A Trust Model Stemmed from the Diffusion Theory for Opinion Evaluation

Hui Fang Jie Zhang[1] Nadia M. Thalmann[2]
[1]School of Computer Engineering [2]Institute for Media Innovation
Nanyang Technological University, Singapore
{hfang1}@e.ntu.edu.sg, {zhangj,nadiathalmann}@ntu.edu.sg

ABSTRACT

In open online communities, everyone can freely express opinions about other entities. As the quality of opinions may vary, it is important for users to evaluate opinions in order to determine how much to rely on. In this paper, we propose a novel trust model stemmed from the diffusion theory in social science (called DiffTrust), to evaluate the opinions of users (referred to as advisors) by modeling their trustworthiness. Specifically, an advisor's trust building among users is considered as a diffusion process. Her trustworthiness perceived by a specific user is influenced by four important factors including the advisor's characteristics directly observed by the user, susceptibility of the user, the contagious influence of other users already having a certain level of trust on the advisor, and the environment. DiffTrust also emphasizes on the dynamics of trust. Experimental results based on four real datasets verify the effectiveness of our model in comparison with state-of-the-art approaches.

Categories and Subject Descriptors

H.1.0 [**Information Systems**]: Models and Principles—*General*; I.2.11 [**Artificial Intelligent**]: Distributed Artificial Intelligence—*Intelligent agents*

Keywords

Trust Modeling; Diffusion Theory; Social Proximity; Intelligent Agent

1. INTRODUCTION

In large and open online communities, users may often encounter other entities which they have no previous experience with or prior knowledge of. In this case, they usually rely on the experience or knowledge of other users (advisors), to choose which entities to interact with. However, in these environments, advisors can freely express their opinions, and the quality of opinions may then vary. For example, some advisors may be dishonest and lie about their experience. They may report their experience with other entities as "completely satisfactory" while the real one is opposite. In addition, an advisor may only be capable of providing reliable opinions under one context but not under another because the advisor lacks knowledge in the second context. For instance, an expert advisor in automobiles may not necessarily be a reliable advisor in recommending doctors. Therefore, it is important for users to evaluate the quality of advisors' opinions in order to determine how much to rely on. One generally adopted approach in the area of multiagent systems [3,4,6,7,13,15,16] is to design intelligent software agents to help users model the trustworthiness of advisors. The basic idea is that a more trustworthy advisor to a user will provide more reliable opinions to the user.

Trust has been recognized as a diffusive concept [14]. When modeling trust, it is crucial to consider the processes through which trust is cultivated in a system. The diffusion theory [12] in social science seeks to explain *how*, *why* and *at what rate* a new *innovation* spreads through a community. It is thus natural to derive a trust model (called DiffTrust) from this well-studied theory by considering an advisor's trustworthiness as an innovation. Specifically, an advisor's trust building among users is considered as a diffusion process. Her trustworthiness perceived by a specific user is influenced by: the advisor's characteristics directly observed by the user, susceptibility of the user, the contagious influence of other users already having a certain level of trust on the advisor, and information about the environment (including both spatial and temporal information). With this model, we can well capture the *dynamics* and *subjectivity* of trust, and its *dependency* on other users and the environment.

Computationally, each user in the system is equipped with a software agent. Each agent first computes its user's direct trust on every other user (advisor) based on their shared interactions. A shared interaction here means that the user and the advisor have previously interacted with a same entity, such as providing a rating to the entity. The direct trust information will be used to form a social (trust) network for all users. Then when a user encounters an advisor's opinions, the user's agent will compute the trustworthiness of the advisor by incorporating the user's direct trust on the advisor in the social network, the user's susceptibility to trust (i.e. the initial trust assigned to the advisor by the user), and trust evaluation on the advisor computed by the agents of other users in the social network. The agent of the user considers another user's trust evaluation on the advisor by also computing the social proximity of that user with respect to its own user, which indicates how much the agent's own user can rely on that user's trust evaluation. In the model, the trustworthiness of an advisor in the view of each user is measured under a specific context, and we assume that there is a finite set of contexts in a specific system. We also take into account both temporal and spatial information in trust computation.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.*

We conduct experiments on real data obtained from eBay (www.ebay.com), FilmTrust (trust.mindswap.org), Epinions (www.epinions.com) and Flixster (www.flixster.com), to verify the effectiveness of our model in comparison with state-of-the-art approaches including TRAVOS [13], the personalized approach [16], CertProp [6] and Shin [7]. The results demonstrate that our model can more accurately model the trustworthiness of advisors than other approaches.

2. RELATED WORK

Different approaches have been proposed to model the trustworthiness of an advisor for a user, some of which are based on shared interactions between the user and the advisor. For example, the TRAVOS model [13] estimates the probability that the advisor's current opinions (ratings) are accurate based on the advisor's previous ratings provided to the user. However, these models are ineffective when there is no or only a very few shared interactions between the user and the advisor. To address this problem, the personalized approach [16] also estimates the public reputation of the advisor by comparing her ratings and other advisors' ratings regarding the same entities. Different from the personalized approach, our DiffTrust model makes use of other users' evaluation of the advisor's trustworthiness, and assesses those users' evaluation based on the social proximity of them with respect to the user, to avoid false evaluation.

The approach of [15] adopts the concept of *referral networks* where users cooperate with each other to find the trustworthiness of advisors by searching a social network. The social network is built upon two dimensions: expertise and sociability. However, this approach suffers from the "rumor problem" that different information about an advisor's trustworthiness received by a user may come from the same source. It also fails to work when new users do not have trust relationships with others, thus is not appropriate for the communities where users are loosely connected. In our model, a user only accepts other users' direct evaluation on an advisor to avoid the rumor problem. It also has no strict requirement on the minimum amount of experience for new users because various types of information (e.g. advisors' profile information) can also be utilized in our model.

