Research Paper

JIS

Journal of Information Science
1–36

DOI: 10.1177/0165551518808191
journals.sagepub.com/home/jis
SAGE

An effective social recommendation method based on user reputation model and rating profile enhancement

Sajad Ahmadian
Department of Computer Engineering, University of Zanjan, Iran

Mohsen Afsharchi
Department of Computer Engineering, University of Zanjan, Iran

Majid Meghdadi
Department of Computer Engineering, University of Zanjan, Iran

Abstract
Trust-aware recommender systems are advanced approaches which have been developed based on social information to provide relevant suggestions to users. These systems can alleviate cold start and data sparsity problems in recommendation methods through trust relations. However, the lack of sufficient trust information can reduce the efficiency of these methods. Moreover, diversity and novelty are important measures for providing more attractive suggestions to users. In this article, a reputation-based approach is proposed to improve trust-aware recommender systems by enhancing rating profiles of the users who have insufficient ratings and trust information. In particular, we use a user reliability measure to determine the effectiveness of the rating profiles and trust networks of users in predicting unseen items. Then, a novel user reputation model is introduced based on the combination of the rating profiles and trust networks. The main idea of the proposed method is to enhance the rating profiles of the users who have low user reliability measure by adding a number of virtual ratings. To this end, the proposed user reputation model is used to predict the virtual ratings. In addition, the diversity, novelty and reliability measures of items are considered in the proposed rating profile enhancement mechanism. Therefore, the proposed method can improve the recommender systems about the cold start and data sparsity problems and also the diversity, novelty and reliability measures. Experimental results based on three real-world datasets show that the proposed method achieves higher performance than other recommendation methods.

Keywords
Diversity; novelty; profile enhancement; reliability; reputation; trust

1. Introduction

With excessive information available on commercial sites, recommender systems have attracted much attention to overcome information overload problem. These systems help users to find the most interesting items among a large number of choices. The basic idea of recommender systems is to collect the historic ratings of target user about items to predict unseen items. Then, a list of the items with highest ratings can be suggested to the target user as recommendation list [1–3].

In recent years, several approaches have been developed for recommender systems, among which collaborative filtering (CF) is one of the most popular and important approaches [4,5]. Generally, the CF-based methods can be classified into two groups including memory-based and model-based approaches. In the memory-based methods, the entire ratings of user-item matrix are used to compute similarity values between users/items. Then, a set of users/items is formed as

Corresponding author:
Majid Meghdadi, Department of Computer Engineering, University of Zanjan, 45371-38791 Zanjan, Iran.
Email: meghdadi@znu.ac.ir

nearest neighbours to predict unseen items for the target users [6,7]. On the other hand, the model-based methods are based on constructing a model using a portion of data as training set to predict ratings that exist in remaining data as test set. These models can be constructed based on a specific method such as clustering [8,9], matrix factorization [10,11], latent semantic models [12,13]. Although the CF method is very popular in recommender systems, it suffers from some shortcomings such as cold start and data sparsity problems. Cold start problem occurs when a user or item in the system has expressed or received a few number of ratings [14,15]. Moreover, data sparsity problem refers to the sparsity of ratings that the recommender systems face, since the number of items is usually millions and users can provide ratings for small portions of these items [16,17]. Therefore, the similarity values between the users/items cannot be calculated correctly and these values have not a high reliability value [18,19]. These problems make to reduce the performance of recommender systems in rating prediction process.

Trust-aware recommender systems have been proposed to overcome the problems of the CF methods using social network information such as trust and friendship among the users [20,21]. The main idea of the systems is to use trust or friend relations as additional information to provide more accuracy and personalised recommendations for users. The trust relations can be represented into two types including explicit and implicit trust statements. The explicit trust refers to the social relations which are explicitly established by the users [22,23]. Moreover, the implicit trust can be extracted implicitly among the users on the basis of their ratings to the items. In other words, the implicit trust networks between the users are identified on the basis of how the users rate the items in the system [24,25].

Although many trust-aware methods have been proposed to achieve improvements for recommender systems, there are still some problems to be considered in these methods. At first, the users may not have sufficient trust relations to use in recommendation process. Therefore, the poor trust networks of the users may lead to reduce the performance of the recommender systems [26]. For instance, suppose that there is no common rating between a target user and other users in the user-item ratings matrix. Moreover, there is no trust relation between the target user and other users in the system. In such condition, the recommender system cannot calculate the similarity values between the target user and others. Therefore, the system cannot provide recommendations for the target user [26]. Second, the diversity and novelty measures of recommendations can be useful to improve the performance of the trust-aware recommender systems for suggesting novel and interesting items to the users. However, most of the trust-aware recommender systems in the literature fail to consider these important measures into the recommendation process [24,27]. Finally, the accuracy of predictions for these systems can be improved by considering reliability measures about the users and items [22].

To address the mentioned problems, we propose a novel reputation-based approach to improve trust-aware recommender systems by enhancing rating profiles of the users who have insufficient ratings and trust information. To this end, a novel user reliability measure is proposed to evaluate the quality of the rating profiles and trust networks of the users in predicting unseen items. This measure helps to find the rating profiles with low reliability in predicting unseen items. Then, the rating profiles of the users who have low user reliability measure are enhanced by adding a number of virtual ratings. These virtual ratings are predicted based on a proposed reputation model for the users. Therefore, the proposed method alleviates data sparsity and cold start problems in trust-aware recommender systems specially when there are no common ratings and also trust relations between the target user and other users in the system. In addition, the diversity, novelty and reliability measures of items are considered for selecting a suitable subset of the virtual ratings to add into the rating profile of the target user. Finally, the similarity values between the users are calculated based on the enhanced rating profiles and the unseen items can be predicted using nearest neighbour set of the target user. Therefore, the proposed method can improve the performance of trust-aware recommender systems in the cases of the cold start and data sparsity problems with considering the diversity, novelty and reliability measures.

The main contributions of the paper are as follows:

1. A novel user reliability measure is proposed to evaluate the quality of the rating profiles and trust networks of the users in predicting unseen items.
2. A novel mechanism is proposed to enhance the rating profiles of the users who have low user reliability measure by adding a number of virtual ratings to their rating profiles. The main advantage of this mechanism is to alleviate cold start and data sparsity problems in recommender systems.
3. A novel user reputation model is introduced based on the combination of the rating profiles and trust networks of the users. This reputation model is used to predict the virtual ratings in the proposed rating profile enhancement mechanism.
4. The diversity, novelty and reliability measures of items are considered to calculate a measurement for selecting an appropriate subset of the items to add into the user's rating profile.

5. To show the improvement of the proposed method in comparison with other methods, a number of experiments are performed on three real-world datasets. It can be concluded from these experiments that the proposed method improves the performance of the recommender systems in terms of several evaluation measures.

The remainder of the paper is organised as follows: Section 2 presents the overview of related works. Section 3 introduces the details of the proposed method. Section 4 presents the experiments results based on three well-known datasets. Finally, Section 5 concludes this article based on future research directions and challenges.

2. Related works

Recently, with the development of online social shopping websites and also growing the amount of information available in these websites, users are forced to spend a lot of time for finding their required information and items. Therefore, recommender systems have been promoted to resolve information overload problem and also enhance customer satisfaction in the ecommerce websites. The main purpose of the recommender systems is to prevent from wasting time of the users and also suggest relevant and interesting items to them [28,29].

Social relations such as trust and friendship have been incorporated in recommender systems to improve the accuracy of the predicted ratings for users [24,30]. In Chen et al. [31], a cold start recommendation method for the new user is proposed which integrates a user model with trust and distrust statements to identify expert users. Then, the recommendation process for the cold start new users is improved by aggregating the suggestions of the expert users. In Guo et al. [32], a clustering-based recommendation method is proposed which the users are clustered from the views of both rating and social trust information. Moreover, a support vector regression model is used to predict unseen items for the users who appear in two different clusters. Also, a probabilistic method is proposed for suggesting recommendations to the cold start users who cannot be clustered due to insufficient data. In Wu et al. [20], the authors proposed a compound recommendation system for social media systems based on social information, item contents and user feedbacks. To this end, the collaborative topic regression model is extended to incorporate trust information, topic modelling and probabilistic matrix factorization. Lee and Ma [33] proposed a hybrid approach based on user ratings and social trust information for making better recommendations to users. In addition, a combination of *k*-nearest neighbours and matrix factorization methods with considering distrust relations between the users is used to maximise the performance of the recommender system.

Several social recommendation methods have been proposed based on matrix factorization techniques [27,34,35]. In Yang et al. [27], a recommendation method is proposed to improve the performance of the CF by integrating twofold sparse information including the conventional rating data and the social information among the users. To this end, a matrix factorization technique is used to map the users into low-dimensional latent feature spaces in terms of their trust statements to reflect users' reciprocal influence on their own opinions more reasonably. Guo et al. [34] analysed the social trust information from four real-world datasets and showed that not only the explicit but also the implicit influence of both ratings and trust should be taken into consideration to improve recommendation process. Therefore, they proposed a trust-based matrix factorization method which involves the explicit and implicit influence of rated items, by further incorporating both of the explicit and implicit influence of trusted users to predict unseen items for a target user. In Jamali and Ester [35], the authors proposed a model-based recommendation method by employing matrix factorization techniques to suggest relevant items to users in social networks. To this end, a mechanism for trust propagation is incorporated into the recommendation process. Moreover, it is shown that the trust propagation can be a crucial phenomenon in the social sciences, social network analysis and trust-based recommendation methods.

Davoudi and Chatterjee [36] proposed a social recommendation method based on similarity, centrality and social relationships. To this end, the probabilistic matrix factorization approach is used to predict user rating for products. Moreover, some centrality metrics are considered in the proposed method including degree, eigen-vector, Katz and PageRank. In Lee et al. [37], a new latent feature is proposed for social recommender systems. Moreover, two novel algorithms are developed based on the proposed latent feature. The main idea of the method is that trustors who follow the same trustee have features in common. In Li et al. [38], the authors proposed a social matrix factorization method based on both user latent feature space and user-item rating space. Moreover, a context-aware model is proposed based on Gaussian mixture model to alleviate data sparsity problem. In Ghavipour and Meybodi [39], a new trust aggregation strategy is proposed based on the standard CF to aggregate trust values of multiple paths. Moreover, a heuristic algorithm based on learning automata is used for discovering reliable paths between two users and inferring the value of trust. Lingam et al. [40] proposed a social recommendation method based on direct and indirect trust values between the users. To this end, a high quality model is considered to estimate utility values with associated weights based on Shannon entropy information gain. Moreover, a trust path selection mechanism is proposed based on learning automata

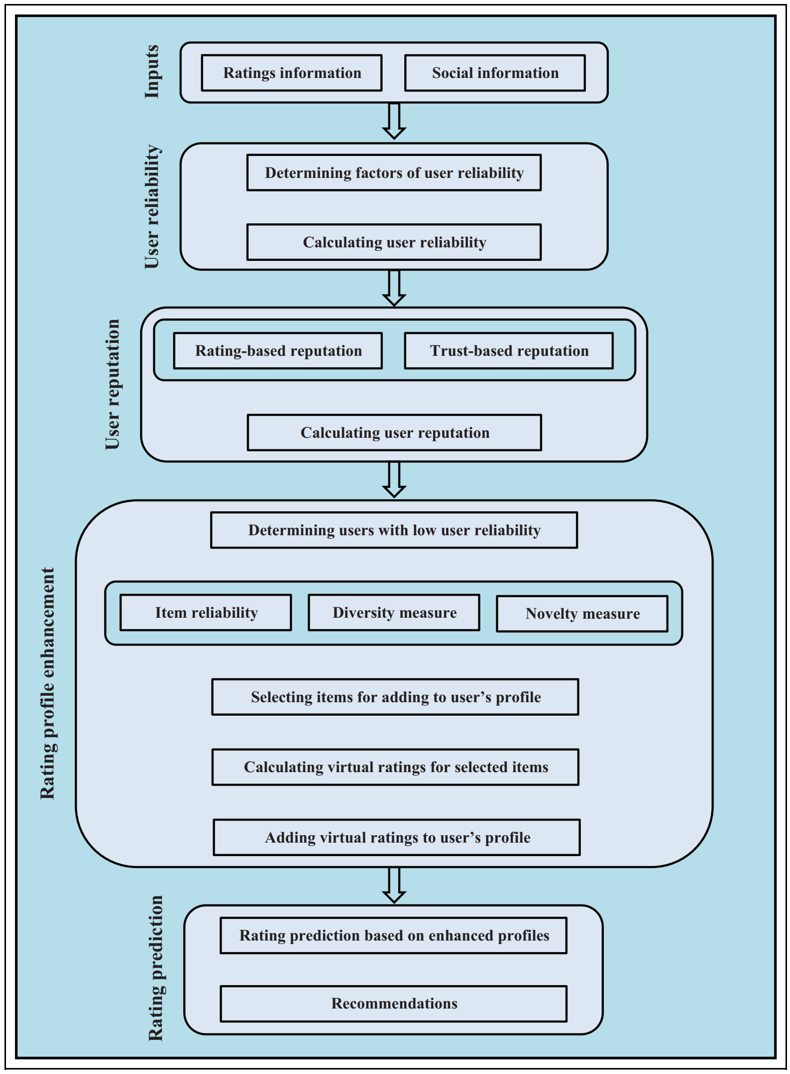

Figure 1. Overview of the proposed method.

to identify multiple recommended trust paths and to determine an aggregate path. In Chen and Gao [41], the authors proposed a trust-based recommendation method which integrates the information of trust relations into the resource-redistribution process. Moreover, a tunable parameter is considered to scale the resources received by trusted users. To achieve the best recommendation accuracy, an optimal scaling parameter is determined for the proposed method. In Faridani et al. [42], a trust-based recommendation method is proposed based on aggregating the trusted neighbours of the target users to alleviate data sparsity problem. To this end, the MoleTrust algorithm is used to provide more similar users into the recommendation process. Li et al. [43] proposed a social recommendation method based on the combination of social tags and trust relations to alleviate data sparsity and cool boot problems. The method uses social trust relations, item tag information and user rating matrix based on probabilistic matrix factorization to connect all the data resources from different dimensions.

Considering diversity and novelty measures into recommendation process can improve users' satisfaction about the recommender systems. Therefore, several methods have been proposed in the literature to improve the diversity and novelty measures of the recommender systems [44–47]. In Liu et al. [44], a novel trust-aware recommender system is proposed which incorporates time factor into similarity function. The main idea of the method is that the users with later creation time of trust can bring more diverse items for suggesting to the other users in their trust networks. A threshold-based recommendation method is proposed in Servajean et al. [45] to return the most relevant and popular items while satisfying content and profile diversity measures. For this purpose, a number of techniques are used to efficiently suggest the relevant and interesting items to users. In Zhang and Hurley [46], the authors proposed a recommendation method to improve the diversity of the retrieved list while considering a binary optimisation problem for maintaining adequate similarity to the user query. Then, a solution strategy is proposed to resolve this optimisation problem by relaxing it to a trust-region problem. A novel graph-based recommendation method is proposed in Lee and Lee [47] which uses only positively rated items in users' profiles to construct an undirected graph. To this end, the items are considered as nodes and positive correlations are considered as edges in the constructed graph. Moreover, the novel and relevant recommendations can be suggested to users using entropy concept and the linked items in the graph.

Many approaches have been proposed to consider reliability and confidence measures in recommender systems [22,26,48,49]. In Guo et al. [26], a novel recommendation method is proposed to address cold start problem by merging the ratings of a user's trusted neighbours to complement and represent the preferences of the user. Moreover, a confidence metric is used in prediction process to measure the quality of the merged ratings. A novel approach is proposed in Zhang et al. [48] to calculate uncertainty of predictions based on two key factors including posterior rating distribution and confidence level of the predicted ratings. The accuracy of recommendations can be improved through incorporating this uncertainty information. In Moradi and Ahmadian [22], the authors proposed a trust-based recommendation method based on a reliability measure to improve the accuracy of predictions. This reliability measure is used to evaluate the predicted ratings and also reconstruct the trust networks of the users. In other words, the trust networks of the users are reconstructed for the predicted ratings which their reliability measures are lower than a threshold value.

3. Proposed method

In this section, we aim to propose a social recommendation method called Reputation-based Trust-Aware Recommender System (in short RTARS) with considering diversity, novelty and reliability measures of items for providing recommendations to users. The overview of the proposed method is shown in Figure 1. In the proposed method, first of all, a user reliability measure is introduced to evaluate the performance of the user's ratings and trust network in predicting unseen items. This measure is based on the combination of similarity values and trust relations between the users and also four different factors. Then, the rating profiles of the users with low user reliability measures are enhanced using a novel rating profile enhancement mechanism. To this end, a number of virtual ratings are added to the user's rating profile. The virtual ratings are calculated using a proposed user reputation model which is based on the combination of user ratings and trust information. Three different measures including diversity, novelty and item reliability are considered for adding the virtual ratings to the user's rating profile. Finally, the similarity values between the users are calculated based on the enhanced rating profiles and the unseen items are predicted using these similarity values. In the following subsections, additional details about the proposed method are discussed.

3.1. User reliability

The users' rating profiles have different abilities to predict unseen items. On the other hand, the users' trust networks may not have a high performance to find nearest neighbours set of the users. Therefore, calculating the performance of the user's rating profile and also user's trust network in predicting unseen items can be helpful to evaluate the quality of these sources. In this section, a user reliability measure is introduced to evaluate the performance of the rating profiles and trust networks of the users. The main idea of the proposed user reliability measure is to identify the users who have the rating profiles and trust networks with low performance to predict unseen items and also propose a mechanism for enhancing the rating profiles of these users. The proposed user reliability measure is based on four different factors which are described in following.

