A trust-based recommendation method using network diffusion processes

Ling-Jiao Chen, Jian Gao *

CompleX Lab, Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China
Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China

HIGHLIGHTS

- A trust-based CosRA+T method is proposed by scaling trusted users' resources before the redistribution.
- The optimal scaling for the best recommendation accuracy is found across several metrics and datasets.
- CosRA+T has a remarkable improvement as suggested by extensive experiments on two real-world datasets.
- Some insights into the mechanisms of the proposed and benchmark methods are suggested.

ARTICLE INFO

ABSTRACT

A variety of rating-based recommendation methods have been extensively studied including the well-known collaborative filtering approaches and some network diffusion-based methods, however, social trust relations are not sufficiently considered when making recommendations. In this paper, we contribute to the literature by proposing a trust-based recommendation method, named CosRA+T, after integrating the information of trust relations into the resource-redistribution process. Specifically, a tunable parameter is used to scale the resources received by trusted users before the redistribution back to the objects. Interestingly, we find an optimal scaling parameter for the proposed CosRA+T method to achieve its best recommendation accuracy, and the optimal value seems to be universal under several evaluation metrics across different datasets. Moreover, results of extensive experiments on the two real-world rating datasets with trust relations, Epinions and FriendFeed, suggest that CosRA+T has a remarkable improvement in overall accuracy, diversity and novelty. Our work takes a step towards designing better recommendation algorithms by employing multiple resources of social network information.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The fast development of information technologies has spawned the emergence of the E-commerce and largely boosted its expansion during the past decades [1,2], especially in China along with its rapid economic growth [3,4]. Recently, a large variety of online stories and services (e.g., online books, music, movies, etc.) have made our lives much easier, however, the tremendous amount of available information in the era of big data has caused a serious problem of information overload [5]. For example, it will be extremely hard for us to review thousands of online stores before choosing a box of favorite chocolate. To address this problem, recommender systems as an information filtering technology have been widely applied by online

* Corresponding author at: CompleX Lab, Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China.
   E-mail addresses: lingjiao.chen@foxmail.com (L.-J. Chen), gaojian08@hotmail.com (J. Gao).

platforms [6,7], where users are provided customized and personalized services. Back to the example, recommender systems will automatically recommend us a few boxes of chocolate that may meet our tastes after analyzing our historical behavior big data including purchase records and search archive [8,9] that are recorded by various socioeconomic platforms [10–12].

As the core of recommender systems, a variety of recommendation algorithms have been proposed and applied to online platforms [2,13,14]. One of the most well-known methods is the collaborative filtering (CF) [15] including the user-based CF (UCF) and the item-based CF (ICF), in which items of potential interest are recommended based on the similar users' opinions and the similarity between items, respectively [16,17]. Later, some dynamical processes borrowed from the field of statistical physics are introduced into the design of diffusion-based recommendation algorithms including the heat conduction (HC) method [18], the mass diffusion (MD) method [19], a hybrid method of HC and MD [20], a weighted HC method [21], and some others [22–25]. The network-based diffusion is indeed a resource-allocation (RA) process [26], for example, MD is essentially a two-step RA process on "user–object" networks [27]. Within this framework, recent literature proposed a method built using the CosRA index [28], which combines advantages of both the cosine similarity and the RA index. In CosRA, resources are initialized for each object and then redistributed via the CosRA-based transformation. Other network-based methods are reviewed by the very recent survey papers [29,30].

Most of these aforementioned recommendation algorithms are designed solely based on the users' rating information [31,32], however, the relationships among users (e.g., trust relations [33,34]) embedded in online social networks are always ignored [35,36]. Yet, in real-world observations, our preference of products or adoption of information could be also affected by our social relationships [37,38] such as the friends in working places and the people connected through social media or by mobile phones [39]. Intuitively, we are more likely to adopt a trusted friend's suggestions than those coming from a stronger in online communities [40]. To this point, some works have integrated trust relations among users into recommender systems [41,42]. For example, Jamali et al. [43] proposed a random walk method that combines trust-based [44] and item-based CF approaches, Ma et al. [45] proposed a social trust ensemble framework that fuses both users' tastes and trusted friends' favors, Shen et al. [46] utilized a trust-combined user-based CF approach by proposing two user trust models, and Guo et al. [47] proposed a trust-based matrix factorization technique that integrates ratings and multiple trust information. However, studies on integrating trust relations into the diffusion-based methods remain still insufficiently [48], which urges further explorations on designing better methods under the network-based diffusion framework by leveraging multiple resources of social information.

In this paper, we propose a trust-based recommendation method, named CosRA+T, by introducing the trust relations among users into the resource-allocation processes of the original CosRA method. Specifically, the amount of the resources received by trusted users are scaled by a tunable parameter before the following redistribution back to their collected objects. Interestingly, we find an optimal value of the scaling parameter for the best recommendation accuracy under several evaluation metrics across different datasets, suggesting the universality of the optimal scaling in implicating the proposed CosRA+T method. Further, we employ two real-world rating and trust datasets, Epinions and FriendFeed, to comprehensively test the performance of CosRA+T and compare it with five benchmark methods. Results suggest that CosRA+T improves the overall performance by giving a highly accurate, encouraging diverse and well novel list of recommendations. Our work highlights the role that social trust relations play in enhancing the algorithmic performance of trust-based recommendation methods.

The remainder of this paper is organized as follows. Section 2 introduces some benchmark recommendation methods and the proposed CosRA+T method. Section 3 introduces the used datasets and performance evaluation metrics. Section 4 presents our main results. Finally, Section 5 provides conclusion remarks and related discussions.

2. Methods

In this section, we will first introduce some basic notations that are traditionally used in describing online rating systems and recommendation algorithms. Then, we will briefly describe the five benchmark methods. Last, we will introduce the proposed CosRA+T method in detail.

2.1. Notations and benchmark methods

Online rating systems can be modeled by a "user–object" bipartite network $G(U, O, E_R)$, where $U = \{u_1, u_2, \ldots, u_m\}$ is the set of users, $O = \{o_1, o_2, \ldots, o_n\}$ is the set of objects, and $E_R = \{e_1, e_2, \ldots, e_l\}$ is the set of rating links. Here, m, n and l are the total number of objects, products and ratings, respectively. The bipartite rating network can be naturally represented by an adjacency matrix A, where the element $a_{i\alpha} = 1$ if there is a link connecting user i and object α, and the element $a_{i\alpha} = 0$ if otherwise. The Greek and Latin letters are used to distinguish object-related and user-related indices, respectively. The key purpose of recommendation algorithms is to provide a list of L objects in the recommendation list o_i^L for the target user i.

Five benchmark recommendation methods will be introduced in the following, including global ranking (GR), user-based collaborative filtering (UCF), heat conduction (HC), mass diffusion (MD), and CosRA. Based on the observations that users prefer popular objects, GR [27] as the most straightforward method recommends objects with the largest degree after sorting all objects in the descending order according to their degrees. In UCF, users will be recommended the objects collected by the users who share the similar tastes, where the user similarity is usually measured by the cosine similarity [49]. Together, there is also the item-based collaborative filtering (ICF), where similar objects to the users' past collected ones will be recommended.

Both MD and HC can be considered as network-based resource-allocation processes. For a given user i, the initial resources of all objects are denoted by the vector $f^{(i)}$, where $f_\alpha^{(i)} = 1$ if user i has collected object α, and $f_\alpha^{(i)} = 0$ if otherwise. Then, these resources are reallocated via the transformation $f'^{(i)} = Wf^{(i)}$, where W is the resource transfer matrix and $f'^{(i)}$ is the vector of the final resources. The resource transfer matrices W in MD and HC are mutually transposed [20]. Specifically, the element of W in MD is given by [27]

$$w_{\alpha\beta} = \frac{1}{k_\beta} \sum_{i=1}^{m} \frac{a_{i\alpha}a_{i\beta}}{k_i},$$ (1)

and the element of W in HC is given by [20]

$$w_{\alpha\beta} = \frac{1}{k_\alpha} \sum_{i=1}^{m} \frac{a_{i\alpha}a_{i\beta}}{k_i},$$ (2)

where k_α and k_β are respectively the degrees of objects α and β, k_i is the degree of user i, and m is the total number of users. So far, there have been many variants of the original resource transfer matrices and their associated recommendation algorithms [29].