Some other trust propagation methods [6, 7, 11] have also been proposed. Sabater and Sierra [11] present a one-level propagation method where a user can make use of other users' direct evaluation on an advisor. This method also considers the user's social relations with the other users, such as competition, cooperation and trade, which are typically difficult to obtain in real-world environments. On the basis of a social network built from users' direct evaluation on each other, Hang et al. [6] design a new algebraic approach called CertProp to propagate trust and address the rumor problem. They formally define three operators: aggregation, concatenation and selection for trust propagation. However, CertProp suffers from the *unreachable witness information* problem where a user may not know anything about another user (called witness) who holds information about an advisor. This problem has further been addressed in their recent work of the *Shin* approach [7] by considering the difference on trust evaluation towards same advisors between the user and the witness. However, both CertProp and Shin, grounded on the trust transitivity theory, are generally criticized by their accuracy, since trust evaluated through long paths endows with high probability to be inaccurate. To analyze further, those trust propagation methods only consider the user's own experience with the advisor if such experience exists, which may lead to inaccurate trust evaluation when the user has only limited experience with the advisor or the advisor dynamically changes her behavior. Conversely, our DiffTrust model flexibly adjusts the weight of the user's own experience and other users' evaluation on the advisor. The quality of the other users' evaluation on the advisor is judged through the concept of social proximity, which has been proven to be effective in the diffusion theory [1]. In addition, social proximity can also be modeled based on other types of viable information provided in different environments other than the buyer's own experience with the other users' evaluation.

3. THE DIFFTRUST MODEL

In this section, we describe in detail the rationale of deriving the DiffTrust model from the diffusion theory, and provide computational procedures for users to model trustworthiness of advisors for evaluating their opinions.

3.1 Trust and the Diffusion Theory

Our DiffTrust model is inspired by the diffusion theory [12] where the home territory is an innovation (e.g. technology, idea or object). The diffusion theory seeks to explain how, why and at what rate a new innovation spreads through a community. Rogers [10] indicates that "diffusion is the process where an *innovation* is communicated through *certain channels* over *time* among the members of a *social system*".

According to [14], trust is more appropriate to be considered as a diffusive concept. In trust modeling, it is critical and more meaningful to explore the context within which the trust is embedded, and to explore the processes through which the trust is cultivated. Besides, a user's trust perception towards another user is not static but dynamically evolves as the surrounding changes. That is, we can consider a user's trustworthiness perceived by other users as a systematic process, involving the purposes of evaluating and using trust of the user (why), the ways and channels of inducing trust towards the user (how), and the induced trust degree (at what rate) varying over time for other users or other contexts. We can clearly see the similarities between trust and the innovation in the diffusion theory. Both of them are *dynamic* and *evolutionary* (varying over time), *subject* to perceiving (adopting) users (i.e. influenced by the intrinsic nature of these users), and *dependent* on the environment. It is thus natural to derive a trust model based on the diffusion theory.

In a system, an advisor's trustworthiness in the view of a specific user, which may be different from the perception of other users, is sensitive to the interactions between the advisor and the user and also subject to the user's own intrinsic nature. Considering this, it is more appropriate to use the individual-level framework of the diffusion theory. We specifically employ Strang and Tuma's individual-level oriented heterogenous diffusion model [12]. This model emphasizes on the heterogeneous characteristics of both spatial and temporal information. Similarly, the diffusion of an advisor's trustworthiness among other users should also consider the spatial heterogeneity among the other users, because different users have different chances of perceiving the advisor's trustworthiness and different strength of affecting others' perception. We are also concerned with the temporal heterogeneity among all historical interactions where a more recent interaction will affect more on trust evaluation. In Strange

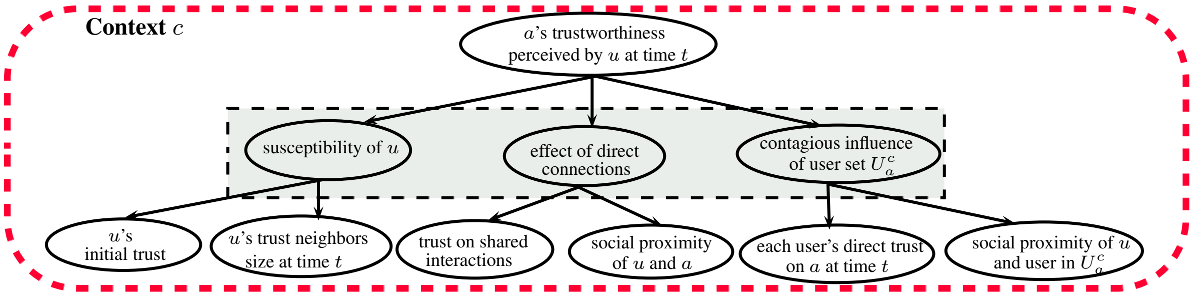


Figure 1: The DiffTrust Model for Evaluating the Trustworthiness of Advisors

and Tuma's model, there are mainly four factors influencing the innovation adoption process of a potential adopter: *the susceptibility of the potential adopter* (**individual tendency**, an adopter's intrinsic nature to adopt a new innovation), *the infectiousness (contagious) of early adopters* and *the proximity of the pairs consisting of one early adopter and the potential adopter* (**contagious influences**, the impact from the users already adopting the innovation), *the characteristics of the innovation* (**direct connections**, the potential adopter's perception about the characteristics of innovations through direct observation or communication), and *temporal variation* as a function of time since adoption. We will explain next how our DiffTrust model is derived based on these factors.