The first factor for calculating the proposed user reliability measure is the number of ratings that active user a has assigned to the items. The more ratings for the active user make a higher performance for rating profile of the user in predicting unseen items. Because, the similarity values between the active user and others can be calculated easily. Therefore, this factor has a positive effect on the user reliability measure and can be calculated as follows [48]

$$f_i(I_a) = 1 - \frac{\bar{a}}{\bar{a} + |I_a|} \tag{1}$$

where $|I_a|$ is the number of ratings that the active user a has assigned to the items and $\bar{a}$ is the median of the values for $|I_a|$.

The number of users in the neighbourhood set of the active user is the second factor that can influence on the user reliability measure. In the proposed method, the neighbourhood set of the active user is formed based on the combination of similarity values and trust relations between the users. The similarity value between a pair of the users is calculated using Pearson coefficient function as follows

$$sim(a,u) = \frac{\sum_{i \in A_{a,u}} (r_i(a) - \bar{r}(a))(r_i(u) - \bar{r}(u))}{\sqrt{\sum_{i \in A_{a,u}} (r_i(a) - \bar{r}(a))^2} \sqrt{\sum_{i \in A_{a,u}} (r_i(u) - \bar{r}(u))^2}} \tag{2}$$

where $r_i(a)$ is the rate of item i given by user a, $\bar{r}(a)$ is the average of the rates given by user a and $A_{a,u}$ is the set of items which are rated by both users a and u.

Moreover, the trust value between a pair of the users can be calculated as follows

$$T_{a,u} = \frac{d_{max} - d_{a,u} + 1}{d_{max}} \tag{3}$$

where $T_{a,u}$ is the trust value between the users a and u, d_{max} is the maximum allowable propagation distance between the users and $d_{a,u}$ shows the trust propagation distance between the users a and u [50].

Finally, a combination of the similarity and trust values between a pair of the users is calculated using equation (4)

$$W_{a,u} = \begin{cases} \dfrac{2 \times sim(a,u) \times T_{a,u}}{sim(a,u) + T_{a,u}} & if\ sim(a,u) + T_{a,u} \neq 0 \\ & and\ sim(a,u) \times T_{a,u} \neq 0 \\ T_{a,u} & else\ if\ sim(a,u) = 0\ and\ T_{a,u} \neq 0 \\ sim(a,u) & else\ if\ sim(a,u) \neq 0\ and\ T_{a,u} = 0 \\ 0 & else \end{cases} \tag{4}$$

where $sim(a,u)$ and $T_{a,u}$ are calculated using equations (2) and (3), respectively.

It should be noted that the neighbourhood set of the active user with further users leads to improve the performance of predicting unseen items. Therefore, the second factor has a positive effect on the proposed user reliability measure and can be calculated as follows

$$f_k(K_a) = 1 - \frac{\bar{k}}{\bar{k} + |K_a|} \tag{5}$$

where $|K_a|$ is the number of users in the neighbourhood set of the active user a, $\bar{k}$ is the median of the values of $|K_a|$ and K_a is the set of neighbours for the active user a which can be calculated using equation (6)

$$K_a = \{u \in U | W_{a,u} \geq \theta\} \tag{6}$$

where U is the set of all users in the system, $W_{a,u}$ is the combined similarity value between the active user a and user u, which is calculated using equation (4), and θ is a threshold value for the combined similarity value.

The number of ratings that the neighbours set of the active user a have assigned to the items is used as the third factor which has a positive effect on the proposed user reliability measure. This means that a high value for this factor leads to increase the user reliability value and vice versa. Therefore, the third factor can be calculated using equation (7) as follows

$$f_{i_k}(I_{K_a}) = 1 - \frac{\overline{i_k}}{\overline{i_k} + |I_{K_a}|} \tag{7}$$

where $|I_{K_a}|$ is the number of ratings that the neighbours set of the active user a have assigned to the items, K_a is the set of neighbours for the active user a (see equation (6)) and $\overline{i_k}$ is the median of the values of $|I_{K_a}|$.

Finally, the fourth factor which is used to calculate the proposed user reliability measure is the summation of the combined similarity values between the active user a and the other users in her or his neighbours set. The higher value of this factor makes a higher performance of the user's rating profile to predict unseen items. Therefore, this factor has a positive effect on the proposed user reliability measure and can be calculated as follows

$$f_s(S_a) = 1 - \frac{\bar{s}}{\bar{s} + S_a} \tag{8}$$

where

$$S_a = \sum_{u \in K_a} W_{a,u} \tag{9}$$

K_a is the set of neighbours for the active user a (see equation (6)), $W_{a,u}$ is the combined similarity value between the active user a and user u (see equation (4)) and $\bar{s}$ is the median of the values of S_a.

The median concept is used to calculate the four factors of the proposed user reliability measure in equations (1), (5), (7) and (8). It should be noted that the used factors have positive effects on the proposed user reliability measure. Moreover, if a factor for the active user is lower than the median value of this factor, the final calculated factor for the active user has a lower value and vice versa. For example, if $|I_a|$ for active user a in equation (1) is lower than the median value of $|I_a|$ for all users (i.e. $\bar{a}$), the final value of this factor (i.e. $f_i(I_a)$) will be decreased and vice versa. On the other hand, the used equations for the proposed factors of the user reliability measure make these factors bound in the range of $[0, 1]$. The proposed user reliability measure can be calculated as the geometric average of the mentioned four factors. It should be noted that the first and second factors (i.e. $|I_a|$ and $|K_a|$) are not dependent on any other factors. So, the weights of these factors in the proposed user reliability measure are considered as with a constant value of 1. On the other hand, the values of the third and fourth factors (i.e. $|I_{K_a}|$ and S_a) are dependent on the value of $|K_a|$. It means that the lower value of $|K_a|$ makes the values of S_a and $|I_{K_a}|$ also be low and vice versa. Therefore, the weights of S_a and $|I_{K_a}|$ in the proposed user reliability measure are considered as the value of $f_k(K_a)$. Finally, the proposed user reliability measure is calculated based on the defined four factors as follows

$$UR_a = \left[f_i(I_a) \cdot f_k(K_a) \cdot f_{i_k}(I_{K_a})^{f_k(K_a)} \cdot f_s(S_a)^{f_k(K_a)} \right]^{\frac{1}{2+2f_k(K_a)}} \tag{10}$$

where UR_a is the proposed user reliability measure for the active user a.

3.2. User reputation model

In this section, a user reputation model is proposed based on the combination of user ratings and trust information. The proposed user reputation model is used to calculate virtual ratings for adding to the rating profiles of the users. In other words, the aim of the proposed rating profile enhancement mechanism is to use the users with higher reputation as expert users for calculating the virtual ratings. Therefore, this mechanism can improve the accuracy of the calculated virtual ratings. To this end, a method described as user reputation in Zhou et al. [51] is used as the rating-based part of the proposed user reputation model. In this method, the user reputation is modelled as the correlation coefficient between the user's rating profile and quality vector of items. The overall steps of this method are summarised as follows:

Step 1: The initial rating-based reputation of user u is calculated as follows

$$RR_u = \frac{|I_u|}{|I|} \tag{11}$$

where $|I_u|$ is the number of ratings that the user u has assigned to the items and $|I|$ is the number of all items in the system.

Step 2: The quality of item i is calculated as follows

$$Q_i = \frac{\sum_{u \in U_i} RR_u r_i(u)}{\sum_{u \in U_i} RR_u} \tag{12}$$

where RR_u is the rating-based reputation of user u which is calculated by equation (11), U_i is the set of users who rated item i and $r_i(u)$ is the rating value of item i provided by user u.

Step 3: The correlation value between user u and quality vector of items is calculated using Pearson coefficient function as follows

$$C_u = \frac{\sum_{i \in I_u} (r_i(u) - \bar{r}(u))(Q_i - \bar{Q}_u)}{\sqrt{\sum_{i \in I_u} (r_i(u) - \bar{r}(u))^2} \cdot \sqrt{(Q_i - \bar{Q}_u)^2}} \quad (13)$$

where $\bar{r}(u)$ is the average of the rates given by user u, I_u is a set of items rated by user u and $\bar{Q}_u$ is the average value of qualities of all items rated by user u. To bound the range of the rating-based reputation of user u into [0, 1], the following equation is used

$$RR_u = \frac{C_u + 1}{2} \quad (14)$$

Step 4: Steps 2 and 3 are iterated until the results satisfy the termination condition as follows

$$\frac{1}{|I|} \sum_{i=1}^{|I|} \left| Q_i^{(n)} - Q_i^{(n-1)} \right| \leq \varepsilon \quad (15)$$

where $|I|$ is the set of all items in the system, $Q_i^{(n)}$ is the quality of item i in iteration n and ε is a constant value which is set to $\varepsilon = 10^{-6}$.

On the other hand, the well-known PageRank algorithm [52] is used to calculate the trust-based reputation for the users. The main idea is that the users who trusted by a large number of trusted users have a higher value of the user reputation. Therefore, the trust-based reputation value for user u can be calculated by a recursive function as follows

$$TR_u = \omega \frac{1}{|U|} + (1 - \omega) \sum_{T_{v,u} \neq 0} \frac{TR_v}{\deg(v)} \quad (16)$$

where $|U|$ is the number of all users in the system, ω is a constant value which is set to $\omega = 0.15$ as suggested by Page et al. [52], $T_{v,u}$ is the trust value between the users u and v which is calculated by equation (3), TR_v is the trust-based reputation value for user v and $\deg(v)$ is the out degree of user v in her or his trust network.

It can be concluded from equation (16) that calculating the trust-based reputation of the users is a recursive procedure, because the reputation value of each user depends on the reputation values of her or his trusted users. Therefore, the trust-based reputation values of the users are randomly initialised by a set of non-negative values. It is shown that the trust-based reputation vector of the users will converge to a unique stationary distribution without depending on the choice of initialised vector [53]. The recursive function (i.e. equation (16)) is iterated until the results satisfy the termination condition as follows

$$TR^{(n)} - TR^{(n-1)} = 0 \quad (17)$$

where $TR^{(n)}$ and $TR^{(n-1)}$ are the trust-based reputation vectors of the users in iterations n and $n - 1$, respectively.

Finally, we propose a novel user reputation model based on the combination of the rating-based and trust-based reputations which are calculated using equations (14) and (16), respectively. For this purpose, the following equation is used to calculate the proposed user reputation model

$$CR_u = \sigma \cdot RR_u + (1 - \sigma) \cdot TR_u \quad (18)$$

where CR_u is the combined reputation model for user u, RR_u and TR_u are respectively the rating-based and trust-based reputations for user u and σ is a parameter in the range of [0, 1] to control the effect of the rating-based and trust-based reputations on the proposed reputation model. The pseudo code of the proposed user reputation model is shown in Figure 2 (i.e. Algorithm 1).

3.3. Rating profile enhancement

A rating profile with a few number of ratings for the items has a low performance in predicting unseen items. Because, the rating profile with insufficient ratings leads to reduce the accuracy of the similarity values between the users. On the other hand, the trust networks of the users may not have a high performance to find nearest neighbours set of the users. Therefore, in this section, a novel mechanism is proposed to enhance the performance of the users' rating profiles in predicting unseen items. To this end, the proposed user reliability measure (i.e. equation (10)) is used to evaluate the performance of the rating profiles and trust networks of the users. It should be noted that the higher value of this measure shows the higher performance of the rating profile and trust network of the user in predicting unseen items. Therefore,

Algorithm 1. The proposed user reputation model

Input: Parameters $\boldsymbol{\varepsilon}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\omega}$.
Output: Combined reputation values for the users.
Begin algorithm:

```
1: Let U be the set of all users;
2: Let I be the set of all items;
3: Rating-based reputation
   3.1: Set initial rating-based reputation for all u ∈ U using Eq. (11);
   3.2: Calculate the quality values for all i ∈ I using Eq. (12);
   3.3: for all u ∈ U do
   3.4:     Calculate the correlation value between user u and quality vector of items using Eq. (13);
   3.5:     Calculate the rating-based reputation for user u (i.e. RR_u) using Eq. (14);
   3.6: end for
   3.7: Check the termination condition using Eq. (15);
   3.8: If the termination condition is not satisfied then go to step 3.2 else go to step 4;
4: Trust-based reputation
   4.1: Set initial trust-based reputation for all u ∈ U randomly by a set of non-negative values;
   4.2: for all u ∈ U do
   4.3:     Calculate the trust-based reputation for user u (i.e. TR_u) using Eq. (16);
   4.4: end for
   4.5: Check the termination condition using Eq. (17);
   4.6: If the termination condition is not satisfied then go to step 4.2 else go to step 5;
5: Combined reputation model
   5.1: for all u ∈ U do
   5.2:     Calculate the combined reputation for user u (i.e. CR_u) using Eq. (18);
   5.3: end for
```

End algorithm.

Figure 2. Pseudo code of the proposed user reputation model.

the rating profiles with low user reliability measure are enhanced by adding a number of virtual ratings. In the proposed mechanism, a threshold value (i.e. r) is used to enhance the rating profiles. In other words, the rating profile of a target user will be enhanced if the user reliability measure of this user is lower than the threshold value r.

Moreover, three different measures including reliability, diversity and novelty of the items are used to enhance the rating profiles of the users. The aim of the proposed rating profile enhancement mechanism is to improve the rating profiles based on the diversity and novelty measures. On the other hand, the reliability of the virtual ratings is considered to improve the accuracy of the predictions. These three measures are merged as a final metric for selecting a suitable subset of the virtual ratings to add into the rating profile of the target user. In the following subsections, additional details about the mentioned measures and calculating virtual ratings are described.

3.3.1. Item reliability.

The first measure which is used in the proposed mechanism to enhance the rating profiles of the users is item reliability measure. This measure is used to calculate the reliability of the items for enhancing the rating profiles of the users. In other words, the aim of the proposed item reliability measure is to improve the accuracy of the predicted virtual ratings. To this end, three different factors are used to calculate the proposed item reliability measure which these factors are described in the following.

The number of ratings which have been assigned to item i is the first factor that can be used to calculate the proposed item reliability measure. It should be noted that the more ratings for item i make the higher reliability value about it. Therefore, this factor has a positive effect on the proposed item reliability measure and can be calculated as follows

$$f_i(I_i) = 1 - \frac{\bar{i}}{\bar{i} + |I_i|} \tag{19}$$

where $|I_i|$ is the number of ratings which have been assigned to item i and $\bar{i}$ is the median of the values of $|I_i|$.

The second factor which is used to calculate the proposed item reliability measure is the standard deviation of ratings that have been assigned to item i. This factor has a negative effect on the item reliability measure, because a larger value of the factor makes a lower item reliability value for item i. Therefore, a decreasing function can be used to calculate this factor as follows [48]

$$f_{sd}(stdev(I_i)) = \frac{max - stdev(I_i)}{max - min} \tag{20}$$

where $stdev(I_i)$ is the standard deviation of ratings that have been assigned to item i; max and min are the maximum and minimum of all values of $stdev(I_i)$, respectively.

Finally, the third factor which is used to calculate the proposed item reliability measure is the summation of the user reliability values for the users who have assigned a rating to item i. The higher value of this factor makes a higher value of the item reliability measure. Therefore, this factor has a positive effect on the proposed item reliability measure and can be calculated as follows

$$f_c(C_i) = 1 - \frac{\bar{c}}{\bar{c} + C_i} \tag{21}$$

where

$$C_i = \sum_{u \in U_i} UR_u \tag{22}$$

where U_i is the set of users which have assigned a rating to item i, UR_u is the user reliability measure for user u which is calculated using equation (10) and $\bar{c}$ is the median of the values of C_i.

It should be noted that higher values of the positive factors make higher values for the proposed item reliability measure. Therefore, the median concept is used for the proposed item reliability measure to calculate the positive factors in equations (19) and (21). Moreover, it makes to bound the positive factors into the range of $[0, 1]$. The proposed item reliability measure can be calculated as the geometric average of the three mentioned factors. It should be noted that the first factor (i.e. $|I_i|$) is not dependent on any other factors. Therefore, the weight of this factor is considered as with a constant value of 1 to calculate the proposed item reliability measure. On the other hand, the values of $stdev(I_i)$ and C_i are dependent on the value of $|I_i|$. It means that the lower value of $|I_i|$ makes the values of $stdev(I_i)$ and C_i also be low and vice versa. Therefore, the weights of $stdev(I_i)$ and C_i are considered as the value of $f_i(I_i)$ in the proposed item reliability measure. Finally, the proposed item reliability measure is calculated based on the three defined factors as follows

$$IR_i = \left[f_i(I_i) \cdot f_{sd}(stdev(I_i))^{f_i(I_i)} \cdot f_c(C_i)^{f_i(I_i)} \right]^{\frac{1}{1 + 2 f_i(I_i)}} \tag{23}$$

where IR_i is the proposed item reliability measure for item i.

3.3.2. Diversity. The second measure which is used in the rating profile enhancement mechanism is the diversity measure of the virtual ratings. This measure is an important measure in the recommendation systems which relates to the internal differences within the set of items recommended to each user. In other words, the rating profiles of the users can be improved by considering the diversification of the results in the recommendation lists to the users. Therefore, one of the purposes of the proposed rating profile enhancement mechanism is to improve the diversity of the recommendation lists to the users. In the proposed method, we use the intra-list diversity measure to calculate the diversity of each item as follows [54]

$$D_i = \frac{1}{|I_a|} \sum_{j \in I_a} (1 - s(i,j)) \tag{24}$$

where D_i is the diversity value of item i based on the rating profile of the active user a, I_a is the set of items that have been rated by the active user a and $s(i,j)$ is the similarity value between items i and j which is calculated using the cosine similarity function as follows

$$s(i,j) = \frac{\sum_{u \in U} r_{ui} \cdot r_{uj}}{\sqrt{\sum_{u \in U} {r_{ui}}^2 \sum_{u \in U} {r_{uj}}^2}} \tag{25}$$

where U is the set of all users and r_{ui} is the rating value of item i which is assigned by user u.

