Using Recency and Relevance to Assess Trust and Reputation

Sarah N. Lim Choi Keung and **Nathan Griffiths** [1]

Abstract. In multi-agent systems, agents must typically interact with others to achieve their goals. Since agents are assumed to be self-interested, it is important to choose reliable interaction partners to maximise the likelihood of success. Sub-standard and failed interactions can result from a poor selection. Effective partner selection requires information about how agents behave in varying situations, and such information can be obtained from others in the form of recommendations as well as through direct experience. In open and dynamic environments, agents face quick and unforeseen changes to the behaviour of others and the population itself. This paper presents a trust and reputation model which allows agents to adapt quickly to changes in their environment. Our approach combines components from several existing models to determine trust using direct experiences and recommendations from others. We build upon previous models by considering the multi-dimensionality of trust, recency of information, and dynamic selection of recommendation providers. Specifically, we take a multi-dimensional approach to evaluating both direct interactions and recommendations. Recommendation sharing includes information about the recency and nature of interactions, which allows an evaluator to assess relevance, and to select recommenders themselves based on trust.

1 Introduction

Trust and reputation have been widely used to attempt to solve some of the issues linked with the uncertainty of interaction. Trust is used to assess the level of risk associated with cooperating with other agents; it is an estimate of how likely another agent is to fulfil its commitments [4, 6, 11]. Trust can be derived from direct interactions between agents and from reputation. Reputation is built from information received from third parties about an agent's behaviour. Based on the reputation information received, agents can make informed decisions about whether or not to interact with others [3].

In a dynamic environment, agents can change behaviour quickly and this must be identified by the agents relying on them, especially if they become less trustworthy in particular aspects of their service. For trust and reputation to be effective in guiding decisions, they must be sensitive to dynamic environments. Therefore, agents should adapt quickly to changes in their environment by selecting appropriate interaction partners and recommenders. In this respect, multi-dimensional trust and reputation allows the original information to be maintained for each service characteristic, such as timeliness and cost, instead of a single aggregated value. Moreover, the sharing of interaction summaries among agents maintains the richness of opinions on a per-characteristic basis and reduces subjectivity. When agents only share calculated trust values, they can be subjectively interpreted in different ways since the evaluator has calculated trust based on its own priorities. Several existing approaches already make use of these aspects, none addresses all of these issues. In this paper we present a model that integrates and extends components from existing approaches to include richer information in decision making and information sharing. The main contributions of our model are: (i) to use the recency of interactions when selecting interaction partners and witnesses, since information can become outdated in dynamic domains, (ii) to ensure that recommendations are accurate and relevant and that they contribute appropriately to the evaluation of reputation, and (iii) to use a richer format of information sharing to reduce subjectivity (including the recency of interactions and the level of witness experience).

[1] University of Warwick, UK, email: {slck, nathan}@dcs.warwick.ac.uk

2 Related Work

Many trust and reputation models have been developed to support agents in soliciting interaction partners. In this section we introduce some of the relevant related work. Marsh's formalism of trust is the basis for many computation approaches, including ours. ReGreT and FIRE are two of the most widely known approaches, while MDT-R and Ntropi introduce features that we build upon in our approach.

2.1 Marsh's Formalism

Marsh's formalism for direct interactions among agents [11], divides trust into *basic trust*, *general trust* and *situational trust*. Basic trust represents an agent's own trusting disposition, derived from its past experiences. An agent's general trust in another depicts how reliable the other is considered, irrespective of the situation. Situational trust is that placed in another agent in a specific situation.

Our model uses these three views of trust when we consider direct trust from direct agent interactions. An agent has an *initial trust* in another agent when it first starts interacting and has had no previous interactions. This is analogous to Marsh's basic trust. Situational trust is used to express an evaluator's trust in a target about a particular task. If the evaluator has interacted with the target but not for the specific task, then general trust is used. General trust is the average trust value calculated from interactions in different situations with the target. Marsh's approach does not take into account reputation and only models trustworthiness from direct experience. This limits the information available for trust evaluation, especially in cases where there are insufficient or no direct interactions. Our model complements direct trust with witness reputation to achieve greater accuracy when predicting agent behaviour. Additionally, we extend Marsh's view by including multidimensionality and agent confidence based on the MDT-R model [6] (described below).

