

Research Article

Recommendation Based on Trust Diffusion Model

Jinfeng Yuan and Li Li

Faculty of Computer and Information Science, Southwest University, Chongqing 400715, China

Correspondence should be addressed to Li Li; lily@swu.edu.cn

Received 4 March 2014; Revised 7 May 2014; Accepted 7 May 2014; Published 9 June 2014

Academic Editor: Yolanda Blanco Fernandez

Copyright © 2014 J. Yuan and L. Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recommender system is emerging as a powerful and popular tool for online information relevant to a given user. The traditional recommendation system suffers from the cold start problem and the data sparsity problem. Many methods have been proposed to solve these problems, but few can achieve satisfactory efficiency. In this paper, we present a method which combines the trust diffusion (DiffTrust) algorithm and the probabilistic matrix factorization (PMF). DiffTrust is first used to study the possible diffusions of trust between various users. It is able to make use of the implicit relationship of the trust network, thus alleviating the data sparsity problem. The probabilistic matrix factorization (PMF) is then employed to combine the users' tastes with their trusted friends' interests. We evaluate the algorithm on Flixster, Moviedata, and Epinions datasets, respectively. The experimental results show that the recommendation based on our proposed DiffTrust + PMF model achieves high performance in terms of the root mean square error (RMSE), Recall, and F Measure.

1. Introduction

With the rapidly growing amount of information available on the WWW, recommender systems become a popular way to help users select relevant information on the Internet. Personalized recommendation adopts knowledge discovery techniques such as data mining and machine learning to discover user interests according to user behavior and then to make recommendations [1, 2]. Typically, collaborative filtering (CF) is the most successful and widely used recommendation technique [3, 4]. CF makes recommendation according to the assumption that users who have the similar performances would like to choose the similar items. Despite its popularity and success, the performance of CF is significantly limited by the “data sparsity” and “cold start” [3, 5].

In view of these limitations, many scholars have recently integrated trust relationship among users into the recommendation system [6–10]. Trust-based recommender systems utilize a social network augmented with trust ratings, known as a trust network, to generate recommendations for users based on people they trust. These systems can deal with the trust relations between users, which cannot be well handled in traditional CF-based recommendations, to support the recommendation process. However, these trust-based methods work to some extent and further have several other

inherent weaknesses. Firstly, the sparsity of trust network. In recommender systems, the number of users is often very large, but in the process of real world recommendations, the number of direct transactions between users is small. So the number of direct trust relationships established by the limited number of transactions is also very little, which leads to the fact that the direct trust relationships can only play a small role in the process of the recommendation. Many trust-based recommendation algorithms only obtain trust values from preexisting social links between users [6–8] and do not consider indirect trust relationships. Secondly, the dynamics of the trust network. In the trust network, the trust relationships between users are not static, but dynamically change over time or other events. The change of the trust relationships may result in change of recommendation results. Therefore, it is more feasible to use dynamic trust network [8, 10, 11]. Thirdly, how to integrate the trust model with the recommendation system. Most recommendation algorithms [12–14] are based on the traditional probabilistic matrix factorization model and fusion between user-item matrix and social relationships by sharing a potential low dimensional user characteristic matrix. These methods can only learn few effective characteristics. This disadvantage not only causes lack of interpretability in the model, but also affects the quality of the recommendation.

Aiming at effectively overcoming the above limitations and modeling recommender systems more accurately this paper presents a new recommendation method based on trust diffusion mechanism (DiffTrust + PMF). The contributions of this paper are as follows. (1) According to the diffusion theory in economics, we improve the trust diffusion model (DiffTrust) suitable for collaborative filtering recommendation systems. DiffTrust makes full use of the direct trust relationship between users, to derive the indirect trust relationship between users through a certain trust propagation rules. (2) We employ the probabilistic matrix factor (PMF) model to put DiffTrust into the process of the recommendation. DiffTrust + PMF has the following characteristics. On one hand, this method matches more trust users for the current user through DiffTrust model and fully digs up more new trust relationships between users and is used for recommendation service, which perfectly solves the data sparsity of the trust network. On the other hand, DiffTrust + PMF combines the users' tastes with their trusted friends' interests through a set of parameters. This consideration can truly reflect the recommendation process in real life. (3) Through experiments on three data sets—Flixster, Moviedata, and Epinions datasets—we found that the proposed DiffTrust + PMF method can improve performance much more than other existing algorithms. In other words, the combination of DiffTrust and probabilistic factor model for recommendation outperforms all others in both the trust-based recommenders and based-PMF (probabilistic matrix factorization) recommenders.

The remainder of the paper is organized as follows. Section 2 discusses our method using DiffTrust + PMF. Section 3 evaluates the procedures experimentally followed by further discussion of relevant issues associated with models used in the paper. The related work is reviewed in Section 4. Finally, Section 5 concludes the paper.

2. New Method for Recommendation

In this section, we first describe recommendation problem based on trust network in Section 2.1 and then provide the solution in Sections 2.2 and 2.3.

2.1. Problem Description. The problem we study in this paper is different from traditional recommender systems, because the latter normally only utilize the information of the user-item rating matrix. In this paper, we will also incorporate the social trust relations among users to improve recommender systems. Figure 1(a) shows the social process of recommendation in the real world. The users are connected with edges, and each edge is associated with a weight in the range $[0, 1]$ to indicate how much user u_i knows or trusts user u_k ; zero means no trust and one means full trust. Each user also rated some items on a 5-point integer scale to express the extent of the favor of each item. The source user u wants a prediction on the target item v . The recommendation scenario includes two central elements: the friends network and the favors of these friends, which can essentially be illustrated by the examples

of social network graph in Figure 1(a) and user-item matrix in Figure 1(b), respectively.

In the real word, we consider an $m \times n$ rating matrix R describing m users' numerical ratings on n items, where r_{ij} represents the rating of user i for item j . In most online systems, r_{ij} is the K-point integer (normally, 1, 2, 3, 4, and 5 represent "very bad," "bad," "neutral," "good," and "very good," resp.).

We define a directed trust-based graph $G = (U, E)$ among users, where the vertex set $U = \{u_i\}_{i=1}^m$ and edge set E correspond to all the users in a social trust network and the trust relations between users, respectively. For a pair of nodes, u_i and u_k , the weight associated with an edge from u_i to u_k is denoted as $t_{ik} \in [0, 1]$, which represents how much a user u_i trusts. Let $T = \{t_{ik}\}$ denote the $m \times m$ trust matrix of G . Note that matrix T is an asymmetric matrix, since in a trust-based social network, user u_i trusting u_k does not necessary indicate user u_k trusts u_i .

