Indirect Recommendations for Improved Trust Assessment

Sarah N. Lim Choi Keung
Department of Computer Science
University of Warwick
Coventry CV4 7AL, UK
slck@dcs.warwick.ac.uk

Nathan Griffiths
Department of Computer Science
University of Warwick
Coventry CV4 7AL, UK
nathan@dcs.warwick.ac.uk

ABSTRACT

Agents in open and dynamic environments face the challenge of uncertainty while interacting with others to achieve their goals. They face quick and unforseen changes to the behaviour of other agents and the population itself as agents join and leave at will. Since agents are assumed to be self-interested, it is essential for them to be able to choose the most reliable interaction partners to maximise the success of their interactions. Efficient agent selection requires information about their behaviour in different situations. This information can be obtained from direct experience as well as from recommendations. This paper presents a trust and reputation model, which allows agents to select interaction partners efficiently by adapting quickly to a dynamic environment. Our approach is built upon a number of components from several existing models to assess trustworthiness from direct interactions and recommendations. We take a multi-dimensional approach to evaluate trust and reputation and include indirect recommendations as another source of trust. This reinforces our previous work on recommendation sharing, which includes information about the recency and relevance of interactions, allowing an evaluator to select recommenders based on trust.

1. INTRODUCTION

Agents in a multi-agent environment seek to maximise the success of their interactions with other agents, while achieving their individual goals. Trust and reputation are popular mechanisms used to help in the selection of the best suited interaction partners by resolving the issues of uncertainty. Trust is an assessment of the likelihood that an agent will cooperate and fulfil its commitments [3, 11]. The reputation of an agent also contributes to its trust assessment and is derived from third party opinions.

Trust assessment in an open and dynamic environment is particularly challenging due to agents joining and leaving the domain at will and to the possible changes in their behaviour. Hence, the objective of a trust assessment mechanism is to allow the best agent selection in the light of the uncertainties linked to dynamic agent behaviour. This paper presents an extension to our previous work [10], which improves trust assessment by considering the recency of interactions and the relevance of recommendations. The model is based on a multi-dimensional representation of trust and reputation and uses an evaluator's past direct interactions with the target, as well as recommendations from witnesses which interacted themselves with the target, to assess the target agent's trustworthiness. However our previous work did not include indirect recommendations as a source of trust information. An indirect recommendation is the information about the target that a witness shares with the evaluator, without any direct interactions between the witness and the target. The witness obtains the recommendation from other agents who have interacted with the target, or who have themselves obtained the information from other agents. We therefore extend our model to enable agents to share indirect recommendations, without having to interact with the target directly. Including indirect recommendations as a source of trust is important for the evaluator to obtain accurate agent information, especially in cases where there is insufficient direct interactions and direct recommendations.

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 direct and indirect 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 [5] (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 [14, 15, 16]. 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. ReGreT considers witness reputation to include both direct and indirect opinions [12]. There is, however, no indication of the nature of the indirect recommendations, whether they are conveyed in original form or if the witness shares a modified opinion based on its own experience. In our model, witnesses share indirect recommendations unchanged, to avoid any subjectivity issues. The *principal recommender* is the first witness in the recommendation chain that the evaluator contacts and it has directly interacted with the evaluator in the past. *Secondary witnesses* are other witnesses along the recommendation chain and are contacted first through the principal recommender, and then by subsequent witnesses. For example, if a principal recommender has had no interactions with the target agent, it will ask secondary witnesses to give their opinion. The principal recommender will then return the opinion obtained to the evaluator and since it has had no interactions with the target about that service, it is not in a position to pass judgement about the secondary witness' recommendation. Moreover, 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. ReGreT uses its ontology of agent behaviours to specify the relevance of certain behaviours when choosing agents. In our model, multi-dimensionality is applied on a per service basis, as agents can behave differently when dealing with different services. The aggregation of trust and reputation values considers not only the different dimensions of trust and reputation, but also the importance of the dimension itself, as well as how direct trust and reputation affects the final value.