2.2 ReGreT

ReGreT is a modular trust and reputation model that combines three dimensions of information to assess reputation: *individual*, *social* and *ontological* dimensions [12, 13, 14]. The individual dimension relates to direct trust resulting from the outcomes of direct interactions between the evaluator and the target. The social dimension complements this by incorporating information on the experiences of other members of the evaluator's group with the target. There are three aspects to the social dimension: the evaluator's experience with its own group, the experience of members of its group with the target, and the view of the evaluator's group regarding the group that the target belongs to. To determine the social dimension of reputation, an evaluator may use three information sources: *witness reputation* calculated using information gathered from other agents; *neighbourhood reputation* based on the social relations between agents; and *system reputation* which is based on knowledge of the target agent's role. Finally, the ontological dimension considers how the various aspects associated with reputation can be combined. For example, the ontological dimension can define how the reputation of being a good seller relates to a reputation for providing a quality product, a reputation for timeliness, and a reputation for appropriate charging.

ReGreT relies heavily on knowledge of the social structure of the system, in terms of the groups to which agents belong, and the roles that they play. It also relies on knowing the ontological structure of reputation in the domain to define how different aspects of reputation relate to each other. The ReGreT model itself does not consider how agents can build knowledge of the social structure of their environment, but assumes that such information is available for a given domain. In open and dynamic domains such information may not be easily available, and may quickly become outdated as agents leave and join. Additionally, the ontological structure of reputation may not be easily available, and furthermore it may change over time as an agent's preferences change about what is important in an interaction. Although the social structure and reputation ontologies are not necessarily fixed in ReGreT, the sophistication of the model makes it hard to deal with any changes. Our approach uses reputation information provided by others in a similar manner to ReGreT, but without requiring knowledge of the social structure of the system or an ontology of reputation aspects, and so we use witness reputation but not neighbourhood or system reputation. In place of knowing the social structure we use the trust of witnesses and an estimation of the accuracy and relevance of their information, and instead of an ontology we use a weighted product model to combine reputation aspects.

2.3 FIRE

FIRE [9, 10] is a modular approach that integrates up to four types of trust and reputation from different information sources, according to availability. *Interaction trust* results from past direct interactions, and adopts the mechanism used in ReGreT's individual dimension of considering the outcomes of direct interactions between the evaluator and the target. *Role-based trust* uses social and role-based relationships between agents to assess trust, for example the power relationships between agents that might influence trust. *Witness reputation* is built from reports of witnesses about the target agent's behaviour. Finally, *certified reputation* is based on rating references from third-parties that are provided to the evaluator by the target agent itself. An extension to FIRE [8] handles possible inaccurate reports from recommending agents by introducing a credibility model.

The modular approach to trust and reputation in FIRE caters for a wide range of situations that can arise in multi-agent systems. In some situations not all components of FIRE can be used, because the required information may not be available. For example, in dynamic open systems it is likely that role-based trust will be of limited use, since roles are likely to be weakly defined and changeable. Similarly, the use of certified reputation is dependent on the existence of a suitable security mechanism, such as a public-key infrastructure [9]. In open and dynamic domains, as considered in this paper, the interaction trust and witness reputation components of FIRE are the most appropriate. As in ReGreT, FIRE enables an evaluator to rate its direct interactions with the target agent according to a number of terms, such as price and delivery date. Trust can then be calculated within these terms, for example an estimate of trust in terms of delivery date can be determined by extracting all available information about delivery dates from the history of interactions. Our approach extends this model, by providing a mechanism in which overall trust is defined as a combination of the various aspects of previous interactions, such that at run-time an agent can combine information about the various aspects according to their current relative importance. In FIRE, witness selection is done by maintaining a list of acquaintances according to their likelihood of providing the required information. FIRE does not consider how this is done, but assumes an application specific method exists [10]. In this paper, we build upon the interaction and witness reputation components of FIRE to use trust as an estimator for the provision of recommendations, removing the need for an application specific mechanism.