The goal we study in this paper is how to effectively and efficiently predict the missing values of the user-item matrix by employing the ratings expressed by trusted friends. The most existing trust-based methods [6–8] only use explicit trust information explicitly indicated by users to generate recommendations, while implicit trust relationship is ignored. We employ diffusion trust model to explore the processes which the trust has cultivated, which can infer the implicit relationship of the trust network. The detailed process will be introduced in the next section.

2.2. Novel Trust Diffusion Model. In order to present the concepts involved for predicting the trustworthiness, we improve the existing DiffTrust method [15] inspired by the individual-level framework of the diffusion theory [16]. According to the diffusion theory in social science, an advisor's trust building among users is considered as a diffusion 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. Our improved trust diffusion model emphasizes the dynamics and evolutionary characteristics of trust. That is, the trustworthiness of an advisor may be perceived differently by different users, which is dependent on the environment and embedded with a specific context. Figure 2 shows the DiffTrust model.

As shown in Figure 2, the DiffTrust models the trustworthiness of users by incorporating three major parts: *intrinsic tendencies*, which refers to user u_k 's intrinsic nature without any direct or indirect experience with user u_i , which includes user u_k 's static initial trust at time t ($T_k^0 \in [0, 1]$), and the number of u_k 's trust neighbors in the social network at a specific time. *Direct connections*, represents u_i 's direct influence on u_k . Two factors are considered: direct trust based on shared interactions between u_i and u_k and social proximity of u_i with respect to u_k , denoted by S_i^k . *Contagion influence*, the impact from the users who have adopted u_i , consists of two factors: each user's direct trust on u_i at time t and social proximity of u_i and user in U_i^c . U_i^c is the set of users who have adopted advisor u_i in their social networks under the context

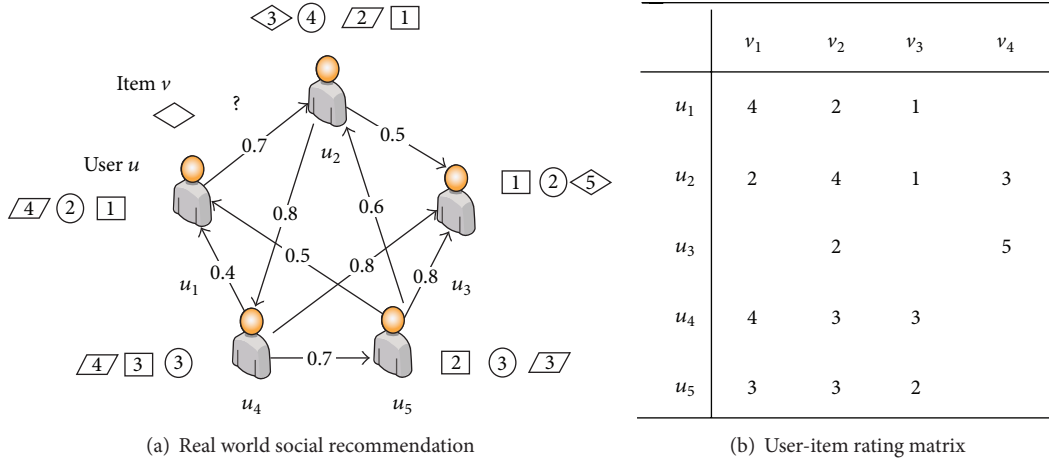


FIGURE 1: Example for trust-based recommendation problems.

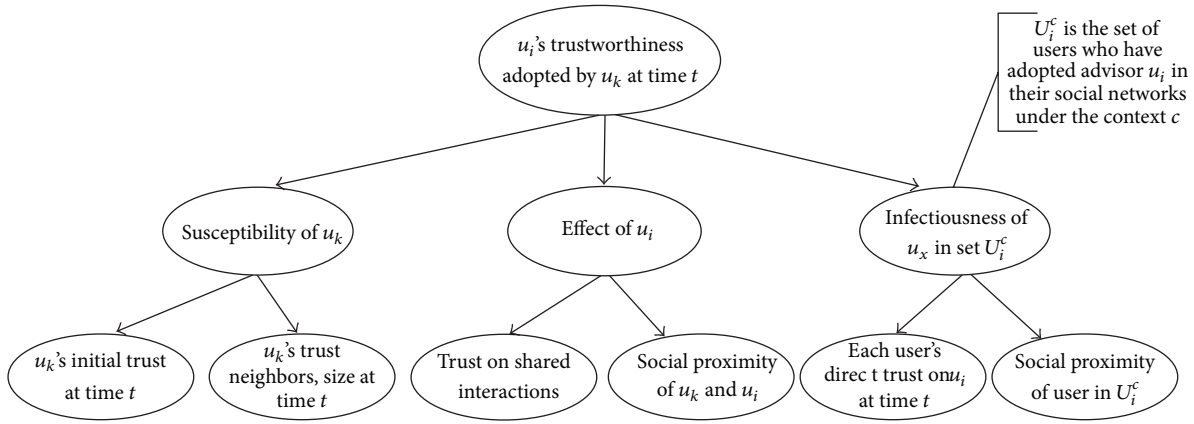


FIGURE 2: Trust Diffusion Model.

c . We employ the existing largest ontology—the LinkedData ontology [17]—to describe the context. The context c ($c \in \{c_1, c_2, \dots, c_m\}$), where m is the number of contexts in the system) is represented by a set of interacted items. Our goal is to compute the u_i 's trustworthiness adopted by u_k at time t and under each context c ($c \in \{c_1, c_2, \dots, c_m\}$), as $T(u_k, u_i, t, c)$.

It's worth noting that we consider trust to be multidimensional. The trust relationships between users take place under certain conditions, which is modeled as multidimensional features—intrinsic tendencies, direct influence, and contagion influence. Users have different levels of expertise in different domains; each user has different opinions about other users' expertise in different conditions. That is, one user may be trustworthy, completely ignorant, or untrustworthy about others. In this paper, we focus on trust and unknown relationships. The discussion of trust (friendship), distrust (enmity), and unknown within the diffusion model was introduced in [18]. Further discussion of trust, distrust, and unknown within the trust diffusion model will be exploited in the future.

The computational steps of trustworthiness are as follows.

