Using Tags To Bootstrap Stereotypes And Trust

(Extended Abstract)

Caroline Player
Department of Computer Science
University of Warwick
Coventry, CV4 7AL, UK
c.e.player@warwick.ac.uk

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

ABSTRACT

Agents joining multi-agent systems (MAS) face two significant problems: they do not know who to trust and others do not know if they are trustworthy. Our contribution extends trust and stereotype approaches to use a comparison of agents' observable features, called *tags*, as an initial indication of expected behaviour. The results show an improvement in agents' rewards in the early stages of their lifetimes, prior to having sufficient information to use trust or stereotype methods.

Keywords

Agents, Trust, Stereotypes, Tags, Reinforcement Learning

1. INTRODUCTION

Trust in another agent is an assessment of their expected behaviour [4]. This is important for choosing an interaction partner, which is challenging in dynamic environments when there is little experience of other agents. Furthermore, agents can get stuck in local maxima when choosing their partners because they continue to interact with the first agent(s) they find who are better than average.

In this paper we bootstrap stereotypes and trust using a technique inspired by tag-based cooperation from evolutionary studies.

2. RELATED WORK

Existing trust and reputation algorithms can accurately evaluate an agent's behaviour if an abundance of past experiences with the agent are available. In this paper, we extend the Beta Reputation System (BRS), a rigorous, mathematical model which underpins many other successful trust algorithms [7, 10, 11], allowing agents to cope with a lack of past experiences.

Stereotypes are correlations between observable features and behaviour [2, 9, 8]. An agent can then predict the behaviour of new agents based on stereotypes despite no interaction data, and so stereotypes can be used to bootstrap trust [1]. In this paper, we draw on the concepts of "relatedness" and selective altruism from evolutionary biology where similar tags amongst agents give rise to cooperation [3, 5]. However, existing tag-based cooperation techniques do not identify features that do not correlate with behaviour or agents who behave well despite having dissimilar tags. We also adopt a reinforcement learning strategy for agents to maximise their use of the information available to them [12].

3. AGENT INTERACTION ENVIRONMENT

The agent environment we consider in this paper is based on Burnett *et al.*'s work [1]. There is a set of agents, $\mathcal{A}$, divided into two subsets, trustors and trustees, connected in a complete bipartite graph where trustors interact with trustees and then assess their expected behaviour.

Agents have observable features, called tags, represented by a vector, $\vec{\tau}$, which are fully and accurately observable. Agents can have *simple behaviours*, where they act the same in every interaction [1], or *biased behaviours* where agents act more favourably to those who are observably similar to them. In an interaction, the trustor receives some utility drawn from the distribution defining the trustee's behaviour.

Partners are chosen with some exploration to prevent agents getting stuck in local maxima. A trustee is chosen with probability $p = e^{\frac{\Delta EB}{T}}$ using simulated annealing, where ΔEB is the difference in expected behaviour between the best known agent and a random trustee. The temperature T encourages the most exploration to occur in the first α percent of an agent's lifetime, a parameter we investigate below.

4. APPROACH

Agent a_i determines trust in a_j using the following method, which calculates belief, b, and uncertainty, u, adopted from BRS [7]:

$$b = \frac{r'_j}{r'_j + s'_j + 2} \tag{1}$$

$$u = \frac{2}{r'_j + s'_j + 2} \tag{2}$$

$$trust_j = b + u \times a \tag{3}$$

Agents maintain tuples $\langle r_j, s_j \rangle$ which store the numbers of good and bad experiences of a_j respectively. A trustor will combine reports on the trustee a_j from other trustors as follows:

$$r'_j = \sum_{tr}^{\mathcal{A}_{tr}} r_j^{tr} \times w, \quad s'_j = \sum_{tr}^{\mathcal{A}_{tr}} s_j^{tr} \times w \tag{4}$$

Once an agent has collected a minimum amount of expe-

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

riences, it will build a decision tree with training data which takes the form $\langle \overrightarrow{\tau_k}, trust_k \rangle$ for each agent a_k that a_i has interacted with, where τ_k is agent a_k's feature vector, thus learning correlations between observable features and trust. The value of a in Equation 3 is then replaced by the output of classifying the trustee's observable features with the decision tree. This is the Stereotype Bootstrap (SB) model. Our proposed Tag Stereotype Bootstrap (TSB) model additionally replaces a before there is sufficient data to use the decision tree.

In the SB model, $w = 1$ to equally weight all reports. In the TSB model w is weighted by how similar the trustor giving the report is to the trustor asking for it. The TSB model tries to account for the different ways a trustee might interact with trustors depending on how similar they are, i.e. the biased behaviour, thus weighting the opinion of a similar trustor higher as they are likely to have been treated the same by the trustee.

4.1 Bootstrapping Using Tags

Agents maintain two $k \times |\overrightarrow{\tau}|$ matrices: one containing Q-values for trust estimates and one for the average error for each Q-value over time. The output of the continuous distance function, $dist$, between two agent's tags is discretized into k equal-width bins. The matrices are then indexed by each tag, $\overrightarrow{\tau_t}$ at each distance interval/bin x:

$$x \leftarrow \left\lfloor dist(\overrightarrow{\tau_t^i}, \overrightarrow{\tau_t^j}) \times k \right\rfloor \quad (5)$$

Where the distance between the t^{th} tag of agents i and j, $dist(\overrightarrow{\tau_t^i}, \overrightarrow{\tau_t^j}) \in [0, 1]$ depends on the type of tag: a binary tag uses Hamming distance, a nominal tag uses the difference in values, and a location tag uses the Euclidean distance.