2.4 Ntropi

Abdul-Rahman and Hailes [1, 2] propose a trust and reputation model in which trust and the outcome of experiences are represented in levels. For instance, the labels for the trust level scale are '*Very Trustworthy*', '*Trustworthy*', '*Moderate*', '*Untrustworthy*', and '*Very Untrustworthy*' [1]. The model uses direct trust and reputation, as well as recommender trust to assess witness credibility, in computing a final trust degree for a target. Ntropi models two types of trust: situational trust and basic trust.

This model represents trust by classifying it into five levels, or *strata*. The disadvantage is that the trust values are coarse-grained, thereby losing both sensitivity and accuracy. Although comparisons are easier, the update of values is more complex than using continuous values [5]. In our approach, trust is stored as continuous values for increased accuracy, both for an evaluator's usage and for information sharing. We use direct trust and recommender trust in a similar way to Ntropi, however, we take a multi-dimensional view of trust and reputation that preserves much of the original meaning of the information gathered. In our model, the selection of witnesses is based on two factors: the accuracy and the relevance of recommendations. This is influenced by how Ntropi uses trust in the context of recommendation [1]. The way in which these factors are incorporated into our model is different to Ntropi due to the difference in the representation of trust values.

2.5 MDT-R

MDT-R [6] is a mechanism of multi-dimensional trust and recommendations. Agents model the trustworthiness of others according to various criteria, such as cost, timeliness or success, depending on which criteria the agent considers important. Agents uses their own direct experience of interacting with others, as well as recommendations. Distinguishing trust and recommendations for individual characteristics is valuable in identifying the service characteristics in which the providing agents perform well, or less well. Trust information in multiple dimensions helps to maintain the original interaction data. Trust values are represented numerically in this approach due to the benefits of accuracy and the easiness of comparisons and updates of values. However, MDT-R stratifies trust into levels (*à la* Ntropi) for ease of comparison. The sharing of information among agents often suffers from subjectivity, due to differences in interpretation. MDT-R deals with this by sharing summaries of relevant past interactions, instead of explicit values for trust.

3 Model Description

Our model is broadly based on MDT-R and adopts the multi-dimensionality of trust and recommendations, as well as the sharing of interaction summaries. We extend MDT-R by including information on recency and the experience of witnesses when sharing interaction summaries. This allows an evaluator to more accurately select witnesses, and thereby providers, as it further reduces the subjectivity of interpretation. Our model also considers the relevance of recommendations to better select recommenders and to assign them appropriate weights when calculating reputation.

3.1 Sources of Trust

As we have seen above, many different sources of information can be used to assess trust. Such sources must be available, relevant and accurate enough to be useful in selecting interaction partners. We view trust from direct interactions and recommendations from third parties as the two most important sources of information, since they are typically available with sufficient relevance and accuracy.

Direct interactions are an evaluator's main source of information about a target, and can be used to assess trust. This type of trust from direct experience is called *direct trust*. The second information source is recommendations from third parties. We assume that witnesses give information about a target only if they have interacted with it. We do not currently consider indirect recommendations due to the added complexity of subjective opinions along a chain of witnesses. Trust from third party information is referred to as *witness reputation*. The term is adopted from FIRE [9, 10] and refers to the same concept, but the way we build the reputation is different from FIRE, due to our use of multiple dimensions for trust and reputation.

Our approach integrates these two types of information in different situations. Witness reputation is especially used when the evaluator has insufficient information from direct experience about a target to make an evaluation. Thus, in the event of insufficient information, the two information sources are combined to increase accuracy. In this paper, we do not consider collusion among agents, where a group of agents cooperate for their mutual benefit but impacting on others in the environment as a result. Any inaccuracies in recommendations arise due to differing circumstances, variations in behaviour of the target towards different witnesses, or malicious witness (giving false information). We will consider collusion in future work, as we aim to first ensure that the basic components of our model are efficiently improving agent interaction in a dynamic environment. We also assume that witnesses freely provide recommendations when requested.