Step 1 (compute the effect of direct connections). Based on the previous shared interactions of user u_k and advisor u_i , at time t_0 , u_k 's direct trust towards u_i under context c is denoted as $DT(u_k, u_i, t_0, c)$. At time t , u_k and u_i have a new shared interaction with the same item, denoted as $v_{u,k}$ under the context $c_{u,k}$ and $v_{u,i}$ under the context $c_{u,i}$, respectively. Then, based on the new shared interaction at time t under context c , we update user u_k 's direct trust towards user u_i , as follows:

$$DT(u_k, u_i, t, c) = \frac{DT(u_k, u_i, t_0, c) \lambda^{(t-t_0)} + DI(v_{u,k}, v_{u,i}, t)}{1 + \lambda^{(t-t_0)}}, \quad (1)$$

where $0 < \lambda \leq 1$ is a time decay factor for user to decrease the effect of old shared interactions between u_k and u_i . $DI(v_{u,k}, v_{u,i}, t)$ is the similarity between $v_{u,k}$ and $v_{u,i}$; the formula is as follows:

$$DI(v_{u,k}, v_{u,i}, t) = (1 - |v_{u,k} - v_{u,i}|) \times S(c_{u,k}, c_{u,i}) \times S(c_{u,k}, c), \quad (2)$$

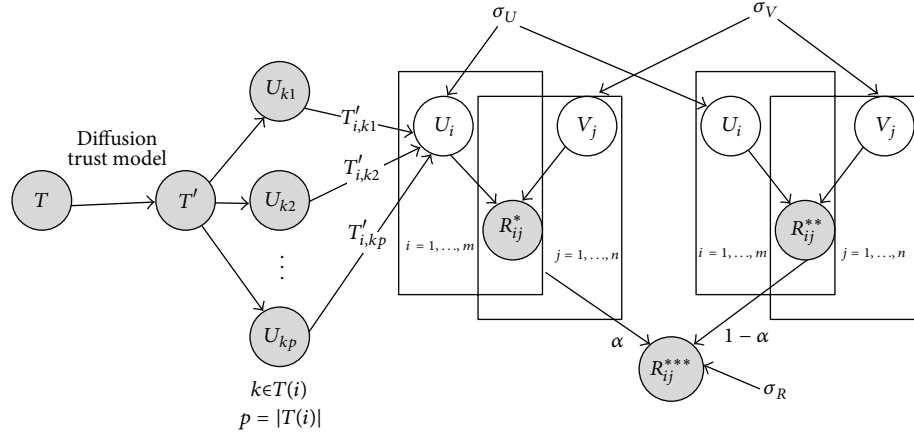


FIGURE 3: The improved probabilistic factor graph based on trust diffusion model.

where $S(c_{u,k}, c_{u,i})$ is the similarity between contexts $c_{u,k}$ and $c_{u,i}$, and $S(c_{u,k}, c)$ is the similarity between contexts $c_{u,k}$ and c . The context $c_{u,k}$ is the set of the items interacted by user u_k , and $c_{u,i}$ is the set of items interacted by advisor u_i . We assume context $c_{u,k} = \langle c_{u,k}^1, c_{u,k}^2, \dots, c_{u,k}^n \rangle$ and context $c_{u,i} = \langle c_{u,i}^1, c_{u,i}^2, \dots, c_{u,i}^n \rangle$, where n is the number of items interacted by user. The context similarity is computed according to paper [19].

Step 2 (computing contagious influence). Based on the diffusion theory, we can employ the weighted average of trust evaluation on user u_i from each user $u_x \in U_i^c$ to model contagious influence of users in U_i^c , denoted as

$$CI(u_k, U_i^c, t) = \frac{\sum_{u_x \in U_i^c} DT(u_k, u_x, t, c) \times S_x^k}{\sum_{u_x \in U_i^c} S_x^k}, \quad (3)$$

where S_x^k is u_x 's social proximity with respect to user u_k , which is computed according to socially spatial information [20].

Step 3 (combining the three factors). By considering the three factors: susceptibility, direct connections, and contagious influence, the u_i 's trustworthiness is adopted by u_k at time t under context c , denoted as $T(u_k, u_i, t, c)$. The previous method [15] computes the $T(u_k, u_i, t, c)$ through the linear combination of three factors.

Consider

$$T(u_k, u_i, t, c) = \omega_1 T_k^0 + \omega_2 DT(u_k, u_i, t, c) + \omega_3 CI(u_k, U_i^c, t), \quad (4)$$

where ω_1 , ω_2 , and ω_3 are the weight of effect from intrinsic tendencies, effect from direct connections, and contagious influence, respectively. The specific calculation of weights is introduced later.

Main Step for Improvement. Our improved method allows exponential decline in the additive model according to the diffusion theory [16], which better accommodates the introduction of temporal heterogeneity.

Consider

$$\begin{aligned} T(u_k, u_i, t, c) &= \exp(\omega_1 T_k^0) \\ &+ \sum_{s \in \varphi(t)} \exp[\omega_2 DT(u_k, u_i, t, c) \\ &+ \omega_3 CI(u_k, U_i^c, t) + \lambda^{(t-t_s)}], \end{aligned} \quad (5)$$

where ω_1 , ω_2 , and ω_3 are same as in (4). λ is same with (1). $\varphi(t)$ contains all spatially relevant individuals adopted before t . t^s is the adoption time of s th member of $\varphi(t)$.

ω_2 can be measured as follows:

$$\omega_2 = \begin{cases} \frac{N_{u_k}^t}{N_{\min}} & \text{if } N_{u_k}^t \leq \frac{N_{u_k}^t}{N_{\min}} \\ 1 & \text{otherwise.} \end{cases} \quad (6)$$

$N_{u_k}^t$ is the number of users who are trusted by user u_k in time t and N_{\min} is the minimum number of neighbors needed for user u_k to be confident about their own evaluation, which can be determined by an acceptable level of error and a confidence measurement in [21] as follows:

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

where ε is the maximal level of error that can be accepted by u_k , and γ is the confidence measure. Note that when $\omega_2 = 1$, the agent consider only direct trust part. In order to more accurately model the advisor's trustworthiness, direct connections and contagious influence will always be used together. That is, in (4) and (5), if $\omega_2 + \omega_3 > 1$, then we change the weights to be $\omega_2 = \omega_2/(\omega_2 + \omega_3)$ and $\omega_3 = 1 - \omega_2$. $\omega_1 = 1 - \omega_2 - \omega_3$.

