Sub-delegation and Trust

(Extended Abstract)

Chris Burnett
University of Aberdeen
King's College
Aberdeen, UK
cburnett@abdn.ac.uk

Nir Oren
University of Aberdeen
King's College
Aberdeen, UK
n.oren@abdn.ac.uk

ABSTRACT

Trust mechanisms can allow an agent to identify the most trustworthy entity to which a task should be delegated. Now this entity may further delegate the task, ultimately resulting in a *delegation chain* representing the sub-delegation process. Such *delegation chains* present a problem for current trust evaluation mechanisms, as they typically which reward or penalise a single agent rather than sharing responsibility among all members of the delegation chain. As a result, decisions made on such incorrect trust values would not be optimal, leading to degraded system performance. In this paper we investigate the effects of sub-delegation on a probabilistic trust model and propose a model of weighting trust updates based on shared responsibility. We evaluate this model in the context of a simulated multi-agent system and describe how different weighting strategies can affect probabilistic trust updates when sub-delegation is possible.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems

General Terms

Theory

Keywords

Trust, Delegation Chains

1. INTRODUCTION

Marsh's seminal thesis [1] identified the existence of an implicit trust relationship in multi-agent systems (MASs), and since then, researchers have investigated mechanisms for computing — and acting based on — different trust levels between agents [3, 4, 6]. Such systems have consistently been shown to improve the overall utility of the MAS, with poorly performing agents quickly garnering a low trust rating, which leads to others minimising their interactions with them, thereby reducing the potential harm such untrustworthy agents can cause to the system.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

Trust is critical for the successful *delegation* of tasks in open MASs, whereby one agent requests that some other agent execute a task on the first agent's behalf. Most existing trust mechanisms assume interactions which affect only the truster and trustee, and ignore the possibility of a task being repeatedly delegated from one agent to another. We refer to this sequence of delegations as a *delegation chain*, with the agent originally desiring the execution of the task at the head of the chain, the agent executing the task at its tail, and other, intermediate delegators between them. Such chains appear in a variety of applications (e.g. virtual organisations [2]). The core question we seek to address in this paper is *how the process of delegation (and sub-delegation) should affect trust measures.*

For example, if the agent at the end of a delegation chain fails to achieve the delegated task, all agents in the chain *should* share some of the blame. However, several intuitive ways of apportioning this blame exist, and we seek to investigate the effects of each of these approaches on the system as a whole. In seeking to answer this question our main contribution is to describe and evaluate a model for updating trust in the presence of delegation. Our approach consists of a weighting scheme which discounts the change in trust placed in an agent based on the outcome of a delegated task and the agent's position in the delegation chain.

As an example, consider the situation where Alice asks Bob to book a hotel for her. Bob, being unfamiliar with hotels, asks Charlie to perform the booking. Charlie delegates this request to Debbie, who books a bad hotel, upsetting Alice. Should Alice ever ask Bob to book a hotel for her again? Intuitively, Bob has done nothing wrong; the delegation means that Alice's trust in Bob should be affected to a lesser degree than Bob and Alice's trust in Charlie, and in turn, by Alice, Bob and Charlie's trust in Debbie.

2. APPROACH

We evaluate different weighting measures over a simple model of delegation. Our system consists of a set of agents, each of which are capable of performing some tasks. Agents also have communication links to other agents, and, if they choose not to perform a task, can request that some agent with which they can communicate, perform the task. Now in order to encourage delegation, we assume that agents have different levels of competence in performing different classes of tasks. When an agent must perform some task, they can deal with it in one of three ways, namely 1) decline the task, 2) delegate the task to another agent, or 3) perform the task. Deciding between these three courses of action is done

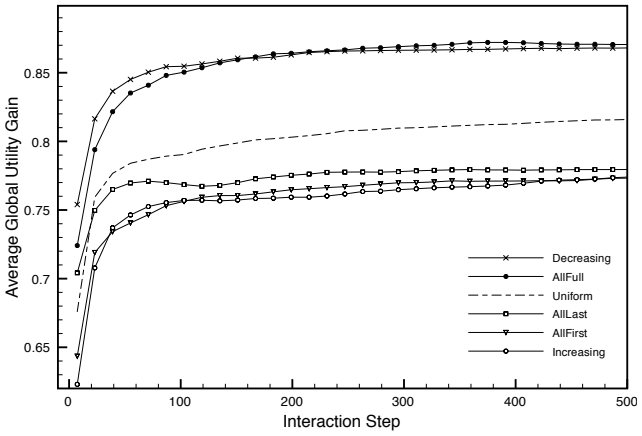

Figure 1: Average global utility gain

based on the agent's own capabilities, and in the trust it has in the agents it communicates with. The same task can be repeatedly delegated, creating a path of agents responsible for its fulfilment, which we refer to as a *delegation chain*.