3.2 Direct Trust

Trust information is captured in multiple dimensions, as in MDT-R [5, 6]. The separation into several dimensions enables information about specific service characteristics to be preserved. The subjectivity of trust, especially from recommendations, is an obstacle to making full use of the information obtained from witnesses. Sharing multi-dimensional trust information within interaction summaries [6], instead of calculated trust values decreases subjectivity. The dimensions correspond to the necessary characteristics that define a service. Any number of dimensions can be used, but for the purpose of illustration in this paper, we consider that an evaluator α models trust in target β along four dimensions [6]:

- success (T^s_β): the likelihood that β will successfully execute the task,
- timeliness (T^t_β): the likelihood that the task will be performed no later than expected,
- cost (T^c_β): the likelihood that the cost of performing the task will not be more than expected, and
- quality (T^q_β): the likelihood that the quality of the task performed by β will be met.

These trust values are derived from the past interactions of α and β. The evaluator stores information about each interaction in which β has performed a task on its behalf. Information about each interaction includes the service characteristics offered by β, as well as the actual values obtained on completion. The derived trust values refer to a specific task and so this is a type of *situational trust*. A successful interaction is one where β delivers results, irrespective of whether the other three characteristics were met. Meanwhile, a positive interaction with respect to the dimensions of timeliness, cost and quality refers to β performing as expected or better. Trust values are calculated when the evaluator needs to make a decision about whom to interact with. The range of the trust values in each dimension is $[-1, +1]$, where -1 means complete distrust and $+1$ means complete trust. The evaluator stores a history of past interactions with each provider for each task type. We denote the set of interactions in the history about provider β for the task type K as HI_{β_K}. The size of the history corresponds to the number of interactions that the evaluator deems relevant. In future work, evaluators should be able to change the size of the history on a per target basis to enable agents to store only the required information to assess trust.

The situational trust value $ST^d_{\beta_K}$ is a function of the history of interactions with target β:

$$ST^d_{\beta_K} = \frac{\sum_{i=1}^{size(HI_{\beta_K})} I^{d+}_{\beta_K} - \sum_{i=1}^{size(HI_{\beta_K})} I^{d-}_{\beta_K}}{\sum_{i=1}^{size(HI_{\beta_K})} I^{d+}_{\beta_K} + \sum_{i=1}^{size(HI_{\beta_K})} I^{d-}_{\beta_K}} \tag{1}$$

where $I^{d+}_{\beta_K}$ is the number of positive interactions of task type K in dimension d, and $I^{d-}_{\beta_K}$ is the number of negative interactions.

The evaluator also stores the *general trust* of each provider it has interacted with, which has no context and applies regardless of the service provided. General trust is used to assess the overall trustworthiness of an agent. It is useful when the evaluator does not have situational trust for a target for a specific task, as it gives an idea of how the target is likely to perform. The general trust GT^β_α of evaluator α for target β is calculated as an average of the situational trust values in the success dimension:

$$GT^\beta_\alpha = \frac{\sum_{i=1}^{allK} ST^s_{\beta_K}}{allK} \tag{2}$$

where $allK$ is the set of task types. We use only the success dimension to simplify calculation, since completing a task successfully has overriding priority when obtaining an agent's overall trustworthiness.

If there are no previous interactions with β, then general trust is equal to α's disposition, referred to as α's *initial trust*.

MDT-R models confidence and trust decay as two important notions an agent should consider when using past experience for trust assessment. In our model, confidence refers to the number of interactions an evaluator has had with a target agent, and is calculated for each dimension, since not all dimensions are relevant in different interactions. C^d_β denotes the confidence level in the trust assessment of the target β for dimension d. Trust decay refers to the trust values becoming outdated when interactions have not recently taken place. The decay function reduces the trust value according to how outdated the trust values are. In our model, we consider the recency of the interaction history. A weight $\omega_{HI_{\beta_K}}$ is assigned to an interaction according to recency; the more recent the interaction, the more weight it has, since more recent interactions give a more accurate reflection. The weight is based on the time since the interaction occurred and the frequency of interaction with β for the task type K. With fewer recent interactions, trust decays towards the initial trust value.

