A Probability-Based Trust Prediction Model Using Trust-Message Passing

Hyun-Kyo Oh, Jin-Woo Kim, Sang-Wook Kim
Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea
rapkyo, racerkim86, wook@agape.hanyang.ac.kr

Kichun Lee
Department of Industrial Engineering, Hanyang University, Seoul, Korea
skylee@hanyang.ac.kr

ABSTRACT

We propose a probability-based trust prediction model based on trust-message passing which takes advantage of the two kinds of information: *an explicit information* and an *implicit information*.

Categories and Subject Descriptors

H.2.8 [**DATABASE MANAGEMENT**]: Database applications—*Data mining*

Keywords

Trust prediction model, Trust propagation, Message passing

1. INTRODUCTION

Trust prediction is the task of inferring trust between pair of users: whether or not a user will build a trust relationship with another user [2]. One of the most important issues in trust prediction research is how to accurately predict trust between such a user pair that is not apparently in a trust relationship [2, 3, 4]. Previous researchers have undertaken the following approaches according to a viewpoint of the kind of information used in the prediction task.

A natural strategy, *an explicit information approach* is to use information on explicit trust relations between pair of users, i.e., taking only into account users explicitly expressing trust on other users [2, 3]. Another approach, *an implicit information approach* which is often applicable to web sites without explicit trust relations, is to infer trust through user interactions [4]. User interactions encompass all activities between a user and others such as evaluating writings and making comments on reviews written by others.

Even though considerable studies on trust prediction have been performed, the collective utilization of both explicit trust and user interactions has attracted little attention in the field of computer science. However, the inclusion of user interactions to explicit trust in a unified model would lead to better trust prediction. Hence this paper aims to propose a new trust prediction model that collectively utilizes the two kinds of information, both explicit trust and user interactions, which have been treated separately.

2. THE PROPOSED MODEL

Let $\mathbf{U}$ denote the set of all users and p_{ij} denote the probability that u_i and u_j build trust. When user u_i decides that user u_j is trustworthy, a trust relationship is formed. In that case, the pair of u_i (called as a trustor) and u_j (called as a trustee) are called *a trust pair*. Let $\mathbf{T}(j)$ be the set of all trustors who trust u_j. Let $\mathbf{R}$ denote the set of all reviews by all users in $\mathbf{U}$. Specifically, let $\mathbf{R}_{ij}$ represent the set of reviews written by user u_j and rated by user u_i. When u_i rates u_j's reviews, the two users are called *a review-rating pair*, i.e., a pair of a review writer and a review rater. Let μ_{ij} indicate the average of all ratings in $\mathbf{R}_{ij}$.

We examine review ratings to estimate *initial trust probabilities* between users. Figure 1 shows the change of the numbers of trust pairs and non-trust pairs according to the increment of $|\mathbf{R}_{ij}|$.

The x coordinate represents $|\mathbf{R}_{ij}|$, and the y coordinate represents the numbers of trust pairs and non-trust pairs in comparison. We notice that, when $|\mathbf{R}_{ij}|$ is less than 20, the number of non-trusted pairs exceeds that of trust pairs. On the contrary, when $|\mathbf{R}_{ij}|$ is greater than 20, the number of trust pairs is greater than that of non-trust pairs. This observation implies that the bigger $|\mathbf{R}_{ij}|$ becomes, the more likely trust formation becomes. Motivated by this observation, we use $|\mathbf{R}_{ij}|$ in estimating initial trust probabilities $p_{ij}^{(0)}$, which are summarized in Table 1.

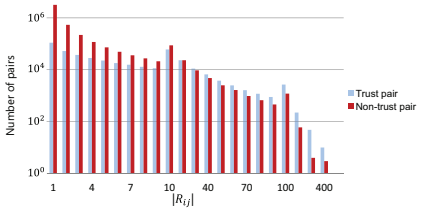

Figure 1: Distribution of trust pairs and non-trust pairs in comparison depending on the number of ratings $|\mathbf{R}_{ij}|$.