We use Jøsang's Subjective Logic based trust model [5]. While more complex trust models exist, the use of a relatively straight-forward model simplifies our experiments and allows us to highlight our contribution. Note that we leave the repetitional dimension for future work – agents do not obtain third-party opinions through communication.

Existing approaches to trust apportion blame without taking delegation into account. That is, the trust of the delegator in the delegatee would be updated without taking into consideration any other agent in the delegation chain. In such cases, we argue that it is appropriate to update our trust in the various agents in a delegation chain to different degrees, to reflect the fact that a particular outcome should not reflect equally on all the agents' responsibilities for this outcome, and therefore on their trustworthiness.

We evaluated several different trust update mechanisms, which allow an agent to update its trust rating in all subsequent agents in the chain. These mechanisms weigh trust based on position in the delegation chain, and are as follows: *Uniform Weighting*, where a responsibility is distributed evenly among all in-chain agents; *All-First/-Last Weighting*, where all weight is applied to the first/last agent; *Increasing/Decreasing Weighting*, where increasing/decreasing) proportions of weight are applied applied to consecutive agents in the chain; and *Full Weighting*, where all agents receive full weight. Intuitively, we seek to weigh trust according to an agent's *responsibility for the final outcome*, and evaluate the performance of different models of responsibility.

3. EVALUATION AND DISCUSSION

The different trust update mechanisms were evaluated via simulation. Agents interact over a number of rounds with partners of varying trustworthiness. Figure 1 shows the performance of the system (with respect to global utility) when using different weighting functions.The *Decreasing* and *AllFull* weighting functions appear to perform best.

While it appears attractive, the *AllFull* model is problematic as it is inherently unfair; intermediate agents are penalised (or rewarded) as if they performed the task alone. While this leads to a rapid convergence of performance, strategically minded agents could collude to abuse this approach. For example, agents pass a task around unnecessarily within a group, before finally delegating to a highly trusted individual, so that each agent in the group receives a full positive trust update without having to perform the task. Using the *Decreasing* mode prevents this possibility, as each sub-delegation reduces the weight applied to each agent in the chain. Investigating such strategic aspects of each weighting function will form one area of future work.

An important feature of our approach is that it places few constraints on the particular trust model used, requiring only that the model permits discounted or weighted update. This is already an important feature of many prominent trust models [4], which use discounting to reduce the impact of older experiences on trust assessments, allowing trust models to cope with dynamic behaviour.

Apart from investigating how repetitional information can be included in our model, as future work we intend to examine how partial observability of the delegation chain can be dealt with. We also intend to investigate more complex delegation behaviours between agents (e.g. task splitting).

4. CONCLUSIONS

This research addresses a new and exciting aspect of trust in multi-agent systems, namely how trust should be updated in the context of delegation. Such an approach has many practical applications. In both human and computational domains, one could ask how contractors should trust each other when tasks may be "outsourced" to other parties, and when trustors may be unable to control or observe this outsourcing process. In such situations, our model allows one to apportion responsibility between individuals in a fine grained manner, leading to improved overall system behaviour. We have shown that the choice of weighting function significantly affects the utility of the system. While our approach goes some way towards addressing delegation, it forms only a first step in investigating this aspect of trust, and many exciting avenues of future research remain open.

5. REFERENCES

[1] S. Marsh. Formalizing Trust as a Computational Concept. PhD thesis, University of Stirling, 1994.

[2] J. Patel, W. T. L. Teacy, N. Jennings, M. Luck, S. Chalmers, N. Oren, T. J. Norman,A. Preece, P. Gray, G. Shercliff, P. Stockreisser, J. Shao, W. Gray, N. Fiddian and S. Thompson. Agent-Based Virtual Organisations for the Grid. *Int. Jnl. of Multi-Agent & Grid Systems*, 1 (4). 237-249, 2005.

[3] J. Sabater and C. Sierra. Review on Computational Trust and Reputation Models. *Artificial Intelligence Review*, 24(1):33–60, 2005.

[4] W. Teacy, J. Patel, N. Jennings, and M. Luck. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.

[5] A. Jøsang, R. Hayward, and S. Pope. Trust network analysis with subjective logic. In *Proc. of the 29th Australasian Comp. Sci. Conf.*, pages 85–94, 2006.

[6] Y. Wang and M. P. Singh. Formal trust model for multiagent systems. In *Proc. of the 20th Int. Joint Conf. on Artificial Intelligence*, pages 1551–1556, 2007.