As proposed in MDT-R, trust values in our model are stratified at the time of comparison. When using numerical values, there is a risk of considering even insignificant differences in values to be important, and stratifying trust reduces this risk. Stratified trust is only used for comparisons and is not communicated to others. In our model, the number of strata used can be specified to allow for different levels of sensitivity. For example, if the number of strata is 10, then trust values in the range $[0.8, 1]$ are taken to be the same. Thus, if two agents β and γ are being compared by situational trust in the success dimension, then if $ST^s_{\beta_K} = 0.85$ and $ST^s_{\gamma_K} = 0.95$ both agents are taken to have similar trust values. A larger number of strata ensures a smoother transition between different strata, especially at the boundary between positive and negative trust [7].

3.3 Witness Reputation

Witness reputation is the trust of a target as communicated by third parties. The reputation of a target is sought when the evaluator has insufficient information to make a decision about whether to cooperate. A lack of information may occur for several reasons. For example, consider an evaluator α who wants to consider agent β for interaction, to perform a task $K1$. In the first case, suppose α has never interacted with β before and thus has no experience of β's behaviour. Alternatively, suppose α has previously interacted with β but for a different task. Another case is when α has had too few interactions with β, or they are too outdated. In all these cases, α can ask the opinions of others who have interacted with β, in order to get a more accurate assessment of β's trustworthiness.

When an evaluator requires recommendations for an agent, it must decide which agents to ask. Such agents might have different kinds of experience with the target, and their opinions might not be useful to the evaluator. To decide who to ask, the evaluator can use *recommendation trust*, which estimates the accuracy and relevance of a witness' recommendation for the evaluator's purposes. Accuracy measures the similarity between the evaluator's own experience and the opinion given by the witness. Meanwhile, relevance relates to how useful the recommendation is based on the recency of the interactions, the experience of the witness, and how trustworthy the witness is in giving recommendations.

FIRE considers whether the witness has sufficient information about the target to give an opinion. An extension to FIRE [8] considers the credibility of the witness in providing opinions about other agents. This enables the evaluator to identify the accuracy of the recommendation by comparing it with its own experience, after an interaction occurs. However, the model does not consider the relevance of a witness' trust information for the evaluator's purposes. In MDT-R, an agent selects witnesses by considering its most trusted interaction partners. However, it does not select witnesses based on the relevance of recommendations and there is no validation of whether the witness has given accurate information. The uncertainty lies in the possible difference in behaviour of the target towards different evaluators. Ntropi considers two factors when dealing with recommendations: (i) the closeness of the witness' recommendation and the evaluator's own judgement about the target, and (ii) the reliability of the witness in giving accurate opinions over time.

Our approach to reputation is influenced by Ntropi's consideration of accuracy and relevance when selecting witnesses. The relevance of recommendations is calculated by taking into account their recency, the experience of the witness, as well as the evaluator's recommendation trust and confidence in the witness. As for the accuracy of opinions, this is done for interactions that have taken place following positive recommendations. The evaluator compares the outcome of the interaction with the recommendation previously obtained to assess how accurate it was. Recommendation trust is updated for each agent that has given recommendations. Initially, witnesses have a recommendation trust value equal to their general trust. This is later updated if the evaluator interacts with the recommended provider.

Witnesses provide the evaluator with interaction summaries for a specific task type where available. The summaries contain information such as the number of interactions the recommendation is based on, the recency of these interactions, and the proportion of positive and negative interactions in each trust dimension. If the witness does not have situational trust information, it provides its general trust in the target. The use of interaction summaries is similar to that in MDT-R with the additional sharing of information about recency and experience, which can improve the evaluator's adaptation to changes in the behaviour of target agents. The evaluator combines the different recommendations by applying weights according to how relevant the witness' experience is, compared to the evaluator's. The weight $\omega_{WRR_{i\beta}}$ is the weight of the witness reputation relevance WRR of witness i in providing a recommendation for target β.

Thus, the witness reputation WR of target β's task type K in the dimension d is a function of the opinions received from witnesses and their respective weights:

$$WR^d_{\beta_K} = \sum_{i=\gamma}^{\epsilon} \left(\frac{I^{d+}_{i\beta} - I^{d-}_{i\beta}}{I^{d+}_{i\beta} + I^{d-}_{i\beta}} \times \omega_{WRR_{i\beta}} \right) \tag{3}$$

where γ to ϵ are the set of selected witnesses for target β. $I^{d+}_{i\beta}$ is the number of interactions of the witness i with the target β, for which β has met expectations for the dimension d, and $I^{d-}_{i\beta}$ is the number where expectations are not met. The weight ascribed to a witness recommendation is dependent on its experience and its relevance. Thus, the evaluator can include the recommendations in each trust dimension of success, timeliness, cost and quality.