2.3. The Improved Probabilistic Factor Model Based on Trust Diffusion Theory. In this section, we present an improved model in Figure 3 (DiffTrust + PMF) based on the social trust

ensemble (RSTE) [8]. RSTE is based on the following three intuitions. (1) Users have their own characteristics, and they have different tastes on different items. (2) Users can be easily influenced by the friends they trust and prefer their friends' recommendations. (3) One user's final decision is the balance between his/her own taste and his/her trusted friends' favors. According to these intuitions, the rating R_{ij} in the user-item matrix is interpreted as the representation mixed by both the user u_i 's taste and his/her trusted friends' tastes on the item v_j . RSTE works to some extent and further has other inherent weaknesses. For example, RSTE only deals with the explicit trust in social trust network. However, trust network from preexisting social links between users is rather sparse, which leads to the fact that direct trust relationships can only play a small role in the process of the recommendation. Different from RSTE, our improved method integrates diffusion trust model into probabilistic factor graph, which will match more trust users for the current user through DiffTrust model, fully dig up more new trust relationships between users, and is used for recommendation service. In other words, our method not only combines the users' tastes with their trusted friends' favors for recommendation, but also studies systematically the implicit relationship of the trust network based on diffusion theory. DiffTrust + PMF can help to alleviate the trust network sparsity problem and will potentially increase the recommendation accuracy. As shown in Figure 3, the left part is trust diffusion model (DiffTrust) and the right part is the social network matrix factorization model. We first employ DiffTrust to explore the processes in which the trust is cultivated. The DiffTrust can be used to infer the implicit relationship of the trust network. The original matrix T is enriched into the dense matrix T' (the detailed process is introduced in Section 2.2). $T(i)$ is the set of users who are trusted by user u_i in the trust network, and $|T(i)|$ is the number of trusted friends of user u_i in the set $T(i)$. Next, we model the problem of social recommendation using matrix factorization model, which makes his/her own taste and his/her trusted friends' favors together. That is, one user's final decision is the balance between his/her own taste and his/her trusted friends' favors. Based on Figure 3, the conditional distribution over the observed ratings and trust relationships is given by

$$\begin{aligned}
 & p(U, V \mid R^{***}, T', \sigma_R^2, \sigma_U^2, \sigma_V^2) \\
 &= \prod_{i=1}^m \prod_{j=1}^n \left[N \left(R_{ij} \mid g \left(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right), \right. \right. \\
 & \quad \left. \left. \sigma_R^2 \right) \right]^{I_{ij}^{R^{***}}} \\
 & \times \prod_{i=1}^m N(U_i \mid 0, \sigma_U^2 I) \times \prod_{j=1}^n N(V_j \mid 0, \sigma_V^2 I),
 \end{aligned} \tag{8}$$

where $N(x \mid \mu, \sigma^2)$ is the probability density function of the Gaussian distribution with mean μ and variance σ^2 , and

$I_{ij}^{R^{***}}$ is the indicator function that is equal to 1 if user u_i rated item v_j and equal to 0 otherwise. The function $g(x) = 1/(1 + \exp(-x))$, which makes it possible to bound the range $U_i^T V_j$ from within the range $[0, 1]$. The parameter α is used to smooth the users' favors and the trusted friends' favors, which controls how much users trust themselves or their trusted friends.

The log of the posterior distribution for the recommendations is given by

$$\begin{aligned}
 & \ln p(U, V \mid R^{***}, T', \sigma_R^2, \sigma_U^2, \sigma_V^2) \\
 &= -\frac{1}{2\sigma_R^2} \\
 & \times \sum_{i=1}^m \sum_{j=1}^n I_{ij}^{R^{***}} \left(R_{ij}^{***} - g \left(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) \right)^2 \\
 & - \frac{1}{2\sigma_U^2} \sum_{i=1}^m U_i^T U_i - \frac{1}{2\sigma_V^2} \sum_{j=1}^n V_j^T V_j \\
 & - \frac{1}{2} \left(\sum_{i=1}^m \sum_{j=1}^n I_{ij}^R \right) \ln \sigma^2 - \frac{1}{2} (ml \ln \sigma_U^2 + nl \ln \sigma_V^2) + C,
 \end{aligned} \tag{9}$$

where C is a constant that does not depend on the parameters. Maximizing the log-posterior over two latent features with hyper-parameters (i.e., the observation noise variance and prior variances) kept fixed is equivalent to minimizing the following sum-of-squared-errors objective functions with quadratic regularization terms:

$$\begin{aligned}
 & L(R, T', U, V) \\
 &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^{R^{***}} \left(R_{ij}^{***} - g \right. \\
 & \quad \left. \times \left(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) \right)^2 \\
 & + \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2,
 \end{aligned} \tag{10}$$

where $T(i)$ is the set of users who are trusted by user u_i in the trust network, $\lambda_U = \sigma^2/\sigma_U^2$, and $\lambda_V = \sigma^2/\sigma_V^2$. In order to reduce the model complexity, in all of the experiments we conduct in Section 4, we set $\lambda_U = \lambda_V$. $\|\cdot\|_F^2$ denotes the Frobenius norm, $\|U\|_F^2 = \sum_{i=1}^m U_i^T U_i$ and $\|V\|_F^2 = \sum_{j=1}^n V_j^T V_j$.

A local minimum of the objective function given by (8) can be found by performing gradient descent in U_i, V_j ,

$$\begin{aligned}
 & \frac{\partial L}{\partial U_i} = \alpha \sum_{j=1}^n I_{ij}^{R^{***}} g' \left(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) V_j \\
 & \times \left(g \left(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) - R_{ij}^{***} \right)
 \end{aligned}$$

$$\begin{aligned}
& (1 - \alpha) \sum_{q \in B(i)} \sum_{j=1}^n I_{qj}^{R^{***}} g' \\
& \times \left(\alpha U_q^T V_j + (1 - \alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) \\
& \times \left(g \left(\alpha U_q^T V_j + (1 - \alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) - R_{qj}^{***} \right) \\
& \times T'_{qj} V_j + \lambda_U U_i, \\
\frac{\partial L}{\partial V_j} &= \sum_{i=1}^m I_{ij}^{R^{***}} g' \left(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) \\
& \times \left(g \left(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} T'_{ik} U_k^T V_j \right) - R_{ij}^{***} \right) \\
& \times \left(\alpha U_i + (1 - \alpha) \sum_{k \in T(i)} T'_{ik} U_k^T \right) + \lambda_V V_j,
\end{aligned} \tag{11}$$

where $g'(x)$ is the derivative of logistic function $g'(x) = \exp(x)/(1 + \exp(x)^2)$, and $B(i)$ is the set of the users who trust user u_i .

3. Experimental Evaluation

In this section, we firstly introduce datasets and evaluation metric. Then we compare performance of our novel method.