3.2 The Trust Model

In DiffTrust, each user is equipped with a software agent being responsible for modeling the trustworthiness of advisors for its user. Each user has a set of past interactions with some entities and provides a rating in the range of $[0, 1]$ for each interaction. We assume that a user u and an advisor a have both previously interacted with a set of the same entities. Based on these shared interactions, the agent computes user u's direct trust on advisor a. The direct trust information of every user will be used to form a social network of all users. Let us assume that the set of users who have adopted advisor a in their social networks under the context c is denoted by the set U_a^c. Our objective is to model the trustworthiness of advisor a perceived by user u at time t. The trustworthiness of the advisor is modeled with respect to each context c ($c \in \{c_1, c_2, \cdots, c_m\}$, where m is the number of contexts in the system), at time t as $\mathrm{T}(u, a, t, c)$.

We adopt the individual-level oriented heterogenous diffusion model [12] to derive the DiffTrust model based on its four factors. More specifically, as shown in Figure 1, the agent of user u models advisor a's trustworthiness by incorporating the following three major parts:

- *Susceptibility of user u trusting advisor a* (**individual tendency**). This refers to user u's intrinsic nature without any direct or indirect experience with advisor a. We consider two factors: user u's static initial trust ($T_0^u \in [0, 1]$), and the number of u's neighbors in the social network at a specific time. A larger number of neighbors implies that user u has already experienced a lot in the system. Therefore, she is more likely to be open to other unknown advisors and confident in her own trust evaluation.
- *Effect of direct connections between user u and advisor a* (**direct connections**), representing user u's direct observations on advisor a. Two factors are involved: direct trust based on shared interactions between u and a, and *social proximity* of a with respect to u, S_a^u, which is treated as a consequence of u's direct observation on a's spatial information.
- *Contagious influence of set U_a^c* (**contagious influences**). We consider the infectiousness of each user in U_a^c, by sharing their evaluation on advisor a derived from their shared interactions. We make use of their trust evaluations by computing the *social proximity* of each specific user in U_a^c with respect to user u. A socially closer user to u indicates a higher probability that u will trust the user's trust evaluation on a.

We also consider the *time decay effect* of shared interactions (temporal variation). The effect of previous shared interactions decreases as time goes by, and thus trust values computed based on long past evidence should also be decreased accordingly.

3.3 Trust Computation

Before introducing the specific computational steps, we first clarify some important concepts related to our model.

3.3.1 Related Concepts

Spatial information is used to model the *social proximity* between users. Social proximity, believed to be able to engender trust [1], can be flexibly modeled according to the consideration of identified spatial information. Two kinds of spatial information are considered in our model, *physically spatial information* and *socially spatial information*. The former refers to users' physical location and identity. It can be obtained from users' profile information. For this kind of information, the method for calculating context similarity in Section 3.3.3 can be used to model the social proximity. The latter refers to users' social identity and status, such as users's position in the social network and their neighbors. Traditional proximity metrics such as the number of common neighbors (CN) (see Equation 1), fraction of common neighbors, or Jaccard coefficient and the Adamic-Adar score can be adopted to measure the social proximity of two users [8]. Note that we prefer to use *socially spatial information* when both of the two kinds of spatial information are available, because social proximity derived from socially spatial information is proven to be more infusive than that from physically spatial information [1].

$$\mathrm{CN} = \frac{1}{2}(|\triangle| + |\triangle'|) \tag{1}$$

where for two users u and v: $\triangle = \Gamma_{\text{out}}(u) \bigcap \Gamma_{\text{in}}(v)$ and $\triangle' = \Gamma_{\text{in}}(u) \bigcap \Gamma_{\text{out}}(v)$. Here, $\Gamma_{\text{out}}(u)$ represents the set of out-neighbors of node u, which in our research corresponds to the set of users having direct trust on u. Similarly, $\Gamma_{\text{in}}(u)$ represents the set of users being directly trusted by u.

Temporal information is bounded with other information, such as trust values and past interactions between users, whose influence would be discounted over time. We mainly consider the time when an interaction happens, when a user is added to another user's social network, and when a trust value is computed.

Context consists of both spatial information and temporal information, according to the *Situation Dependency Theory* [5]. In our work, context is usually mentioned together with an interaction or an advisor's trustworthiness, *i.e.*, the context of a specific *interaction*, and advisor a's *trustworthiness* perceived by user u under a specific context.

3.3.2 Computational Steps

We now elaborate the computational steps in detail.

Step 1: Model direct trust based on shared interactions. As indicated in Section 3.2, user u models the direct trustworthiness of advisor a based on their shared interactions. This is a continuous process. Whenever a new shared interaction happens between u and a, u's trust on a needs to be updated based on the new observation. We assume that based on the previous shared interactions of user u and advisor a, at time t_0, u's direct trust towards a under context c is denoted as $\mathrm{DT}(u, a, t_0, c)$. At time t, u and a have a new shared interaction with a same entity, denoted as i_u under the context c_u and i_a under the context c_a respectively. Note that contexts c_u, c_a and c may not be the same. The problem can be identified as to update u's direct trust towards a at time t as $\mathrm{DT}(u, a, t, c)$ by considering the effect of the new shared interaction i_u and i_a.