The relevance of the recommendation of witness i about target β $WRR_{i\beta}$ is calculated as:

$$\begin{aligned} WRR_{i\beta} \quad = \quad & \left(\frac{t_{curr} - t_{median(HI_{\beta_K})}}{t_{curr}} \right) \\ & + \frac{max_{WI}}{total_{WI}} + RT^i_\alpha + \omega_{C_{RT^i_\alpha}} \end{aligned} \tag{4}$$

where t_{curr} denotes the current time and $t_{median(HI_{\beta_K})}$ is the recorded time of the median interaction as provided by the witness i

for interaction with target β about task K. The inclusion of time in the calculation indicates the recency of the interactions on which the recommendation is based. The maximum number of interactions that the witnesses have used when giving recommendations is max_{WI}, and $total_{WI}$ is the total number of interactions actually used in that recommendation. The confidence of the evaluator α in its recommendation trust in the witness i is denoted as RT^i_α and the confidence weight $\omega_{C_{RT^i_\alpha}}$ shows the amount of influence this recommendation has compared to others.

The evaluator collects information about the witness from direct interactions and from previous recommendations the witness has provided. We do not assess the reliability of witnesses by collecting information from other agents because of the subjectivity of evaluating a witness' ability to provide recommendations to different agents.

3.4 Aggregation of Trust Sources

The evaluator α makes use of direct trust and witness reputation when assessing the trustworthiness of several potential providers for a task, and selects the best provider. The performance value of each provider is calculated as in MDT-R [6], with some changes to cater for the additional information when evaluating witness reputation.

The performance value for each potential provider is calculated as:

$$PV(\beta) = \prod_{i=1}^{n} (f_{\beta_i})^{\mu_i} \tag{5}$$

where there are n factors and f_{β_i} is the value for agent β in terms of the i'th factor and μ_i is the weighting given to the i'th factor in the selection of the agent's preferences. To assess trust using only direct trust, the values are stratified and the performance value is:

$$\begin{aligned} PV(\beta) \quad = \quad & (max_c + 1 - \beta_c)^{\mu_c} \times \beta_q^{\mu_q} \\ & \times stratify(T^s_\beta)^{\mu_{ts}} \times stratify(T^t_\beta)^{\mu_{tt}} \\ & \times stratify(T^c_\beta)^{\mu_{tc}} \times stratify(T^q_\beta)^{\mu_{tq}} \end{aligned} \tag{6}$$

where β_c and β_q are β's advertised cost and quality respectively, max_c is the maximum advertised cost of the agents being considered, μ_c and μ_q are the weightings given to the advertised cost and quality, and μ_{ts}, μ_{tt}, μ_{tc}, μ_{tq} are the weightings for the trust dimensions of success, timeliness, cost and quality respectively.

The calculation of the performance value, considering both direct trust and witness reputation is as follows:

$$\begin{aligned} PV(\beta) \quad = \quad & (max_c + 1 - \beta_c)^{\mu_c} \times (\beta_q)^{\mu_q} \\ & stratify(T^s_\beta)^{\mu_{ts}} \times stratify(T^c_\beta)^{\mu_{tc}} \\ & stratify(T^t_\beta)^{\mu_{tt}} \times stratify(T^q_\beta)^{\mu_{tq}} \\ & stratify(WR^s_\beta)^{\mu_{rs}} \times stratify(WR^c_\beta)^{\mu_{rc}} \\ & stratify(WR^t_\beta)^{\mu_{rt}} \times stratify(WR^q_\beta)^{\mu_{rq}} \end{aligned} \tag{7}$$

where WR^d_β is the witness reputation for target β in the dimension *d*, and μ_{rs}, μ_{rc}, μ_{rt}, μ_r are the weightings for the witness reputation in the dimensions of success, timeliness, cost and quality respectively. (Note that the weights μ_i must sum to 1.)