CosRA is a recently proposed recommendation method based on the CosRA similarity index, which combines both the cosine index and the resource-allocation (RA) index [26]. Specifically, for a pair of objects α and β, the CosRA index is given by [28]

$$S_{\alpha\beta}^{CosRA} = \frac{1}{\sqrt{k_\alpha k_\beta}} \sum_{i=1}^{m} \frac{a_{i\alpha}a_{i\beta}}{k_i}.$$ (3)

In CosRA, the resources of object α are initialized as $f_\alpha^{(i)} = a_{i\alpha}$ for a given user i. Then, these resources are redistributed via the transformation $f'^{(i)} = S^{CosRA}f^{(i)}$, where $f'^{(i)}$ is the n-dimensional vector recording all the final resources that located on each object. After sorting all objects by their final resources $f'^{(i)}$, the top-L uncollected objects are recommended to the user i.

2.2. Trust-based recommendation method

The trusted-based CosRA+T method is proposed by introducing the trust relations into the network-based diffusion processes of the original CosRA method. The intuition behind the new method is that not only the similarity among objects or users can help predict users' preferences to objects, but also the friendship or trust relations could potentially affect users' decisions on choosing objects. For example, if two users have the similar tastes as indicated by their past ratings, the performance of recommendations to either of them may be further improved if they are close friends who trust each other and have similar demands in the same living environment. The trust relations among users are usually collected from online social platforms such as Epinions and FriendFeed, where not only users' ratings but also their trust relations could be recorded.

To explore the role that trust relations play in enhancing or suppressing the recommendation performance, we propose the CosRA+T method by taking the "user–user" trust network into consideration. The trust from user i to j defines a link from node i to j. Formally, the trust network can be represented by a adjacency matrix B, where the element $b_{ij} = 1$ if there is a link from node i to j, and $b_{ij} = 0$ if otherwise. Together, Fig. 1 illustrates the "user–object" bipartite network and the "user–user" trust network, where circles and squares correspond to users and objects, respectively. Note that, two circles in the same row connected by the dashed lines correspond to the same user. The trust relations are presented by solid directed and unweighted links from users (circles) in the right column to their trusted users (circles) in the left column.

The CosRA+T method works in three steps: First, for a given user i, the resources that object α is initially assigned is given by

$$f_\alpha^{(i)} = a_{i\alpha},$$ (4)

where $a_{i\alpha} = 1$ if user i has collected object α, and $a_{i\alpha} = 0$ if otherwise. Second, user i's neighboring users receive resources diffused through the bipartite network from their collected objects. Formally, the resources that user j receives are given by

$$f_j^{(i)} = \sum_{\alpha=1}^{n} \frac{a_{j\alpha}}{\sqrt{k_j k_\alpha}} f_\alpha^{(i)},$$ (5)

where k_α and k_j are respectively the degree of object α and user j, and n is the total number of objects. Third, users redistribute their resources back to all objects after considering the trust relations from the target user i. Specifically, the amount of user j's resources are scaled by a tunable parameter θ before the redistribution if user j is trusted by the user i, otherwise the
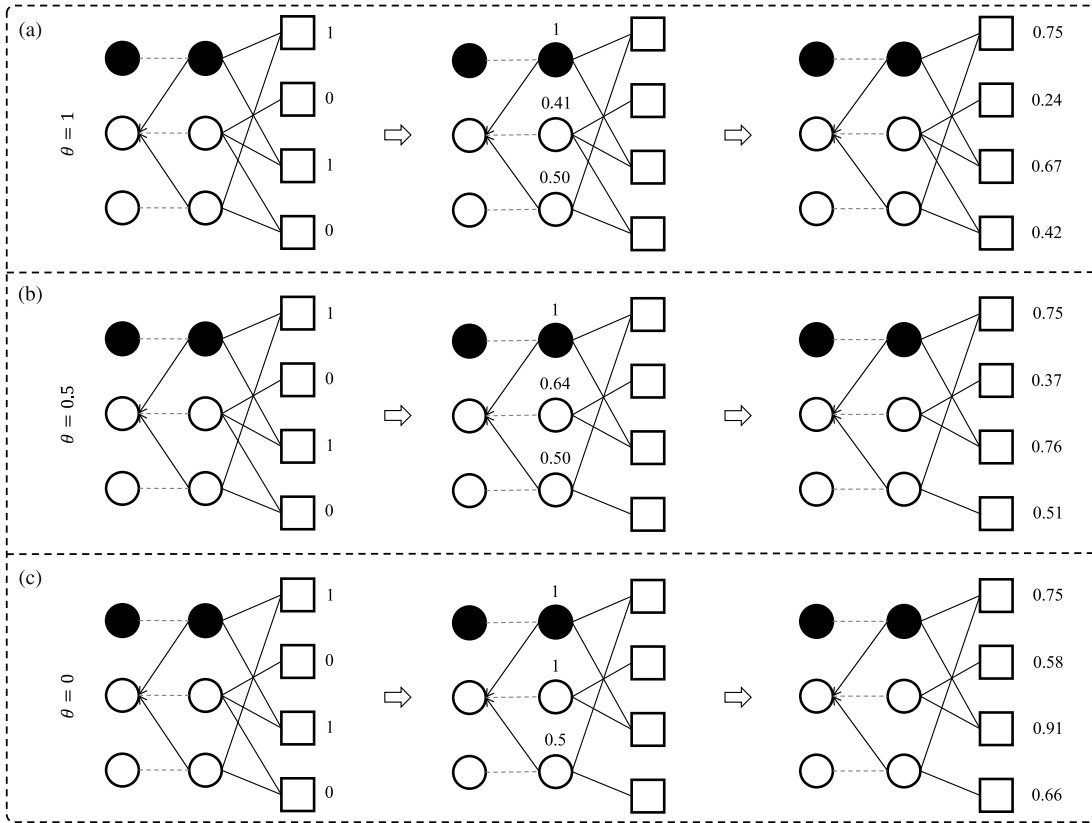
Fig. 1. Illustrations of the CosRA+T method. Initially, for a target user (colored black), the resources of each object are assigned by Eq. (4). Then, users receive resources diffused from the objects they collected according to Eq. (5). Finally, users redistribute resources to their collected objects after considering the trust relations regarding the target user. Only for users (the second user in this example) trusted by the target user (the first user), their resources are scaled by a tunable parameter θ before redistributed by Eq. (6). Values besides the nodes are the amount of received resources. Panel (a) illustrates the case of $\theta = 1$, where the CosRA+T method degenerates into the original CosRA method. Panels (b) and (c) illustrate the case of $\theta = 0.5$ and the case of $\theta = 0$, respectively.

resources of user j are directly redistributed back to the collected objects. Formally, the final resources that object β receives regarding the target user i is given by

$$f'^{(i)}_\beta = \sum_{j=1}^{m} \frac{a_{j\beta}}{\sqrt{k_j k_\beta}} \left(b_{ij} f_j^{(i)\theta} + (1 - b_{ij}) f_j^{(i)} \right), \tag{6}$$

where $b_{ij} = 1$ if user i trusts user j and $b_{ij} = 0$ if otherwise, θ is a tunable scaling parameter, and m is the total number of users. After sorting all objects by their final resources $f'^{(i)}$, the top-L uncollected objects are the recommendations to the target user i.

The motivation behind the introduction of the scaling parameter θ is to explore whether the trust relations affect the recommendation accuracy. As the value of the scaling parameter θ increases from 0 to 1, the effects of the trust relations in CosRA+T gradually diminish. In the case of $\theta = 0$ as illustrated in Fig. 1(a), the resources of the trusted users become 1 no matter how many resources they receive while the resources of the untrusted users remain unchanged before the redistribution. In the case of $\theta = 0.5$ as illustrated in Fig. 1(b), the resources of the trusted users are squared rooted and thus increased before the redistribution. In the case of $\theta = 0$ as illustrated in Fig. 1(c), the trust relations have no effects on the resource-allocation processes, and CosRA+T degenerates into the original CosRA method. There could be many other ways to integrate the trust relations when designing recommendation algorithms, while we only introduce one straightforward way here for the illustration.

