MetaTrust: Discriminant Analysis of Local Information for Global Trust Assessment

(Extended Abstract)

Liu Xin
School of Computer Engineering
NTU, Singapore
liu_xin@pmail.ntu.edu.sg

Gilles Tredan
TU-Berlin/T-Labs
gilles@net.t-labs.tu-berlin.de

Anwitaman Datta
School of Computer Engineering
NTU, Singapore
anwitaman@ntu.edu.sg

ABSTRACT

A traditional approach to reasoning about the trustworthiness of a transaction is to determine the trustworthiness of the specific agent involved, based on its past behavior. As a departure from such traditional trust models, we propose a transaction centered trust model (MetaTrust) where an agent uses its previous transactions to assess the trustworthiness of a potential transaction based on associated meta-information, which is capable of distinguishing successful transactions from unsuccessful ones. This meta information is harnessed using a machine learning algorithm (namely, discriminant analysis) to extract relationships between the potential transaction and previous transactions.

Categories and Subject Descriptors

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

General Terms

Algorithms, Security

Keywords

trust, discriminant analysis, meta data, large-scale systems

1. INTRODUCTION

Traditional trust approaches [2, 4], while effective when the necessary information is available, often rely upon knowledge that may not actually be available locally to the assessor. For instance, they require to find a trust path between trustor and the target agent, which is not trivial in large systems, and suffers the "weakest link phenomenon" [1]. We thus explore a new trust model (MetaTrust), which is capable of harnessing meta-information which is generally not considered in existing trust models, and may be available locally. This new model, given its use of different kind of information, is meant to complement traditional models.

MetaTrust relies on discriminant analysis (DA) [3] to exploit the agent's local knowledge. DA is a well known family of methods for dimensionality reduction and classification. DA methods take as input a set of events belonging to k (≥ 2) different classes and characterized by various features, and find a combination of the features (a classifier) that separates these k classes of events.

In MetaTrust, a user's past transactions are described by a set of meta-information and classified according to their outcome: successful or unsuccessful (without loss of generality, we consider linear DA over two classes). Each transaction information is stored locally by the user. The user then performs a linear DA on this data to obtain a linear classifier that allows him to estimate whether a potential transaction is likely to be successful or not.

2. OUR APPROACH

Consider a scenario where a customer a_x encounters a potential service provider a_y and a_x has no prior experience with a_y. We assume that a_x can obtain meta information about this potential transaction Θ_{a_x,a_y}. We denote such meta information of Θ_{a_x,a_y} by $M_{\Theta_{a_x,a_y}} = \{m^1_{\Theta_{a_x,a_y}}, m^2_{\Theta_{a_x,a_y}}, ..., m^d_{\Theta_{a_x,a_y}}\}$. So the potential transaction is represented by vector $p = (m^1\ m^2\ m^3\ \dots m^d)$.

We assume that a_x has recorded n historical transactions with other agents. To estimate reliability of this potential transaction, based on transaction outcome, a_x classifies its historical transactions into two disjoint groups, the successful (G_s) and the unsuccessful transaction group (G_u), which are represented as:

$$\mathbf{G_{s/u}} = \begin{pmatrix} m^1_{\Theta_{a_x,a_1}}(s/u) & ... & m^d_{\Theta_{a_x,a_1}}(s/u) \\ \vdots & \vdots & \vdots \\ m^1_{\Theta_{a_x,a_{n_s}}}(s/u) & ... & m^d_{\Theta_{a_x,a_{n_s}}}(s/u) \end{pmatrix} \quad (1)$$

The two transaction groups contain respectively n_s and n_u transactions ($n = n_s + n_u$).

Agent a_x performs linear discriminant analysis to classify the potential transaction as belonging to successful or unsuccessful transaction group to decide whether or not to transact with the corresponding service provider. Let h_x be a $x \times 1$ (column) vector of ones.

This work was funded by A-STAR grant 072 134 0055.

Cite as: MetaTrust: Discriminant Analysis of Local Information for Global Trust Assessment (Extended Abstract), X. Liu, G. Tredan and A. Datta, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1071-1072.

Agent a_x first calculates the centroid of each group: $c_s = \frac{1}{n_s} \cdot h_{n_s}^T G_s$ and $c_u = \frac{1}{n_u} \cdot h_{n_u}^T G_u$.

Similarly, the global centroid is calculated by averaging each type of meta information across all past transactions:

$$c = \frac{1}{n} \cdot h_n^T \begin{bmatrix} G_s \\ G_u \end{bmatrix} \tag{2}$$

In LDA, the internal variance (within-class scatter matrix) and external variance (between-class scatter matrix) are used to indicate the degree of class separability, i.e., to what extent can the successful transactions be distinguished from the unsuccessful transactions. The internal variance, which is the expected covariance of each group is obtained by $S_w^s = \frac{1}{n_s}(G_s - h_{n_s} c_s)^T (G_s - h_{n_s} c_s)$ and $S_w^u = \frac{1}{n_u}(G_u - h_{n_u} c_u)^T (G_u - h_{n_u} c_u)$. So the overall within-class scatter matrix is calculated as the weighted sum of each group's internal variance, where the weight is fraction of transactions regarding the corresponding group: $S_w = \frac{1}{n}(n_s S_w^s + n_u S_w^u)$.