4 Experimental Results

To validate our approach we have built a simulation environment, and have obtained a number of initial experimental results. Although more experimentation is required, our initial results are promising and demonstrate how trust and reputation can be used to facilitate more effective partner selection.

4.1 Effects of Size of Interaction History

We have investigated how our model behaves when agents change behaviour dynamically. Using a population of 50 agents we observe specific agent interactions. Half of the agents are malicious, and do not always complete the tasks. The remaining agents can be dishonest, and for instance, may charge more than advertised. We have simulated agent interactions over 1500 cycles, where one cycle allows every agent to have part of its tasks performed and to carry out tasks for others. We select one provider for a specific type of task and observe the evaluator's assessment of trust and performance of that provider.

The evaluator uses a history of interactions for each provider task type to predict that provider's likely future behaviour. We observe how the size of the history window affects the evaluator's decision making when others' behaviour changes. Tables 1 and 2 show the average number of cycles the evaluator takes to reach the updated behaviour of the target agent.

Table 1. Reliable to unreliable

Size	Average Duration
5	58.2
10	145.4
20	162.2
30	348.0

Table 2. Unreliable to reliable

Size	Average Duration
5	395.8
10	425.6
20	757.2
30	831.0

Tables 1 and 2 show that it takes longer for the evaluator to notice a change in provider behaviour with larger interaction window sizes. From these results, we expect that fewer failures will occur when the window size is smaller. In experiments where provider behaviour oscillates between good and bad , we also found that for smaller window sizes, the evaluator reacts faster to changes. Figure 1 shows the proportion of failed tasks for each window size.

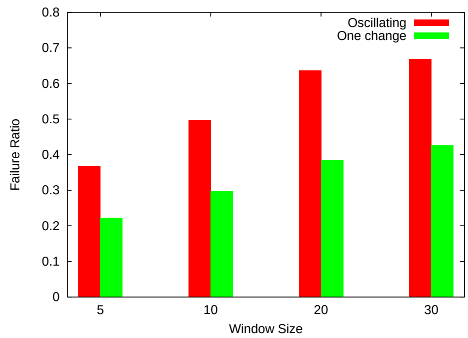

Figure 1. Failure ratio where provider behaviour oscillates between good and bad, compared to one change from good to bad

A malicious provider can however exploit an evaluator's small window to fail some interactions, while keeping the number of successful ones high enough for the evaluator to predict high reliability for that provider. We have set up an experiment where the malicious provider fails an interaction with the evaluator every 6 interactions. For window sizes 5, 10 and 20, the failure ratios are similar at around 0.16, while for the larger window size 30, we observe a slight decrease in failure of around 3%. Compared to Figure 1, smaller window sizes are not beneficial in recognising some behaviour changes. Hence, the evaluator needs to find the right balance between adaptation speed and guarding against such malicious behaviour.

4.2 Comparison of Evaluation Mechanisms

We have compared the effectiveness of using trust, and trust with reputation, against using single service characteristics in a number of settings. Again, we use a population of 50 agents, half of which are malicious. The simulation ran for 500 cycles with individual task execution taking several cycles, depending on execution speed and task duration. The set of agents offers the same task types over the simulation runs, but agent behaviour varies in terms of honesty.

The experiments are set up to observe the performance of evaluator $a1$. Agent $a1$ has a set of tasks to be performed and there are several alternative providers. We look at three evaluation mechanisms that $a1$ might use to assess providers: cost, trust and trust with reputation. We consider the number of tasks generated that have been successful, unsuccessful or incomplete. These are presented as a ratio of the total number of $a1$'s tasks. If the evaluator adds a new task type later in the simulation, it will have no previous interactions for this task and so will ask for recommendations.

Figures 2 and 3 show representative results for the distribution of task performance, where new task types are introduced during the simulation. The ratio of success (Success), execution failure (Failed-U), declined tasks (Failed-D) and any remaining tasks (Remaining) is shown. The evaluation mechanisms are denoted as C, T and TR for cost, trust and trust with reputation respectively. The results are affected firstly by the nature of the population, with more honest populations giving higher success rates, as expect. In the case of Figure 2 the evaluator was situated in a more cooperative environment. The results also show that using trust or trust and reputation improve the success rate compared to using the service characteristics (in this case, cost). In cooperative environments there is a small improvement, while in less honest populations the improvement is more significant (Figure 3). Our results also show that depending on the environment, trust or trust and reputation may give the best result. We are conducting ongoing experiments to identify the conditions that determine which method is best.