We first model the direct influence ($\mathrm{DI}(i_u, i_a, t)$) of the new shared interaction on the trustworthiness of advisor a. Intuitively, in the shared interaction, if advisor a's opinion is more similar to that of user u towards the same entity, a can be considered as more trustworthy. Thus, $\mathrm{DI}(i_u, i_a, t)$ is formulated as the similarity between i_u and i_a:

$$\mathrm{DI}(i_u, i_a, t) = (1 - |i_u - i_a|) \times \mathrm{S}(c_u, c_a) \times \mathrm{S}(c, c_u) \qquad (2)$$

where $\mathrm{S}(c, c_u)$ is the similarity between contexts c and c_u, and $\mathrm{S}(c_u, c_a)$ is the similarity between contexts c_u and c_a. Calculation of context similarity will be specified in Section 3.3.3. In this way, we align the user and advisor's interactions with the same entity to context c.

We then update user u's direct trust towards advisor a based on the new shared interaction at time t under context c, by combining the new shared interaction (i_u and i_a) with the previous trust values $\mathrm{DT}(u, a, t_0, c)$ at t_0, as follows:

$$\mathrm{DT}(u, a, t, c) = \frac{\mathrm{DT}(u, a, t_0, c)\lambda^{(t-t_0)} + \mathrm{DI}(i_u, i_a, t)}{1 + \lambda^{(t-t_0)}} \qquad (3)$$

where $0 < \lambda \leq 1$ is a time decay factor for user u to decrease the effect of old shared interactions between u and a.

Note that two special scenarios need to be considered: 1) if there is no new shared interaction after t_0, we consider the time decay effect of previously computed trust value at t_0 such that $\mathrm{DT}(u, a, t, c) = \mathrm{DT}(u, a, t_0, c)\lambda^{(t-t_0)}$; and 2) if there is no shared interaction between u and a till time t, the direct connections between u and a can be modeled by the social proximity of u and a instead, which could be considered as user u's direct observation of advisor a's characteristics, that is $\mathrm{DT}(u, a, t, c) = S_a^u$.

Step 2: Model contagious influence. According to the diffusion theory, the contagious influence of users in U_a^c, denoted as $\mathrm{CI}(u, U_a^c, t)$, can be modeled as the weighted average of trust evaluation on advisor a from each user $x \in U_a^c$, denoted as $\mathrm{DT}(x, a, t, c)$. Each weight corresponds to user x's *social proximity* with respect to user u denoted as S_x^u. The way of computing social proximity can vary depending on information available in the environment. In our experiments, we use Equation 1 to compute social proximity. Higher social proximity means that two users are more similar/closer according to their spatial information, and user x's evaluation on advisor a is pitched to resonate more in the mind of user u. The contagious influence of users U_a^c can then be formalized as follows:

$$\mathrm{CI}(u, U_a^c, t) = \frac{\sum_{x \in U_a^c} \mathrm{DT}(x, a, t, c) \times S_x^u}{\sum_{x \in U_a^c} S_x^u} \qquad (4)$$

Step 3: Combine the three factors. We now evaluate the trustworthiness of advisor a with regard to user u at time t under context c, $\mathrm{T}(u, a, t, c)$, by considering the three factors: direct connections, contagious influence and susceptibility (initial trust) of u, as follows:

$$\begin{aligned}\mathrm{T}(u, a, t, c) = \omega_1 \mathrm{DT}(u, a, t, c) + \omega_2 \mathrm{CI}(u, U_a^c, t) \\ + (1 - \omega_1 - \omega_2) T_0^u\end{aligned} \qquad (5)$$

where ω_2 is the weight of the contagious influence of the set of users in U_a^c and can be modeled as the average social proximity of users in U_a^c with respect to user u under the context c, and ω_1 is the weight of the direct trust derived from shared interactions. The weight ω_1 is influenced by user u's confidence in the system. It relates to the number of user u's neighbors. The larger size of u's neighbors implies that u is more experienced and more confident to rely on her own evaluations. We calculate ω_1 as follows:

$$\omega_1 = \begin{cases} \frac{N_u^t}{N_{min}} & \text{if } N_u^t \leq N_{min}; \\ 1 & \text{otherwise.} \end{cases} \qquad (6)$$

where N_u^t is the number of u's trust neighbors in the social network at time t, and N_{min} is the minimum number of neighbors needed for user u to be totally confident about her own evaluation. We adopt the method in [16] to compute the value of N_{min} according to an acceptable level of error ε (for u) and a confidence measure γ, as follows:

$$N_{min} = -\frac{1}{2\varepsilon^2} \ln \frac{1-\gamma}{2} \qquad (7)$$

Note that $\omega_1 = 1$ is only used in direct trust part, and in order to more accurately model the advisor's trustworthiness, the agent will always consider direct trust and contagious influence together. That is to say, in Equation 5, if $\omega_1 + \omega_2 > 1$, then we change the weights to be: $\omega_1 = \frac{\omega_1}{\omega_1 + \omega_2}$ and $\omega_2 = 1 - \omega_1$.

When there is no shared interaction between user u and advisor a, a will not be included into u's social network. If a is a newcomer of the system, we consider only the social proximity between u and a and u's initial trust (without contagious influence). An advisor will be added into user u's social network only when shared interactions between the advisor and the user are identified. However, if an advisor's trustworthiness from the view of user u based on shared interactions is equivalent to the initial trust T_0^u, the advisor will be immediately excluded from u's social network. Intuitively, it represents that user u's previous shared interactions with the advisor have been forgotten and thus lost influence on trust evaluation on the advisor. Then, only new shared interactions between the advisor and the user will be taken into consideration in the trust computation.