3.1. Dataset Description. We evaluate our method on Epinions [22], Flixster [23], and Douban [24] datasets.

Epinions dataset [22] was given directly by Epinions staff to Paolo Massa. Note that it is not a typical collaborative filtering dataset, since the ratings are about the articles and not about items: the ratings represent how much a certain user rates a certain textual article written by another user, that is, a review. We randomly select 2,100 users and 3,648 articles. These users issued 6408 article ratings and 5,392 statements (4,575 trusts and 817 distrusts).

Flixster [23] is a social movie site allowing users to share movie ratings, discover new movies and meet others with similar movie taste. This contains the friendship network crawled in December 2010 by Javier Parra (Javier.Parra@asu.edu). We randomly select 1,000 users and 2,913 items from the Flixster dataset. These users issued 8,127 ratings and 3,714 relation statements.

Douban dataset, crawled by Ma et al. [24], contains 16,830,839 ratings of 129,490 users on 58,541 movies and 1,692,952 friend links between these users. The Douban dataset used in our experiments consists of 3,160 users who have rated at least one of a total of 6,529 different items. The total number of ratings is 12,963 and the total number of friend links is 7,152.

The statistical information of these three datasets is summarized in Table 1.

3.2. Evaluation Metrics. We measure the prediction quality of our proposed method in terms of the root mean square error (RMSE), Recall, and *F* Measure.

The metrics RMAE is defined as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i,j} (r_{i,j} - \hat{r}_{i,j})^2}, \tag{12}$$

where $r_{i,j}$ and $\hat{r}_{i,j}$ denote the actual and recommended rating user i gave to item j , respectively. N denotes the number of tested ratings. The smaller the value of RMSE, the more precise the recommendation.

Recall is the percentage of $\langle user, item \rangle$ pairs in the test set for which we can compute a recommendation.

Consider

$$Recall = \frac{tp}{tp + fn}, \tag{13}$$

where tp is the number of actual $\langle user, item \rangle$ pairs predicted by recommendation method and fn is the number of actual $\langle user, item \rangle$ pairs not predicted by recommendation method.

We compute the *F* Measure by combining RMSE and Recall into a single evaluation as follows:

$$F \text{ Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{14}$$

where the precision is defined as follows:

$$\text{Precision} = 1 - \frac{RMSE}{4}. \tag{15}$$

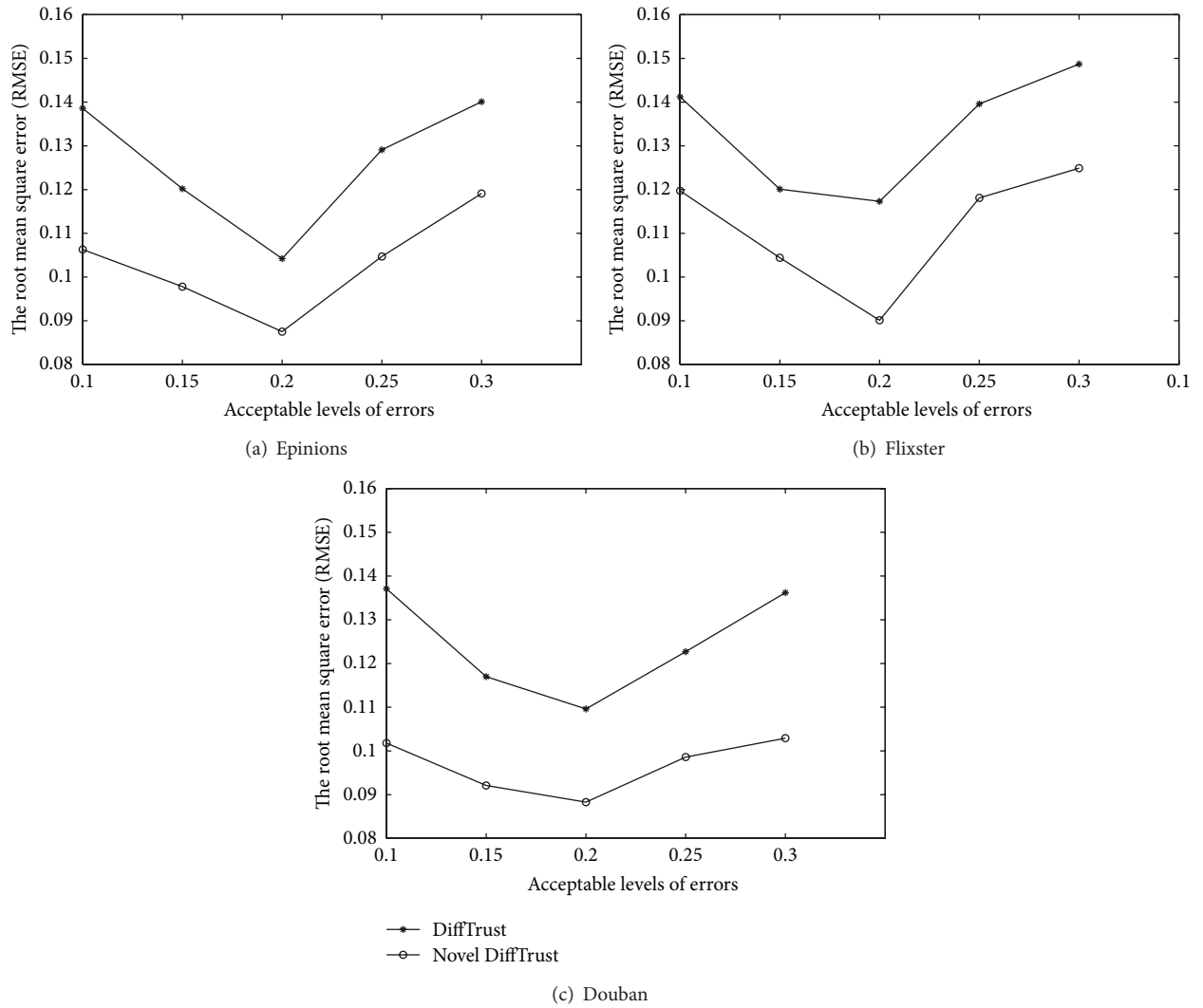
In this equation, 4 is the maximum possible error since the values of ratings are in the range $[1, 5]$.

3.3. Comparison of the Primary DiffTrust and Novel DiffTrust. To validate and gain insights about the usefulness of our novel DiffTrust introduced in Section 2.2. For all three datasets, we use the metrics, the root mean square error (RMSE), between predicted trust value for each user pair and the actual trust relationship, to measure the prediction quality of our method in comparison with the existing DiffTrust [15]. In all the experiments conducted in the following sections, we adopt the method in [21] to set the parameter $\gamma = 0.8$. It means that the confidence value should be no less than 0.8 in order for the users to be confident about their own evaluation values of advisors. In addition, we set the time decay factor $\lambda = 0.8$ (in (1) and (5)) according to the paper [25]; when $\lambda = 0.8$, the model reaches its best performance. Table 2 lists different acceptable levels of errors, correspondent weights from direct connections, and different results of RMSE. Figure 4 shows the corresponding charts.