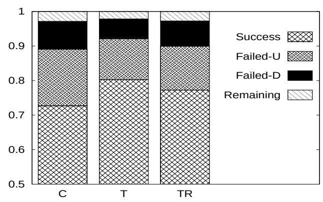

Figure 2. Population set 1

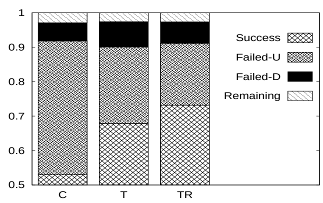

Figure 3. Population set 2

5 Conclusions and Future Work

From our experiments we observe that using trust and trust with reputation to select providers gives better results in most cases, than using service characteristics only. In some situations, the use of trust together with reputation is an improvement over the use of trust only. However, further experimentation is required to determine the circumstances in which the improvement is significant. The ability to recognise different situations can help an agent to better decide which evaluation mechanism to use for the maximum benefits. We have also considered how our model performs when agents change behaviour. Our aim is to enable an evaluator to quickly identify behaviour changes and adapt its strategy to maintain a high success rate. A smaller interaction window size enables the evaluator to reassess trust quickly. However, in certain cases, malicious agents can exploit this by periodically failing. The development of our model, and our initial results, highlight many questions that must be answered for effective use of trust and reputation. One important question is how to balance the potentially conflicting features that an evaluator needs, such as the compromise between the speed of adaptivity to behaviour changes and guarding against malicious behaviour. Future work will consider how agents can achieve this balance, and will investigate the circumstances under which trust or trust and reputation should be used.

REFERENCES

[1] A. Abdul-Rahman. *A Framework for Decentralised Trust Reasoning*. PhD thesis, Department of Computer Science, University College London, UK, 2005.

[2] A. Abdul-Rahman and S. Hailes. Supporting trust in virtual communities. In *Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS 2000)*, page 6007. IEEE Computer Society, 2000.

[3] V. Buskens. Social networks and the effect of reputation on cooperation. ISCORE Paper No. 42, Utrecht University, 1998.

[4] D. Gambetta, editor. *Trust: Making and Breaking of Cooperative Relations*. Department of Sociology, University of Oxford, 2000.

[5] N. Griffiths. Task delegation using experience-based multi-dimensional trust. In *Proceedings of the 4th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005)*, pages 489–496, New York, NY, USA, 2005. ACM Press.

[6] N. Griffiths. Enhancing peer-to-peer collaboration using trust. *International Journal of Expert systems with Applications*, 31(4):849–858, 2006.

[7] N. Griffiths. A fuzzy approach to reasoning with trust, distrust and insufficient trust. In *Cooperative Information Agents X*, volume 4149 of *Lecture Notes in Computer Science*, pages 360–374. Springer-Verlag, 2006.

[8] T. D. Huyhn. *Trust and Reputation in Open Multi-Agent Systems*. PhD thesis, Electronics and Computer Science, University of Southampton, June 2006.

[9] T. D. Huynh, N. R. Jennings, and N. Shadbolt. Developing an integrated trust and reputation model for open multi-agent systems. In *Proceedings of the 7th International Workshop on Trust in Agent Societies*, pages 65–74, New York, USA, 2004.

[10] T. D. Huynh, N. R. Jennings, and N. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Journal of Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, 2006.

[11] S. Marsh. *Formalising trust as a computational concept*. PhD thesis, Department of Computer Science, University of Stirling, 1994.

[12] J. Sabater and C. Sierra. REGRET: reputation in gregarious societies. In Jörg P. Müller, Elisabeth Andre, Sandip Sen, and Claude Frasson, editors, *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 194–195, Montreal, Canada, 2002. ACM Press.

[13] J. Sabater and C. Sierra. Social ReGreT, a reputation model based on social relations. *ACM SIGecom Exchanges*, 3(1):44–56, 2002.

[14] J. Sabater and C. Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24:33–60, 2005.