Journal of Computers Vol. 31 No. 5, 2020, pp. 261-276
doi:10.3966/199115992020103105020

A Trusted Hybrid Recommendation Method of Web Services Considering Timeliness

Hongguo Zhang[1], Xiaoqi Liu[1], Chao Ma[2*], Zhou Fang[3], Huazhu Chen[2], Haining Yu[4]

[1] School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang 150040, China

[2] School of Software & Microelectronics, Harbin University of Science and Technology, Harbin, Heilongjiang 150040, China

[3] Heilongjiang Province Cyberspace Research Center, Harbin, Heilongjiang 150001, China

[4] Research Center of Computer Network and Information Security Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China
{zhg07, xiaoqi102, machao8396, fangzhou8396, huazhu0634}@163.com, yuhaining@hit.edu.cn

Received 7 July 2019; Revised 14 December 2019; Accepted 20 January 2020

Abstract. The rapid development of the Internet has made the number of candidate Web services satisfying users' needs grow rapidly. How to effectively find Web services that meet the needs of users' Mashup from such a large-scale Web service collection has become a major research issue. Therefore, aiming at how to recommend appropriate Web services to build high-quality Mashup application, this paper proposes a trusted hybrid recommendation method of Web services considering timeliness combining rating timeliness, trust and tag information. First, in the use of rating, the time of service rating is considered to solve the problem of user preference transfer. Secondly, considering the trust relationship between users in the social network, the trust degree between users is calculated to solve the malicious recommendation. Finally, so as to alleviate the issues of data sparsity, the tag information has imported and utilized to improve the recommendation method of Web services. The experimental results show that compared with other existing methods, the method could improve the accuracy and credibility of service recommendation and have a good ability to resist malicious recommendation attacks.

Keywords: hybrid recommendation, tag information, timeliness, trust, Web services

1 Introduction

With the rapid development of technologies such as cloud computing, mobile Internet, and Internet of Things, a large number of Internet applications of Service-Oriented Architecture (SOA) have been created. The number and service functions of Web services are growing rapidly, which makes users have to spend a lot of time and effort to choose appropriate Web services. Users meet their needs from a large number of candidate Web services. In recent years, a lightweight service composition model "Mashup", has emerged on the Internet. It combines two or more Web services to create a brand new Web application [1]. Mashup technology is mainly to help software developers to combine Web services on the Internet to build a composite service that meets the complex needs of users. At present, under the idea of service computing "Servitization", the number of Web services has multiplied. Therefore, how to choose the most appropriate Web service from a large number of candidate services with similar functions to join the Mashup application is an urgent problem to be solved, and it is also the main content of this paper [2].

* Corresponding Author

In response to this problem, some researchers use service recommendations to select Web services. Service recommendation is to recommend a more appropriate service for user according to the target user's own characteristics and interest preferences, and it is an important method of service selection. At present, the following methods are mainly included in the service recommendation method: content-based service recommendation [3], collaborative filtering service recommendation [4], and hybrid service recommendation [5]. These recommendation methods greatly improve the effectiveness of service recommendations, but also have some problems.

In recent years, more and more scholars have done more research on service recommendation. Through a lot of experiments, it is found that applying collaborative filtering technology to the recommendation system can produce good results. The application of the collaborative filtering technology to the service recommendation is mainly based on the rating of the service used by the target user. It determines the user whose service rating behavior is similar to the target user and the similar users are regarded as neighbor users. The target user's rating of the services is predicted by calculating the rating of the different services used by the neighbors, and the service recommendation of the target user is implemented according to the predicted rating of these services. The existing recommendation methods have the following problems: First of all, it is not accurate to find similar users, but only calculate the similarity based on users' initial rating of the service, and the data of the rating matrix is relatively sparse. Secondly, most of the recommendation methods are based on reliable information. But in reality, users may viciously degrade or enhance a service. Existing recommendation methods are not good at resisting malicious recommendations. Therefore, aiming at the problems of data sparseness and malicious recommendation existing in the collaborative filtering recommendation, this paper proposes a trusted hybrid recommendation method for Web services considering timeliness. The main research contents are as follows:

(1) The method in this paper takes full account of individual rating habits and timeliness when calculating similarity using user rating information. It solves the problem of shifting user preferences.