Then a_x calculates external variance, which is actually the covariance of the two groups, each of which is represented by its mean vector: $S_b = \frac{1}{n}(n_s(c_s - c)^T(c_s - c) + n_u(c_u - c)^T(c_u - c))$.

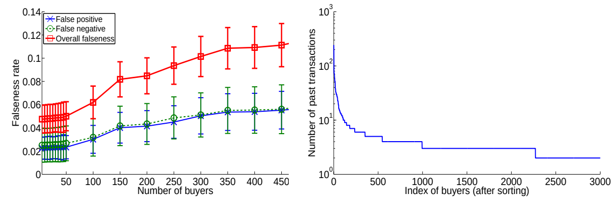
LDA aims to find a projection direction (a transformation) v that maximizes the inter class variance and minimizes the intra class variance. Formally, the criterion function $J(v) = \frac{v^T S_b v}{v^T S_w v}$ is to be maximized.

The projection direction v is found as the eigenvector associated with the largest eigenvalue of $S_w^{-1} S_b$. We then transform the two groups of transactions using v. Similarly, the potential transaction $p = (m^1\ m^2\ m^3 \ldots m^d)$ is also transformed and classified by measuring the distances between transformed potential transaction and the two groups (i.e., centroid), which are calculated as $D_s = v^T p - v^T c_s$ and $D_u = v^T p - v^T c_u$. If $D_u > D_s$, then transaction p is predicted as successful, otherwise it is predicted as unsuccessful.

Note that we try to collect as much meta information as possible, and the MetaTrust model filters out the not-so-relevant variables for us. That is to say, the meta information which is more capable of distinguishing successful transactions from unsuccessful ones will have more impact on the final classification result.

3. EVALUATION

We use real dataset collected from an Internet auction site Allegro to conduct experiments. The Allegro dataset contains 10,000 sellers, 10,000 buyers, more than 200,000 transactions and over 1.7 million comments. In the experiments, a transaction is considered successful if its feedback is positive, otherwise, it is considered unsuccessful. We extract three kinds of meta information from Allegro data: M_1: category of the item; M_2: price of the item and M_3: number of items already sold by the seller when the transaction occurs. We evaluate performance of MetaTrust by studying its capability of detecting Internet auction fraud. When a buyer encounters a potential transaction, which is conducted by an unknown seller, it will gather meta information regarding the item (i.e., M_1, M_2 and M_3) and then perform MetaTrust to estimate the trustworthiness of this transaction with respect to the buyer's past transactions that belong to the same category M_1



(a) Performance. (b) ♯ of buyers' past transaction.

Figure 1: Experiments using Allegro dataset.

We first rank the 10,000 buyers according to number of their past transactions, i.e., the first buyer has the most past transactions. We select subset U_b of these buyers starting from the first one. Each buyer evaluates 100 randomly selected transactions (50% are successful and 50% are unsuccessful). We vary the size of U_b to investigate effect of local knowledge volume.

Fig. 1(a) demonstrates how average rates of various falseness evolve when U_b varies from 5 to 500. As expected, all falseness rates increase when U_b grows. This shows the impact of local knowledge on MetaTrust: when U_b is small, it contains only experienced agents, that all have enough past transactions to allow MetaTrust to issue accurate predictions. As U_b grows, it contains more and more inexperienced agents, for which MetaTrust predictions are less accurate.

Fig. 1(b) shows the distribution of numbers of individual buyers' past transactions (only first 3000 are shown). Note the logarithmic scale for y-axis: the number of past transactions is quickly decreasing. Estimating the minimal number of transactions that allow MetaTrust to be precise is challenging, since not all transactions have the same importance. However, in this set of experiments, we estimate empirically that when numbers of transactions is over 6, the potential transaction can be relatively reliably predicted (i.e., the overall falseness rate is smaller than 0.1).

4. CONCLUSION

Unlike many existing trust models [2, 4], which rely on specific agent's historical information to predict its future behavior, MetaTrust only uses trustor's local knowledge. Using DA, MetaTrust analyzes characteristics of interactions' meta information to obtain a classifier that helps estimate whether the potential interaction is likely to get classified in the successful group or not. Evaluation using real dataset demonstrates efficacy of MetaTrust in detecting Internet auction fraud.

5. REFERENCES

[1] A. Datta, M. Hauswirth, and K. Aberer. Beyond "web of trust": Enabling p2p e-commerce. In *Proceedings of the IEEE CEC*, pages 24–27, 2003.

[2] A. Jøsang and R. Ismail. The beta reputation system. In *Proceedings of the 15th Bled Conference on Electronic Commerce*, 2002.

[3] G. J. McLachlan. *Discriminant Analysis and Statistical Pattern Recognition*. Wiley-Interscience, Augest 2004.

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

www.allegro.pl