Selecting Service Providers from Noisy Reputations

Sandip Sen
Department of Math & CS
University of Tulsa
sandip@kolkata.mcs.utulsa.edu

Neelima Sajja
Department of Math & CS
University of Tulsa
sajjane@ens.utulsa.edu

Categories and Subject Descriptors

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

General Terms

Experimentation, Performance, Economics

Keywords

reputation and trust, task allocation

1. INTRODUCTION

Trust can be a critical parameter in interaction decisions of autonomous agents [2, 4]. We believe that in dynamic, open societies, agents will have to routinely interact with other entities about which they have little or no information. For example, agents will have to select one or few of several such less familiar entities or agents for economic transactions. The decision to interact or enter into partnerships can be critical both to the short term utility and in some cases long term viability of agents in open systems.

The current work builds on our previous work to which we refer the reader for a more complete problem description [3]. We consider the problem of user agents selecting service providers to process tasks. We assume that the performance of different service providers providing a given service can differ significantly and the performance of any service provider varies around some average value. To select a service provider, an user agent queries other user agents for their rating of available service providers. We assume that there are a known percentage of "unreliable" users, who frequently report erroneous estimates about the performance of service providers. In practice, unreliable users may be providing inconsistent erroneous estimates due to self interest, envy, processing constraints, etc. We develop a trust mechanism that determines the number of users to query given a target guarantee threshold likelihood of choosing high-performance service providers in the face of such "noisy" reputations. A key component of our approach is a probabilistic estimation mechanism used to separate reliable and unreliable reputations. We evaluate the robustness of this reputation-based trusting mechanism over varying environmental parameters like percentage of unreliable users, performance difference and variances for high and low-performing agents, learning rates, etc.

AAMAS'03, July 14–18, 2003, Melbourne, Australia.
ACM 1-58113-683-8/03/0007.

2. PROCESSING MIXED REPUTATIONS

We assume the following framework of interaction of the user agent group: a population of N user agents (of which $l \leq \frac{N}{2}$ are unreliable, i.e., provide erroneous recommendations) and P service providers, g is the required probabilistic guarantee for our provider selection procedure given l and N, b is the number of user agents to whom the performance of a provider agent is broadcasted when the latter performs a task for any user agent (the observations differ somewhat from the actual performance which is conveyed accurately only to the user agent whose task was performed).

The performance of a service provider is drawn from a Gaussian distribution with a provider-specific performance mean, μ_p, $\mu_L \leq \mu_p \leq \mu_H$, and a standard deviation common to all providers, σ_p. Let v be the value of service obtained by a user agent from a service provider in some interaction. The b other observers to this event observe performance values drawn from a Gaussian distribution with mean v and standard deviation σ_o. Each user agent updates its rating of a service provider every time it either directly obtaining services from it, or by observing the service provider providing services to another user.

When a user agent i queries another user agent about the performance of a given provider agent j, the queried agent returns its updated performance value r_{ij}^t, which is obtained after t direct interactions or observations. We have considered both consistent and intermittent erroneous estimates being reported by the unreliable users. The procedure for selecting a service provider involves the following major steps:

- The algorithm has to decide how many other users to request reputations from. This is calculated as the minimum integer such that the probability that the majority of the users provide correct reputation estimates is larger than the required guarantee, g [3]:

$$\sum_{i=\max(\lceil \frac{q}{2} \rceil, \lfloor \frac{q}{2} \rfloor + 1)}^{q} \frac{\binom{N-l}{i}\binom{l}{q-i}}{\binom{N}{q}} \geq g.$$

- In contrast to our previous work where other users provided boolean feedback, in the current context we process real-valued ratings from other users. This provides a significantly different research problem and requires the use of more powerful learning techniques to be applied for separating the received ratings into two groups, one for the reliable estimates and the other for the unreliable ones. This separation is possible up to a certain extent under assumptions about the source of erroneous ratings, e.g., all erroneous ratings have similar causes. The separation of the groups can be done without consideration of which group corresponds to the correct/incorrect reputations. To subsequently label the correct reputation group, we have to assume that a sufficient number of users have been queried. To form the two groups and to find the means and standard deviations of the two respective Gaussian distributions, we use the EM algorithm [1][1]. The mean of the ratings from the larger group is used as an approximation of the provider's actual performance. Based on our procedure to select the number of users to query, the larger group will correspond to reliable users only up to a fraction, g, of the number of time the algorithm is executed.