From Figures 4(a), 4(b), and 4(c), it is clear that our novel DiffuTrust performs quite well compared with the primary DiffTrust. This is probably because our improved method allows exponential decline in the additive model, which better accommodates the introduction of temporal heterogeneity. In addition, we observe that the value of ϵ impacts the results of RMSE significantly. No matter our

TABLE 1: Statistical information about the three real datasets.

Datasets	Epinions	Flixster	Douban
Statistics	2,100 users	1,100 users	3,160 users
	3,648 articles	2,913 movies	6,529 movies
	6,408 article ratings	8,127 movie ratings	12,963 movie ratings
	5,392 relations statements	3,714 relation statements	7,152 relation statements
Data types	users \rightarrow articles	users \rightarrow movies	users \rightarrow movies
	users \rightarrow users	users \rightarrow users	users \rightarrow users
Rating scale	1, 2, 3, 4, 5	0.5, 1, 1.5, ..., 5	1, 2, 3, 4, 5
Trust between users	1	1	1

FIGURE 4: Impacts of ϵ on RMSE.

novel DiffuTrust or primary DiffTrust, as ϵ increases, RMSE decreases (prediction accuracy increases) at first, but when ϵ surpasses a certain threshold, RMSE increases (prediction accuracy decreases) with further increase of the value of ϵ . This phenomenon confirms the intuition that purely using the effect from intrinsic tendencies or purely using the effect

from direct connections or purely using contagious influence to model the advisor's trustworthiness cannot generate better performance than fusing these three favors together. It should be noted that both the two methods achieve the best performance when $\epsilon = 0.2$, while smaller values like $\epsilon = 0.15$ or larger values like $\epsilon = 0.25$ can potentially degrade the

TABLE 2: Shows the impacts of ε on RMSE.

Datasets	ε	0.1	0.15	0.2	0.25	0.3
	N_{\min}	115	51	29	19	13
Epinions	RMSE (DiffTrust)	0.1386	0.1202	0.1042	0.1291	0.1401
	RMSE (Novel DiffTrust)	0.1063	0.0978	0.0875	0.1047	0.1191
Flixster	RMSE (DiffTrust)	0.1412	0.1201	0.1173	0.1396	0.1487
	RMSE (Novel DiffTrust)	0.1197	0.1044	0.0901	0.1181	0.1249
Douban	RMSE (DiffTrust)	0.1371	0.1170	0.1096	0.1227	0.1362
	RMSE (Novel DiffTrust)	0.1018	0.0921	0.0883	0.0986	0.1029

TABLE 3: Shows all comparisons on three datasets.

Datasets	Metrics	PMF	BPMF	RSTE	Our method
Epinions	RMSE	0.8812	0.8461	0.8219	0.7146
	Precision	0.7797	0.7885	0.7945	0.8213
	Recall	0.7196	0.7468	0.8725	0.9603
	F Measure	0.7484	0.7671	0.8317	0.8854
Flixster	RMSE	0.8919	0.8749	0.8362	0.7313
	Precision	0.7770	0.7813	0.7910	0.8172
	Recall	0.6817	0.7364	0.8449	0.9415
	F Measure	0.7262	0.7582	0.8170	0.8749
Douban	RMSE	0.8815	0.8625	0.8317	0.7207
	Precision	0.7796	0.7844	0.7921	0.8200
	Recall	0.7024	0.7409	0.8616	0.9571
	F Measure	0.7390	0.7620	0.8254	0.8832

model performance. These experiments clearly demonstrate different acceptable levels of errors (for u_k) can impact the advisor's trustworthiness.

3.4. Results and Analysis. In this section, in order to show the performance improvement of our DiffTrust + PMF method, we compare our method with the following methods.

PMF [12]—probabilistic matrix factorization, it is a probabilistic linear model with Gaussian observation noise and only uses user-item matrix for the recommendations.

BPMF [13]—Bayesian probabilistic matrix factorization, it provides a predictive distribution instead of just a single number, allowing the confidence in the prediction to be quantified and taken into account when making recommendations using the model.

RSTE [8]—recommendation with social trust ensemble (RSTE), it is proposed by Ma et al. and employs opinions of the trusted friends in the social trust network to make recommendations for the users.

The parameter setting of our method is $\alpha = 0.4$, and in all the experiments conducted in the following sections, we set all of the parameters $\lambda_U = \lambda_V = 0.001$. Table 3 shows the RMSE, Coverage, and F Measure for all comparisons. Figures 5 and 6 show the charts comparing the results of different methods according to each of the three evaluation measures separately.

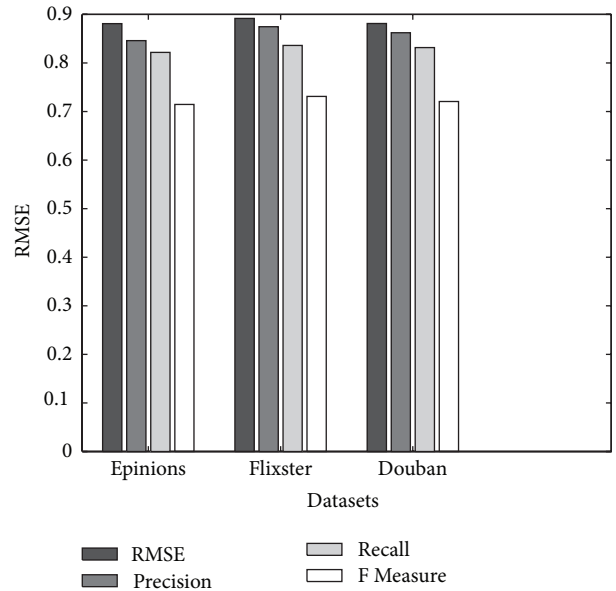


FIGURE 5: Comparison of RMSE for different methods.