3. Data and metrics

In this section, we first introduce the two rating datasets with the information of the trust relations among users, based on which we implement our CosRA+T method. Then, we introduce the evaluation metrics for testing the performance of recommendation methods.

Table 1
Summary statistics of the two rating and trust datasets.

Dataset	Users (m)	Objects (n)	Rating Links (l_R)	Sparsity (S_R)	Trust Links (l_T)	Sparsity (S_T)
Epinions	4066	7649	154,122	4.96×10^{-3}	217,071	1.31×10^{-2}
FriendFeed	4148	5700	96,942	4.10×10^{-3}	386,804	2.25×10^{-2}

Notes: The $S_R = l_R/(m \times n)$ is the rating network sparsity, and the $S_T = l_T/(m \times m)$ is the trust network sparsity.

3.1. Rating and trust datasets

Two benchmark rating datasets, namely, Epinions and FriendFeed, are used to test the recommendation performance. Epinions is a general consumer review website where people can review products by writing subjective reviews while FriendFeed is a social networking and bookmarking website where people can rate and share customized feeds. Both datasets use a 5-point rating scale from 1 to 5 (i.e., worst to best). For building the "user–object" rating network, ratings are converted to binary links by assigning 1 if the rating value is no less 3 and 0 if otherwise. After the coarse-graining, the Epinions dataset contains 4066 users, 7649 objects and 154,122 rating links with the network sparsity $S_R = 4.96 \times 10^{-3}$, and the FriendFeed dataset contains 4148 users, 5700 objects and 96,942 rating links with the network sparsity $S_R = 4.10 \times 10^{-3}$.

The two datasets contains also information of social networks, which can be used to estimate trust relations among users. On the two platforms (Epinions and FriendFeed), users can be friends by following each other in addition to rating objects. For building the "user–user" trust network, a directed binary link from node i to j is assigned 1 if user i follows user j and 0 if otherwise. Epinions dataset contains 217,071 trust links with the network sparsity $S_T = 1.31 \times 10^{-2}$, and FriendFeed dataset contains 386,804 trust links with the network sparsity $S_T = 2.25 \times 10^{-2}$. Basic statistics of the Epinions and FriendFeed datasets are summarized in Table 1.

3.2. Evaluation metrics

We employ a 10-folder cross-validation strategy to evaluate the algorithmic performance in each independent realization. Specifically, all ratings are divided into 10 equal sized subsamples. Then, one subsample is left out as the testing set and the remaining 9 subsamples are used as the training set. Next, the cross-validation process is repeated 10 times, making sure that each subsample serves as the testing set once. Finally, the 10 results are averaged to produce one single result for this independent realization. To quantitatively compare the recommendation performance, we apply eight commonly used evaluation metrics, including five accuracy metrics (AUC, Ranking Score, Precision, Recall, and F_1), two diversity metrics (Hamming distance and Intra-similarity), and one novelty metric (Popularity). These metrics are respectively introduced in the following.

Accuracy is the most important metric for the recommendation performance evaluation. We start by introducing two accuracy metrics that are independent of the recommendation list's length L. One is the AUC (area under the ROC curve) [50]. Given the ranks of all objects in the testing set, the value of AUC corresponds to the probability that a randomly chosen collected object is ranked higher than a randomly chosen un-collected object. After N times independent comparisons of the resources received by a pair of collected and un-collected objects, the AUC value is calculated by [51]

$$AUC = \frac{1}{m} \sum_{i=1}^{m} \frac{(N_1^{(i)} + 0.5 N_2^{(i)})}{N^{(i)}}, \tag{7}$$

where $N_1^{(i)}$ denotes the times that the collected object has more resources than the uncollected object, and $N_2^{(i)}$ denotes the times that they have the same amount of resources for user i. Larger AUC value suggests higher recommendation accuracy. The other one is the Ranking Score (RS) [27]. For a given user, RS measures the relative ranking of a relevant object in the recommendation list. Formally, the value of RS is calculated by [27]

$$RS = \frac{1}{|E^P|} \sum_{(i,\alpha) \in E^P} \frac{p_\alpha}{l_i}, \tag{8}$$

where $|E^P|$ is the size of the testing set, p_α is the position of object α in the ranking list of the recommendation, and l_i is the number of uncollected objects of user i in the training set. Smaller RS value suggests higher accuracy of a recommendation algorithm.

We then introduce three L-dependent accuracy metrics: Precision, Recall, and F_1 [52]. For all user, the average value of Precision $P(L)$ is calculated by

$$P(L) = \frac{1}{m} \sum_{i=1}^{m} \frac{d_i(L)}{L}, \tag{9}$$

where $d_i(L)$ is the number of recommended objects appeared in the testing set, and L is the total number of recommended objects. The average value of Recall $R(L)$ is calculated by

$$R(L) = \frac{1}{m} \sum_{i=1}^{m} \frac{d_i(L)}{D(i)}, \tag{10}$$

where $D(i)$ is the total number of objects in the testing set. Larger Precision and Recall suggest higher recommendation accuracy, however, the two measures are usually antagonistic since $P(L)$ usually decreases with L while $R(L)$ usually grows with L. To balance both Precision and Recall, the F_1 metric is introduced. The average value of $F_1(L)$ is calculated by

$$F_1(L) = \frac{1}{m} \sum_{i=1}^{m} \frac{2P_i(L)}{P_i(L) + R_i(L)}, \tag{11}$$

where $P_i(L)$ is the value of Precision and $R_i(L)$ is the value of Recall for user i. Larger F_1 value suggests higher recommendation accuracy.

Diversity is an important metric in evaluating the variety of recommendations. One diversity metric is Hamming distance [53]. The average value of Hamming distance $H(L)$ is calculated by

$$H(L) = \frac{1}{m(m-1)} \sum_{i=1}^{m} \sum_{j=1}^{m} (1 - \frac{C(i,j)}{L}), \tag{12}$$

where $C(i,j) = |o_i^L \cap o_j^L|$ is the number of common objects in the lists of two users i and j with L recommended objects. Larger $H(L)$ value suggests higher diversity. The other diversity metric is Intra-similarity [54], which can be measured by the similarity between objects in the recommendation list. The average value of Intra-similarity $I(L)$ is calculated by

$$I(L) = \frac{1}{mL(L-1)} \sum_{i=1}^{m} \sum_{\substack{o_\alpha, o_\beta \in o_i^L \\ \alpha \neq \beta}} S_{\alpha\beta}^{Cos}, \tag{13}$$

where $S_{\alpha\beta}^{Cos}$ is the cosine similarity between objects α and β in the recommendation list o_i^L of user i with list's length L. Smaller $I(L)$ value suggests higher diversity of recommendations.