- The algorithm has to determine the summary estimates for the providers from the groups of correct estimates, and finally selecting a provider to interact with. Our agent requests service from the service provider with the highest estimated rating.

3. EXPERIMENTAL RESULTS

We assume that $\forall i, j$, $r_{ij}^0 = 0.5$, i.e., user agents start off with neutral estimates of provider agents. We performed a series of experiments with 40 users and 40 service providers by varying the number of unreliable users for different guarantee thresholds, the spread between the performance means of high-quality (μ_H) and low-quality (μ_L) service providers, the standard deviation in performance σ_p, the number of agents who can observe an interaction, and the estimation error of the number of unreliable users in the population, i.e., the querying user agent believes there are less unreliable users in the population than the actual number. The performance metric we have used is the average performance delivered by the selected service providers over all such selections. Results presented are averaged over 50 runs.

Figure 1 presents the average performance over all interactions when the guarantee threshold is varied for different number of unreliable users. For a guarantee threshold of 0.95 the agents appear to be able to withstand the increase in unreliable user population until they become so numerous that the required number of agents to query increases beyond the population size. This happens at around $l = 16$. The same trend is observed for other plots as well. Note that with lower threshold guarantees, fewer users need to be queried for the same number of unreliable users in the population. Hence, as seen from Figure 1 experiments could be run for a larger number of unreliable user population. These plots demonstrate that the selection procedure prescribed in this paper works well and maintains a steady performance even with increasing unreliable user population. The robustness of our simple probabilistic scheme was surprising and encouraging at the same time.

[1]Space restrictions prevent us from presenting details of this technical core of our work. We would be happy to entertain any questions the reader has about this process

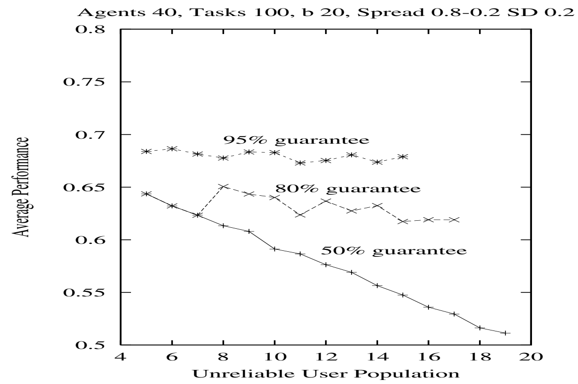

Figure 1: Performance variation with different probabilistic guarantee thresholds.

Our work is in some sense simpler than some of the social reputation mechanisms [2, 4], but addresses a complementary problem of providing a probabilistic guarantee of selection of service providers given only summary statistics of the population distribution. As elaborate long-term modeling is not required, new agents to the community can immediately start using the reputation-based trust mechanisms without maintaining extensive history and knowledge about the social network. Whereas performance can be improved by modeling the trustworthiness of recommending agents, the current work will enable user agents to make prudent selections in volatile groups as long as the percentage mix of unreliable and reliable user agents is only approximately known. The model presented here is simple. It can easily be enhanced to model the nature of user agents, e.g., whether they can be trusted or not, etc. Also, agents may use preconceived model about the distribution of provider performances to identify erroneous estimates if they significantly differ from the prior model. But each of these extensions may limit the applicability of this mechanism, e.g., agents must be in a system for some time before they can effectively rate other agents or a new agent may not be effective in a domain where their prior expectations are wrong.

Acknowledgments: This work has been supported in part by an NSF CAREER award IIS-9702672.

4. REFERENCES

[1] T.M. Mitchell. *Machine Learning*. WCB McGraw-Hill, Boston, MA, 1997.

[2] Michael Schillo, Petra Funk, and Michael Rovatsos. Using trust for detecting deceiptful agents in artificial societies. *Applied Artificial Intelligence*, 14:825–848, 2000.

[3] Sandip Sen and Neelima Sajja. Robustness of reputation-based trust: Boolean case. In *Proceedings of the First Intenational Joint Conference on Autonomous Agents and Multiagent Systems*, pages 288–293, New York, NY, 2002. ACM Pres.

[4] Bin Yu and Munindar P. Singh. Towards a probabilistic model of distributed reputation management. In *Proceedings of the Fourth Workshop on Deception, Fraud, and Trust in Agent Societies*, pages 125–137, 2001.