As shown in Figure 6, our method and RSTE have lower error than the other two methods. This is probably because our method and RSTE take into account the users' tastes with their trusted friends' favors for recommendation. In addition, the RMSE of our method's RMST (0.7146) is lower than that of RSTE (0.8219), which means that considering

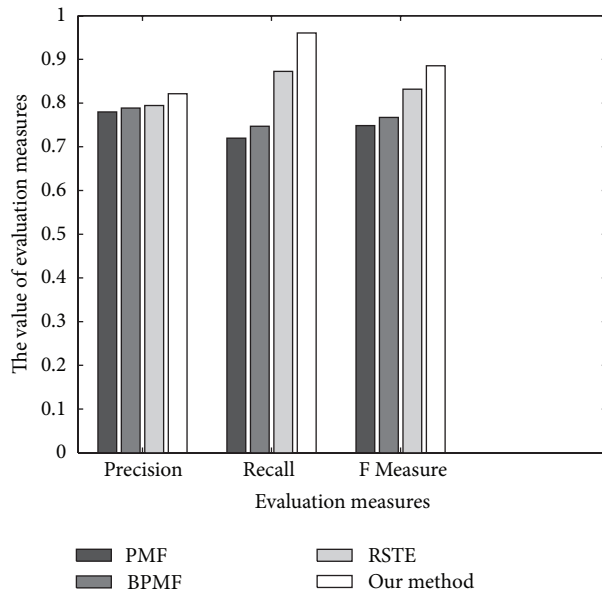


FIGURE 6: Comparison of evaluation measures for different methods on Epinions Dataset.

the possible diffusions of trusts between various users can reduce the error. Figure 6 shows the F Measure together with precision and coverage for all methods on Epinions dataset. Our method outperforms all other methods in terms of F Measure. It should be noted that our method not only clearly has a lower error, but also has a better coverage than other all methods on the three datasets.

In summary, our method clearly outperforms matrix factorization for recommendation (both PMF and BPMF) in terms of coverage because of exploiting the trust-network. Moreover, our method substantially improves the precision of existing trust-based method (RSTE). This improvement is achieved by considering trust diffusion processes. This consideration will help to alleviate the data sparsity problem and will potentially increase the prediction accuracy.

4. Related Work

Here we categorize the related works into two areas: trust diffusion and recommender systems.

Our recommendation method is based on trust diffusion. Trust diffusion, which is also known as trust propagation, is about predicting the trust worthiness of nonadjacent agents by combining trust values through distinct indirect paths. Trust propagation is widely studied in [26–28]. Teacy et al. [26] develop TRAVOS, which models an agent's trust in an interaction partner. Specifically, trust is calculated using probability theory taking account of past interactions between agents. Hang et al. [27] model trust as a binary event. They define three operators for concatenating trust along a path, aggregating trust from distinct paths from the same witness, and selecting the most trustworthy path among all witnesses, respectively. Teacy and Hang only consider the user's own experience with the advisor, which may lead to

inaccurate trust evaluation when the user has only limited experience with the advisor or the advisor dynamically changes her behavior. To address this problem, Fang et al. [28] propose a trust diffusion model; trustworthiness perceived by a specific user is measured under a specific context. Their method can flexibly model adjusts the weight of the user's own experience and other users' evaluation on the advisor. In order to better accommodate the introduction of temporal heterogeneity, we improve the method in paper [28]. Our improved DiffTrust allows exponential decline in the additive model, the experimental results show that our novel DiffTrust performs quite well compared with the primary DiffTrust. Some other trust propagation methods [18, 29, 30] have also been proposed.

Now we discuss some related work of recommender systems. Recommender system is an indispensable technique in the field of information filtering. Trust-based recommendation assumes the additional knowledge of a trust network among users and can better deal with cold start users. Most related to our method, Ma et al. [8] utilized a novel probabilistic factor analysis framework to fuse the users' tastes and their trusted friends' favors together by an ensemble parameter: their method achieves an improvement over the very large datasets since it scales linearly with the number of observations. Different from their method, we mainly consider the possible diffusions of trusts between various users by combining DiffTrust with PMF for recommendation, which can help to alleviate the trust network sparsity problem and will potentially increase the prediction accuracy. Other similar works are also employing trust relationship among users to solve recommendation problems by [11, 31–33]. For example, [11] proposed a novel implicit trust aware recommendation model (iTARS) based on the small-worldness of the implicit trust network, in which the implicit trust is generated from the user similarities. Wang et al. [31] develop a method to generate trust between users based on their common tastes. By first grouping items into different clusters, the authors use a frequency measure of the number of ratings each user asserts to different groups of items, and thus build up a personalized taste set for user and infer trust based on the common taste sets between users. Shambour and Lu [32] propose a fusion-based recommendation approach that fuses the trust and semantic information of users and items within the CF framework to achieve yet more effective results in terms of recommendation accuracy and coverage, especially when dealing with “data sparsity” and “cold start” problems. To sum up, most of these trust-based recommendation algorithms have exposed one major limitation: they do not consider trust network dynamic and evolutionary (varying over time or interaction context). This problem has further been addressed in our work.

Other algorithms based matrix factorization have been proposed for recommendation, which assume that the preference of a user can be represented by a small number of unobserved features, such as probabilistic matrix factorization (PMF) [12], variational Bayesian matrix factorization (VBMF) [34], Bayesian probabilistic matrix factorization (BPMF) [13], and general probabilistic matrix factorization (GPMF) [35]. However, these methods suffer from the “data

sparsity” problem and the “cold start” problem. These problems have further been addressed in our work by fusing social relations among users with rating data, which can help to improve the performance of recommender systems.

5. Conclusion

In our study of recommendation method, we found two prevalent problems which must be resolved in order to make a more useful recommendation. The first problem regards dealing with the possible diffusions of trust between various users and the second is the data sparsity problem. As discussed earlier, the challenge arises in finding the ideal tradeoff between the accuracy and efficiency, which is the focus in this paper.

In this study, we have presented the DiffTrust + PMF method and achieved good results in recommendation. We have (1) significantly studied the implicit relationship of the trust network based on diffusion theory and alleviated the data sparsity problem, and (2) made use of the users’ tastes and their trusted friends’ favors for recommendation by integrating diffusion trust model into probabilistic factor graph. Experimental results based on real datasets showed that our method of combining DiffTrust with the PMF can achieve high performance.

In this paper, we only use the trust information, while distrust statements are also provided in many online social networks. The future work intends to apply distrust information to recommendation systems. An important problem worthy of future research is how the distrust relations affect the user preference.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Authors would like to thank anonymous reviewers for their valuable suggestions that help us improve the quality of the paper. This work is supported by National Natural Science Foundation of China (no. 61170192) and Natural Science Foundation of CQ (no. CSTC2012JJB 40012). Li Li is the corresponding author for the paper.