3.3.3. Novelty. Novelty is the third measure which is used in the proposed mechanism to enhance the rating profiles of the users. This measure indicates the degree of difference between the items recommended to and known by the user which can increase the satisfaction of the users in the system. Therefore, the novelty measure of the recommendations for the users will be increased by considering this measure in the proposed rating profile enhancement mechanism. To this end, we use the inverse user frequency (IUF) measure in the proposed mechanism [55] to calculate the novelty value of each item as follows

$$N_i = -P(i)\log_2 P(i) \tag{26}$$

where N_i is the novelty value of item i and $P(i)$ is the probability of item i being drawn from the recommendation lists which can be calculated as follows

$$P(i) = \frac{|I_i|}{|U|} \tag{27}$$

where $|I_i|$ is the number of ratings which have been assigned to item i and $|U|$ is the number of all users in the system.

3.3.4. Adding virtual ratings. In this step, the rating profiles of the users who their user reliability measures (i.e. equation (10)) are lower than a threshold value (i.e. r) are enhanced by adding a number of virtual ratings. The number of these ratings depends on the user reliability values. It means that a user with a higher value of the user reliability measure needs a lower number of the virtual ratings for adding to her or his profile, and vice versa. Therefore, we propose the following equation to calculate the number of virtual ratings for adding to the rating profile of the active user a

$$m' = (1 - UR_a) \times m \tag{28}$$

where UR_a is the user reliability measure for the active user a (i.e. equation (10)) and m is the maximum number of the virtual ratings that can be added to the user's rating profile.

In the proposed rating profile enhancement mechanism, we consider the reliability, diversity and novelty measures of the items to enhance the rating profiles of the users. To this end, a linear combination of these measures is used to calculate the final measure for selecting the virtual ratings that can be added to the rating profiles of the users. Therefore, the following equation is used to calculate the combination of the mentioned measures as the final measure for the items

$$F_i = \alpha \cdot IR_i + \beta \cdot N_i + (1 - \alpha - \beta) \cdot D_i \tag{29}$$

where F_i is the final measure for item i, IR_i is the item reliability measure for item i (i.e. equation (23)), N_i is the novelty measure for item i (i.e. equation (26)), D_i is the diversity measure for item i (i.e. equation (24)) and α and β are two parameters to control the effects of the used measures.

The calculated final measures of the items are used to select an appropriate subset of the items for adding to the rating profiles of the users. In other words, a number of the items (i.e. m') with highest values of the final measure (see equation (29)) are selected for using in the proposed rating profile enhancement mechanism. Then, the virtual ratings of these selected items are calculated based on the proposed user reputation model as follows

$$VR_i = \frac{\sum_{u \in U_i} CR_u \cdot r_i(u)}{\sum_{u \in U_i} CR_u} \tag{30}$$

where VR_i is the virtual rating of item i, U_i is the set of users which have assigned a rating to item i, CR_u is the user reputation model for user u which is calculated using equation (18) and $r_i(u)$ is the rate of item i given by user u.

After calculating the virtual ratings of the items, these ratings can be added to the rating profile of the active user. It should be noted that the added virtual ratings are completely different from the real ratings of the active user. Moreover, the real ratings of the active user have been considered without any changes into the recommendation process. The main advantage of this step is to alleviate cold start and data sparsity problems of the recommender systems. In other words, the proposed mechanism makes a denser user-item matrix than the original matrix by adding a number of virtual ratings to the rating profiles of the users. Therefore, the calculation of the similarity values between the users performs simply and also the system can be able to find better neighbours set for the users. In addition, three different measures including reliability, diversity and novelty of the items are considered in the proposed rating profile enhancement mechanism to improve the performance of trust-aware recommender systems.

Algorithm 2. Reputation-based Trust-Aware Recommender System (RTARS)

Input: Parameters $\boldsymbol{\theta}$, $\boldsymbol{r}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{m}$.
Output: Predicted ratings for active users.
Begin algorithm:

Let $\boldsymbol{U}$ be the set of all users;
for all $\boldsymbol{a} \in \boldsymbol{U}$ **do**
    Calculate the four factors of the user reliability measure using Eqs. (1)-(9);
    Calculate the user reliability measure $\boldsymbol{UR_a}$ for the active user $\boldsymbol{a}$ using Eq. (10);
end for
Apply algorithm 1 to calculate user reputation values of all $\boldsymbol{a} \in \boldsymbol{U}$;
Let $\boldsymbol{I}$ be the set of all items;
for all $\boldsymbol{i} \in \boldsymbol{I}$ **do**
    Calculate the three factors of the item reliability measure using Eqs. (19)-(22);
    Calculate the item reliability measure $\boldsymbol{IR_i}$ for the item $\boldsymbol{i}$ using Eq. (23);
    Calculate the item novelty measure $\boldsymbol{N_i}$ for the item $\boldsymbol{i}$ using Eq. (26);
end for
for all $\boldsymbol{a} \in \boldsymbol{U}$ **do**
    if $(\boldsymbol{UR_a} < \boldsymbol{r})$ **then**
        Let $\boldsymbol{I_a}$ be the set of items that have been rated by the active user $\boldsymbol{a}$;
        Set $\boldsymbol{I'_a} = \boldsymbol{I} - \boldsymbol{I_a}$;
        Calculate the item diversity measure $\boldsymbol{D_i}$ for all $\boldsymbol{i} \in \boldsymbol{I'_a}$ using Eq. (24);
        Calculate the final measure $\boldsymbol{F_i}$ for all $\boldsymbol{i} \in \boldsymbol{I'_a}$ using Eq. (29);
        Sort $\boldsymbol{I'_a}$ descending based on their final measures $\boldsymbol{F_i}$;
        Calculate $\boldsymbol{m'}$ using Eq. (28);
        Select $\boldsymbol{top\_m'}$ items from $\boldsymbol{I'_a}$ as $\boldsymbol{\tilde{I}_a}$;
        for all $\boldsymbol{i} \in \boldsymbol{\tilde{I}_a}$ **do**
            Calculate virtual rating of the item $\boldsymbol{i}$ using Eq. (30);
            Add the calculated virtual rating to the rating profile of the active user $\boldsymbol{a}$;
        end for
    end if
end for
for all $\boldsymbol{a} \in \boldsymbol{U}$ **do**
    Calculate the similarity values between the active user $\boldsymbol{a}$ and other users using Eq. (4) based on the enhanced rating profiles;
    Calculate the nearest neighbors set of the active user $\boldsymbol{a}$ using Eq. (6);
end for
Predict the unseen items for the active users using Eq. (31);

End algorithm.

Figure 3. Pseudo code of the proposed method.

3.4. Rating prediction

In this step, the enhanced rating profiles of the users are used to calculate the similarity values between the users. These similarity values can be calculated using equation (4) and a set of neighbours for the active user a is formed using equation (6). Finally, the rating of unseen item i for the active user a can be predicted using the following equation

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u \in K_{a,i}} W_{a,u} \cdot (r_{u,i} - \bar{r}_u)}{\sum_{u \in K_{a,i}} W_{a,u}} \tag{31}$$

where $\bar{r}_a$ is the average of the ratings for active user a, $K_{a,i}$ is the set of neighbours for active user a that have rated item i, $r_{u,i}$ is the rate of item i given by user u and $W_{a,u}$ is the similarity value between the users a and u which is calculated using equation (4). After predicting unseen items for the active user, a number of items with higher ratings will be recommended to the active user as recommendations list. The pseudo code of the proposed method is shown in Figure 3 (i.e. Algorithm 2).

Table 1. The example user-item matrix consisting of three users and five items.

	i_1	i_2	i_3	i_4	i_5
u_1	2	–	–	**?**	5
u_2	–	4	1	3	–
u_3	–	3	2	2	–

(?) shows the unknown rating which should be predicted.

Table 2. The example trust matrix consisting of three users.

	u_1	u_2	u_3
u_1	–	–	–
u_2	–	–	1
u_3	–	1	–

3.5. An illustrated example

In this section, an example is represented to describe the general process of predicting an unknown rating using the proposed method. To this end, a small user-item matrix with three users and five items is used which the ratings are shown in Table 1. The rating values in the matrix are in the range of 1 (min) to 5 (max). The purpose of the example is to predict the rating of item i_4 for the target user u_1. Moreover, the trust relations among the users are shown in Table 2. It can be seen form Tables 1 and 2 that there are no common items and also trust relations between the target user u_1 and others. Therefore, the similarity values between the target user and other users cannot be calculated by the recommender system. One of the most important advantages of the proposed method is to resolve this problem by adding virtual ratings into the recommendation process. The virtual ratings are calculated using the proposed user reputation model. Suppose that the reputation values of the users u_2 and u_3 are equal to $CR_{u_2} = 0.6$ and $CR_{u_3} = 0.7$ which can be calculated using equation (18). Based on the rating profile of the user u_1, the virtual ratings can be calculated for items i_2 and i_3. Therefore, the virtual ratings of items i_2 and i_3 for the target user u_1 are equal to $VR_{i_2} = 3.46$ and $VR_{i_3} = 1.54$ which are calculated using equation (30). After calculating the virtual ratings of the items i_2 and i_3, the similarity values between the target user u_1 and the users u_2 and u_3 are calculated using equation (4). These similarity values are equal to $W_{u_1,u_2} = 0.77$ and $W_{u_1,u_3} = 0.42$. Finally, the rating of the unknown item i_4 for the target user u_1 is calculated using equation (31). The predicted rating for the unknown item is equal to $P_{u_1,i_4} = 3.59$.

4. Evaluation and results

In this section, several experiments are conducted based on three real-world datasets to determine how the proposed recommendation method performs in terms of different evaluation measures. Therefore, in the following subsections, the used datasets are described at first. Then, the evaluation measures which are used to evaluate the effectiveness of the proposed method are detailed in the next subsection. Then, a subsection is provided to discuss about the parameter settings for the proposed method. In addition, the results of the experiments are reported to compare the proposed method with the other state-of-the-art methods. Sensitivity analysis of the parameters which are used in the proposed method is discussed to show the effects of different values of the parameters on the performance of the proposed method. Finally, the execution time of the proposed method is discussed in the final subsection.

4.1. Datasets

In this article, three well-known datasets including Epinions,[1] Flixster[2] and FilmTrust[3] are used in the experiments to compare the proposed method with the other methods. Epinions dataset contains the opinions of the users about the items (such as movies, books) which are numerical ratings in the range 1 (min) to 5 (max). Also, this dataset includes explicit trust statements between the users which the values of them are 0 or 1. Moreover, there are 49,290 users in the Epinions dataset who rated at least once among 139,738 items. On the other hand, Flixster is a social movie site in which the users are able to rate the existing movies in the range of 0.5 (min) to 4.0 (max) with step 0.5. Moreover, the friend

Table 3. The statistics of the evaluation datasets.

Dataset	#Users	#Items	#Ratings	#Trust	Sparsity (%)
Epinions	10,000	117,000	385,000	288,000	99.97
Flixster	10,000	6,000	55,000	89,000	99.92
FilmTrust	1986	2071	35,497	1853	99.14

relationships between the users are used in this dataset as the trust statements. In our experiments, we sampled two subsets of the original Epinions and Flixster datasets for simplicity by randomly selecting 10,000 users with their corresponding ratings and trust statements. Finally, FilmTrust is a movie recommendation site in which the users can rate the items (i.e. movies) in the range 0.5 (min) to 4.0 (max) with step 0.5. The used dataset includes 1986 users, 2071 items and 35,497 ratings. Moreover, the link information among the users is used as the explicit trust statements which a trust value is 1 if a link exists between two users and otherwise the value is 0. The statistics and descriptions of the datasets are presented in Table 3.

4.2. Evaluation measures

Several evaluation measures are used to compare the effectiveness of the proposed method with the other recommendation methods. These measures include mean absolute error (MAE), root mean square error (RMSE), diversity, novelty, precision, recall and normalised discounted cumulative gain (NDCG). The MAE and RMSE measures are used to evaluate the accuracy of the predictions. Therefore, the predicted ratings are compared with the real ratings and the mean value of their differences is calculated as the final prediction error. To this end, the MAE and RMSE measures are calculated respectively using equations (32) and (33) as follows

$$MAE = \frac{\sum_{i=1}^{n} |r_i - p_i|}{n} \tag{32}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (r_i - p_i)^2} \tag{33}$$

where r_i and p_i are real and predicted ratings of item i, respectively. Moreover, n is the total number of ratings that can be predicted by a recommendation method.

Moreover, the diversity and novelty measures are used to compare the proposed method with others. The diversity measure refers to the differences between the items which are recommended to the active user as the recommendations list. On the other hand, the novelty measure about an item generally refers to the difference between it and other items which are previously experienced by the active user. In the experiments, the intra-list diversity of a set of recommended items is used to calculate the diversity measure as follows [54]

$$Diversity = \frac{\sum_{u=1}^{|U|} diversity_u}{|U|} \tag{34}$$

where

$$diversity_u = \frac{1}{|L_u|(|L_u| - 1)} \sum_{i \in L_u} \sum_{j \in L_u, j \neq i} [1 - s(i,j)] \tag{35}$$

and U is the set of all users, L_u is the recommendations list to user u and $s(i,j)$ indicates the cosine similarity value between items i and j which is calculated using equation (25).

Finally, the Shannon entropy [56] is used to calculate the novelty measure in the experiments. Therefore, this measure can be calculated as follows [18]

$$Novelty = -\sum_{i \in I} p(i|s) \log_2 p(i|s) \tag{36}$$

where $p(i|s)$ is the probability of the item i being drawn from the recommendation lists for the users which are generated using the system s. Therefore, this probability value can be calculated as follows

$$p(i|s) = \frac{|\{u \in U | i \in L_u\}|}{\sum_{j \in I} |\{u \in U | j \in L_u\}|} \tag{37}$$

where U is the set of all users, I indicates the set of all items and L_u is the recommendations list to the user u.

Precision and recall are two common evaluation measures to evaluate the performance of the recommendation systems. Precision refers to the ratio of relevant items chosen by the recommendation method over the total list of recommended items. On the other hand, recall refers to the ratio of relevant items in the recommendation list over the total number of relevant items. It should be noted that the items are considered as relevant if their ratings are higher than the average of ratings provided by the target user. The following equations are used to calculate the precision and recall measures [57]

$$Precision = \frac{|relevant\ items\ recommended|}{|all\ items\ retrieved\ and\ recommended|} \tag{38}$$

$$Recall = \frac{|relevant\ items\ recommended|}{|all\ relevant\ items\ retrieved\ and\ not\ recommended|} \tag{39}$$

NDCG is a measure of ranking quality to evaluate the usefulness of an item based on its rank in a recommendation list. The main idea of the measure is that the relevant items should have a higher rank in the recommendation list. Therefore, NDCG can be calculated as follows [57]

$$NDCG = \frac{DCG}{DCG_{max}} \tag{40}$$

where

$$DCG = rel_1 + \sum_{i=2}^{|L|} \frac{rel_i}{\log_2(i+1)} \tag{41}$$

and rel_i indicates the relevancy of item i in the recommendation list which $rel_i = 1$ if item i is a relevant item; otherwise $rel_i = 0$. Moreover, DCG_{max} is the maximum value of discounted cumulative gain which is calculated using the following equation

$$DCG_{max} = 1 + \sum_{i=2}^{|L|} \frac{1}{\log_2(i+1)} \tag{42}$$

4.3. Parameter settings

There are several parameters in the proposed method that need to be initialised. The parameter d_{max} (see equation (3)) is the maximum allowable propagation distance between the users which can be calculated as follows [58]

$$d_{max} = \frac{\ln(n)}{\ln(k)} \tag{43}$$

where n and k are respectively the size and the average degree of the trust network in a specific recommender system. The parameter θ is a threshold value in equation (6) which is set to $\theta = 0.5$ for the Epinions, Flixster and FilmTrust datasets. The parameter σ is used to calculate the proposed user reputation model based on equation (18) that is set to $\sigma = 0.5$ for the used datasets. Moreover, some parameters are used in the proposed rating profile enhancement mechanism including r, m, α and β (see Section 3.3.4). The parameter r is used as the user reliability threshold which is set to $r = 0.7$ for all of the used datasets. In addition, the parameter m is used in equation (28) as the maximum number of the virtual ratings that can be added to the user's rating profile. The value of this parameter is set to $m = 10$ for all of the Epinions, FilmTrust and Flixster datasets. Finally, the parameters α and β are used in equation (29) which are set to $\alpha = 0.4$ and $\beta = 0.3$ for all of the used datasets. The fivefold cross-validation approach is applied for comparing the results of the recommendation methods. In this approach, each dataset is randomly divided into five folds and in each run, four folds are used as the training set and the remaining fold as the test set. Five runs are performed for testing all of the folds and the average of these results are considered as the final result of the fivefold cross-validation. This procedure is performed based on 10 independent runs and in each run, the fivefold cross-validation is applied on the datasets. Finally, the average results of these 10 independent runs are reported as the final results of the recommendation methods.