Novelty is a key metric for quantifying an algorithm's ability to generate novel (i.e., unpopular) and unexpected results [2]. Here, we use the Popularity of the recommended objects to estimate the novelty of recommendations. The average Popularity $N(L)$ is calculated by

$$N(L) = \frac{1}{mL} \sum_{i=1}^{m} \sum_{o_\alpha \in o_i^L} k_\alpha, \tag{14}$$

where k_α is the degree of object α in the recommendation list. Smaller $N(L)$ value suggests higher novelty.
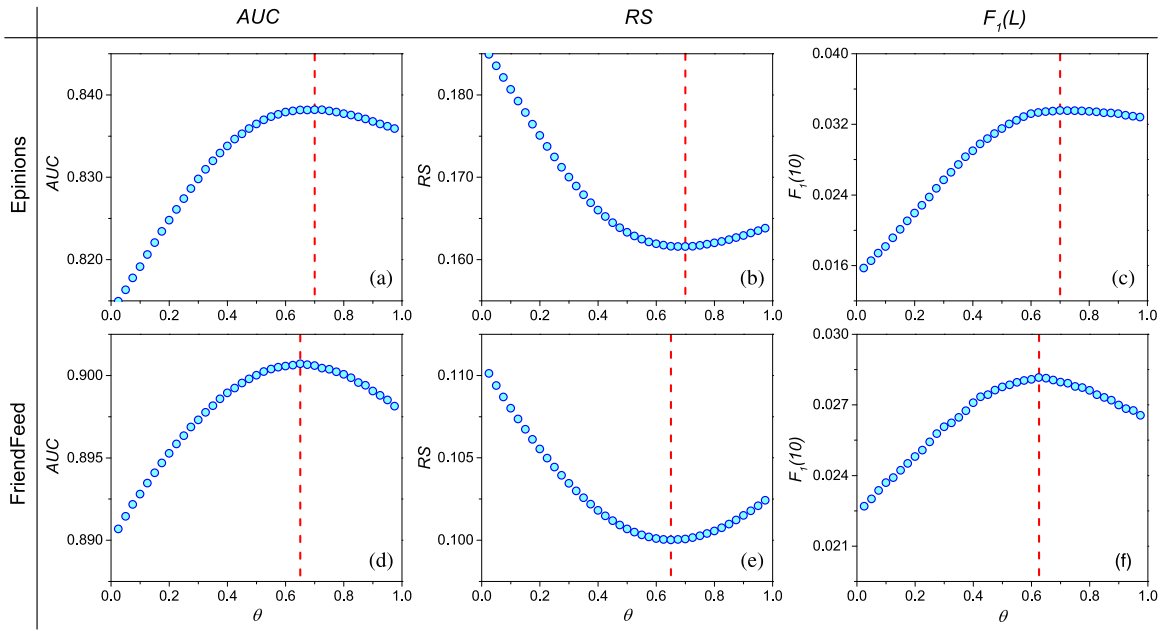
4. Results

In this section, we will first present the results on analyzing the effects of the scaling parameter in the proposed CosRA+T method. Then, we will show the results on the recommendation performance of the CosRA+T method and the five benchmark methods. Next, we study how the length of recommendation list affects the algorithmic performance. Finally, we presents the results to help understand the mechanisms of these methods.

4.1. The impact of scaling parameter

In the proposed CosRA+T method, a tunable scaling parameter θ is used to scale the resources that are received by the trusted users before the redistribution. The effects of users' trust relations on the resources-allocation processes increases as the scaling parameter θ decreases from 1 to 0. To explore the impact of the scaling parameter on the recommendation performance of the CosRA+T method, we vary θ from 0 to 1 and evaluate it on the Epinions and FriendFeed datasets by employing the three accuracy metrics: AUC, RS, and $F_1(L)$. Fig. 2 presents the corresponding results.

We observe that there seems to be an optimal value of the parameter θ for the best recommendation accuracy. This observation shows consistency on both Epinions and FriendFeed datasets under all the three accuracy metrics. Specifically, the values of AUC and $F_1(L)$ gradually increase as the increase of θ while slightly decrease after θ crosses its optimal value θ^*, which is around 0.70 for Epinions (see Fig. 2(a) and 2(c) for AUC and $F_1(L)$, respectively) and around 0.65 for FriendFeed (see Fig. 2(d) and 2(f) for AUC and $F_1(L)$, respectively). The value of RS decreases (i.e., the recommendation accuracy increases) as θ increases before θ crossing its optimal value $\theta^* \approx 0.70$ for Epinions (see Fig. 2(b)) and $\theta^* \approx 0.65$ for FriendFeed (see Fig. 2(e)).

Fig. 2. Results on the impact of the scaling parameter. (a–c) and (d–f) are for Epinions and FriendFeed, respectively. In the vertical-axis, three accuracy metrics are used, including AUC, RS, and $F_1(L)$ (from the left to the right column). In the horizontal-axis, the scaling parameter θ varies from 0 to 1. The vertical dashed red lines make the optimal value θ^*, where the accuracy metrics reach their maximum. The length of the recommendation list is set as $L = 10$ in calculating the $F_1(L)$ metric. The results are based on a 10-fold cross-validation and averaged over 20 independent realizations.

To investigate the universality of the optimal value θ^* and to better determine its critical value, we additionally employ the two L-dependent accuracy metrics, namely, Precision and Recall. By varying the recommendation list's length L, we explore how the optimal value θ^* changes with respect to $P(L)$, $R(L)$ and $F_1(L)$ on both datasets and present the results in Fig. 3. We find that the optimal value θ^* of the scaling parameter are not sensitive to the recommendation list's length L. The optimal values θ^* are around 0.70 and 0.65 for Epinions and FriendFeed as indicated by the horizontal trends in Fig. 3(a) and 3(d), respectively. Similar trends are also observed for F_1 on both datasets (see Fig. 3(b) for Epinions and Fig. 3(e) for FriendFeed).

Further, we determine the optimal value θ^* by averaging the results under different L for the three L-dependent accuracy metrics and present the results in Fig. 3(a) and 3(d) for Epinions and FriendFeed, respectively. Together, the results for the L-independent accuracy metrics (AUC and RS) are also included. We found that the optimal values ($\langle\theta^*\rangle \approx 0.70$ for Epinions and $\langle\theta^*\rangle \approx 0.65$ for FriendFeed) are very close to each other under several accuracy metrics on different datasets, suggesting the universality of the optimal scaling in CosRA+T for its best performance. These results showed that the user trust relations really matter, which supports the hypothesis that information from trusted users could help predict users' preferences to objects. Furthermore, the presence of the optimal scaling θ^* suggests that integrating the trust relations can enhance the recommendation performance, however, relying too much on it may result in the opposite. In other words, the integration of the trust relations is like the two sides of one coin, when it comes to the improvement of recommendation accuracy.

4.2. Performance of recommendation

We further provide a more comprehensive evaluation of the proposed CosRA+T method and compare its performance with the five benchmark methods, namely, GR, UCF, HC, MD, and CosRA. We apply five accuracy evaluation metrics including the two L-independent metrics (AUC and RS) and the three L-dependent metrics ($P(L)$, $R(L)$ and $F_1(L)$). We also employ two L-dependent diversity metrics ($H(L)$ and $I(L)$) and one popularity metric ($N(L)$). In the experiments, we set $L = 10$ for all L-dependent metrics and analyze the impact of L in the next section. In the comparisons with benchmark methods, CosRA+T method uses the optimal value θ^* for each dataset across all the evaluation metrics. Results of the algorithmic performance are summarized in Table 2.