References

- [1] S. Zhu, K. Yu, and Y. Gong, “Stochastic relational models for large-scale dyadic data using MCMC,” in *Proceedings of the 22nd Annual Conference on Neural Information Processing Systems (NIPS '08)*, vol. 21, pp. 1993–2000, December 2008.
- [2] M. Jahrer, A. Töschner, and R. Legenstein, “Combining predictions for accurate recommender systems,” in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 693–701, Washington, DC, USA, July 2010.
- [3] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.
- [4] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, *Recommender Systems Handbook*, Springer, New York, NY, USA, 2011.
- [5] J.-G. Liu, T. Zhou, and B.-H. Wang, “Research progress in personalized recommendation system,” *Progress in Natural Science*, vol. 19, pp. 1–15, 2009.
- [6] J. Golbeck, “Generating predictive movie recommendations from trust in social networks,” in *Trust Management*, K. Stølen, W. H. Winsborough, F. Martinelli, and F. Massacci, Eds., pp. 93–104, Springer, Berlin, Germany, 2006.
- [7] C. S. Hwang and Y. P. Chen, “Using trust in collaborative filtering recommendation,” in *New Trends in Applied Artificial Intelligence*, pp. 1052–1060, Springer, Berlin, Germany, 2007.
- [8] H. Ma, I. King, and M. R. Lyu, “Learning to recommend with social trust ensemble,” in *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09)*, pp. 203–210, Boston, Mass, USA, July 2009.
- [9] J. O'Donovan and B. Smyth, “Trust in recommender systems,” in *Proceedings of the International Conference on Intelligent User Interfaces (IUI '05)*, pp. 167–174, January 2005.
- [10] P. Victor, M. de Cock, and C. Cornelis, “Trust and recommendations,” in *Recommender Systems Handbook*, pp. 645–675, Springer, New York, NY, USA, 2011.
- [11] W. Yuan, L. Shu, H.-C. Chao, D. Guan, Y.-K. Lee, and S. Lee, “iTARS: trust-aware recommender system using implicit trust networks,” *IET Communications*, vol. 4, no. 14, pp. 1709–1721, 2010.
- [12] R. Salakhutdinov and A. Mnih, “Probabilistic matrix factorization,” in *Proceedings of the 21st Annual Conference on Neural Information Processing Systems (NIPS '08)*, vol. 20, pp. 1257–1264, 2008.
- [13] R. Salakhutdinov and A. Mnih, “Bayesian probabilistic matrix factorization using markov chain Monte Carlo,” in *Proceedings of the 25th International Conference on Machine Learning*, pp. 880–887, ACM, July 2008.
- [14] H. Shan and A. Banerjee, “Generalized probabilistic matrix factorizations for collaborative filtering,” in *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM '10)*, pp. 1025–1030, IEEE, December 2010.
- [15] H. Fang, J. Zhang, and N. M. Thalmann, “A trust model stemmed from the diffusion theory,” in *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '13)*, pp. 805–812, Saint Paul, Minn, USA, 2013.
- [16] D. Strang and N. B. Tuma, “Spatial and temporal heterogeneity in diffusion,” *American Journal of Sociology*, vol. 99, no. 3, pp. 614–639, 1993.
- [17] C. Bizer, T. Heath, and T. Berners-Lee, “Linked data—the story so far,” *International Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.
- [18] J. Yuan, L. Li, and F. Tan, “Dealing with trust, distrust and ignorance,” in *Knowledge Science, Engineering and Management*, M. Wang, Ed., vol. 8041 of *Lecture Notes in Computer Science*, pp. 551–560, Springer, Berlin, Germany, 2013.
- [19] K. Radinsky, S. Davidovich, and S. Markovitch, “Learning causality for news events prediction,” in *Proceedings of the 21st Annual Conference on World Wide Web (WWW '12)*, pp. 909–918, April 2012.
- [20] 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*, vol. 56, no. 8, pp. 1219–1241, 2010.
- [21] J. Zhang and R. Cohen, “A comprehensive approach for sharing semantic web trust ratings,” *Computational Intelligence*, vol. 23, no. 3, pp. 302–319, 2007.
 - [22] P. Massa and P. Avesani, “Trust-aware bootstrapping of recommender systems,” in *Proceedings of the ECAI Workshop on Recommender Systems*, 2006.
 - [23] <http://www.sfu.ca/~sja25/datasets/>.
 - [24] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, “Recommender systems with social regularization,” in *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM '11)*, pp. 287–296, ACM, February 2011.
 - [25] C. Xia, X. Jiang, S. Liu, Z. Luo, and Y. Zhang, “Dynamic item-based recommendation algorithm with time decay,” in *Proceedings of the 6th International Conference on Natural Computation (ICNC '10)*, pp. 242–247, August 2010.
 - [26] 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*, vol. 12, no. 2, pp. 183–198, 2006.
 - [27] 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 Conference on Autonomous Agents and Multiagent Systems (AAMAS '09)*, pp. 1025–1032, IFAAMAS, Columbia, SC, USA, 2009.
 - [28] H. Fang, J. Zhang, and N. M. Thalmann, “A trust model stemmed from the diffusion theory,” in *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '13)*, pp. 805–812, Saint Paul, Minn, USA, 2013.
 - [29] R. Guha, P. Raghavan, R. Kumar, and A. Tomkins, “Propagation of trust and distrust,” in *Proceedings of the 13th International World Wide Web Conference Proceedings (WWW '04)*, pp. 403–412, May 2004.
 - [30] J. Leskovec, D. Huttenlocher, and J. Kleinberg, “Predicting positive and negative links in online social networks,” in *Proceedings of the 19th International World Wide Web Conference (WWW '10)*, pp. 641–650, ACM, April 2010.
 - [31] J. Wang, J. Yin, Y. Liu, and C. Huang, “Trust-based collaborative filtering,” in *Proceedings of the 8th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD '11)*, pp. 2650–2654, IEEE, July 2011.
 - [32] Q. Shambour and J. Lu, “A trust-semantic fusion-based recommendation approach for e-business applications,” *Decision Support Systems*, vol. 54, no. 1, pp. 768–780, 2012.
 - [33] C. C. Chen, Y.-H. Wan, M.-C. Chung, and Y.-C. Sun, “An effective recommendation method for cold start new users using trust and distrust networks,” *Information Sciences*, vol. 224, pp. 19–36, 2013.
 - [34] Y. Lim and Y. Teh, “Variational bayesian approach to movie rating prediction,” in *Proceedings of the KDD Cup and Workshop*, pp. 15–21, Citeseer, 2007.
 - [35] H. Shan and A. Banerjee, “Generalized probabilistic matrix factorizations for collaborative filtering,” in *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM '10)*, pp. 1025–1030, December 2010.