4.4. Comparison with different methods

In this section, the performance of the proposed method is compared with the other recommendation methods based on different evaluation metrics. The brief descriptions of the compared methods are provided in the following:

- User-based collaborative filtering (UCF): This method is the traditional CF which is based on computing similarity values between the users using the Pearson coefficient function (i.e. equation (2)) to determine the nearest neighbours set of the active users.
- Item-based collaborative filtering (ICF): This method calculates the similarity values between the items and uses these values to predict the ratings of the unseen items based on the weighted averaging of the ratings that have been rated by the target user [59].
- TARS: This approach is the basic model of Trust-Aware Recommender System [50] which propagates the trust statements over the trust network and estimates a trust weight that can be used in place of the similarity weight.
- Merge: This recommendation method is based on incorporating trust information in providing suggestions to the users for alleviating the cold start and data sparsity problems. Moreover, a confidence model is used to measure the quality of predictions which is based on a positive and a negative factor [26].
- Reliability-based trust-aware collaborative filtering (RTCF): This recommendation approach uses a reliability measure to improve the performance of trust-aware recommender systems in predicting ratings for the users. Moreover, a reconstruction mechanism is proposed to form trust networks for the users with high reliability and accuracy in predicting unseen items [22].
- TrustMF: This is a model-based recommendation method which uses matrix factorization technique to form low-dimensional latent factors spaces for the users in terms of their trust statements [27].
- SocialMF: This recommendation method incorporates trust propagation into a matrix factorization technique to provide suggestions for the users in social networks [35].
- TrustSVD: This method is based on incorporating trust statements into matrix factorization technique which involves the explicit and implicit influences of rated items. Moreover, both of the explicit and implicit influences of trusted users on the prediction of items are considered in this method [34].
- RTARS: This is the proposed method in this article which uses a user reliability measure and a user reputation model to improve trust-aware recommender systems. In addition, the diversity, novelty and reliability measures of the items are considered for producing relevant and novel recommendations to the users.

The experiments are performed for two different views of data including all users and also cold start users (i.e. the users who have less than five ratings). The results of the experiments based on the MAE and RMSE measures are reported in Tables 4–6 for the Epinions, Flixster and FilmTrust datasets. It can be concluded from Table 4 that the proposed method obtains better results based on both of the MAE and RMSE measures for all users and also cold start users in comparison with the other methods. Moreover, the results of Tables 5 and 6 indicate that the proposed method has the best performance compared with the other methods based on the MAE and RMSE measures for both of the all users and cold start users. Therefore, these results show that the proposed method can improve the performance of the recommender systems based on all of the used datasets in terms of the accuracy metrics.

In addition, the results of the experiments based on the diversity and novelty measures for all users view are reported in Tables 7–9. It should be noted that the diversity and novelty measures depend on the length of the recommendations list to the users (see equations (35) and (37)). Therefore, the experiments are performed for different lengths of the recommendations list including $L = 5,\ 10,\ \text{and}15$. Table 7 shows the diversity and novelty evaluation on the Epinions dataset for all users view. As you can see from these results, the proposed method obtains better performance than the other methods based on both of the diversity and novelty measures for all lengths of the recommendations list. Moreover, it can be concluded that the diversity and novelty measures for the proposed method increase, when the value of parameter L is increased. The values of the diversity measure for the proposed method are 0.528, 0.531 and 0.532 for $L = 5$, $L = 10$ and $L = 15$, respectively. Also, the values of the novelty measure for the proposed method are 8.874, 8.982 and 9.033 for $L = 5$, $L = 10$ and $L = 15$, respectively.

Table 4. Experiment results on the Epinions dataset for MAE and RMSE measures.

Algorithms	All users		Cold users	
	MAE	RMSE	MAE	RMSE
UCF	0.865	1.165	1.063	1.357
ICF	0.824	1.136	1.102	1.426
TARS	0.826	1.143	0.864	1.158
Merge	0.798	1.032	0.837	1.123
RTCF	0.638	0.925	0.725	0.997
TrustMF	0.804	1.048	0.835	1.114
SocialMF	0.813	1.054	0.846	1.127
TrustSVD	0.786	1.021	0.825	1.102
RTARS	**0.577**	**0.811**	**0.619**	**0.873**

MAE: mean absolute error; RMSE: root mean square error; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

Table 5. Experiment results on the Flixster dataset for MAE and RMSE measures.

Algorithms	All users		Cold users	
	MAE	RMSE	MAE	RMSE
UCF	0.956	1.289	1.213	1.371
ICF	0.912	1.203	1.268	1.412
TARS	0.873	1.134	0.952	1.154
Merge	0.885	1.157	0.965	1.173
RTCF	0.725	0.952	0.814	0.996
TrustMF	0.864	1.126	0.891	1.125
SocialMF	0.756	0.978	0.877	1.086
TrustSVD	0.719	0.935	0.823	1.028
RTARS	**0.609**	**0.827**	**0.744**	**0.939**

MAE: mean absolute error; RMSE: root mean square error; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

Table 6. Experiment results on the FilmTrust dataset for MAE and RMSE measures.

Algorithms	All users		Cold users	
	MAE	RMSE	MAE	RMSE
UCF	0.703	0.884	0.744	0.916
ICF	0.698	0.871	0.786	1.084
TARS	0.763	0.924	0.827	1.098
Merge	0.696	0.858	0.745	0.914
RTCF	0.648	0.852	0.729	0.925
TrustMF	0.631	0.810	0.674	0.867
SocialMF	0.638	0.837	0.680	0.907
TrustSVD	0.607	0.787	0.661	0.853
RTARS	**0.548**	**0.692**	**0.599**	**0.774**

MAE: mean absolute error; RMSE: root mean square error; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

Moreover, the results of the experiments based on the diversity and novelty measures are reported for the Flixster dataset in Table 8. These results show that the proposed method has better performance in comparison with the other recommendation methods for all users view and different lengths of recommendations list to the users. Also, it can be

Table 7. Diversity and novelty evaluation on the Epinions dataset for all users and different lengths of recommendations list (*L*).

Algorithms	*L* = 5		*L* = 10		*L* = 15	
	Diversity	Novelty	Diversity	Novelty	Diversity	Novelty
UCF	0.379	6.545	0.386	6.671	0.392	6.815
ICF	0.382	6.728	0.394	6.861	0.405	6.981
TARS	0.401	6.906	0.415	7.068	0.419	7.127
Merge	0.492	8.012	0.503	8.176	0.513	8.283
RTCF	0.458	7.601	0.463	7.784	0.471	7.958
TrustMF	0.426	7.105	0.435	7.236	0.439	7.354
SocialMF	0.437	7.334	0.441	7.459	0.445	7.612
TrustSVD	0.473	7.894	0.482	7.968	0.493	8.027
RTARS	**0.528**	**8.874**	**0.531**	**8.982**	**0.532**	**9.033**

UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

Table 8. Diversity and novelty evaluation on the Flixster dataset for all users and different lengths of recommendations list (*L*).

Algorithms	*L* = 5		*L* = 10		*L* = 15	
	Diversity	Novelty	Diversity	Novelty	Diversity	Novelty
UCF	0.221	4.203	0.236	4.354	0.258	4.623
ICF	0.243	4.486	0.262	4.784	0.287	4.913
TARS	0.265	4.979	0.283	5.230	0.301	5.472
Merge	0.329	5.224	0.353	5.425	0.394	5.816
RTCF	0.273	4.985	0.291	5.259	0.317	5.517
TrustMF	0.291	5.011	0.302	5.312	0.331	5.593
SocialMF	0.309	5.105	0.318	5.394	0.352	5.638
TrustSVD	0.358	5.237	0.374	5.689	0.412	5.927
RTARS	**0.412**	**5.891**	**0.437**	**6.232**	**0.485**	**6.539**

UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

Table 9. Diversity and novelty evaluation on the FilmTrust dataset for all users and different lengths of recommendations list (*L*).

Algorithms	*L* = 5		*L* = 10		*L* = 15	
	Diversity	Novelty	Diversity	Novelty	Diversity	Novelty
UCF	0.201	4.267	0.187	4.378	0.176	4.397
ICF	0.214	4.624	0.198	4.748	0.184	4.813
TARS	0.222	4.979	0.219	5.043	0.211	5.084
Merge	0.269	5.614	0.263	5.724	0.254	5.932
RTCF	0.243	5.129	0.229	5.213	0.218	5.268
TrustMF	0.248	5.217	0.235	5.279	0.227	5.304
SocialMF	0.256	5.358	0.251	5.362	0.246	5.412
TrustSVD	0.271	5.847	0.267	5.897	0.262	5.968
RTARS	**0.302**	**6.406**	**0.298**	**6.433**	**0.296**	**6.472**

UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

concluded that the values of the diversity and novelty measures for the proposed method increase, when the lengths of the recommendations list are increased. Table 9 reports the diversity and novelty evaluation on the FilmTrust dataset for all users and different lengths of the recommendations list to the users. These results indicate that the proposed method

Table 10. Diversity and novelty evaluation on the Epinions, Flixster and FilmTrust datasets for cold start users: length of recommendations list is equal to 5 (L = 5).

Algorithms	Epinions		Flixster		FilmTrust	
	Diversity	Novelty	Diversity	Novelty	Diversity	Novelty
UCF	0.209	5.278	0.195	3.259	0.055	3.255
ICF	0.228	5.545	0.209	3.436	0.063	3.369
TARS	0.242	5.912	0.228	3.559	0.069	3.636
Merge	0.307	7.586	0.281	3.893	0.082	4.867
RTCF	0.264	6.861	0.237	3.647	0.072	3.952
TrustMF	0.247	6.243	0.245	3.712	0.075	4.124
SocialMF	0.253	6.489	0.263	3.768	0.079	4.318
TrustSVD	0.289	7.125	0.296	3.976	0.084	4.956
RTARS	**0.349**	**8.409**	**0.326**	**4.214**	**0.096**	**5.538**

UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

obtains better performance than the other methods based on the diversity and novelty measures. The diversity measure of the proposed method decreases when the length of the recommendations list is increased. On the other hand, the novelty measure of the proposed method increases when the length of the recommendations list is increased. The values of the diversity measure for the proposed method are 0.302, 0.298 and 0.296 for $L = 5$, $L = 10$ and $L = 15$, respectively. In addition, the values of the novelty measure for the proposed method are 6.406, 6.433 and 6.472 for $L = 5$, $L = 10$ and $L = 15$, respectively.

The experiments are repeated for cold start users based on the diversity and novelty measures and the results are reported in Table 10 for all of the Epinions, FilmTrust and Flixster datasets. It should be noted that the cold start users in the experiments are those of the users who have less than five ratings. Therefore, the experiments on the cold start users are performed only for $L = 5$ as the length of the recommendations list to the cold start users. As you can see from Table 10, the proposed method outperforms the other methods under both of the diversity and novelty measures for the Epinions, FilmTrust and Flixster datasets. Therefore, the proposed method can provide better recommendations for the cold start users than the other recommendation methods with considering diversity and novelty measures of the recommendations.

The results of the experiments based on the precision, recall and NDCG measures for the Epinions, Flixster and FilmTrust datasets are shown in Tables 11–13. To this end, the results are reported based on different lengths of the recommendations list including $L = 5$, $L = 10$ and $L = 15$. These results indicate that the proposed method obtains better results in comparison with the other recommendation methods based on all of the used datasets and also different lengths of the recommendations list. Moreover, the values of the precision measure for the proposed method decrease when the lengths of the recommendations list are increased. On the other hand, the recall values of the proposed method increase when the lengths of the recommendations list are increased. The experiments are repeated on the Epinions, Flixster and FilmTrust datasets for the cold start users and the results are shown in Table 14. It should be noted that the length of the recommendations list is set to $L = 5$, because the cold start users are those of the users with less than five ratings. It can be concluded from Table 14 that the proposed method outperforms other methods based on precision, recall and NDCG measures for the cold start users. Therefore, the proposed method can alleviate the cold start problem in social recommender systems.

4.5. Impact of the parameters

In this section, several experiments are performed to evaluate the performance of the proposed method based on different values of the input parameters including θ, σ, r, m, α and β. The parameter θ is used as a threshold value for the combined similarity values between the users in equation (6). The effect of different values of the parameter θ based on the MAE and RMSE measures for all users and cold start users is shown in Figure 4. Moreover, Figure 5 reports the results of the experiments for different values of the parameter θ on the diversity and novelty measures. As you can see from these results, the values of the MAE, RMSE, diversity and novelty measures decrease in most cases when the value of the parameter θ is increased. Therefore, the higher values of the parameter θ can improve the performance of the proposed method on the MAE and RMSE measures. On the other hand, it leads to reduce the performance of the proposed method in terms of the diversity and novelty measures. Figure 6 shows the effects of different values of the parameter θ

Table 11. Precision, recall and NDCG evaluation on the Epinions dataset for all users and different lengths of recommendations list (*L*).

Algorithms	Precision			Recall			NDCG		
	L = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15
UCF	0.104	0.045	0.041	0.148	0.195	0.215	0.247	0.214	0.175
ICF	0.126	0.044	0.041	0.174	0.245	0.253	0.261	0.195	0.163
TARS	0.146	0.078	0.061	0.234	0.263	0.284	0.268	0.234	0.182
Merge	0.359	0.338	0.294	0.534	0.568	0.598	0.395	0.364	0.328
RTCF	0.342	0.315	0.281	0.526	0.542	0.571	0.371	0.335	0.294
TrustMF	0.253	0.231	0.204	0.489	0.497	0.543	0.289	0.256	0.237
SocialMF	0.298	0.283	0.265	0.512	0.572	0.596	0.334	0.292	0.265
TrustSVD	0.367	0.352	0.324	0.548	0.597	0.613	0.421	0.386	0.369
RTARS	**0.417**	**0.396**	**0.365**	**0.593**	**0.645**	**0.697**	**0.465**	**0.412**	**0.387**

NDCG: normalised discounted cumulative gain; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System.
The best results are shown as bold face.

Table 12. Precision, recall and NDCG evaluation on the Flixster dataset for all users and different lengths of recommendations list (*L*).

Algorithms	Precision			Recall			NDCG		
	L = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15
UCF	0.321	0.262	0.214	0.402	0.422	0.465	0.374	0.318	0.286
ICF	0.354	0.241	0.205	0.439	0.473	0.498	0.395	0.312	0.274
TARS	0.406	0.297	0.261	0.493	0.521	0.531	0.432	0.356	0.315
Merge	0.652	0.581	0.537	0.671	0.698	0.735	0.548	0.507	0.462
RTCF	0.638	0.564	0.527	0.652	0.686	0.713	0.542	0.491	0.431
TrustMF	0.584	0.452	0.438	0.607	0.642	0.681	0.495	0.412	0.354
SocialMF	0.628	0.503	0.487	0.687	0.701	0.724	0.517	0.465	0.378
TrustSVD	0.673	0.531	0.512	0.716	0.754	0.786	0.556	0.487	0.409
RTARS	**0.694**	**0.607**	**0.563**	**0.765**	**0.812**	**0.843**	**0.573**	**0.528**	**0.491**

NDCG: normalised discounted cumulative gain; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System.
The best results are shown as bold face.

Table 13. Precision, recall and NDCG evaluation on the FilmTrust dataset for all users and different lengths of recommendations list (*L*).

Algorithms	Precision			Recall			NDCG		
	L = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15	*L* = 5	*L* = 10	*L* = 15
UCF	0.409	0.289	0.195	0.428	0.581	0.598	0.342	0.315	0.261
ICF	0.431	0.289	0.203	0.471	0.553	0.563	0.361	0.307	0.273
TARS	0.458	0.292	0.208	0.473	0.556	0.589	0.369	0.324	0.282
Merge	0.491	0.375	0.282	0.552	0.651	0.685	0.458	0.397	0.356
RTCF	0.483	0.352	0.257	0.537	0.645	0.673	0.445	0.382	0.332
TrustMF	0.478	0.284	0.217	0.490	0.592	0.627	0.426	0.304	0.297
SocialMF	0.478	0.315	0.212	0.508	0.634	0.643	0.413	0.358	0.284
TrustSVD	0.471	0.302	0.213	0.517	0.614	0.620	0.395	0.336	0.295
RTARS	**0.518**	**0.394**	**0.317**	**0.626**	**0.685**	**0.738**	**0.472**	**0.431**	**0.374**

NDCG: normalised discounted cumulative gain; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System.
The best results are shown as bold face.

Table 14. Precision, recall and NDCG evaluation on the Epinions, Flixster and FilmTrust datasets for cold start users: length of recommendations list is equal to 5 (L = 5).

Algorithms	Epinions			Flixster			FilmTrust		
	P@5	R@5	N@5	P@5	R@5	N@5	P@5	R@5	N@5
UCF	0.056	0.112	0.108	0.254	0.315	0.211	0.122	0.390	0.243
ICF	0.082	0.149	0.125	0.296	0.376	0.234	0.147	0.416	0.251
TARS	0.097	0.201	0.129	0.379	0.436	0.251	0.171	0.449	0.282
Merge	0.292	0.486	0.342	0.605	0.638	0.496	0.265	0.588	0.394
RTCF	0.274	0.465	0.314	0.589	0.617	0.462	0.246	0.574	0.387
TrustMF	0.198	0.472	0.213	0.571	0.591	0.416	0.215	0.678	0.335
SocialMF	0.215	0.493	0.276	0.594	0.637	0.473	0.223	0.703	0.356
TrustSVD	0.321	0.517	0.361	0.625	0.685	0.507	0.224	0.615	0.361
RTARS	**0.363**	**0.554**	**0.382**	**0.657**	**0.718**	**0.524**	**0.336**	**0.723**	**0.425**

NDCG: normalised discounted cumulative gain; UCF: user-based collaborative filtering; ICF: item-based collaborative filtering; TARS: Trust-Aware Recommender System; RTCF: Reliability-based Trust-aware Collaborative Filtering; RTARS: Reputation-based Trust-Aware Recommender System. The best results are shown as bold face.