As shown in the first five columns of Table 2, the proposed CosRA+T method outperforms all the five benchmark methods on both Epinions and FriendFeed datasets. Specifically, CosRA+T has remarkable advantages to GR, HC and UCF, as suggested by its larger AUC value (0.8382 on Epinions and 0.9007 on FriendFeed). For the L-dependent metrics, CosRA+T has even better performance than these benchmark methods, for example, CosRA+T gives remarkably larger $P(L)$ and $R(L)$ values on both datasets. Also, CosRA+T gives competitive $F_1(L)$ value (0.0335) compared to that given by UCF (0.0259) and MD (0.0286)
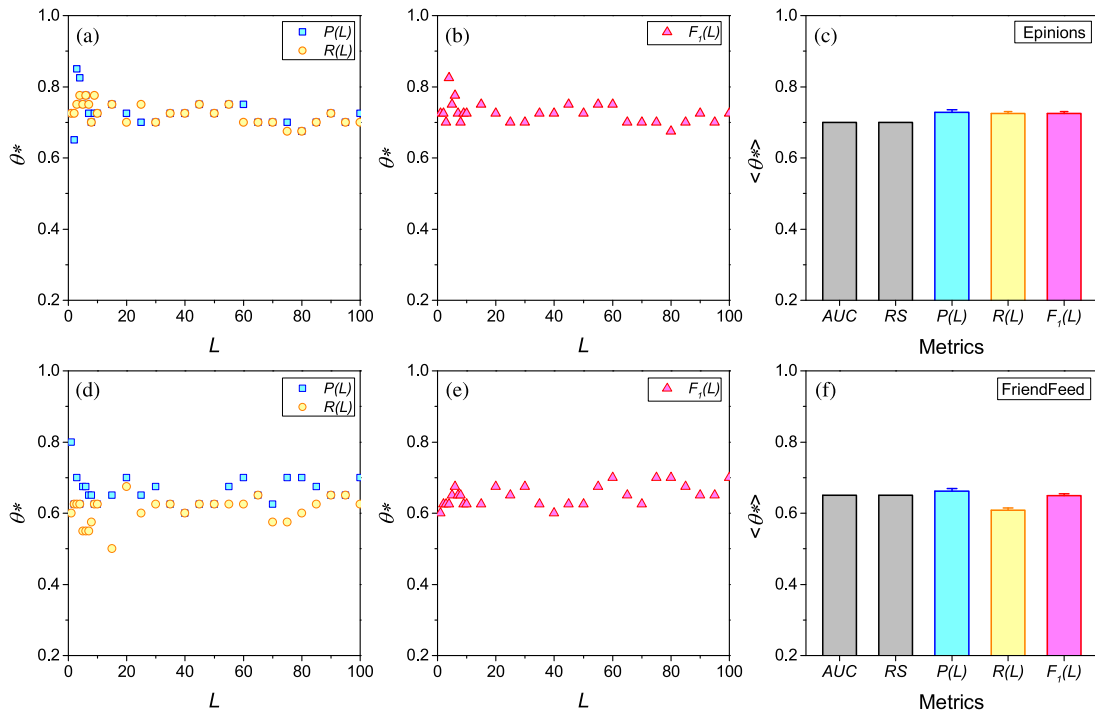
Fig. 3. Results on the optimal value of the scaling parameter. (a–c) and (d-f) are for Epinions and FriendFeed, respectively. In (a) and (d), the optimal values $\theta*$ in the vertical-axis are as a function of the recommendation list's length L in the horizontal-axis for both $P(L)$ and $R(L)$. In (b) and (e), the results are for $F_1(L)$. In (c) and (f), the average optimal value $\langle\theta*\rangle$ for each accuracy metric is presented with the error bar showing the standard error. The results are based on a 10-fold cross-validation and averaged over 20 independent realizations.

Table 2
Results of the performance evaluation metrics after applying different recommendation methods on the Epinions and FriendFeed datasets.

Methods	AUC	RS	$P(L)$	$R(L)$	$F_1(L)$	$H(L)$	$I(L)$	$N(L)$
Epinions								
GR	0.6974	0.3006	0.0094	0.0315	0.0144	0.1338	0.1389	308
HC	0.7845	0.2161	0.0052	0.0153	0.0077	**0.9742**	**0.0245**	**5**
MD	0.8256	0.1735	0.0189	0.0590	0.0286	0.6753	0.1140	235
UCF	0.8141	0.1844	0.0170	0.0537	0.0259	0.5748	0.1317	262
CosRA	0.8356	0.1641	0.0221	0.0629	0.0327	0.9472	0.0900	107
CosRA+T	**0.8382**	**0.1616**	**0.0226**	**0.0651**	**0.0335**	0.9544	0.0917	101
FriendFeed								
GR	0.6058	0.3921	0.0050	0.0215	0.0081	0.0739	0.0935	172
HC	0.8833	0.1182	0.0088	0.0370	0.0142	**0.9907**	**0.0542**	**11**
MD	0.8925	0.1077	0.0163	0.0683	0.0263	0.9422	0.1195	73
UCF	0.8869	0.1133	0.0155	0.0661	0.0252	0.8857	0.1616	92
CosRA	0.8978	0.1028	0.0167	0.0633	0.0265	0.9895	0.0890	35
CosRA+T	**0.9007**	**0.1000**	**0.0175**	**0.0693**	**0.0280**	0.9899	0.1008	35

Notes: The length of the recommendation list is set as $L = 10$. The scaling parameter θ in CosRA+T is set as its optimal value for each dataset. The results are based on a 10-fold cross-validation and averaged over 20 independent realizations. The best result of each metric is in bold.

on Epinions. Moreover, the values of *RS* given by CosRA+T are the smallest on both datasets, indicating that CosRA+T performs the best in recommendation accuracy among all the considered methods.

The diversity metrics are shown in the sixth and seventh columns of Table 2 for Hamming distance $H(L)$ and Intra-similarity $I(L)$, respectively. We notice that the recommendations given by HC have the best diversity as measured by diversity metrics on both Epinions and FriendFeed. Even though, CosRA+T still outperforms the other four benchmark methods as it gives larger $H(L)$ value (0.9544 on Epinions and 0.9899 on FriendFeed) and smaller $I(L)$ value (0.0917 on Epinions and 0.1008 on FriendFeed). The inferior of CosRA+T to HC is indeed very small, for example, the differences between $H(L)$ values are only 2.04% and 0.08% on Epinions and FriendFeed, respectively. Regarding the novelty of the recommendations, the performance of CosRA+T is also encouraging as it gives smaller $N(L)$ value than most benchmark methods except for HC (see the last column of Table 2). These results suggest that CosRA+T has overall larger accuracy, more diversity and better novelty.
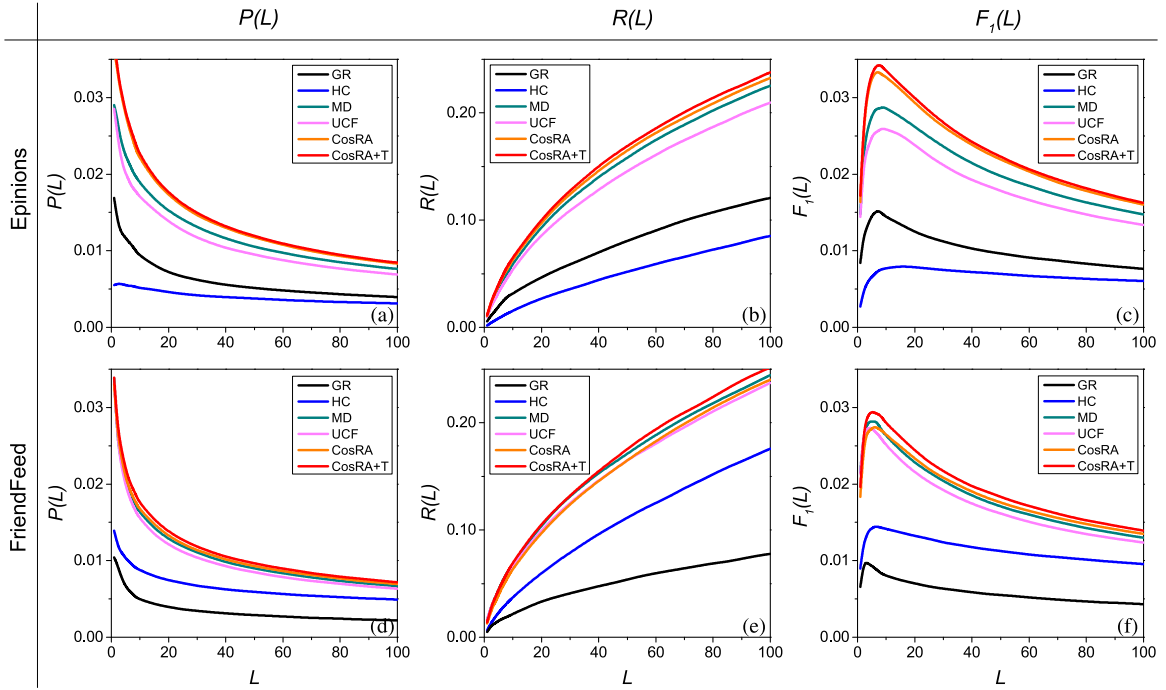
Fig. 4. Results of the recommendation accuracy affected by the list's length. (a–c) and (d–f) are for Epinions and FriendFeed, respectively. In the vertical-axis, three accuracy metrics are respectively shown. In the horizontal-axis, the length of the recommendation list L increases from 1 to 100. The scaling parameter θ in CosRA+T is set as its optimal value for each dataset. The results are based on a 10-fold cross-validation and averaged over 20 independent realizations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3. The impact of recommendation list's length

The length of recommendation list L may affect the evaluation of recommendation performance under L-dependent metrics including three accuracy metrics ($P(L)$, $R(L)$ and $F_1(L)$), two diversity metrics ($H(L)$ and $I(L)$), and one novelty metric ($N(L)$). To this point, we explore how the length L affects the performance of recommendation methods by varying L from 1 to 100. The method of interest is CosRA+T, and benchmark methods are considered for comparison.