Table 1: Statistical Information about the Four Real Datasets

Datasets	eBay	FilmTrust	Flixster	Epinions
Statistics	186,957 buyers 3,046 sellers 214,115 transactions	1,508 users 2,071 items 35,497 ratings	1,000 users 2,867 items 7,905 ratings	500 users 44,288 items 66,287 ratings
Data types	buyers→sellers	users→items users→users	users→items users→users	users→items users→users
Vertices in social network	5,531 (buyers)	1,508 (users)	1,000 (users)	500 (users)
Edges in social network	1,144,675	1,632	3,337	23,848
Rating scale	-1, 0, 1	0.5, 1, 1.5,···, 4	0.5, 1, 1.5, ···, 5	1, 2, 3, 4, 5
Trust between users	-	1	1	1
Ratio of having shared interactions of all users	7.5%	81.9%	1.5%	21.8%
Maximal number of shared interactions between users	21	105	444	121

3.3.3 Context Similarity

The context information in our work is represented by an ontology, and we employ the existing largest ontology - the LinkedData ontology [2] to present the concepts involved for describing the contexts in a specific system. We assume two contexts c_i and $c_j \in \{c_1, c_2, \cdots, c_m\}$ represented by a set of concepts (objects) in our universe, and relations between the concepts are generalization relationships (e.g. isA and $partOf$). Contexts $c_i = \langle c_i^1, c_i^2, \cdots, c_i^n \rangle$ and $c_j = \langle c_j^1, c_j^2, \cdots, c_j^n \rangle$, where n is the number of concepts describing a context. We adopt the definition of concept similarity in [9] that two concepts (e.g. apple and pear) are similar if they relate to a third concept (e.g. fruit). More formally, the distance between two concepts is defined as the length of the shortest generalization path between the two concepts. For example, the distance of apple and pear equals to 1 ($apple \overset{isA}{\to}$ **fruit**, and $pear \overset{isA}{\to} fruit$), while the distance between apple and rice is 2 ($apple \overset{isA}{\to} fruit \overset{isA}{\to}$ **plant**, and $rice \overset{isA}{\to} vegetable \overset{isA}{\to}$ **plant**). We then define the similarity of two contexts c_i and c_j as:

$$\mathrm{S}(c_i, c_j) = \frac{1}{\sum_{y=1}^{n} \mathrm{dist}(c_i^y, c_j^y) + 1} \tag{8}$$

where $\mathrm{dist}(c_i^y, c_j^y)$ is the distance between concepts c_i^y and c_j^y defined previously using the shortest generalization path. Note that social proximity from physically spatial information can also be modeled in this way because physically spatial information can also be described by the ontology.

4. EMPIRICAL EVALUATION

We carry out two kinds of experiments to evaluate the performance of our DiffTrust model. The first kind is to verify whether the trustworthiness of advisors modeled by DiffTrust can be used to accurately model the trustworthiness of entities in reputation systems. These experiments are conducted on real data collected from eBay, modeling trustworthiness of sellers for buyers to predict the outcomes of their future transactions with the sellers. The second kind is to verify whether the modeled trustworthiness of advisors is the same as what have been explicitly indicated by users. We conduct these experiments using real datasets extracted from FilmTrust, Epinions and Flixster. We also compare our model with several competing models, including TRAVOS [13], the personalized approach [16], CertProp [6], Shin [7], and a baseline approach which judges a seller's trustworthiness for a buyer by aggregating the ratings from advisors without considering their trustworthiness.

4.1 Data Acquisition

For eBay, we first randomly select $26,922$ sellers selling products in different categories and collect all their past transactions (including $969,213$ positive ratings, 1914 neutral ratings and 3590 negative ratings). The data is crawled from April 10, 2000 to June 4, 2011. Then, we randomly select $23,630$ buyers from all the buyers who have previously interacted with at least one of the selected sellers. We collect all past transactions of these buyers. Each transaction consists of buyer ID, seller ID, rating provided by the buyer, and time of the transaction. We then particularly select $3,046$ target sellers who have at least one unsuccessful (negative or neutral) past transaction to fully test the performance of different models. Correspondingly, the number of buyers in the buyer set is $5,531$. We then predict the outcomes of past transactions that buyers have conducted with sellers using the leave-one-out strategy. More specifically, to predict the outcome of a transaction between a buyer b and a seller s at time t, we select ratings provided by other buyers (advisors) towards seller s before time t. Trustworthiness of advisors is modeled using different approaches based on all ratings provided before time t. The trustworthiness of seller s from the view of buyer b can then be evaluated by aggregating ratings from advisors weighted by their trustworthiness. In the end, the outcome of the unknown transaction can be predicted using the computed trustworthiness of seller s.

FilmTrust dataset consists of $1,508$ users, $2,071$ movies and $35,497$ movie ratings issued by the users. There are $1,632$ trust relationships explicitly identified by users, which are used as test data in our experiments. These trust relationship are directed. That is, a user a trusting another user b does not imply b also trusting a. On the basis of shared interactions (commonly rated items) between users, we model the trust value between two users using different models. The same evaluation method has been used on the Flixster and Epinions datasets. We randomly select $1,000$ users from the Flixster dataset[1], and 500 users from the Epinions dataset[2]. The statistical information of these four datasets is summarized in Table 1.