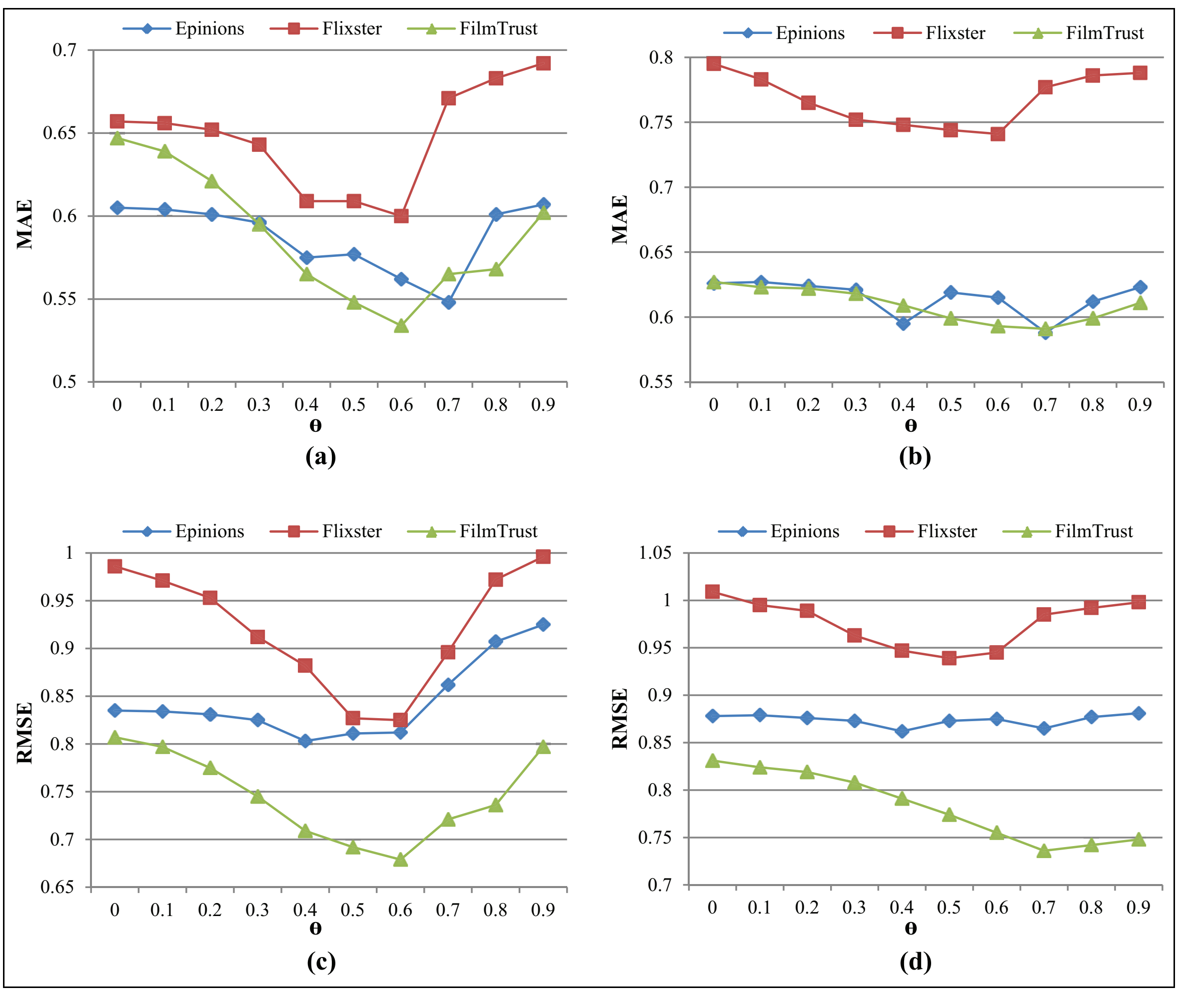

Figure 4. The effect of parameter θ on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

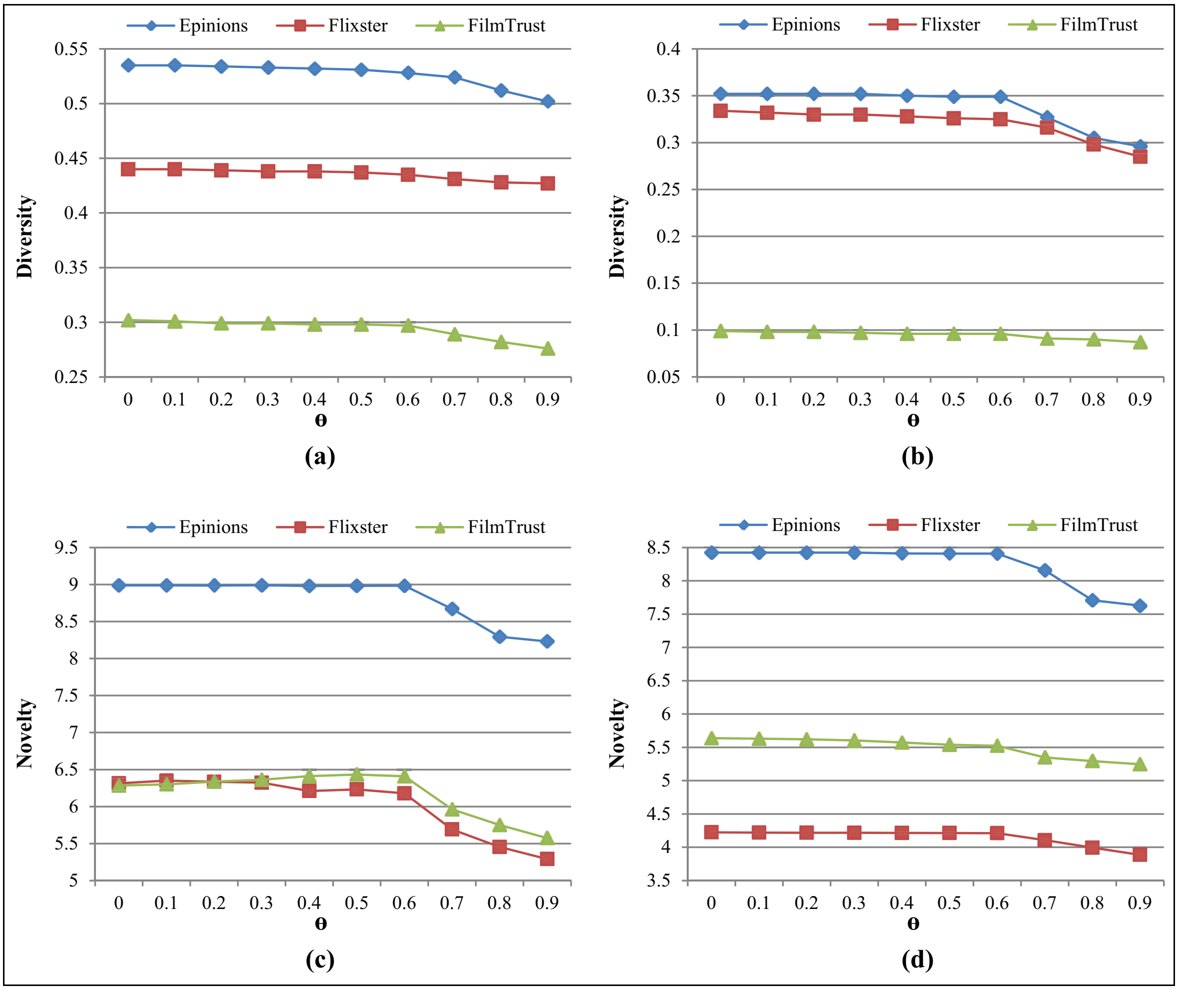

Figure 5. The effect of parameter θ on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

on the precision, recall and NDCG measures. These results indicate that the higher values of the parameter θ have a positive effect in most cases on the performance of the proposed method.

The parameter σ is a parameter in the range of $[0, 1]$ to control the effect of the rating-based and trust-based reputations on the proposed user reputation model in equation (18). The sensitivity analysis of the parameter σ is shown in Figures 7–9 based on different evaluation measures and also different views of the users for the Epinions, Flixster and FilmTrust datasets. These results indicate that the performance of the proposed method improves in most cases when the value of the parameter σ is increased. However, the performance of the proposed method will be reduced when the value of the parameter σ is higher than a specific value. For example, the diversity and novelty values of the proposed method decrease in most cases when the value of the parameter σ is higher than 0.7 (see Figure 8). Moreover, the MAE and RMSE values of the proposed method increase when the value of the parameter σ is higher than 0.5 (see Figure 7). Figure 9 indicates that in most cases, the higher values of the parameter σ lead to improve the performance of the proposed method based on the precision, recall and NDCG measures.

Figure 10 shows the results of the experiments based on different values of the parameter r for the MAE and RMSE measures. In addition, the results of the experiments for different values of the parameter r on the diversity and novelty measures are shown in Figure 11. This parameter is used in the proposed method as a threshold value for the user reliability measure to enhance the rating profiles of the users (see Section 3.3.4). It can be concluded from Figures 10 and 11 that the performance of the proposed method improves based on all of the evaluation measures while the value of the parameter r is increased. The results of the experiments based on different values of the parameter r are shown in Figure 12

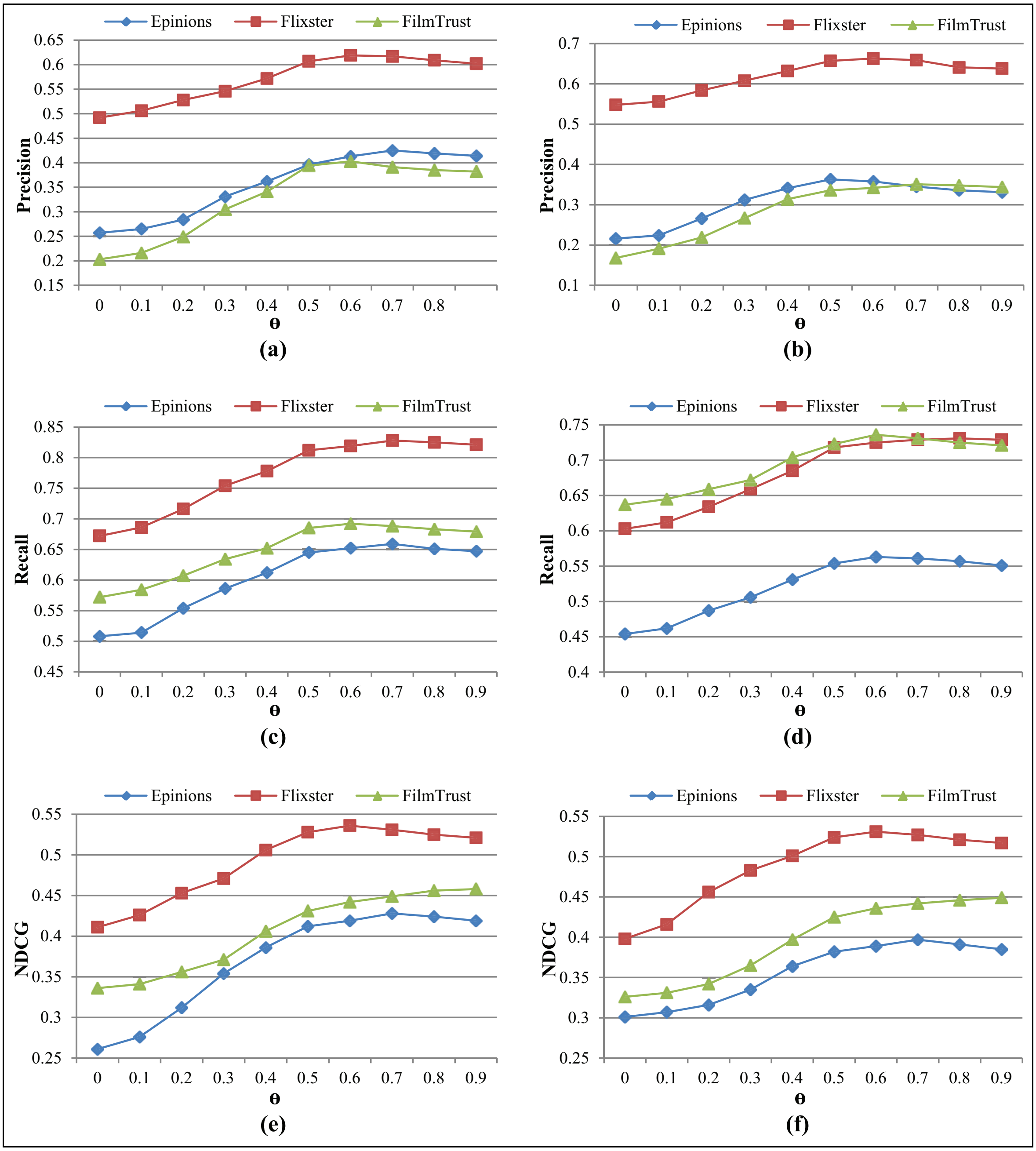

Figure 6. The effect of parameter θ on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

for the precision, recall and NDCG measures. These results indicate that the higher values of the parameter r have a positive effect on the performance of the proposed method.

Figures 13–15 report the results of the experiments for different values of the parameter m on the Epinions, Flixster and FilmTrust datasets. This parameter is used in equation (28) as the maximum number of virtual ratings that can be added to the user's rating profile. To this end, the values of the parameter m change from 5 to 20 with step 5. These results indicate that the higher value of the parameter m leads to improve the performance of the proposed method in terms of the diversity and novelty measures (see Figure 14). On the other hand, the MAE and RMSE values of the proposed method for the FilmTrust dataset increase when the value of the parameter m is increased (see Figure 13). Figure 15 indicates that the higher values of the parameter m lead to improve the performance of the proposed method in most

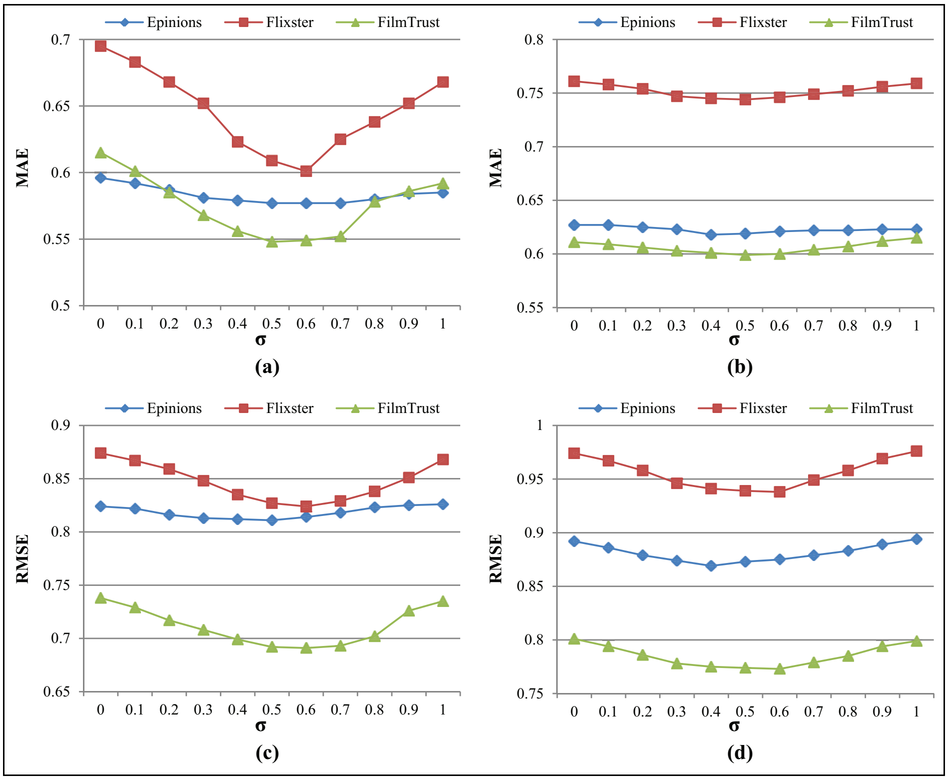

Figure 7. The effect of parameter σ on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

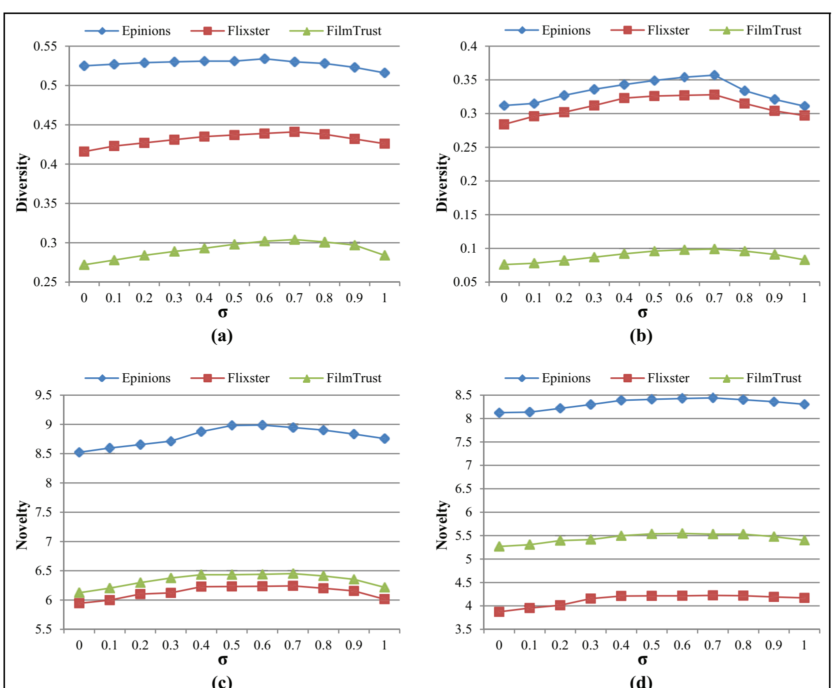

Figure 8. The effect of parameter σ on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