Fig. 4 presents the results regarding the three accuracy metrics on both Epinions and FriendFeed datasets. We notice that, as the recommendation list's length L increases, the values of $P(L)$ decreases (see Fig. 4(a) and 4(d)) while the values of $R(L)$ increases (Fig. 4(b) and 4(e)) for all the considered methods. The values of F_1 first increases and then decreases as L increases, where the values of $F_1(L)$ reach their maximum at $L \approx 10$ on Epinions (see Fig. 4(c)) and $L \approx 5$ on FriendFeed (see Fig. 4(f)). Moreover, we find that the performance of CosRA+T is relatively better than other benchmark methods under different L, and its advantages are even remarkable when L is around its optimal value, for example, $L^* \approx 10$ for Epinions. Further, we notice that CosRA+T has a relative improvement compared to the original CosRA, MD and UCF are competitive to each other, and HC and GR have the lowest accuracy.

Fig. 5 presents the results regarding the two diversity metrics and the novelty metric. For the Hamming distance, the values of $H(L)$ slightly decrease as the increasing of L on both Epinions and FriendFeed except for MD and UCF, whose $H(L)$ values first decrease but then increase as L increases on the Epinions dataset (see Fig. 5(a)). HC always gives the best results, followed close by CosRA and CosRA+T. UCF is remarkably inferior to MD, and GR performs the worst. For the Intra-similarity, the values of $I(L)$ first increase rapidly but then decreases slowly as L increases (see Fig. 5(b) and 5(e)). UCF performs the worst as indicated by its largest $I(L)$ values. The performance of CosRA+T is ranked the second on Epinions but the middle on FriendFeed. HC outperforms all the other methods on both datasets as it gives the smallest $I(L)$ values. For the Novelty, the values of $N(L)$ decrease strongly as L increases at the beginning but the decreases become slow thereafter (see Fig. 5(c) and 5(f)). We notice that CosRA+T outperforms most of the benchmark methods in the novelty metric except for HC, and the result is not sensitive to L. These results suggest that CosRA+T has relatively higher diversity and better novelty in the recommendations of uncollected objects.

4.4. The analysis of mechanisms

In order to better understand the mechanisms of the CosRA+T method, we further focus on the degree distributions of the recommended objects for all users. For the purpose of comparison, three benchmark methods (HC, MD and CosRA) are also
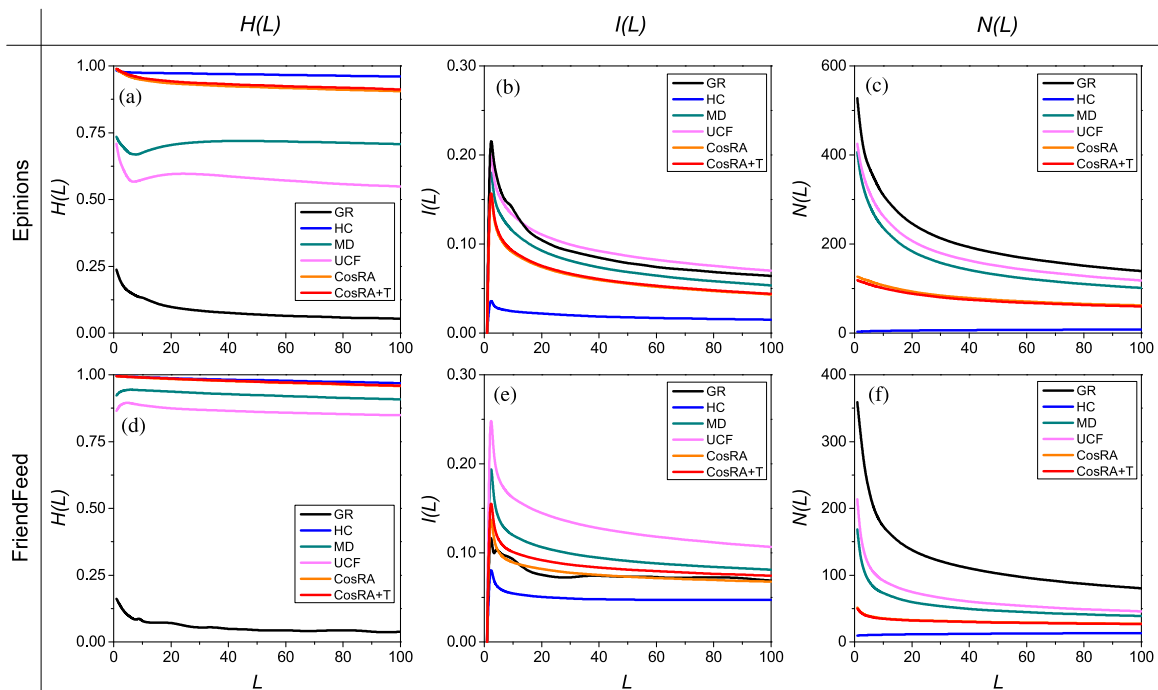
Fig. 5. Results of the recommendation diversity and novelty affected by the list's length. (a–c) and (d–f) are for Epinions and FriendFeed, respectively. In the vertical-axis, two diversity metrics and one novelty metric are respectively shown. In the horizontal-axis, the length of the recommendation list L increases from 1 to 100. The scaling parameter θ in CosRA+T is set as its optimal value for each dataset. The results are based on a 10-fold cross-validation and averaged over 20 independent realizations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

considered in the study. Two different recommendation list's lengths ($L = 10$ and $L = 50$) are used. The scaling parameter θ in CosRA+T is set as its optimal value for each dataset. The results are presented in Fig. 6(a–c) for Epinions and in Fig. 6(d–f) for FriendFeed, respectively.

We notice that small-degree objects have a high probability to be recommended by HC (see Fig. 6(a) and 6(e) for Epinions and FriendFeed, respectively), as suggested by the relatively small degrees of the recommended objects. By comparison, there is a large change for large-degree objects being recommended by MD (see Fig. 6(b) and 6(f)) as we can observe that the degrees of its recommended objects are relatively large. These results suggest the strong trends and the potential bias of both HC and MD. When it comes to the recommendations, this issue may result in their remarkable disadvantages, for example, the low accuracy of HC and the poor novelty of MD.

By adopting a novel similarity index, CosRA fortunately finds a balance among the recommendation diversity and accuracy by recommending objects of large degree and small degree at the same time (see Fig. 6(c) and 6(g) for Epinions and FriendFeed, respectively). The main reasons are that, in the calculation of the CosRA index, the effects of popular objects with large degrees are restricted, and the effects of small-degree users are enhanced. Further, by introducing the scaling parameter to enlarge the resources received by trusted users before the redistribution processes, CosRA+T relatively improves the algorithmic performance by recommending both large-degree and small-degree objects (see Fig. 6(d) and 6(h)), especially when the recommendation list's length L is around its optimal value.