The first matrix, M^q, maintains Q-values for each tag t in distance bin x at $M^q[x][t]$. The Q-values are updated using the function defined in Equation 6, where the new trust value is the expected behaviour of the partner after the interaction has occurred. The discount factor is set to $\gamma = 1$, and the learning rate, $\lambda = 0.3$. Different values for λ were not found to significantly vary the results, therefore this is not explored further.

The second matrix, M^ϵ, maintains at index $M^\epsilon[x][t]$, a running average error between the original Q-value at that index, and the newly calculated trust after the interaction has occurred. This is defined in Equation 7, where m_i^t is the number of times this tag-distance pair has been updated in order to calculate an average. This attempts to identify irrelevant features as a continually high error for certain tags implies it is likely uncorrelated with behaviour.

$$M^Q[x][t] \leftarrow M^Q[x][t] + \lambda(new\_trust - \gamma(M^Q[x][t])) \quad (6)$$

$$M^\epsilon[x][t] \leftarrow M^\epsilon[x][t] + \frac{|trust - M^Q[x][t]| - M^\epsilon[x][t]}{m_i^t} \quad (7)$$

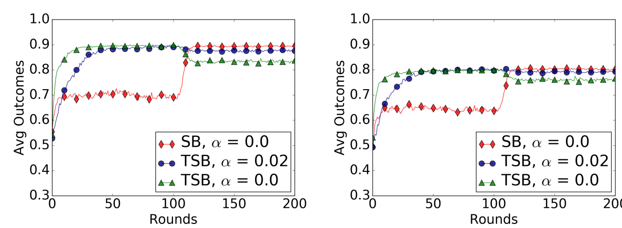
Using these two matrices, we calculate the *a priori* using a weighted sum, given in Equation 8. The intuition is that each tag is associated with an expected reward, and we weight the influence this has on the overall expected behaviour calculation by the error associated with it.

$$a \leftarrow \frac{\sum_t^{|\overrightarrow{\tau}|} M^q[x][t] \times (1 - M^\epsilon[x][t])}{|\overrightarrow{\tau}|} \quad (8)$$

The values in M^Q belonging to trustor a_i are initialised to how a_i would behave towards theoretical partners with each of those tags at each of the distance intervals from their own tags. With simple behaviours, this means that all tags in every interval are initialised to their own behaviour. When behaviours are biased, the same tag will be initialised to a less cooperative value the larger the distance between them. The error matrix is initialised to 0 at each index.

5. EVALUATION AND DISCUSSION

The TSB model significantly improves an agent's performance in the early stages of its life while it is unable to use stereotyping, as shown in Figure 1. Our results also demonstrate a need for exploration, because otherwise the TSB model quickly identifies and selects the best agents, never accurately learning the behaviour of other agents. This causes the stereotype model to train on inaccurate data and perform worse later on. As exploration increases, agents are forced to interact with those who have been, potentially, misinterpreted. Figure 1 demonstrates the trade off between immediate early rewards and attaining accuracy for later rewards. The results between simple and biased behaviours are not directly comparable as the same profile yields different outcomes in each behaviour type.



(a) Agents use Simple behaviours.

(b) Agents use Biased behaviours

Figure 1: Comparison of SB and TSB models

The problem of lack of experience data is especially prominent in dynamic populations where agents rarely have time to collect sufficient experiences. As the TSB model can potentially offer high immediate reward, it is particularly useful in environment that requires agents to become established quickly, such as online marketplaces [6].

The TSB model performs at least as well as the SB model when we reduce the number of experiences required to build the stereotype model. This simplistic environment maintains a static relationship between the tag values and behaviour, so stereotyping data will never go out of date. Realistically, this relationship may change, and the appropriate learning interval to uncover new correlations will depend on the pace of change. However, the TSB model reduces the necessity to know this.

6. ACKNOWLEDGEMENTS

This research is funded by the EPSRC.

REFERENCES

[1] C. Burnett, T. J. Norman, and K. Sycara. Bootstrapping trust evaluations through stereotypes. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, pages 241–248, 2010.

[2] C. Burnett, T. J. Norman, and K. Sycara. Stereotypical trust and bias in dynamic multiagent systems. *ACM Transactions on Intelligent Systems and Technology*, 4(2):26, 2013.

[3] R. Dawkins. *The Selfish Gene*. Oxford University Press, 1976.

[4] D. Gambetta. Can we trust trust? *Trust: Making and breaking coooperative relations*, 13:213–237, 2000.

[5] W. D. Hamilton. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7(1):1–16, 1964.

[6] S. D. Hunt. Resource-advantage theory: An evolutionary theory of competitive firm behavior? *Journal of Economic Issues*, 31(1):59–78, 1997.

[7] A. Jøsang and R. Ismail. The beta reputation system. *Proceedings of the 15th Bled Electronic Commerce Conference*, 5:2502–2511, 2002.

[8] X. Liu, A. Datta, K. Rzadca, and E.-P. Lim. Stereotrust: a group based personalized trust model. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 7–16, 2009.

[9] M. Şensoy, B. Yilmaz, and T. J. Norman. Stage: Stereotypical trust assessment through graph extraction. *Computational Intelligence*, 32(1):72–101, 2016.

[10] L. Teacy, M. Luck, A. Rogers, and N. Jennings. An efficient and versatile approach to trust and reputation using hierarchical bayesian modelling. *Artificial Intelligence*, 193:149–185, 2012.

[11] L. 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.

[12] C. J. C. H. Wakins. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.