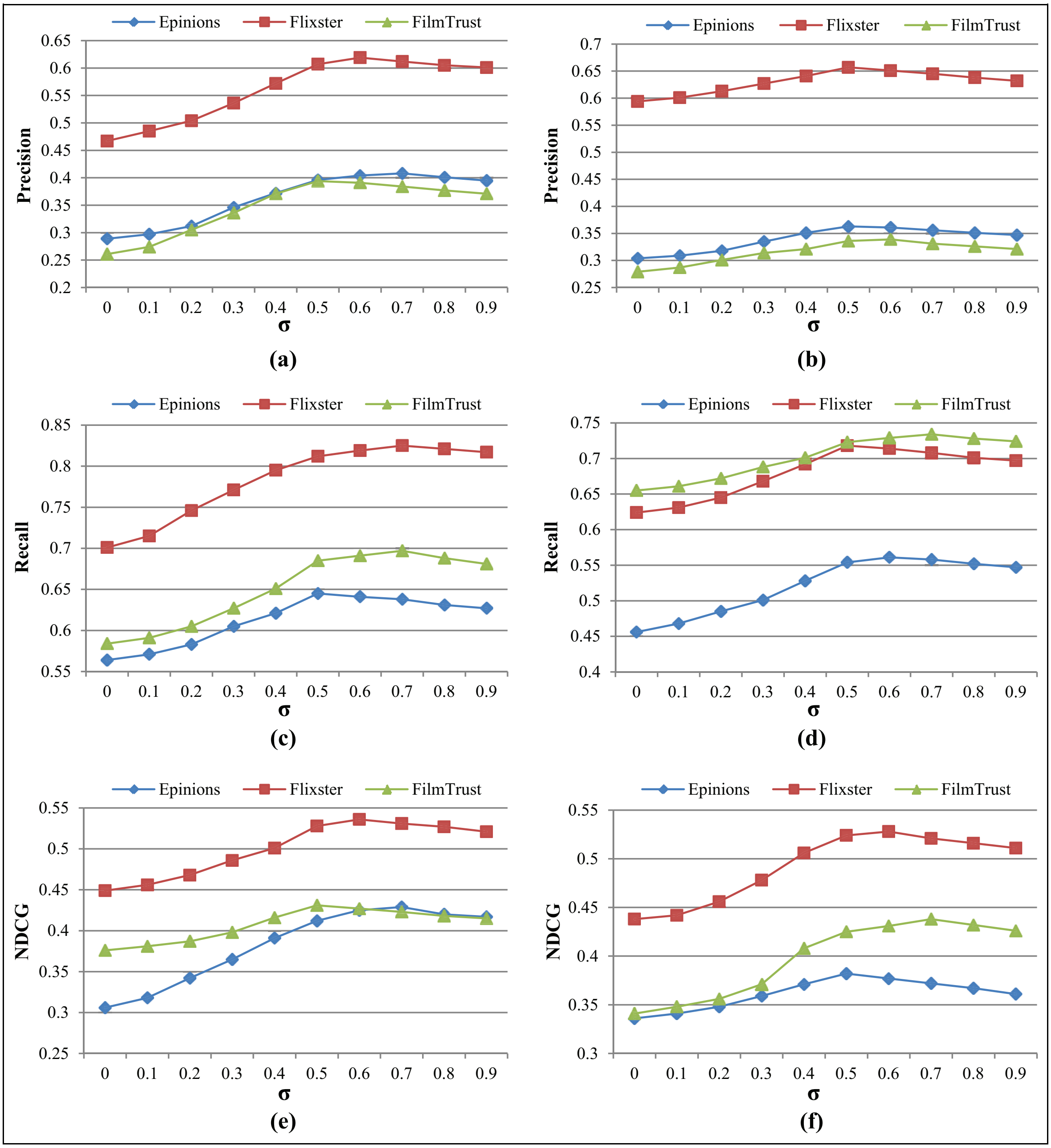

Figure 9. The effect of parameter σ on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

cases based on the Epinions and Flixster datasets. However, the performance of the proposed method decreases when the values of the parameter m are increased for the FilmTrust dataset.

Finally, α and β are two parameters which are used in the proposed method to control the effects of the reliability, novelty and diversity measures (see equation (29)). The effects of different values of the parameter α are shown in Figure 16 for the MAE and RMSE measures and also in Figure 17 for the diversity and novelty measures. As you can see from these figures, the values of the evaluation measures decrease in most cases when the values of the parameter α are increased. Therefore, the performance of the proposed method based on the MAE and RMSE measures will be improved in most cases when the value of the parameter α increases (see Figure 16). On the other hand, the higher value of the parameter α has a negative effect on the performance of the proposed method in terms of the diversity and novelty measures (see Figure 17). The experiments are repeated based on the precision, recall and NDCG measures and the

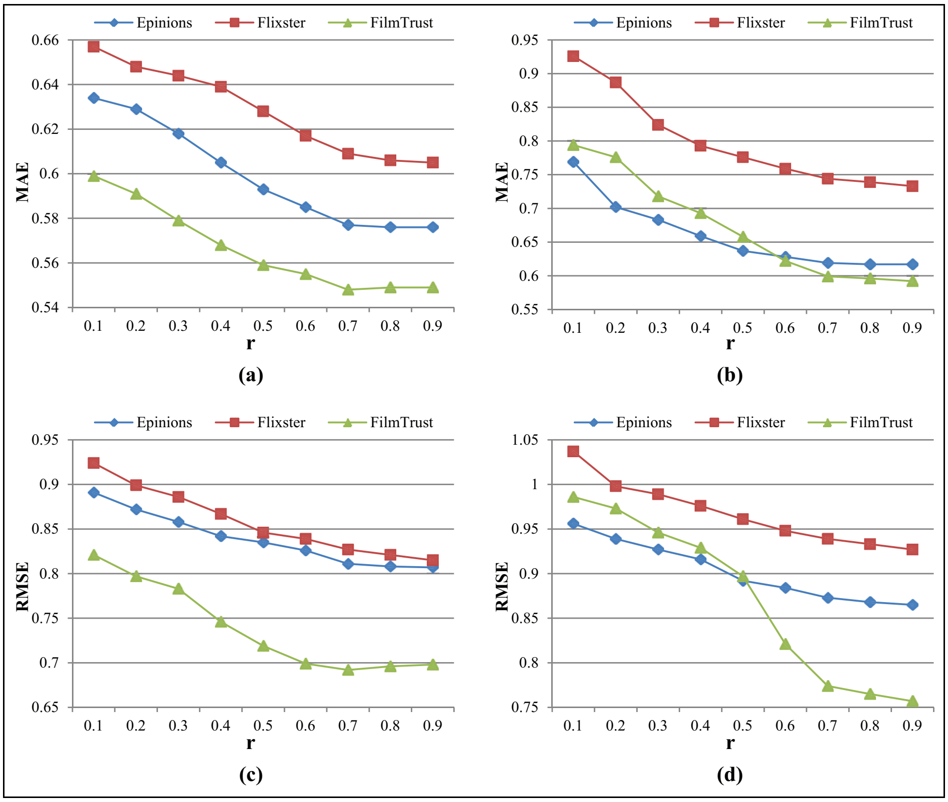

Figure 10. The effect of parameter *r* on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

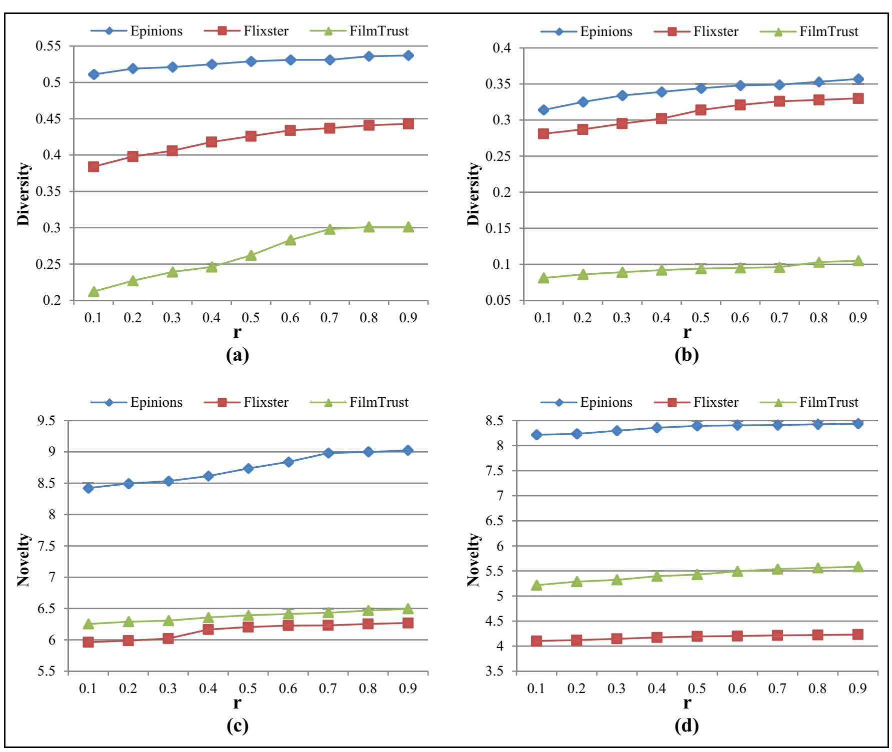

Figure 11. The effect of parameter *r* on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

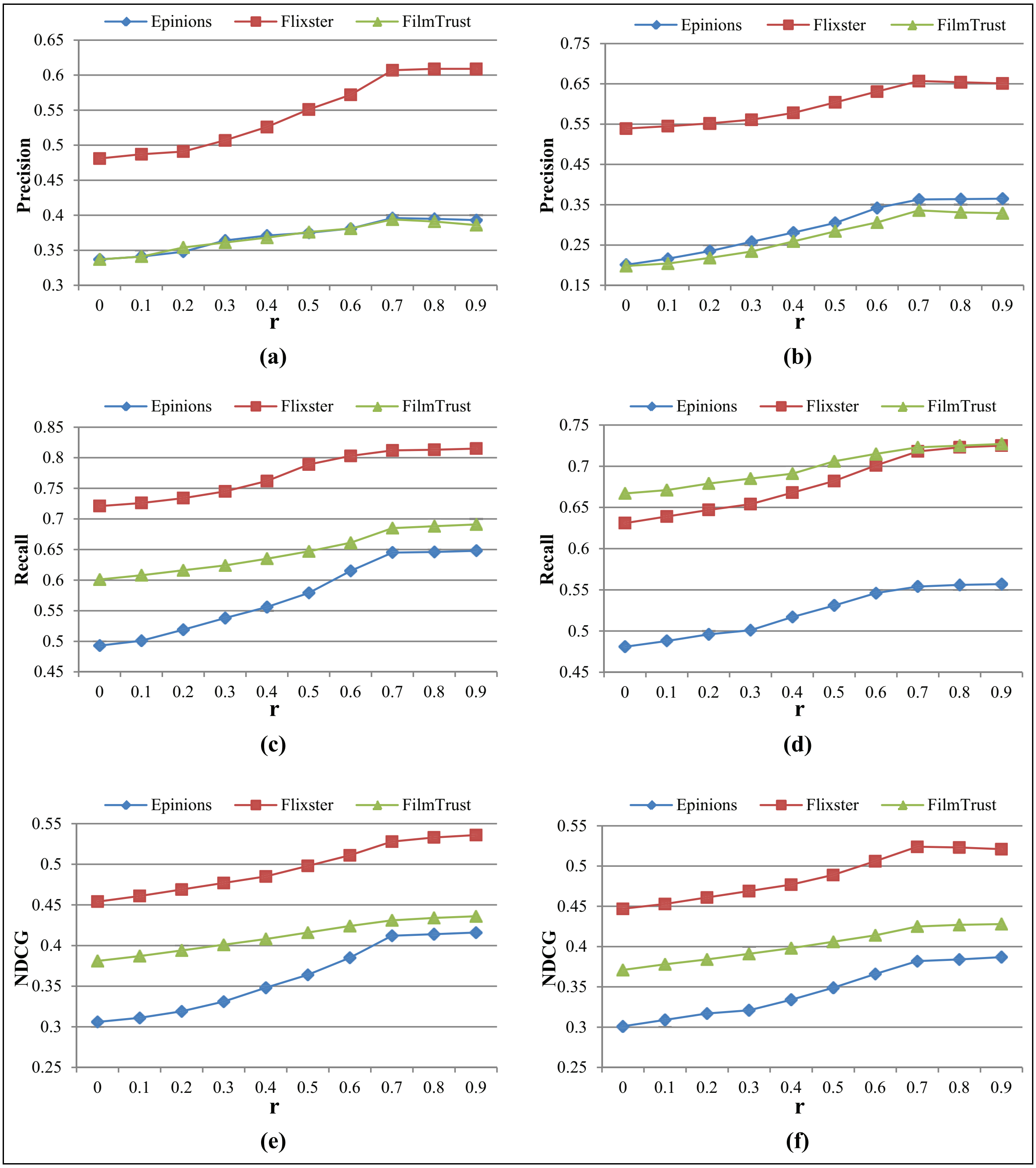

Figure 12. The effect of parameter *r* on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

results are shown in Figure 18. It can be concluded from these results that the higher values of the parameter α lead to improve the performance of the proposed method.

The experiments are repeated for different values of the parameter β and the results are shown in Figure 19 for the MAE and RMSE measures and also in Figure 20 for the diversity and novelty measures. As you can see from Figure 19, the MAE and RMSE measures decrease in most cases when the value of the parameter β is increased for the cold start view of the Epinions and FilmTrust datasets and also for the all users view of the Epinions dataset. Moreover, the performance of the proposed method will be decreased when the values of the parameter β are increased for the all users and cold start views of the Flixster dataset and also for the all users view of the FilmTrust dataset. On the

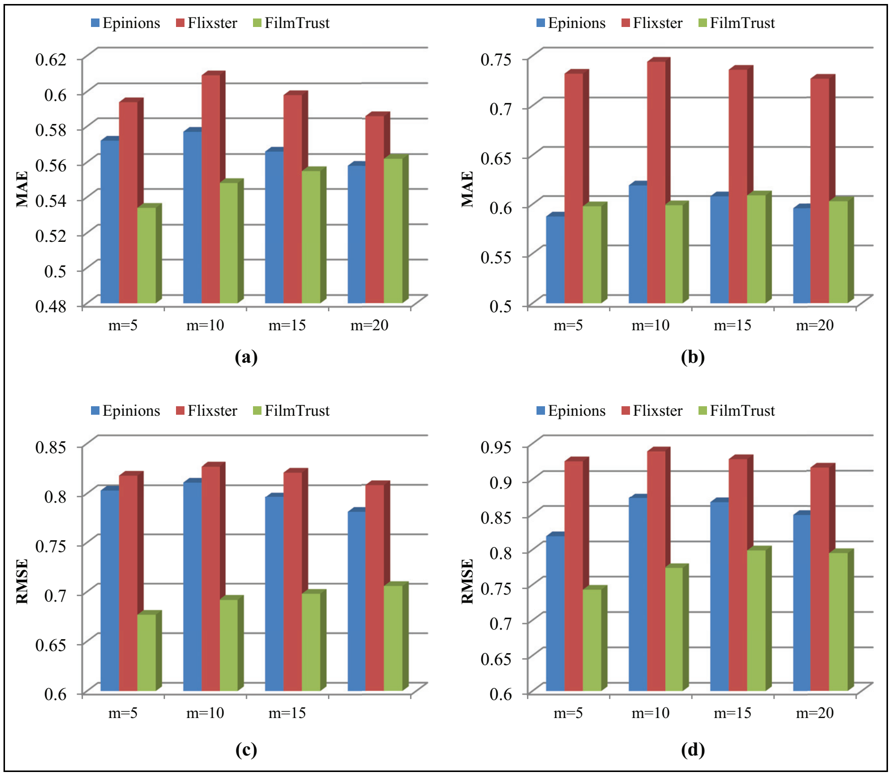

Figure 13. The effect of parameter *m* on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

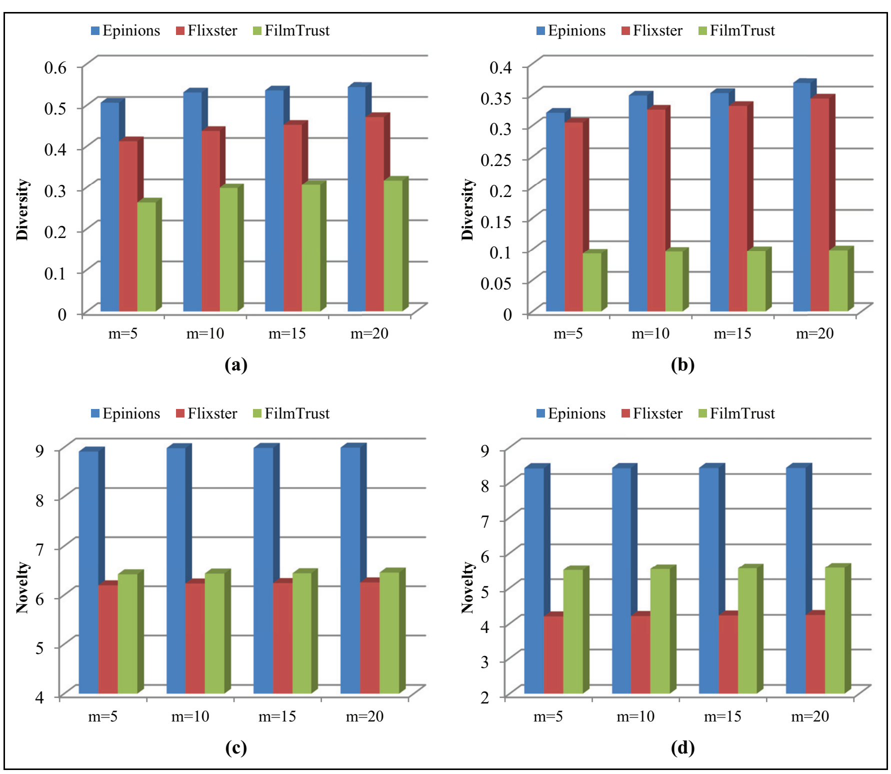

Figure 14. The effect of parameter *m* on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