5. Conclusion and discussions

In this paper, we explored the role of the trust relations among users in improving the performance of recommendation under the framework of network-based diffusion processes. Specifically, by introducing the trust relations into the original CosRA method, we proposed a trust-based recommendation method, named CosRA+T, in which the resources received by the trusted users are scaled by a tunable scaling parameter before being redistributed back to their collected objects in the networked resource redistribution process. We found an optimal scaling parameter for the proposed CosRA+T method, under which the method achieves the best accuracy in the recommendation. Interestingly, the optimal scaling values are very close to each other under different accuracy evaluation metrics on different rating and trust datasets. The result suggests the universality of the optimal scaling parameter in the proposed CosRA+T method for easy implementations. These findings thus showed that the user trust relations really matter, in support of the hypothesis that the information from the trusted users could be helpful in predicting users' preferences to objects.
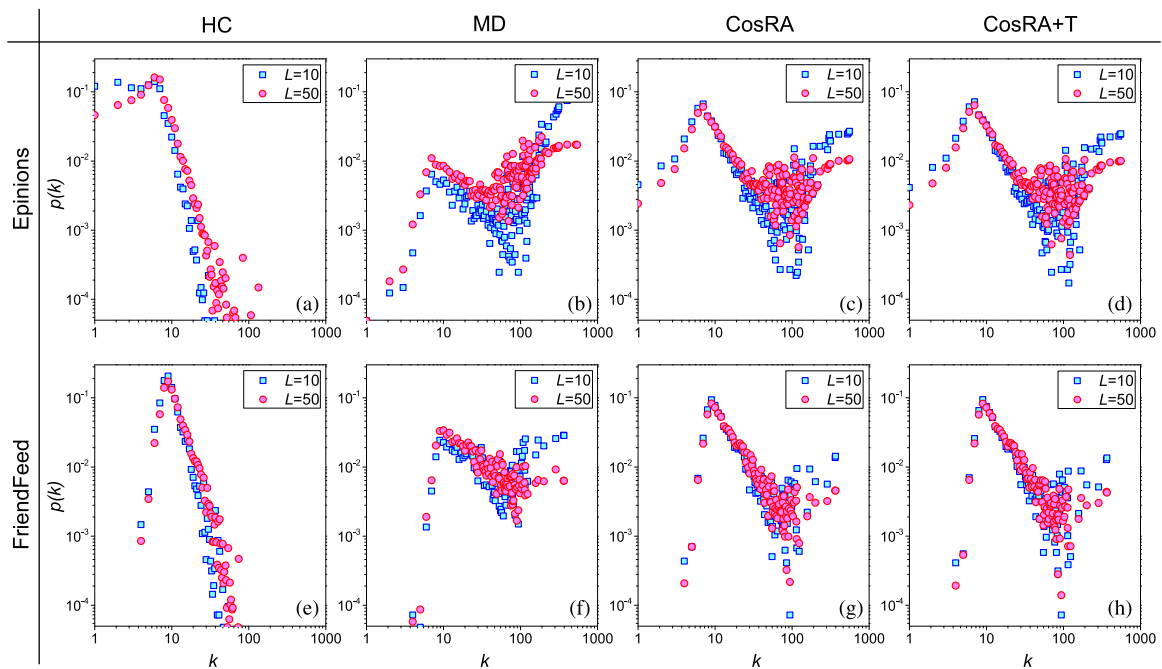
Fig. 6. Degree distribution of the recommended objects after applying HC, MD, CosRA and CosRA+T methods on the Epinions and FriendFeed datasets. Results are shown for one realization on each dataset in log–log plot. Blue squares and red circles correspond to results under the recommendation list's length $L = 10$ and $L = 50$, respectively.

Results of extensive experiments based on the two real-world rating and trust datasets, Epinions and FriendFeed, showed that the proposed CosRA+T method outperforms benchmark methods by giving a higher accuracy, an encouraging diversity and a well novelty in recommendations. Regarding the effects of the recommendation list's length on the performance evaluations under some parameter-dependent metrics (e.g., Precision and Recall), we found that the optimal lengths of the recommendation list are nearly the same on the same dataset for different methods while the optimal lengths may differ on different datasets. Finally, we tried to provide some insights to the mechanisms of some considered methods through presenting the degree distributions of their recommended objects. Results suggested that CosRA and CosRA+T balance well both small-degree and large-degree objects, leading to their better performance. Our work provides a promising step towards enhancing the recommendation performance by additionally considering users' social trust relations besides the traditionally used users' ratings.

The presented results are under some limitations on the datasets and the modeling process, which call for further improvements towards designing a better method that deals well with accuracy, diversity and novelty. The evaluation of algorithmic performance uses two rating datasets with trust relations, which are only represents of numerous real-world online rating platforms and socioeconomic systems [11]. It would be an improvement if recommendation methods could be comprehensively tested and compared on various datasets and even on real platforms, focusing on how different recommendation algorithms affect the long-term evolution of online systems [55]. Moreover, the proposed method uses a simple way to enhance the resources received by trust users by a scaling parameter before the redistribution in the network-based diffusion processes. Yet, other possible realizations of introducing the trust relations into the rating-based methods under different frameworks could also be considered [20,27], and the effects of different similarity measures on the recommendation performance could be further explored [56]. Besides, not only social relationships (e.g., trust relations among users) but also users' features (e.g., online reputation of users [57,58]) are critical to web-based recommender systems. As future works, a promising step is to combine social trust information and users' reputations [59,60] to further improve the performance of personalized recommendation algorithms for real-world applications.

Acknowledgments

References

[1] J.B. Schafer, J. Konstan, J. Riedl, Recommender systems in e-commerce, in: Proceedings of the 1st ACM Conference on Electronic Commerce, ACM Press, New York, NY, USA, 1999, pp. 158–166.

[2] L. Lü, M. Medo, C.H. Yeung, Y.-C. Zhang, Z.-K. Zhang, T. Zhou, Recommender systems, Phys. Rep. 519 (1) (2012) 1–49.

[3] M.D. Clemes, C. Gan, J. Zhang, An empirical analysis of online shopping adoption in Beijing, China, J. Retail. Consumer Serv. 21 (3) (2014) 364–375.

[4] J. Gao, T. Zhou, Quantifying China's regional economic complexity, Physica A 492 (2018) 1591–1603.

[5] Y.-C. Chen, R.-A. Shang, C.-Y. Kao, The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment, Electron. Commerce Res. Appl. 8 (1) (2009) 48–58.

[6] J.-G. Liu, T. Zhou, B.-H. Wang, Research progress of personalized recommendation system, Progr. Natur. Sci. 19 (1) (2009) 1–15.

[7] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, Recommender systems survey, Knowl.-Based Syst. 46 (1) (2013) 109–132.

[8] C.-J. Zhang, A. Zeng, Behavior patterns of online users and the effect on information filtering, Physica A 391 (4) (2012) 1822–1830.

[9] L. Jin, Y. Chen, T. Wang, P. Hui, Understanding user behavior in online social networks: A survey, IEEE Commun. Mag. 51 (9) (2013) 144–150.

[10] C.L.P. Chen, C.-Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, Inform. Sci. 275 (2014) 314–347.

[11] J. Gao, T. Zhou, Big data reveal the status of economic development, J. Univ. Electron. Sci. Technol. China 45 (4) (2016) 625–633.

[12] X. Yang, J. Gao, J.-H. Liu, T. Zhou, Height conditions salary expectations: Evidence from large-scale data in China, Physica A 501 (2018) 86–97.

[13] Z.-K. Zhang, C. Liu, Y.-C. Zhang, T. Zhou, Solving the cold-start problem in recommender systems with social tags, Europhys. Lett. 92 (2) (2010) 28002.

[14] L. Yu, J. Huang, G. Zhou, C. Liu, Z.-K. Zhang, TIIREC: A tensor approach for tag-driven item recommendation with sparse user generated content, Inform. Sci. 411 (2017) 122–135.

[15] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, Grouplens: An open architecture for collaborative filtering of netnews, in: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, ACM Press, New York, NY, USA, 1994, pp. 175–186.

[16] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: Proceedings of the 10th International Conference on World Wide Web, ACM Press, New York, NY, USA, 2001, pp. 285–295.

[17] Q. Guo, W.-J. Song, J.-G. Liu, Ultra-accurate collaborative information filtering via directed user similarity, Europhys. Lett. 107 (1) (2014) 18001.

[18] Y.-C. Zhang, M. Blattner, Y.-K. Yu, Heat conduction process on community networks as a recommendation model, Phys. Rev. Lett. 99 (15) (2007) 154301.

