Trust and Reputation Model for Agent-Based Systems

Sarah Lim Choi Keung

Intelligent and Adaptive Systems Group

Warwick Postgraduate Colloquium in Computer Science
(WPCCS 2008)
30 June & 01 July 2008

WARWICK

Context

- Dealing with the issue of uncertainty of agent interactions.
    - Agents exist in open and dynamic environments.
    - Their behaviour is difficult to predict.

WARWICK

Solution with Trust and Reputation

- Trust
    - Measure of the level of risk associated with cooperating with other agents.
    - Derived from direct interactions and reputation.
- Reputation
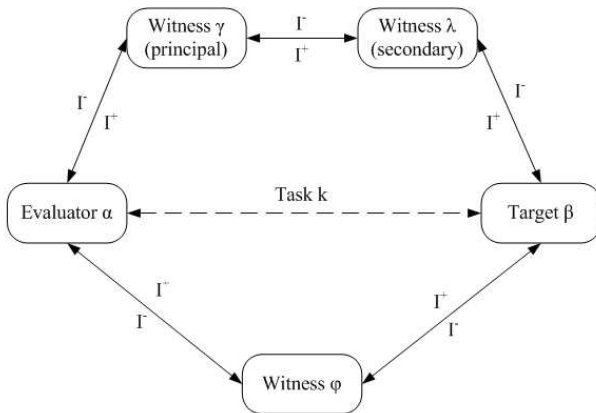   - Built from information received by third-parties about an agent's behaviour.

Figure: Agent interactions and terminology used

Proposed Model

- ▶ A trust and reputation model that allows agents to quickly adapt to their dynamic environment.
- ▶ Approach combines components from several existing models.
- ▶ Builds upon aspects of multi-dimensionality of trust and reputation, recency of information and dynamic selection of recommendation providers.
- ▶ Includes the use of both direct and indirect recommendations for witness reputation.

Marsh's Formalism [1]

- ▶ Formalism of trust from direct interactions, divided into 3 types:
    - ▶ Basic trust
    - ▶ General trust
    - ▶ Situational trust
- ▶ Our approach:
    - ▶ uses the 3 types of trust for direct interactions.
    - ▶ Witness reputation complements direct trust to achieve greater accuracy when predicting agent behaviour.

---

[1]Marsh 1994

ReGreT [2]

- ReGreT is a modular trust and reputation model with 3 dimensions of information:
  - Individual dimension
  - Social dimension
  - Ontological dimension
- Our approach:
  - uses the trust of witnesses and an estimation of the accuracy and relevance of their information
  - uses a weighted product model to combine reputation aspects

---

[2]Sabater 2002, 2003

WARWICK

FIRE [3]

- ▶ Modular approach that integrates 4 types of trust and reputation information sources:
  - ▶ Interaction trust
  - ▶ Role-based trust
  - ▶ Witness reputation
  - ▶ Certified reputation
- ▶ Our approach:
  - ▶ considers the interaction trust and witness reputation components.
  - ▶ uses trust in multiple dimensions as a estimator for the provision of recommendations.

---

[3]Huyhn 2006

WARWICK

Ntropi [4]

- ▶ Trust and reputation model in which trust and the outcome of experiences are represented in levels.
  - ▶ Direct trust is used: basic and situational.
  - ▶ Models reputation.
  - ▶ Recommender trust is used to assess witness credibility.
- ▶ In our model:
  - ▶ trust is stored as continuous values, while levels are only used to compare similar values.
  - ▶ we use direct trust and recommender trust in multiple dimensions.
  - ▶ witnesses are selected according to accuracy and relevance of recommendations

---

[4]Abdul-Rahman 2000, 2005

WARWICK

MDT-R [5]

- ▶ Mechanism of multi-dimensional trust and recommendations:
  - ▶ Agents model trustworthiness according to various criteria important to them, such as timeliness, cost.
  - ▶ Trust values are numerical but trust is stratified into levels for ease of comparison.
  - ▶ Sharing of information is done through interaction summaries of past interactions.

---

[5] Griffiths 2006

Proposed Model Overview

- ▶ Our model is broadly based on MDT-R with extensions to include information on recency and the experience of witnesses when sharing interaction summaries.

- ▶ We also consider the relevance of recommendations to better select witnesses and give them appropriate weights when calculating reputation.

- ▶ We use indirect recommendations as an additional source of trust information to direct trust and direct recommendations.

Sources of Trust

- ▶ Direct trust from direct interactions.
- ▶ Witness reputation as recommendations from third parties.
- ▶ The 2 types of trust information are used in different situations; witness reputation being used especially when the evaluator has insufficient direct experience.
- ▶ Witness reputation is built from both direct and indirect recommendations from third parties.

Direct Trust: Multiple Dimensions

▶ The separation into several dimensions preserves information about specific service characteristics.

▶ Sharing of multi-dimensional trust information decreases subjectivity.

▶ Any number of dimensions can be used, for purposes of illustration, 4 dimensions are modelled: success ($T_{\alpha\beta}^{s}$), timeliness ($T_{\alpha\beta}^{t}$), cost ($T_{\alpha\beta}^{c}$), and quality ($T_{\alpha\beta}^{q}$).