2.3 FIRE

FIRE [8, 9] 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 [7] 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 [8]. 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 [9]. 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. Witnesses in FIRE return referrals to those acquaintances they believe will provide relevant information when they cannot provide direct recommendations. In our model, the principal recommender makes the request to secondary witnesses and returns the original ratings to the evaluator. In this paper, we limit the level of indirection to one, such that the principal recommender only obtains indirect recommendation for the evaluator from its most trusted secondary recommender. This ensures that the evaluator obtains more reliable information from an agent which is closer to the target in terms of the level of indirection.

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 [4]. 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's direct and recommender contexts. However, we take a multi-dimensional view of trust and reputation that preserves much of the original meaning of the information gathered. Moreover, in our model, recommender trust is based on how reliable the principal witness is in sharing both direct and indirect recommendations. The evaluator does not need to know who gave the indirect recommendation or the number of levels of indirectness involved. Meanwhile, Ntropi looks at recommendation chains individually and takes into account the intermediate recommenders. Our approach bases the selection of witnesses 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. We use a similar approach to indirect recommendation as Ntropi, in that insufficient direct interactions will drive the evaluator to seek for recommendations and although direct recommendations are preferred, a chain of recommendations is also considered.

2.5 MDT-R

MDT-R [5] 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. We further extend MDT-R's mechanism of obtaining recommendations by also including indirect recommendations.

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. We also build upon our previous work to include indirect recommendations as a source of trust.

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. Witnesses give information about a target from their own previous experience or from indirect recommendations from trusted agents. Trust from third party information is referred to as *witness reputation*, irrespective of whether is based on direct or indirect recommendations. The term is adopted from FIRE [8, 9] 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. We assume that witnesses freely provide recommendations when requested. Additionally, we assume that the principal recommender will provide the best possible indirect recommendation as this impacts how the evaluator regards it as a trustworthy recommender. 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.

3.2 Direct Trust

Trust information is captured in multiple dimensions, as in MDT-R [4, 5]. 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 [5], 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 [5]:

- success ($T^s_{\alpha\beta}$): the likelihood that β will successfully execute the task,
- timeliness ($T^t_{\alpha\beta}$): the likelihood that the task will be performed no later than expected,
- cost ($T^c_{\alpha\beta}$): the likelihood that the cost of performing the task will not be more than expected, and
- quality ($T^q_{\alpha\beta}$): 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 $HI_{\beta K}$. 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 $ST^d_{\alpha\beta K}$ is a function of the history of interactions of evaluator α with target β:

$$ST^d_{\alpha\beta K} = \frac{I^{d+}_{\alpha\beta K} - I^{d-}_{\alpha\beta K}}{I^{d+}_{\alpha\beta K} + I^{d-}_{\alpha\beta K}} \quad (1)$$

where $I^{d+}_{\alpha\beta K}$ is the number of positive interactions agent α has experienced with target β, of task type K in dimension d, and $I^{d-}_{\alpha\beta K}$ 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 $GT_{\alpha\beta}$ of evaluator α for target β is calculated as an average of the situational trust values in the success dimension:

$$GT_{\alpha\beta} = \frac{\sum_{k=1}^{allK} ST^s_{\alpha\beta K}}{allK} \quad (2)$$

where $allK$ is the size of 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, in the cases where past experience in specific task types are not available. If there are no previous interactions with β, then general trust is equal to α's disposition, referred to as α's *initial trust*, denoted as $initialT_\alpha$. The initial trust of an agent is based on its disposition to trust, that is, its success disposition when an interaction is considered to have or result in a positive outcome.

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. The decay of the situational trust value can thus be defined as a function of the current situational trust value, the recency weight and the initial trust value, as in the following equation:

$$decay(ST^d_{\alpha\beta K}) = f(ST^d_{\alpha\beta K}, \quad initialT_\alpha, \quad \omega_{HI_{\beta K}}) \quad (3)$$

where $ST^d_{\alpha\beta K}$ is the current situational trust of the evaluator in the target β in the dimension d for the task type K. The term $initialT_\alpha$ expresses the initial trust value of the evaluator α. In this equation, if the situational trust is not available, the general trust is used instead.