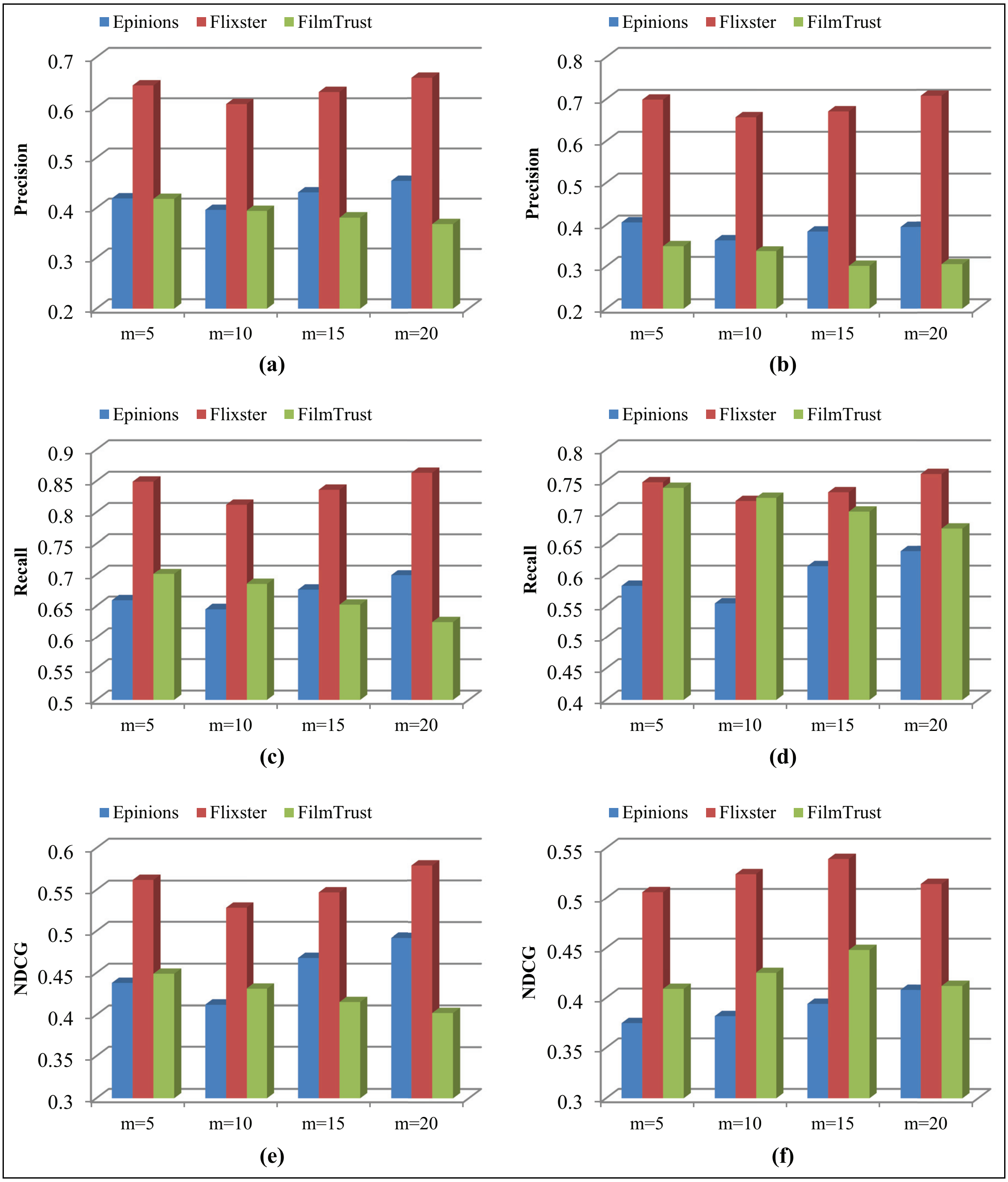

Figure 15. The effect of parameter *m* on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

other hand, Figure 20 indicates that the diversity and novelty measures of the proposed method improve when the value of the parameter β is increased. These results are expected, because the effect of the novelty measure increases for higher values of the parameter β (see equation (29)). The results of the experiments based on different values of the parameter β are shown in Figure 21 for the precision, recall and NDCG measures. It can be concluded from these results that the higher values of the parameter β have a negative effect on the performance of the proposed method.

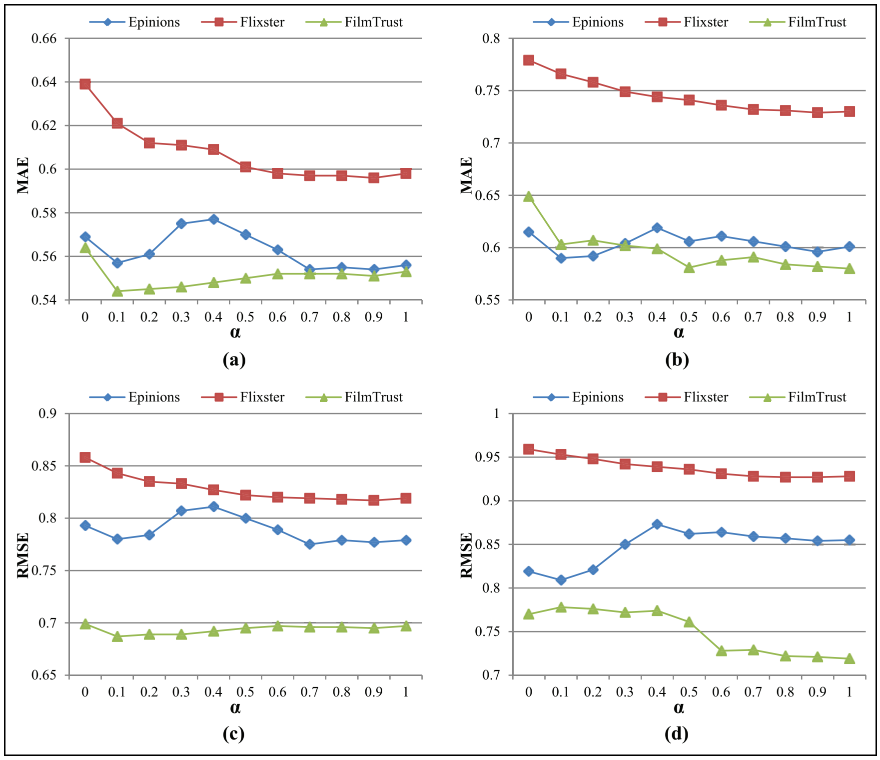

Figure 16. The effect of parameter α on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

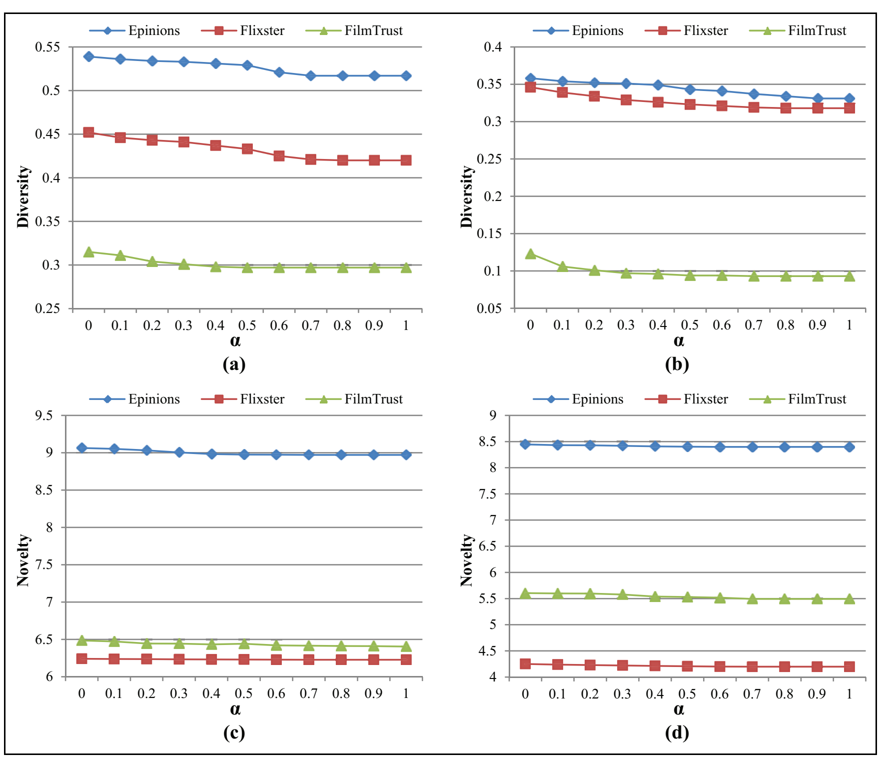

Figure 17. The effect of parameter α on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

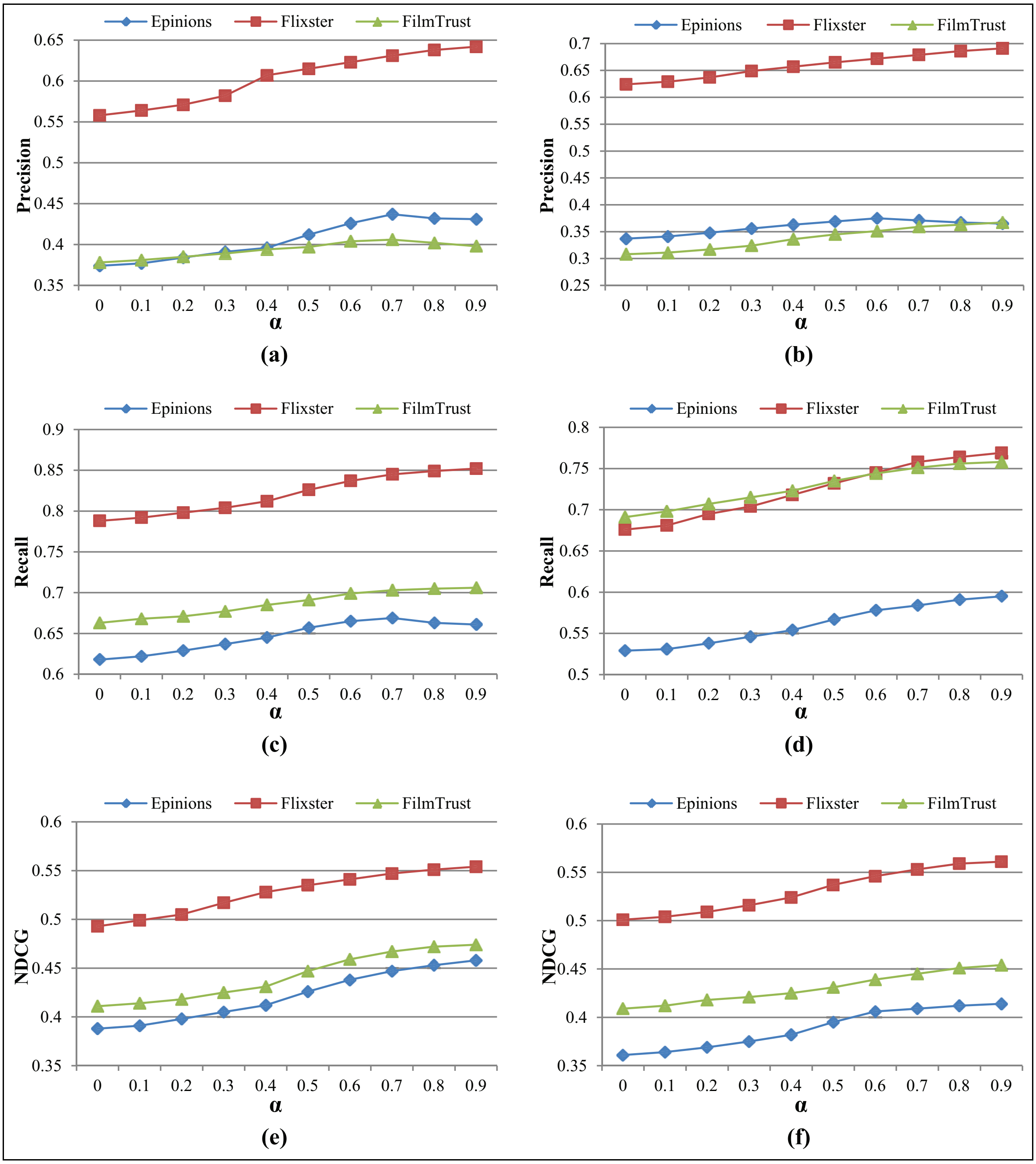

Figure 18. The effect of parameter α on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

4.6. Execution time of the proposed method

In this section, some experiments are performed to evaluate the execution time of the proposed method. The experiments are performed on a machine with a 3.1 GHz Intel Core-i5 CPU and 4 GB of RAM using C# programming language. Figure 22 indicates the results of the experiments based on the Epinions, Flixster and FilmTrust datasets. As you can see from these results, the execution time of the proposed method based on the FilmTrust dataset is less than the Epinions and Flixster datasets. These results are expected, because the number of users and items in the FilmTrust dataset is less than other datasets.

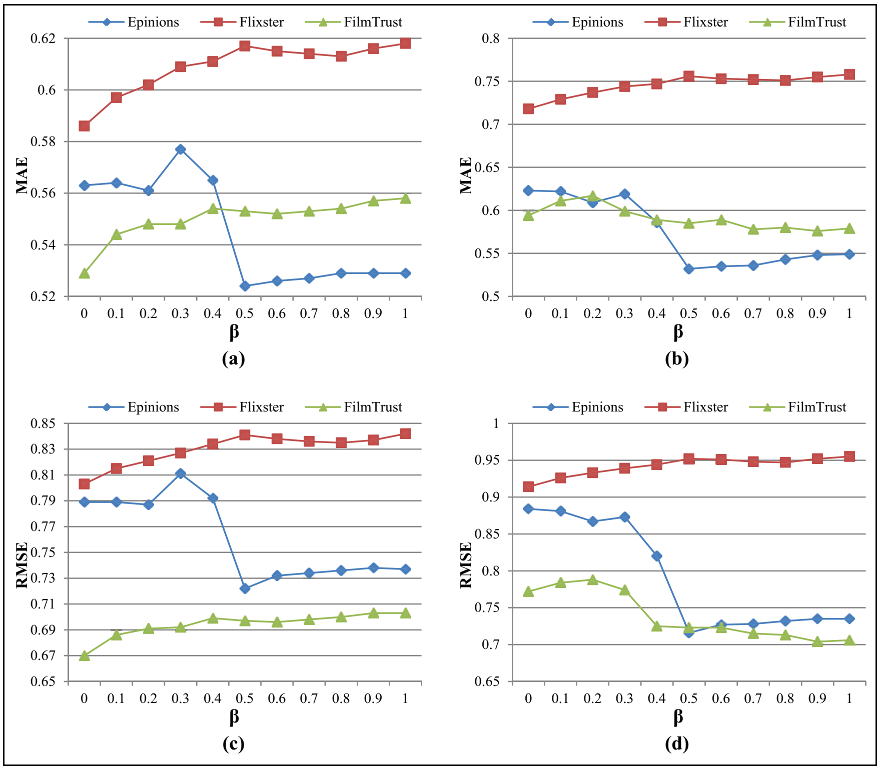

Figure 19. The effect of parameter β on the system performance: (a) MAE for all users, (b) MAE for cold users, (c) RMSE for all users and (d) RMSE for cold users.

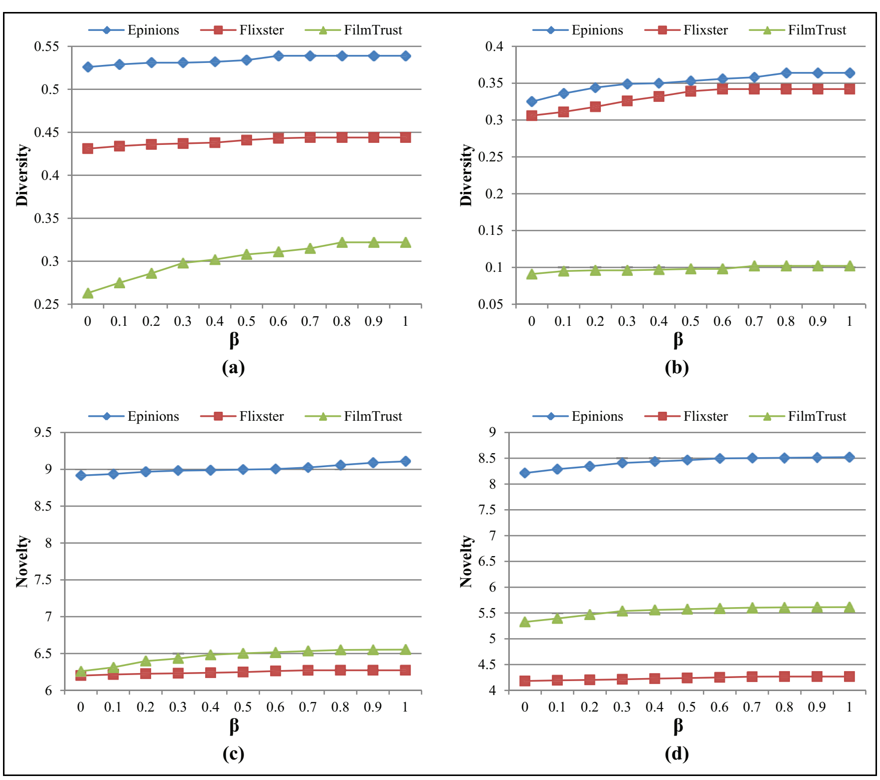

Figure 20. The effect of parameter β on the system performance: (a) diversity for all users, (b) diversity for cold users, (c) novelty for all users and (d) novelty for cold users.