(2) In the Web 2.0 era, tags as an important information resource not only express the main features of services, but also reflect users' preferences. In this paper, the time when users use tags and the degree of preference for tags are taken into consideration to calculate the similarity of tags. Tags are used to expand the neighbor users of target user, so as to alleviate the problem of sparse data.

(3) With the development of the Internet, a variety of social software has become popular, and recommendations from friends tend to have high credibility. Therefore, the introduction of trust relationship in social networks in recommendation methods can better resist malicious recommendations. This paper considers two aspects of direct trust and indirect trust comprehensively. Direct trust is achieved through the interaction between users and the position of recommended users in social networks, while the user's indirect trust can be obtained through the length, quantity and direct trust of the relationship path between users.

The following contents of this paper are as follows. Section 2 introduces the background and related work of this paper; Section 3 gives the calculation of similarity and trust among users in detail, screens out the trusted users of target users, and gives a trusted hybrid recommendation method for Web services considering timeliness; Section 4 proves the feasibility of the method by experimental results; Finally, the paper summarizes and prospects.

2 Background and Related Work

In recent years, service recommendation has become a more active research direction in the field of service computing. Many scholars have conducted in-depth research on it. Among them, the service recommendation method introduced collaborative filtering technology has achieved good results and improved the accuracy of the recommendation. Collaborative filtering technology is mainly divided into the following two categories [6]: (1) Neighborhood-based collaborative filtering, which is to recommend related items to target users based on similar users or items; (2) Collaborative filtering based on matrix decomposition is mainly to facilitate the calculation by reducing the rating matrix into two low-order matrices, and using two low-order matrices to recommend the target users. But collaborative filtering technology still faces some challenges of cold start and data sparsity.

Traditional collaborative filtering algorithms mostly rely on rating information to calculate similarity while ignoring tag information. Social tagging, as an important information resource of Web2.0, can reflect user's thoughts and preferences. Therefore, many scholars have introduced tag information into the service recommendation method based on collaborative filtering. Jaschke et al. incorporated tag information into the recommendation process [7]. The user-item-tag triple is broken down into user-item and user-tag duals. The method selects user neighbors from two aspects to generate recommendations. In 2010, Kim et al. used relatively dense tag information to determine the set of tags that the user was interested in [8]. Then it combined with the Bayesian probability model to generate project recommendations. The methods proposed in the above literature can alleviate the sparsity and user cold start problems to some extent. However, a single tag is used as the data information of user's preference will ignore the impact of the score on the tag's preference. Zhang et al. proposed a new recommendation framework [10]. It mainly uses user interest expansion strategies based on social tag information and enhances the diversity of user preferences by extending the size and category of the original user-project interaction record. A traditional recommendation model is used to generate a list of recommendations. Zhang et al. proposed a collaborative filtering recommendation algorithm based on tag optimization [11]. When using user-based collaborative filtering algorithm to calculate user similarity to find neighbors, the tag information is used to find extended neighbors. It can alleviate the problem of data sparsity. The above literature uses tag information to extend user preferences. However, the timeliness of the tag is not considered so as to affect the accuracy of the recommendation.

The social recommendation algorithm [12] has become one of the important research directions in the field of recommendation systems in academia and industry. With the advent of the Web2.0 network era, social networks have developed rapidly, especially the popularity of online social networks. Introducing the trust relationship information into the recommendation algorithm has become a hot research direction of the recommendation system. Yu et al. proposed a collaborative filtering algorithm based on social networks [13]. It integrates the user's trust and preferences based on social network information to find the nearest neighbor of the target user. Forsati R et al. proposed a matrix social recommendation model [14]. The model incorporates trust and distrust information between users to improve the quality of recommendations. The above literature has improved the accuracy of recommendations and mitigated malicious recommendation. Because the indirect trust relationship is not mentioned when considering the trust relationship, the measure of trust is not accurate enough. Wang et al. proposed a service recommendation method based on Trusted Alliance [4]. This method introduced the recommended attribute characteristics of the service, and built the trusted alliance of neighbor users in combination with the trust degree of service recommendation. It improves the credibility of the service recommender while improving the accuracy of the service recommendation. Lin Xiao et al. [15] proposed a new user preference model. The model takes into account the visibility of the project and social relationships, and it coordinates the two types of information in a unified model inspired by the transfer learning philosophy. It solves the problem of recommendation system data sparseness and obtains better recommendation results. Li et al. proposed a collaborative filtering algorithm based on trust network enhancement [16]. The algorithm uses the rating behavior to establish the trust relationship between users on the original rating matrix. It fully considers the transitivity of trust and builds a user trust network, which alleviates the problem of data sparsity and improves the recommendation accuracy to some extent. The above literature has added trust relationships to improve the accuracy of recommendations and the cold start problem. They build trust networks taking into account the transitivity of trust. However, other features in the trust relationship and interactions between users are ignored. Therefore, a more comprehensive consideration of the trust relationship between users is still a problem worth studying.