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 $ST^s_{\alpha\beta K} = 0.85$ and $ST^s_{\alpha\gamma K} = 0.95$ 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 [6].

3.3 Witness Reputation

Witness reputation is the trust of a target as communicated by third parties and can be built from either direct or indirect recommendations. The reputation of a target is sought when the evaluator has insufficient information from it own past experience 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 of type K_1. 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 type, such as K_2. 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. Direct and indirect recommendations can provide useful information about the trustworthiness of the target in meeting its commitments.

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 [7] 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. The same mechanism applies to direct and indirect recommendations as the evaluator does not differentiate between the two sources of recommendation. The evaluator's recommendation trust in the principal recommender relies on how reliable it is in providing accurate and relevant opinions. 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. The evaluator does not distinguish between direct and indirect recommendation trust and therefore the recommendation trust value represents the trustworthiness of the witness in providing any type of recommendation. 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. The update function is outlined in Equations 4 to 7. The evaluator keeps a record of all the recommenders for a task and it updates its recommendation trust in each of them after the interaction with the target.

Equation 4 shows the evaluator α's update of its recommendation trust RT of witness i when $STdiff < 0.2$. $STdiff$ is the difference between the new situational trust value resulting from the interaction and the value recommended by witness i. For small differences, the recommendation trust increases, as it suggests that the recommendation is accurate and reliable enough.

$$\begin{aligned} update(RT^i_\alpha) &= RT^i_\alpha + posIncrement \\ & \quad if\, STdiff < 0.2 \end{aligned} \tag{4}$$

$$\begin{aligned} posIncrement &= \left(\frac{max_{RT} - STdiff}{|max_{STdiff}|}\right) \times \omega_{opinion} \\ & \quad \times disposition_s \times distToMaxRT \end{aligned} \tag{5}$$

where max_{STdiff} is the maximum difference in value between the resulting situational trust and the recommended value. The term $disposition_s$ represents the success disposition of the evaluator, which is an indication of its behaviour as a result of a successful interaction, while $distToMaxRT$ is the difference between the current recommendation trust and the maximum value.

The next two equations 6 and 7 show how the recommendation trust is updated if the recommendation is further from the actual interaction.

$$\begin{aligned} update(RT^i_\alpha) &= RT^i_\alpha - negIncrement \\ & \quad if\, STdiff >= 0.2 \end{aligned} \tag{6}$$

$$\begin{aligned} negIncrement &= \left(\frac{STdiff}{|max_{STdiff}|}\right) \times \omega_{opinion} \\ & \quad \times disposition_f \times distToMinRT \end{aligned} \tag{7}$$

where $disposition_f$ is the failure disposition of the evaluator, which is an indication of its behaviour as a result of a failed interaction, while $distToMinRT$ is the difference between the current recommendation trust and the minimum recommendation trust value.

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, as viewed by evaluator α is a function of the opinions received from witnesses and their respective weights:

$$WR^{d}_{\alpha\beta K} = \sum_{i=\gamma}^{\epsilon} \left(\frac{I^{d+}_{i\beta K} - I^{d-}_{i\beta K}}{I^{d+}_{i\beta K} + I^{d-}_{i\beta K}} \times \omega_{WRR_{i\beta}} \right) \tag{8}$$

where γ to ϵ are the set of selected witnesses for target β. $I^{d+}_{i\beta K}$ is the number of interactions of the witness i with the target β for tasks of type K, for which β has met expectations for the dimension d, and $I^{d-}_{i\beta K}$ 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 $WRR_{i\beta K}$ is calculated as:

$$\begin{aligned} WRR_{i\beta K} \quad = \quad & \left(\frac{t_{curr} - t_{median(HI_{\beta K})}}{t_{curr}} \right) \\ & + \frac{max_{WI}}{total_{WI}} + RT^{i}_{\alpha} + \omega_{C_{RT^{i}_{\alpha}}} \end{aligned} \tag{9}$$

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.

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 [5], with some changes to cater for the additional information when evaluating witness reputation.