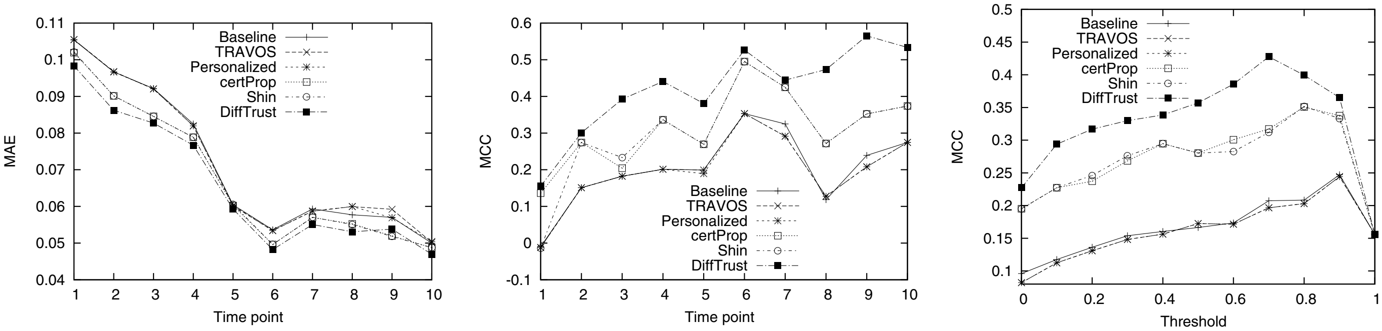
4.2 Evaluation Metrics

We use three metrics to measure the performance of different approaches. One is the Matthews Correlation Coefficient (MCC) for the eBay dataset. MCC is a measure for the quality of binary classifications. It is generally regarded

[1] http://www.cs.sfu.ca/~sja25/personal/datasets/.
[2] http://www.trustlet.org/datasets/downloaded_epinions.

Table 2: Performance Comparison on the eBay Dataset

Dataset	All Data		Cold Start Buyers				Sellers Non-consistently Perform			
			$N_e \leq 2$		$N_e \leq 10$		$R_n \geq 0.01$		$R_n \geq 0.1$	
	MCC	**MAE**	**MCC**	**MAE**	**MCC**	**MAE**	**MCC**	**MAE**	**MCC**	**MAE**
DiffTrust	**0.327**	**0.0648**	**0.007734**	**0.0769**	**0.1601**	**0.0699**	**0.3431**	**0.0890**	**0.5479**	**0.3010**
Baseline	0.166	0.0708	-0.001773	0.0819	0.0820	0.0738	0.1722	0.0971	0.3229	0.3664
TRAVOS	0.156	0.0710	-0.002861	0.0817	0.0093	0.0784	0.1630	0.0974	0.3121	0.3679
Personalized	0.161	0.0710	-0.002861	0.0817	0.0093	0.0785	0.1684	0.0973	0.3270	0.3656
CertProp	0.270	0.0673	-0.002453	0.0812	0.0574	0.0735	0.2801	0.0922	0.4932	0.3225
Shin	0.269	0.0670	-0.002453	0.0813	0.0222	0.0732	0.2790	0.0922	0.4932	0.3225



(a) **(b)** **(c)**

Figure 2: Performance Comparison on the eBay Dataset by: (a) Varying the Time Slots (MAE); (b) Varying the Time Slots (MCC); (c) Varying the Trust Threshold (MCC)

as a balanced measure which can be used even if the classes have distinct sizes. Thus, it is very suitable for our eBay dataset, where over 95% of the historical transactions are positive. MCC can be calculated as follows:

$$MCC = \frac{t_p t_n - f_p f_n}{\sqrt{(t_p + f_p)(t_p + f_n)(t_n + f_p)(t_n + f_n)}} \qquad (9)$$

where t_p t_n, f_p and f_n refer to an actual positive transaction predicted to be successful, an actual negative transaction predicted to be unsuccessful, an actual negative transaction predicted to be successful, and an actual positive transaction predicted to be unsuccessful, respectively. A MCC value of 1 represents a perfect prediction, 0 an average random prediction and -1 the worst possible prediction. For the trust value prediction in FilmTrust, Epinions and Flixster, there is only information about the relationship that a user trusts an advisor. Precision is thus used as the evaluation metric, which refers to the actual fraction of successful prediction of those trust relationships. More specifically, given the fact that a user trusts an advisor as explicitly specified by the user, if the modeled trustworthiness of the advisor is greater than a threshold, its trust value is successfully predicted to be 1. Otherwise, this is an unsuccessful prediction. For all four datasets, we also use the third metric the mean absolute error (MAE) between predicted rating of each transaction (or predicted trust value for each user pair in FilmTrust, Epinions and Flixster) and the real rating of the transaction (or the actual trust relationship).

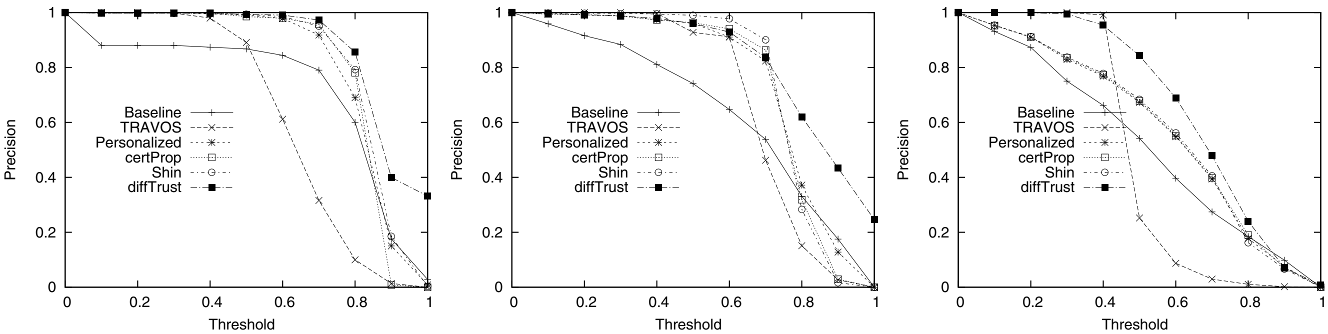
4.3 Results and Discussion

Here, we present the performance of our model and the competing approaches on the four datasets, respectively. We also examine these approaches in detail by varying the time when a transaction happened and trust threshold for predicting whether a transaction is successful on the eBay dataset, and the trust threshold for predicting whether a user indicated trust on another user (advisor) on the other three datasets. Table 2 and Figure 2 summarize the performance comparison of DiffTrust with other approaches in terms of MCC and MAE on eBay data in different scenarios. Table 3 and Figure 3 present the results of Precision and MAE on the FilmTrust, Epinions and Flixster datasets.