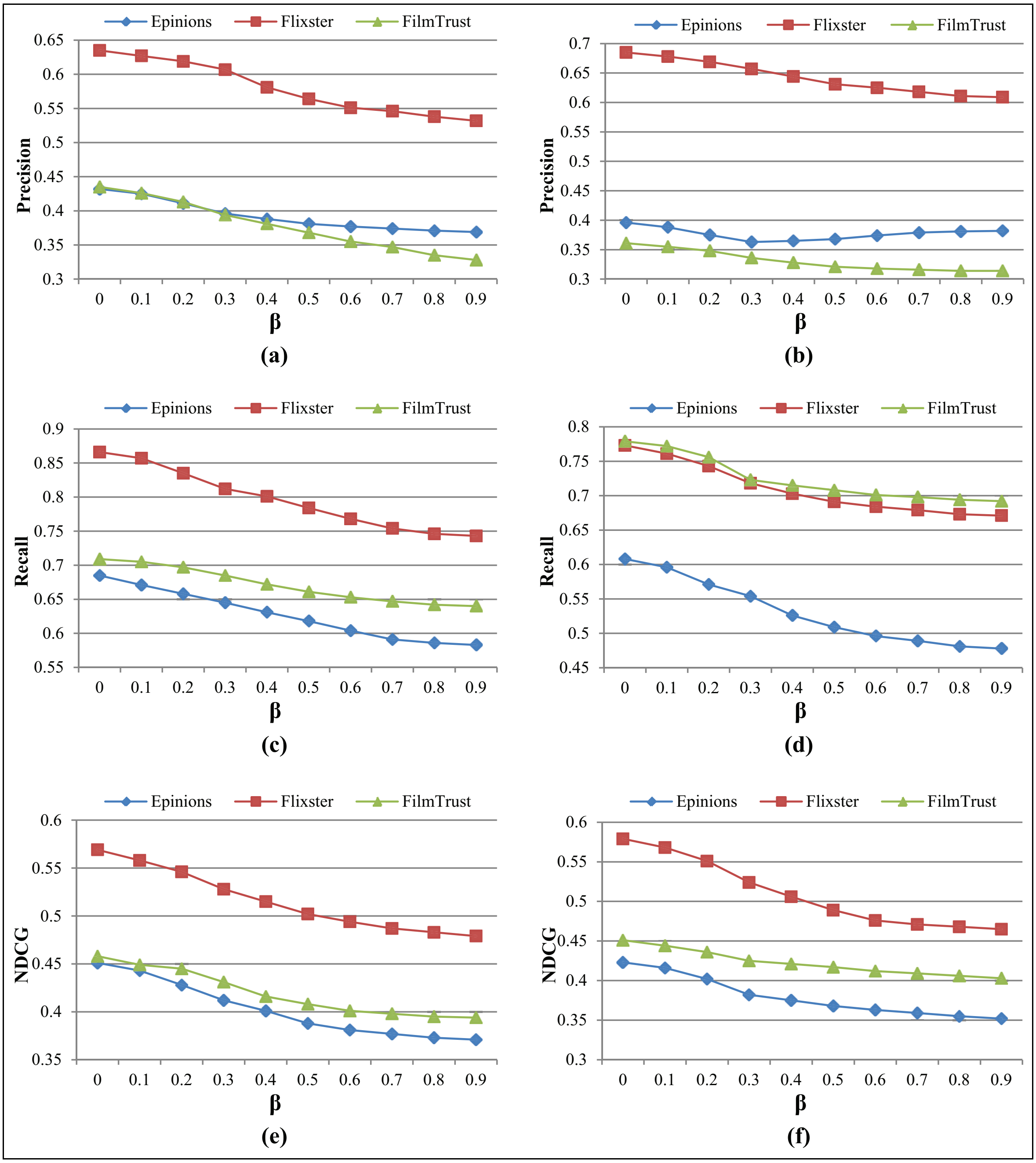

Figure 21. The effect of parameter β on the system performance: (a) precision for all users, (b) precision for cold users, (c) recall for all users, (d) recall for cold users, (e) NDCG for all users and (f) NDCG for cold users.

5. Conclusion and future work

Social network information such as trust relations between the users can be useful to improve the performance of the recommender systems. In particular, the trust information can alleviate cold start problem about the users who their rating profiles have not sufficient ratings to use in the recommendation system. However, the lack of trust relations between the users reduces the performance of the trust-aware recommender systems. Moreover, considering diversity and novelty measures of the recommendations can enhance user's satisfaction about the system. Therefore, a novel method is proposed in this article to improve trust-aware recommender systems with considering the reliability, diversity and novelty measures. To this end, a user reliability measure is introduced to evaluate the performance of the rating profiles and trust

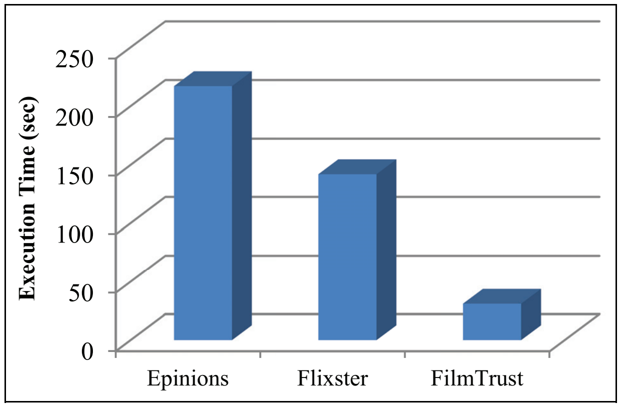

Figure 22. The execution time (in second) of the proposed method.

networks of the users. In addition, a novel mechanism is proposed to enhance the rating profiles of the users who their user reliability measures are lower than a threshold value. These rating profiles are enhanced by adding a number of virtual ratings which are predicted based on a novel user reputation model. In the proposed rating profile enhancement mechanism, the diversity, novelty and reliability measures of the items are considered to improve the recommendation results of the system. Experiments results based on three real-world datasets showed that the proposed method outperformed other state-of-the-art approaches significantly.

There are some directions to improve the proposed method which can be considered as future works. First, using additional social information such as distrust relations between the users may improve the efficiency of the present work. Second, incorporating temporal effect into recommendation process to capture the changes of user's interests over time can be useful to improve the recommendation lists of the users. Third, the weights of the reliability, diversity and novelty measures for the items can be mathematically determined as a novel approach to enhance the proposed method based on these measures. Finally, proposing a noise detection and correction method may improve the accuracy of the predicted ratings by detecting the noisy virtual ratings.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. http://www.trustlet.org/datasets/downloaded_epinions
2. http://www.flixster.com
3. http://trust.mindswap.org/FilmTrust/

References

[1] Beladev M, Rokach L and Shapira B. Recommender systems for product bundling. *Knowl-Based Syst* 2016; 111: 193–206.
[2] Domingues MA, Jorge AM and Soares C. Dimensions as virtual items: improving the predictive ability of top-N recommender systems. *Inform Process Manag* 2013; 49: 698–720.
[3] Shimon DB, Rokach L and Shapira B. An ensemble method for top-N recommendations from the SVD. *Expert Syst Appl* 2016; 64: 84–92.
[4] Braida F, Mello CE, Pasinato MB et al. Transforming collaborative filtering into supervised learning. *Expert Syst Appl* 2015; 42: 4733–4742.

[5] Hernando A, Moya R, Ortega F et al. Hierarchical graph maps for visualization of collaborative recommender systems. *J Inform Sci* 2014; 40: 97–106.
[6] Ghazarian S and Nematbakhsh MA. Enhancing memory-based collaborative filtering for group recommender systems. *Expert Syst Appl* 2015; 42: 3801–3812.
[7] Shi Y, Larson M and Hanjalic A. Unifying rating-oriented and ranking-oriented collaborative filtering for improved recommendation. *Inform Sciences* 2013; 229: 29–39.
[8] Salah A, Rogovschi N and Nadif M. A dynamic collaborative filtering system via a weighted clustering approach. *Neurocomputing* 2016; 175: 206–215.
[9] Moradi P, Ahmadian S and Akhlaghian F. An effective trust-based recommendation method using a novel graph clustering algorithm. *Physica A* 2015; 436: 462–481.
[10] Nilashi M, Ibrahim O and Ithnin N. Multi-criteria collaborative filtering with high accuracy using higher order singular value decomposition and Neuro-Fuzzy system. *Knowl-Based Syst* 2014; 60: 82–101.
[11] Ortega F, Hernando A, Bobadilla J et al. Recommending items to group of users using matrix factorization based collaborative filtering. *Inform Sciences* 2016; 345: 313–324.
[12] Sun J, Wang G, Cheng X et al. Mining affective text to improve social media item recommendation. *Inform Process Manag* 2015; 51: 444–457.
[13] Lee JW, Kim HJ and Lee SG. A probability-based unified framework for semantic search and recommendation. *J Inform Sci* 2013; 39: 608–628.
[14] Peng F, Lu J, Wang Y et al. N-dimensional Markov random field prior for cold-start recommendation. *Neurocomputing* 2016; 91: 187–199.
[15] Ocepek U, Rugelj J and Bosnić Z. Improving matrix factorization recommendations for examples in cold start. *Expert Syst Appl* 2015; 42: 6784–6794.
[16] Xie F, Chen Z, Shang J et al. Grey forecast model for accurate recommendation in presence of data sparsity and correlation. *Knowl-Based Syst* 2014; 69: 179–190.
[17] Hoseini E, Hashemi S and Hamzeh A. SPCF: a stepwise partitioning for collaborative filtering to alleviate sparsity problems. *J Inform Sci* 2012; 38: 578–592.
[18] Ricci F, Rokach L and Shapira B. *Recommender systems handbook*. New York: Springer, 2011.
[19] Xie H and Lui JCS. Mathematical modeling and analysis of product rating with partial information. *ACM T Knowl Discov D* 2015; 9: 26.
[20] Wu H, Yue K, Pei Y et al. Collaborative topic regression with social trust ensemble for recommendation in social media systems. *Knowl-Based Syst* 2016; 97: 111–122.
[21] Seo YD, Kim YG, Lee E et al. Personalized recommender system based on friendship strength in social network services. *Expert Syst Appl* 2017; 69: 135–148.
[22] Moradi P and Ahmadian S. A reliability-based recommendation method to improve trust-aware recommender systems. *Expert Syst Appl* 2015; 42: 7386–7398.
[23] Liu Y, Lin Z, Zheng X et al. Incorporating social information to perform diverse replier recommendation in question and answer communities. *J Inform Sci* 2016; 42: 449–464.
[24] Fang H, Guo G and Zhang J. Multi-faceted trust and distrust prediction for recommender systems. *Decis Support Syst* 2015; 71: 37–47.
[25] Mazhari S, Fakhrahmad SM and Sadeghbeygi H. A user-profile-based friendship recommendation solution in social networks. *J Inform Sci* 2015; 41: 284–295.
[26] Guo G, Zhang J and Thalmann D. Merging trust in collaborative filtering to alleviate data sparsity and cold start. *Knowl-Based Syst* 2014; 57: 57–68.
[27] Yang B, Lei Y, Liu D et al. Social collaborative filtering by trust. In: *Proceedings of the 23th international joint conference on artificial intelligence (IJCAI)*, Beijing, China, 3–9 August 2013, pp. 2747–2753. New York: IEEE.
[28] Liu CL and Wu XW. Large-scale recommender system with compact latent factor model. *Expert Syst Appl* 2016; 64: 467–475.
[29] Véras D, Prota T, Bispo A et al. A literature review of recommender systems in the television domain. *Expert Syst Appl* 2015; 42: 9046–9076.
[30] Lai CH, Liu DR and Liu ML. Recommendations based on personalized tendency for different aspects of influences in social media. *J Inform Sci* 2015; 41: 814–829.
[31] Chen CC, Wan YH, Chung MC et al. An effective recommendation method for cold start new users using trust and distrust networks. *Inform Sciences* 2013; 224: 19–36.
[32] Guo G, Zhang J and Smith NY. Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems. *Knowl-Based Syst* 2015; 74: 14–27.
[33] Lee WP and Ma CY. Enhancing collaborative recommendation performance by combining user preference and trust-distrust propagation in social networks. *Knowl-Based Syst* 2016; 106: 125–134.
[34] Guo G, Zhang J and Smith NY. TrustSVD: collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In: *Proceedings of the 29th AAAI conference on artificial intelligence*, Austin, TX, 25–30 January 2015, pp. 123–129. New York: IEEE.

[35] Jamali M and Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: *Proceedings of the 4th ACM conference on recommender systems (RecSys)*, Barcelona, 26–30 September 2010, pp. 135–142. New York: IEEE.

[36] Davoudi A and Chatterjee M. Social trust model for rating prediction in recommender systems: effects of similarity, centrality, and social ties. *Online Soc Netw Media* 2018; 7: 1–11.

[37] Lee J, Noh G, Oh H et al. Trustor clustering with an improved recommender system based on social relationships. *Inform Syst* 2018; 77: 118–128.

[38] Li J, Chen C, Chen H et al. Towards Context-aware Social Recommendation via Individual Trust. *Knowl-Based Syst* 2017; 127: 58–66.

[39] Ghavipour M and Meybodi MR. Trust propagation algorithm based on learning automata for inferring local trust in online social networks. *Knowl-Based Syst* 2018; 143: 307–316.

[40] Lingam G, Rout RR and Somayajulu D. Learning automata-based trust model for user recommendations in online social networks. *Comput Electr Eng* 2018; 66: 174–188.

[41] Chen LJ and Gao J. A trust-based recommendation method using network diffusion processes. *Physica A* 2018; 506: 679–691.

[42] Faridani V, Jalali M and Jahan MV. Collaborative filtering-based recommender systems by effective trust. *Int J Data Sci Anal* 2017; 3: 297–307.

[43] Li H, Zhang S, Hu Y et al. Research of social recommendation based on social tag and trust relation. In: Hariri S (ed.) *Cluster computing*. New York: Springer, 2017, pp. 1–11.

[44] Liu H, Bai X, Yang Z et al. Trust-aware recommendation for improving aggregate diversity. *New Rev Hypermedia M* 2015; 21: 242–258.

[45] Servajean M, Akbarinia R, Pacitti E et al. Profile diversity for query processing using user recommendations. *Inform Syst* 2015; 48: 44–63.

[46] Zhang M and Hurley N. Avoiding monotony: improving the diversity of recommendation lists. In: *Proceedings of the 2nd ACM conference on recommender systems (RecSys)*, Lausanne, 23–25 October 2008, pp. 123–130. New York: ACM.

[47] Lee K and Lee K. Escaping your comfort zone: a graph-based recommender system for finding novel recommendations among relevant items. *Expert Syst Appl* 2015; 42: 4851–4858.

[48] Zhang M, Guo X and Chen G. Prediction uncertainty in collaborative filtering: enhancing personalized online product ranking. *Decis Support Syst* 2016; 83: 10–21.

[49] Ahmadian S, Meghdadi M and Afsharchi M. A social recommendation method based on an adaptive neighbor selection mechanism. *Inform Process Manag* 2018; 54: 707–725.

[50] Massa P and Avesani P. Trust-aware recommender systems. In: *Proceedings of the ACM conference on recommender systems, (RecSys)*, Minneapolis, MN, 19–20 October 2007. New York: ACM.

[51] Zhou YB, Lei T and Zhou T. A robust ranking algorithm to spamming. *Europhys Lett* 2011; 94: 48002.

[52] Page L, Brin S, Motwani R et al. The pagerank citation ranking: bringing order to the web. *Stanford Infolab*, 1999, http://ilpubs.-stanford.edu:8090/422/

[53] Erkan G and Radev DR. LexRank: graph-based lexical centrality as salience in text summarization. *J Artif Intell Res* 2004; 22: 457–479.

[54] Hurley N and Zhang M. Novelty and diversity in top-N recommendations analysis and evaluation. *ACM T Intern Techn* 2011; 10: 1–29.

[55] Breese JS, Heckerman D and Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. In: *Proceedings of the 14th conference on uncertainty in artificial intelligence*, Madison, 24–26 July, pp. 43–52. San Francisco, CA: Morgan Kaufmann Publishers Inc.

[56] Shannon CE. A mathematical theory of communication. *ACM SIGMOBILE Mob Comput Commun Rev* 2001; 5: 3–55.

[57] Herlocker JL, Konstan JA, Terveen LG et al. Evaluating collaborative filtering recommender systems. *ACM T Inform Syst* 2004; 22: 5–53.

[58] Yuan W, Guan D, Lee YK et al. Improved trust-aware recommender system using small-worldness of trust networks. *Knowl-Based Syst* 2010; 23: 232–238.

[59] Sarwar B, Karypis G, Konstan J et al. Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the 10th international conference on World Wide Web*, Hong Kong, 1–5 May 2001, pp. 285–295. New York: ACM.