Table 1: Initial trust probabilities

$\lvert\mathbf{R}_{ij}\rvert$	1-2	3-5	6-9	10-24	25-40	41-100	101-300	301-500
$p_{ij}^{(0)}$	0.050	0.187	0.375	0.431	0.532	0.600	0.700	0.910

To construct *trust propagation strategies* used for the proposed model, a trust network needs to be established. We regard users as nodes and the existence of either review ratings between users or trust relations as nodes. The direction of an edge is determined by the direction from a review rater to a reiview writer or from a trustor to a trustee. Initially, each edge has a vector of probabilities, denoted by $\mathbf{p}_{ij}$, with the initial probability of u_i building trust to u_j, denoted by $p_{ij}^{(0)}$, and that of not building trust, $1 - p_{ij}^{(0)}$. When trust exists between two users, we assign to 1 due to the obvious trust relationship. When only review ratings are available, we set $p_{ij}^{(0)}$ as Table 1 suggests.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

We describe two primary trust propagation strategies in which the proposed new *message-passing mechanisms*, motivated from the idea of belief propagation [1], facilitate the integration of explicit trust & review ratings and the prediction of the probability of an edge. We define *a source link* to be an edge from which a trust-message originates and *a target link* to be an edge to which a trust-message is delivered. Thus, we estimate trust degrees of target links after applying the trust-message passing mechanisms. For the sake of simplicity, the status of a target link in the following figures is represented by p_{ij}.

As the first primary strategy, *direct propagation* is a trust propagation strategy in which, when u_i trusts u_k and then u_k trusts u_j, the probability that u_i trusts u_j increases by way of u_k. Figure 2(a) explains the direct propagation strategy. When both p_{ik} and p_{kj} are greater than 0.5, we increase the probability that u_i trusts u_j, p_{ij}, to reflect p_{kj}.

As the second primary strategy, when u_i trusts u_k and also u_j trusts u_k, *transpose trust* involves with increasing the probability p_{ij} that u_i trusts u_j by way of u_k. Figure 2(b) explains the strategy of transpose trust. When $p_{ik} > 0.5$ and $p_{jk} > 0.5$, p_{ij} increases by receiving trust-message from p_{jk}. When $p_{jk} \leq 0.5$, the trust relationship between u_j and u_k does not influence p_{ij}.

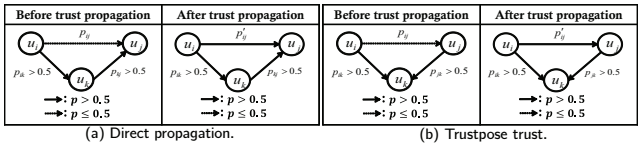

Figure 2: Propagation strategies.

The following equations (1) and (2) describe the computation of evolving trust degree through the trust-messages passing for the two primary strategies, respectively. The trust-messages passing, in effect, is realized by a propagation matrix ψ, shown in Table 2, connecting the status of a source link to that of a target link. Propagation matrix ψ can control the level of connectedness between the two links by parameter ϵ.

The equations (3) describe the computation of trust probabilities for the proposed models *ITD*. The initial trust probabilites are denoted by I and direct propagation and transpose trust are denoted by D and T, respectively. In the equations, n represents general constants to normalize the message vector because the sum of the elements should be equal to one. The parameter ϵ in the propagation matrix is preset through experiments.

Table 2: Initial trust probabilities

Target link	Source link	
	$p > 0.5$	$p \leq 0.5$
$p_{ij} > 0.5$	ϵ	$1-\epsilon$
$p_{ij} \leq 0.5$	$1-\epsilon$	ϵ

$$\mathbf{m}^{d}_{ik,kj\rightarrow ij} = \psi \times \mathbf{p}_{kj}\,, p_{ik} > 0.5 \text{ and } p_{kj} > 0.5 \quad (1)$$

$$\mathbf{m}^{t}_{ik,jk\rightarrow ij} = \psi \times \mathbf{p}_{jk}\,, p_{ik} > 0.5 \text{ and } p_{jk} > 0.5 \quad (2)$$

$$ITD: \mathbf{p}_{ij} = \alpha_I \mathbf{p}^{(0)}_{ij} + (\alpha_T \times n \prod \mathbf{m}^{t}_{ik,jk\rightarrow ij}) + (\alpha_D \times n \prod \mathbf{m}^{d}_{ik,kj\rightarrow ij}) \quad (3)$$