The performance value for each potential provider is calculated as:

$$PV(\beta) = \prod_{i=1}^{n} (f_{\beta_i})^{\mu_i} \tag{10}$$

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:

$$\begin{aligned} PV(\beta) \quad = \quad & (max_c + 1 - \beta_c)^{\mu_c} \times \beta_q^{\mu_q} \\ & \times \mathit{stratify}(ST^{s}_{\alpha\beta K})^{\mu_{ts}} \times \mathit{stratify}(ST^{t}_{\alpha\beta K})^{\mu_{tt}} \\ & \times \mathit{stratify}(ST^{c}_{\alpha\beta K})^{\mu_{tc}} \times \mathit{stratify}(ST^{q}_{\alpha\beta K})^{\mu_{tq}} \end{aligned} \tag{11}$$

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 general trust is used if the situational trust is not available.

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

$$\begin{aligned} PV(\beta) \quad = \quad & (max_c + 1 - \beta_c)^{\mu_c} \times (\beta_q)^{\mu_q} \\ & \times \mathit{stratify}(ST^{s}_{\alpha\beta K})^{\mu_{ts}} \times \mathit{stratify}(ST^{c}_{\alpha\beta K})^{\mu_{tc}} \\ & \times \mathit{stratify}(ST^{t}_{\alpha\beta K})^{\mu_{tt}} \times \mathit{stratify}(ST^{q}_{\alpha\beta K})^{\mu_{tq}} \\ & \times \mathit{stratify}(WR^{s}_{\alpha\beta K})^{\mu_{rs}} \times \mathit{stratify}(WR^{c}_{\alpha\beta K})^{\mu_{rc}} \\ & \times \mathit{stratify}(WR^{t}_{\alpha\beta K})^{\mu_{rt}} \times \mathit{stratify}(WR^{q}_{\alpha\beta K})^{\mu_{rq}} \end{aligned} \tag{12}$$

where $WR^{d}_{\alpha\beta K}$ is the evaluator α's witness reputation for target β for task type K 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.)

3.5 Recommender's Perspective

The previous sections have described our model from the point of view of an evaluator. The evaluator builds the reputation of a target agent in the same way, whether the recommendations provided are direct or indirect. It assesses the principal witness on its reliability and accuracy of providing recommendations, using recommendation trust, irrespective of the source. In future work, we will consider using two separate recommendation trust values for direct and indirect recommendations from the principal recommender.

The principal recommender is the agent from whom the evaluator requests information about a target and it is selected from the evaluator's trusted recommenders or providers. It first searches for any direct task interactions with the target in its interaction history. Past experience with the target is shared with the evaluator in the form of an interaction summary. If there are insufficient or no direct task interactions, the principal recommender requests the opinion of its most trusted recommender. In this version of our model, we consider one level of indirection as this reduces the possibility of inaccuracies. Future work will look into how to apply an efficient way of obtaining indirect opinions along a recommendation chain, whilst maintaining accuracy and relevance.

The secondary recommender returns direct task interaction information with the target to the principal witness as an interaction summary. If it has had interactions about different task types, the secondary witness shares its overall agent recommendation about

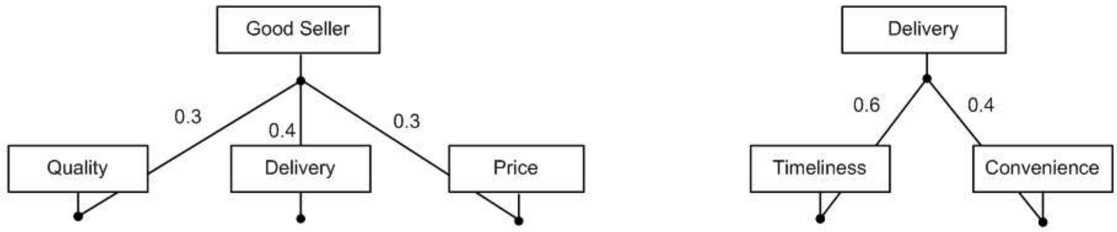

Figure 1: Ontological Structure in ReGreT

the target. If the principal witness has interacted with the target in a different task type as requested by the evaluator, it return its own agent recommendation, rather than the indirect agent recommendation from the secondary recommender. The principal witness does not update its recommendation trust in the secondary witness as it is only passing on the indirect opinion and there has been no effect on its own tasks.

