophy and design of Socially Adept Technologies*, pages 57–60, Minneapolis, USA, 2002.
2. J. Carbo, J. Molina, and J. Davila. Trust management through fuzzy reputation. *International Journal of Cooperative Information Systems*, 12(1):135–155, 2003.
3. N. Griffiths and M. Luck. Coalition formation through motivation and trust. In *Proceedings of the Second International Joint Conference on AAMAS*, pages 17–24, Melbourne, Australia, 2003.
4. C. Jones, W.S. Hesterly, and S.P. Borgatti. A General Theory of Network Governance: Exchange Conditions and Social Mechanisms. *Academy of Management Review*, 22:911–945, 1997.
5. T. Norman, A. Preece, S. Chalmers, N. R. Jennings, M. Luck, V. Dang, T. Nguyen, V. Deora, J. Shao, W. Gray, and N. Fiddian. Agent-based formation of virtual organisations. *International Journal of Knowledge Based Systems*, 17(2–4):103–111, 2003.
6. E. Oliveira and A. Rocha. Agents advanced features for negotiation in electronic commerce and virtual organisations formation processes. In *Agent Mediated Electronic Commerce, the European AgentLink Perspective*, volume 1991 of *Lectures Notes in Artificial Intelligence*, pages 77–96, 2000.
7. J. Sabater and C. Sierra. Reputation and social network analysis in multi-agent systems. In *Proceedings of the First International Joint Conference on AAMAS*, pages 475–482, Bologna, Italy, 2002.
8. Carl Shapiro. Consumer information, product quality, and seller reputation. *The Bell Journal of Economics*, 13:20–35, 1982.
9. G. Zacharia and P. Maes. Trust management through reputation mechanisms. *Applied Artificial Intelligence*, 14(8):881–907, 2000.