[19] Y.-C. Zhang, M. Medo, J. Ren, T. Zhou, T. Li, F. Yang, Recommendation model based on opinion diffusion, Europhys. Lett. 80 (6) (2007) 68003.

[20] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J.R. Wakeling, Y.-C. Zhang, Solving the apparent diversity-accuracy dilemma of recommender systems, Proc. Natl. Acad. Sci. USA 107 (10) (2010) 4511–4515.

[21] J.-G. Liu, Q. Guo, Y.-C. Zhang, Information filtering via weighted heat conduction algorithm, Physica A 390 (12) (2011) 2414–2420.

[22] L. Lü, W. Liu, Information filtering via preferential diffusion, Phys. Rev. E 83 (2) (2011) 066119.

[23] C. Liu, W.-X. Zhou, Heterogeneity in initial resource configurations improves a network-based hybrid recommendation algorithm, Physica A 391 (22) (2012) 5704–5711.

[24] A. Zeng, A. Vidmer, M. Medo, Y.-C. Zhang, Information filtering by similarity-preferential diffusion processes, Europhys. Lett. 105 (5) (2014) 58002.

[25] W.-J. Song, Q. Guo, J.-G. Liu, Improved hybrid information filtering based on limited time window, Physica A 416 (2014) 192–197.

[26] Q. Ou, Y.-D. Jin, T. Zhou, B.-H. Wang, B.-Q. Yin, Power-law strength-degree correlation from resource-allocation dynamics on weighted networks, Phys. Rev. E 75 (2) (2007) 021102.

[27] T. Zhou, J. Ren, M. Medo, Y.-C. Zhang, Bipartite network projection and personal recommendation, Phys. Rev. E 76 (4) (2007) 046115.

[28] L.-J. Chen, Z.-K. Zhang, J.-H. Liu, J. Gao, T. Zhou, A vertex similarity index for better personalized recommendation, Physica A 466 (2017) 607–615.

[29] F. Yu, A. Zeng, S. Gillard, M. Medo, Network-based recommendation algorithms: A review, Physica A 452 (2016) 192–208.

[30] Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, Y.-C. Zhang, Dynamics of information diffusion and its applications on complex networks, Phys. Rep. 651 (2016) 1–34.

[31] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, H.W. Lauw, Detecting product review spammers using rating behaviors, in: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, ACM Press, New York, NY, USA, 2010, pp. 939–948.

[32] L. Yu, C. Liu, Z.-K. Zhang, Multi-linear interactive matrix factorization, Knowl.-Based Syst. 85 (2015) 307–315.

[33] P. Massa, P. Avesani, Trust-aware recommender systems, in: Proceedings of the 2007 ACM Conference on Recommender Systems, ACM Press, New York, NY, USA, 2007, pp. 17–24.

[34] H. Ma, M.R. Lyu, I. King, Learning to recommend with trust and distrust relationships, in: Proceedings of the Third ACM Conference on Recommender Systems, ACM Press, New York, NY, USA, 2009, pp. 189–196.

[35] N.B. Ellison, C. Steinfield, C. Lampe, The benefits of facebook "friends": Social capital and college students' use of online social network sites, J. Comput.-Mediat. Commun. 12 (4) (2007) 1143–1168.

[36] J. Yuan, Q.-M. Zhang, J. Gao, L. Zhang, X.-S. Wan, X.-J. Yu, T. Zhou, Promotion and resignation in employee networks, Physica A 444 (2016) 442–447.

[37] D. Centola, The spread of behavior in an online social network experiment, Science 329 (5996) (2010) 1194–1197.

[38] J. Gao, T. Zhou, Y. Hu, Bootstrap percolation on spatial networks, Sci. Rep. 5 (2015) 14662.

[39] Q. Wang, J. Gao, T. Zhou, Z. Hu, H. Tian, Critical size of ego communication networks, Europhys. Lett. 114 (5) (2016) 58004.

[40] H. Liao, G. Cimini, M. Medo, Measuring quality, reputation and trust in online communities, in: L. Chen, A. Felfernig, J. Liu, Z.W. Raś (Eds.), Foundations of Intelligent Systems, Springer, Berlin, Heidelberg, 2012, pp. 405–414.

[41] J. O'Donovan, B. Smyth, Trust in recommender systems, in: Proceedings of the 10th International Conference on Intelligent User Interfaces, ACM Press, New York, NY, USA, 2005, pp. 167–174.

[42] F.E. Walter, S. Battiston, F. Schweitzer, A model of a trust-based recommendation system on a social network, Autonom. Agents Multi-Agent Syst. 16 (1) (2008) 57–74.

[43] M. Jamali, M. Ester, Trustwalker: A random walk model for combining trust-based and item-based recommendation, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM Press, New York, NY, USA, 2009, pp. 397–406.

[44] R. Andersen, C. Borgs, J. Chayes, U. Feige, A. Flaxman, A. Kalai, V. Mirrokni, M. Tennenholtz, Trust-based recommendation systems: An axiomatic approach, in: Proceedings of the 17th International Conference on World Wide Web, ACM Press, New York, NY, USA, 2008, pp. 199–208.

[45] H. Ma, I. King, M.R. Lyu, Learning to recommend with explicit and implicit social relations, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 1–19.

[46] X. Shen, H. Long, C. Ma, Incorporating trust relationships in collaborative filtering recommender system, in: 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, IEEE Press, 2015, pp. 1–8.

[47] G. Guo, J. Zhang, N. Yorkesmith, A novel recommendation model regularized with user trust and item ratings, IEEE Trans. Knowl. Data Eng. 28 (7) (2016) 1607–1620.

[48] X. Wang, Y. Liu, G. Zhang, F. Xiong, J. Lu, Diffusion-based recommendation with trust relations on tripartite graphs, J. Stat. Mech. Theory Exp. 2017 (8) (2017) 083405.

[49] R.-R. Liu, C.-X. Jia, T. Zhou, D. Sun, B.-H. Wang, Personal recommendation via modified collaborative filtering, Physica A 388 (4) (2009) 462–468.

[50] J.A. Hanley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (roc) curve, Radiology 143 (1) (1982) 29–36.

[51] T. Zhou, L. Lü, Y.-C. Zhang, Predicting missing links via local information, Eur. Phys. J. B 71 (4) (2009) 623–630.

[52] J.L. Herlocker, J.A. Konstan, L.G. Terveen, J.T. Riedl, Evaluating collaborative filtering recommender systems, ACM Trans. Inf. Syst. 22 (1) (2004) 5–53.

[53] T. Zhou, L.-L. Jiang, R.-Q. Su, Y.-C. Zhang, Effect of initial configuration on network-based recommendation, Europhys. Lett. 81 (5) (2008) 58004.

[54] T. Zhou, R.-Q. Su, R.-R. Liu, L.-L. Jiang, B.-H. Wang, Y.-C. Zhang, Accurate and diverse recommendations via eliminating redundant correlations, New J. Phys. 11 (12) (2009) 123008.

[55] D.-D. Zhao, A. Zeng, M.-S. Shang, J. Gao, Long-term effects of recommendation on the evolution of online systems, Chin. Phys. Lett. 30 (11) (2013) 118901.

[56] J.-G. Liu, L. Hou, X. Pan, Q. Guo, T. Zhou, Stability of similarity measurements for bipartite networks, Sci. Rep. 6 (2016) 18653.

[57] J. Gao, Y.-W. Dong, M.-S. Shang, S.-M. Cai, T. Zhou, Group-based ranking method for online rating systems with spamming attacks, Europhys. Lett. 110 (2) (2015) 28003.

[58] X.-L. Liu, J.-G. Liu, K. Yang, Q. Guo, J.-T. Han, Identifying online user reputation of user-object bipartite networks, Physica A 467 (2017) 508–516.

[59] H. Liao, A. Zeng, R. Xiao, Z.-M. Ren, D.-B. Chen, Y.-C. Zhang, Ranking reputation and quality in online rating systems, PLoS One 9 (5) (2014) e97146.

[60] J. Gao, T. Zhou, Evaluating user reputation in online rating systems via an iterative group-based ranking method, Physica A 473 (2017) 546–560.