Direct Trust: Situational and General

- Situational trust is a function of the history of interactions of evaluator α with target β:

$$ST_{\alpha\beta K}^d = \frac{I_{\alpha\beta K}^{d+} - I_{\alpha\beta K}^{d-}}{I_{\alpha\beta K}^{d+} + I_{\alpha\beta K}^{d-}} \tag{1}$$

- General trust in a target applies regardless of the service provided:

$$GT_{\alpha\beta} = \frac{\sum_{k=1}^{allK} ST_{\alpha\beta K}^s}{allK} \tag{2}$$

Direct Trust: Decay and Confidence

- ▶ Trust decay occurs when trust values become outdated due to lack of fresh interactions. Trust decays towards the initial trust value.

- ▶ Confidence refers to the number of interactions between the evaluator and the target, in each dimension.

Witness Reputation: Witness Selection

▶ Recommendations involve the selection of witnesses.

▶ Recommendation trust estimates the accuracy and relevance of the witness recommendation:

  ▶ Accuracy measures similarity of experiences.

  ▶ Relevance relates to the usefulness of the recommendation, based on recency, witness experience and trustworthiness of witness.

▶ Witnesses are selected from the evaluator's most trusted interaction partners.

▶ The evaluator combines different recommendations by applying weights according to their relevance.

Recommendations: Evaluator's View

▶ The evaluator does not distinguish between direct and indirect recommendations.

▶ Recommendation trust represents the trustworthiness of the witness to provide any type of recommendation.

▶ Future work will look into potential benefits of having different recommendation trust values for direct and indirect recommendations.
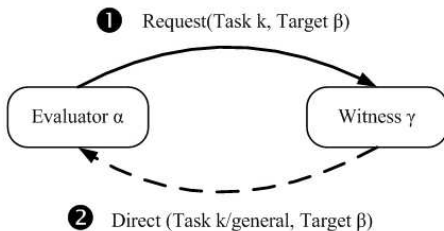
Example: Direct Recommendations

Figure: Interactions between the Evaluator and the Witness

Recommendations: Principal Recommender's View

- ▶ An evaluator requests information about a target from the principal recommender.
- ▶ The principal recommender first considers its own direct interactions with the target.
- ▶ In cases of insufficient or no direct interactions, the principal recommender asks the opinion of its most trusted recommender.
- ▶ We use one level of indirection in this version of our model.
- ▶ Future work will look into how to apply an efficient way of obtaining indirect opinions along a recommendation chain.

Recommendations: Secondary Recommender's View

- ▶ The secondary recommender provides direct task interaction information to the principal recommender.
- ▶ If it has had interactions about different task types than requested, it shares its recommendation about the target's general trust.
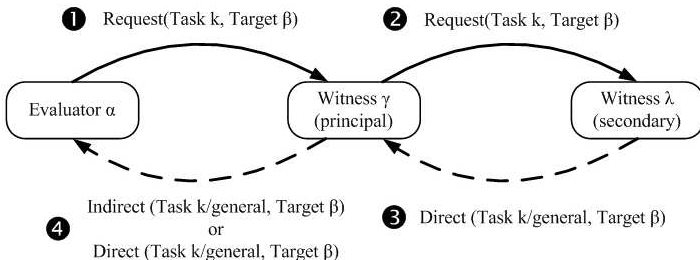
Example: Indirect Recommendations

Figure: Interactions between the Evaluator, Principal and Secondary Witnesses

Witness Reputation: Calculation

▶ The witness reputation WR of target β's task type K in the dimension d is a function of the opinions received from witnesses and their respective weights:

$$WR_{\alpha\beta K}^d = \sum_{i=\gamma}^{\epsilon} \left(\frac{I_{i\beta K}^{d+} - I_{i\beta K}^{d-}}{I_{i\beta K}^{d+} + I_{i\beta K}^{d-}} \times \omega_{WRR_{i\beta}} \right) \qquad (3)$$

▶ $\omega_{WRR_{i\beta}}$ is the weight of the witness reputation relevance WRR of witness i in providing a recommendation for target β.

WARWICK

Aggregation of Trust Sources

▶ The evaluator α uses direct trust and witness reputation to assessing the trustworthiness of several potential providers for a task, and selects the best provider by comparing each provider's performance value:

$$PV(\beta) = \prod_{i=1}^{n}(f_{\beta_i})^{\mu_i} \qquad (4)$$

where there are n factors and f_{β_i} is the value for agent β in terms of the i'th factor and μ_i is the weighting given to the i'th factor in the selection of the agent's preferences.

WARWICK

Conclusion and Future Work

- ▶ We have presented our trust and reputation model based on a number of trust sources: direct interactions, direct and indirect recommendations.
- ▶ Initial experiments on our model show that trust and trust with reputation for selecting providers gives mostly better results than using service characteristics only.
- ▶ Further experimentation on the added benefits of indirect recommendations for the assessment of trust.
- ▶ Future work will focus on how to balance the potentially conflicting features that an evaluator needs to consider.
- ▶ We will also look into how the decay function for trust relates to the interaction history size.
- ▶ Collusion among agents adds to the challenge of accurately predicting agent behaviour.