More specifically, Table 2 shows the experimental results on 1) the whole eBay dataset with $5,331$ buyers and $3,046$ sellers; 2) the dataset considering only cold start buyers where $N_e \leq 2$ and $N_e \leq 10$ respectively, and N_e refers to the number of other users (advisors) with whom a buyer has shared interactions with; and 3) the dataset considering only the sellers who do not perform consistently well and $R_n \geq 0.01$ and $R_n \geq 0.1$ for sellers respectively, where R_n refers to the ratio of sellers' transactions received negative ratings from buyers. From all these results, we can see that our DiffTrust model achieves more consistent and better performance both in MAE and MCC than other approaches. To be more specific, the performance of TRAVOS and the personalized approach is very close to the baseline, mainly because there are only a very few shared interactions between buyers, which are needed for TRAVOS to work well. Besides, the dataset of eBay is extremely skewed and almost over 99% of the transaction outcomes are positive. Thus, the personalized approach, relying on the majority of other users' ratings to evaluate advisors' trustworthiness, is more likely to underestimate the trustworthiness of advisors who provide negative ratings. In consequence, it is difficult for the personalized approach to predict the unsuccessful transactions. In our model, we adopt other buyers' direct trust evaluations on an advisor derived from their shared interactions with the advisor by considering these buyers' social proximity with respect to the current buyer. In so doing, we can assure a more accurate trust evaluation of the advisor because we are more likely to receive different trust evalu-

Table 3: Performance Comparison on the FilmTrust, Epinions and Flixster Datasets

Dataset	FilmTrust		Epinions		Flixster	
	Precision	MAE	Precision	MAE	Precision	MAE
DiffTrust	**0.867**	**0.1154**	**0.8163**	**0.2064**	**0.6616**	**0.3229**
Baseline	0.711	0.2136	0.6365	0.3428	0.5193	0.4757
TRAVOS	0.682	0.3008	0.6796	0.2943	0.4884	0.464
Personalized	0.793	0.1798	0.7406	0.2266	0.5754	0.414
CertProp	0.791	0.1854	0.7327	0.2417	0.5781	0.416
Shin	0.809	0.1576	0.7413	0.2314	0.5783	0.4123



(a) (b) (c)

Figure 3: Performance Comparison by Varying the Trust Threshold: (a) FilmTrust; (b) Epinions; (c) Flixster

ations on the advisor, especially on the advisor having ever provided negative ratings. Trust propagation approaches of Shin and CertProp are able to achieve consistently much better results than the baseline approach in most scenarios, but worse than our DiffTrust model. This is mainly because CertProp only considers the reachable buyers' direct trust on an advisor, but overlooks the unreachable buyers' (see Section 2 for details). This may result in the loss of some important trust evaluations. Shin considers trust propagation for reachable witnesses (buyers's direct trust evaluations on the advisor), and use similarity of the buyer and unreachable witnesses' direct trust evaluations on the same entities to model those unreachable witnesses. This setting is not very reasonable and fair for the eBay dataset. The weight of considering other buyers' direct trust evaluations on the advisor is more likely to approach 1 since most of the transactions on eBay are positive, but the weight for unreachable witnesses is more likely to be less than 1 according to the similarity metric. In contrast, in DiffTrust, we equally treat other buyers' direct trust evaluations towards the advisor by considering the social proximity. This can avoid the above mentioned "unfair" problem. It is worth noting that in the cold start case, only our model can obtain a positive MCC value and relatively smaller MAE value when the number of other buyers with whom a buyer has shared interactions is very sparse. In this scenario, the performance of other approaches closely approaches that of the baseline approach. This demonstrates that our model can effectively address the cold start problem, assist new buyers to accurately evaluate advisors' trustworthiness, and thus assesses the quality of advisors' opinions to make correct decisions for buyers.

Figure 2 illustrates the performance of different approaches by varying the time when each predicted transaction was conducted, and varying the trust threshold so that if predicted trustworthiness of a seller at specific time is higher than the threshold, the seller's transaction conducted at that time will be predicted to be successful. As shown in Figures 2(a) and 2(b), we divide the time range of April 10, 2000 to June 4, 2011 into 10 time slots of equal range. We can see that the performance of all the approaches improves gradually in terms of both the MAE and MCC as time goes by. This is mainly because users gain more experience and can better model the trustworthiness of advisors. As can be seen in Figure 2(b), our model performs consistently better than other approaches, and the performance gap between our model and each other approach increases gradually as buyers gain more experience in the system. Figure 2(c) demonstrates that the best trust threshold for our model is 0.7, while 0.8 for Shin and CertProp and 0.9 for Baseline, TRAVOS and the personalized approach, respectively. Our model outperforms the other approaches for all thresholds.