3. EXPERIMENTS

In our experiments, we used Epinions dataset[1] which has 131,828 users, 841,372 trust relations, and 13,668,319 review ratings. We demonstrated the excellence of *ITD* by the performance comparison of two previously proposed methods, *ABIT_L* [4] and *MoleTrust* [3]. We empirically found the parameter ϵ, 0.7 and 0.8 for the proposed strategies of direct propagation and transpose trust, respectively, in such a way that accuracy for the training data set is maximized.

To assess the performance of *ITD*, we basically followed [4]; We randomly selected 1,000 user pairs which are trust pairs and review-rating pairs at the same time, denoted by Answer set. As for the training, among pairs of users providing information on review ratings, we selected 2,000 pairs: 1,000 pairs in Answer set and random 1,000 pairs of users who have not built trust. Then, using each of the trust prediction models, we estimated trust probabilities for the selected 2,000 pairs.

In the training phase, the initial probabilities, provided in Table 1, were assigned for the 2,000 pairs. For the user pairs other than the users from the selected 2,000 pairs, $p^{(0)}_{ij} = 1$ is assigned if explicit trust formation exists. Otherwise, the initial probabilities as in Table 1 were assigned.

For each model, the predicted probabilities for the 2,000 pairs are sorted in descending order, and then trust probabilities of the top 1,000 pairs from the sorted list are selected. By comparing the estimated results of the top 1,000 pairs with the 1,000 pairs in Answer set, we computed the accuracy of the model, which is a ratio of the number of correct pairs to 1000. For *ITD*, *ABIT_L*, and *MoleTrust*, the above-stated testing was repeated five times, and the average was used for the model's accuracy.

Table 3 shows the accuracy of the trust prediction models as well as the used weights in *ITD*, α_I, α_T, and α_D, which maximized the accuracy. The accuracy of *ITD* is higher than that of *ABIT_L* and *MoleTrust* by 12.5% and 29.1%, respectively. The result is not surprising because, while the *ABIT_L* model relies on user interactions and the *MoleTrust* model depends on information on explicit trust, the proposed *ITD* model takes into account information on both user interactions and explicit trust.

Table 3: Accuracy of the proposed model in comparison with the previous models

Model (weights)	Accuracy
ITD ($\alpha_I = 0.2, \alpha_T = 0.7, \alpha_D = 0.1$)	**0.932**
ABIT_L	0.797
MoleTrust	0.609

4. CONCLUSIONS

This paper has proposed a new model for trust prediction. In experiments using real-life data, we have demonstrated that the proposed model *ITD* outperformes *ABIT_L* and *MoleTrust* in accuracy by 12.5% and 29.1%, respectively.

5. ACKNOWLEDGEMENTS

This research was supported by (1) Basic Science Research Program through NRF (No. 2012R1A1A2007817), (2) IT/SW Creative Research Program supervised by NIPA (NIPA-2012-H0503-12-1018), and (3) Convergence-ITRC Support Program supervised by NIPA (NIPA-2013-H0401-13-1001).

6. REFERENCES

[1] D. H. Chau, S. Pandit, and C. Faloutsos. Detecting Fraudulent Personalities in Networks of Online Auctioneers. *ECML/PKDD*, pages 103-114, 2006.

[2] R. V. Guha, R.Kumar, P.Raghavan, and A. Tomkins. Propagation of Trust and Distrust. *WWW*, pages 403-412, 2004.

[3] P. Massa and P. Avesani. Controversial Users Demand Local Trust Metrics: an Experimental Study on Epinions.com Community. *AAAI*, pages 121-126, 2005.

[4] V.-A. Nguyen, E. P. Lim, J. Jiang, and A. Sun. To Trust or Not To Trust? Predicting Online Trusts using Trust Antecedent Framework. *ICDM*, pages 896-901, 2009.

[1] http://www.Epinions.com