

Trust and Reputation Model for Agent-Based Systems

Sarah N. Lim Choi Keung

slck@dcs.warwick.ac.uk

Introduction

Interactions among software agents occur when individual agents are unable to achieve their goals alone. In decentralised multi-agent systems, agents possess limited information about their environment and information obtained from self-interested agents cannot be expected to always be truthful. Each agent aims to achieve its individual goals, and the choice of who to interact with determines, in part, whether goals are fully achieved, partially achieved, or fail. Therefore, the selection of appropriate and reliable interaction partners is crucial to the success of subsequent interactions. In open and dynamic environments, this selection is made more complex due to potential rapid and unforeseen changes in agent behaviour and the agent population itself. Our research aims to minimise the risk associated with the uncertainty of agent interactions. We adopt the commonly used concepts of trust and reputation in managing this uncertainty. Trust is defined as the level of risk associated with cooperating with another agent and it estimates how likely the agent is to fulfil its commitments. Trust can be derived from direct interactions among agents and from reputation, which is built from information received from third parties. Agents can thus make more informed decisions about whether to interact with others, based on trust and reputation.

Related Work

We propose a model of trust and reputation that enables agents to adapt quickly in dynamic environments by judiciously choosing interaction partners. Our approach combines components from several existing models and we build upon these to more efficiently determine trust from direct experiences and recommendations. We take a multi-dimensional approach for evaluating trust and reputation and include a richer set of recommendation information sharing, to allow for the recency and nature of interactions to be communicated. Moreover, agents can better select recommenders based on the relevance of their opinions for the evaluating agent's purposes. Some of the most relevant related work includes Marsh's trust formalism, which considers direct interactions among agents. We complement direct trust with witness reputation to achieve greater accuracy in predicting agent behaviour. Some relevant approaches that also combine both trust and reputation are ReGreT, FIRE, Ntropi, and MDT-R.

Proposed Model

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 a richer set of information on recency and the experience of witnesses when sharing information. This

allows the evaluator to more accurately select witnesses, and thereby providers. We use two main sources of trust information in our model, direct trust from interactions between the evaluator and the target, and witness reputation, which is built from direct and indirect recommendations from third parties. Trust is represented as a continuous value between -1 and +1, which allows us to maintain both sensitivity and accuracy. However, as in Ntropi, we use discrete levels to compare trust values to give a simple comparison method and avoid overfitting. Recommendation trust is used by the evaluator to assess the accuracy of witnesses in their recommendation. The evaluator uses this, together with witness experience and the relevance of opinions based on the recency of interactions to associate weight with recommendations. Reputation is then computed from the recommendations obtained, with more relevant recommendations having greater weight. Direct trust and witness reputation are aggregated to compute a performance value for each potential provider. Various factors contribute to this assessment, according to the importance of direct trust, witness reputation, and advertised service characteristics.

Conclusion and Future Work

We have built a simulated environment to validate our approach. Initial experimental results demonstrate that the use of both trust and reputation in assessing trustworthiness can facilitate more effective interaction partner selection. We have compared the effectiveness of using trust, and trust with reputation, against using single service characteristics in a number of settings. From our experiments, we observe that using trust and reputation gives better results in most cases, compared to using service characteristics only. For instance, in an experiment where the population contains dishonest agents, the success rate was higher when trust or trust with reputation is used by an average of 18%, compared to using cost as service characteristic. Further experimentation is needed to determine the circumstances in which the improvement of using trust with reputation is significantly better than using trust only. Future work will consider how agents can achieve a compromise between quick adaptation and guarding against malicious behaviour. We will also look at the issue of collusion among agents and how agents can improve their success rate when interacting with other agents, in detecting collusive behaviour and consequently avoiding potential failures. More details of our model and related work can be found in [1].

References

- [1] S. N. Lim Choi Keung and N. Griffiths. Indirect recommendations for improved trust assessment. In *Proceedings of 11th International Workshop on Trust in Agent Societies*, pages 33–40. IFAAMAS, 2008.