4. DISCUSSION AND CONCLUSIONS

In this paper, we have presented our mechanism of trust and reputation, which is based on a number of trust sources, including direct interactions, as well as direct and indirect recommendations. Combining these sources aims to ensure that the evaluator can more accurately assess the trustworthiness of a potential interaction partner, in different situations. Insufficient direct interactions and direct recommendations can be complemented by including indirect recommendations from trusted agents. Our approach also represents trust and reputation in multiple dimensions to maintain the richness of the original information so as to make more accurate decisions based on the various service characteristics and agent behaviour.

In our model, we make the assumptions that agents have weights associated with the relevance of each service dimension, trust and reputation in the computation of the final performance value that is used to compare several potential interaction agents. We also assume that agents in the domain agree on a set of dimensions that characterise the important features any service should have. As we have illustrated in this paper, agents can consider a set of dimensions, including success, timeliness, cost and quality. Agents are free to use subsets of the agreed dimensions to characterise the interactions they conduct with others.

The trust and reputation model which closely relates to ours is ReGreT [12, 14]. ReGreT also assumes the use of weights by agents when considering the various behaviours in combining the different sources of trust and reputation in their ontological dimension. Figure 1 illustrates an ontological structure of the reputation of a good seller, as used in ReGreT. The reputation of a good seller is related to the reputation of its delivery, the price and quality of its product. As illustrated in the diagram, the delivery aspect can be further broken down into the aspects of timeliness and convenience. In ReGreT, the calculation of the reputation of each aspect involves calculating the reputation of the related aspects which can be in subgraphs [13]. Although the model handles complex behaviours of agents, changes in the weights in any of the subgraphs would involve a recalculation of the reputation in the related aspects, in order to obtain the most reputation value that reflects the agent's current behaviour.

In our approach, agents agree on a fixed set of dimensions that characterise the services in the domain. For example, as illustrated in Figure 2 , a service can be characterised in four dimensions, which have a weight associated to represent their importance. Even if the agents update the weights of the different dimensions to reflect their preferences, this does not affect how the trust and reputation values are calculated. Furthermore, the agreed set of dimensions makes the sharing of information more flexible among agents as all agents use dimensions within the set. The values for each dimension is still subjective, but the sharing of information about the aspects of a service is easier, as compared to ReGreT, where the ontology used can vary among agents and for different aspects that they represent. The different ways of expressing these aspects in ReGreT makes the translation of the meanings among different agents more difficult and is prone to larger extent to the subjectivity problem.

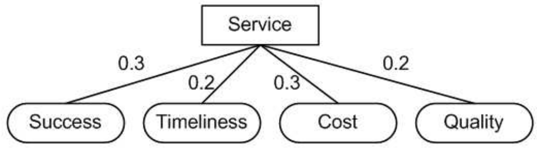

Figure 2: Multiple service dimensions

4.1 Ongoing Work and Future Directions

We are currently developing a simulation to observe the features of our model when used by agents in an open and dynamic domain. We aim to discover the circumstances in which our approach gives an improvement on agent performance. For example, we believe that indirect recommendations are useful especially when an evaluator starts interacting with other agents for new services. Ongoing work is looking at how an agent can recognise the different environmental conditions and choose the most appropriate assessment mechanism to maximise successful interactions.

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

[4] 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.

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

[6] 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.

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

[8] 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.

[9] 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.

[10] S. N. Lim Choi Keung and N. Griffiths. Using recency and relevance to assess trust and reputation. In *Proceedings of AISB 2008 Symposium on Behaviour Regulation in Multi-Agent Systems (to appear)*, 2008.

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

[12] J. Sabater. *Trust and Reputation in Agent Societies*. PhD thesis, Universitat Aùtomata de Barcelona (UAB), Spain, 2003.

[13] J. Sabater and C. Sierra. A reputation model for gregarious societies. In *Fourth Workshop on Deception Fraud and Trust in Agent Societies*, pages 61–70, 2001.

[14] 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.

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

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