Table 3 depicts the results of Precision and MAE on the FilmTrust, Epinions and Flixster datasets. We can see that our model performs consistently better than other approaches. As the ratio of shared interactions between users increases (FilmTrust > Epinions > Flixster), the performance of all the approaches also improves. The performance of TRAVOS is very close to Baseline and sometimes worse than Baseline. TRAVOS works only for binary ratings. Thus, in these datasets, we first map the ratings to either positive or negative by choosing the middle of the rating scale in our experiments, which may lead to inaccurate modeling of advisors' trustworthiness according to shared interactions. For the Flixster dataset, as there are a very limited number of shared interactions between users (0.015 in Table 1), the performance gap between our DiffTrust model and other approaches regarding to both Precision and MAE is relatively larger than that on FilmTrust and Epinions. This is mainly because in Flixster, of all the users having shared interactions, the average number of shared interactions is relatively larger (i.e. 3.2774). That is to say, if there are any shared interactions between a user and an advisor, the user can make an accurate trust evaluation on the advisor. Through social proximity, our model can maximally make use of each individual user's trust evaluation on the advisor. This demonstrates that our model, compared to other approaches, are more suitable to address the trustworthiness of advisors in the systems where users have a few shared interactions (the same trend found for eBay). It is worth noting that on these datasets, the performance of the personalized approach is much better than Baseline and TRAVOS (compared to that on the eBay dataset), implying that even in the environments where users are having a relatively larger number of commonly rated items, the public information is still worth considering when addressing the trustworthiness of advisors.

Figure 3 presents the performance of different approaches by varying the trust threshold so that if a predict trust value of advisor a in the view of a user u is larger than that threshold, user u trusts advisor a. It shows that in general our DiffTrust model outperforms the other approaches. It consistently achieves high precision, demonstrating its effectiveness in modeling the trustworthiness of advisors.

5. CONCLUSION AND FUTURE WORK

Aiming at evaluating the quality of opinions of users (advisors) in open online environments such as e-marketplaces, we propose a novel trust model called DiffTrust, stemmed from the diffusion theory in social science, to model the trustworthiness of users. Specifically, an advisor's trust building among users is considered as a diffusion process. Her trustworthiness perceived by a specific user is influenced by four important factors including the advisor's characteristics directly observed by the user, susceptibility of the user, the contagious influence of other users already having a certain level of trust on the advisor, and the environment. Our DiffTrust model emphasizes on the dynamics and evolutionary characteristics of trust, and highlights that the trustworthiness of an advisor may be perceived differently by different users, dependent on the environment, and embedded with a specific context. We compare our model with a baseline approach, TRAVOS, the personalized approach, CertProp and Shin, on four real datasets of eBay, FilmTrust, Epinions and Flixster. Experimental results demonstrate that DiffTrust can consistently perform better in both loosely-connected and well-connected environments. Besides, it can also help model the trustworthiness of advisors for users who are new to the system.

Our current work focuses on modeling the trustworthiness of advisors. For future work, we will derive a trust model also from the diffusion theory to model the trustworthiness of entities to which the advisors provide opinions. Together, we will offer a unified trust model for users to make informed decisions about which entities to interact with. In addition, due to limitations of the obtained datasets, we did not consider context information and physically spatial information in our experiments. In the future, we will continue to fully demonstrate the effectiveness of our DiffTrust model by exploring more relevant information, and design an alternative method to compute context similarity without ontologies.

6. ACKNOWLEDGEMENT

This work has been made possible thank to the Institute for Media Innovation at Nanyang Technological University who has given a scholarship to the first author, and supported by the MOE AcRF Tier 1 Grant (M4010265 RG15/10) awarded to Dr. Jie Zhang.

7. REFERENCES

[1] C. M. Angst, R. Agarwal, V. Sambamurthy, and K. Kelley. Social contagion and information technology diffusion: The adoption of electronic medical records in U.S. hospitals. *Management Science*, 56(8):1219–1241, 2010.

[2] C. Bizer, T. Heath, and T. Berners-Lee. Linked data - the story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.

[3] H. Fang, J. Zhang, M. Sensoy, and N. Thalmann. A generalized stereotypical trust model. In *Proceedings of 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 698–705, 2012.

[4] H. Fang, J. Zhang, M. Sensoy, and N. M. Thalmann. SARC: Subjectivity alignment for reputation computation. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1365–1366, 2012.

[5] S. Figge. Situation-dependent services - a challenge for mobile network operators. *Journal of Business Research*, 57(12):1416–1422, 2004.

[6] C.-W. Hang, Y. Wang, and M. P. Singh. Operators for propagating trust and their evaluation in social networks. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1025–1032, 2009.

[7] C.-W. Hang, Z. Zhang, and M. Singh. Generalized trust propagation with limited evidence. *IEEE Computer*, 2012. doi:10.1109/MC.2012.116.

[8] K. Lerman, S. Intagorn, J.-H. Kang, and R. Ghosh. Using proximity to predict activity in social networks. In *Proceedings of the 21st International Conference Companion on World Wide Web*, pages 555–556, 2012.

[9] K. Radinsky, S. Davidovich, and S. Markovitch. Learning causality for news events prediction. In *Proceedings of the 21st international conference on World Wide Web*, pages 909–918, 2012.

[10] E. M. Rogers, editor. *Diffusion of Innovations*, page 5. Simon and Schuster, 4th edition, 1995.

[11] J. Sabater and C. Sierra. Reputation and social network analysis in multi-agent systems. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 475–482, 2002.

[12] D. Strang and N. B. Tuma. Spatial and temporal heterogeneity in diffusion. *American Journal of Sociology*, 99(3):614–639, 1993.

[13] 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(2):183–198, 2006.

[14] O. E. Williamson. Calculativeness, trust, and economic organization. *Journal of Law and Economics*, 36(1):453–486, 1993.

[15] B. Yu and M. P. Singh. Searching social networks. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 65–72, 2003.

[16] J. Zhang and R. Cohen. A comprehensive approach for sharing semantic web trust ratings. *Computational Intelligence*, 23(3):302–319, 2007.