It can be seen from the above literature that the existing service recommendation methods have some problems. (1) In the tag-based recommendation method, tag information is used as a single data information in literature [7] and [8], and tag information is used as an extension in literature [10] and [11]. Compared with the method in this paper, calculation of similarity of these literatures is not accurate enough because they integrate the tag information into methods to calculate the similarity, ignoring the timeliness of the tag and the user's preference for the tag. (2) In the socialization recommendation method, literature [12] and [13] introduced trust relationship into recommendation, while literature [14] and [15] considered trust transitivity and other issues. But they do not take into account the full extent of the calculation of trust. There are many important factors in social network that can reflect the trust relationship between users. This paper fully considers these factors, such as interaction time, intimacy between users, and the core degree of users in social relations. At the same time, in addition to the direct trust relationship between users, there is also an indirect relationship, which is also a very important part.

Therefore, the indirect trust relationship between users can be added to the calculation of trust degree, which can achieve more accurate recommendation.

3 A Trusted Hybrid Recommendation Method for Web Services Considering Timeliness

In this paper, a trusted hybrid recommendation method for Web services considering timeliness is proposed, which integrates user similarity based on timeliness, user similarity based on timeliness tag and trust among users in social networks. The model of service recommendation process is shown in Fig. 1.

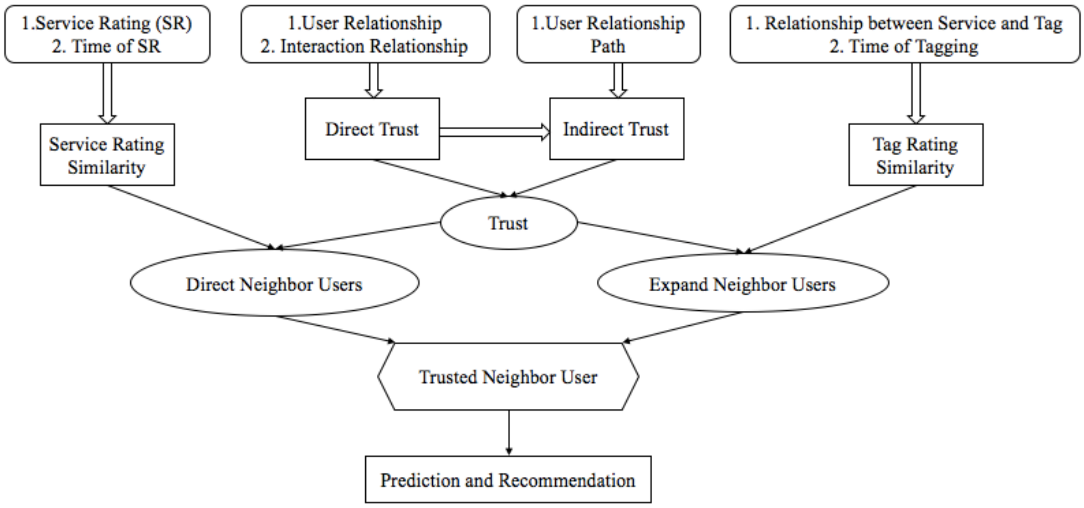

Fig. 1. Trusted hybrid recommendation model for Web services considering timeliness

3.1 User Similarity Calculation Considering Timeliness

The traditional user-based collaborative filtering recommendation method mainly calculates the similarity based on the user-service rating matrix. However, this rating is the most original user's evaluation of the Web service, and there is a problem of data sparsity. On this basis, we need to consider personal rating habits and timeliness to more accurately calculate the user's rating of a Web service. Thereby it can calculate the similarity between users more accurately.

Let the user's set is U, $U = \{u_1, u_2, \ldots, u_m\}$, The set of all Web service is S, $S= \{s_1, s_2, \ldots, s_n\}$, $r_{i,k}$ represents rating between the user u_i and the service s_k. Since the ratings in the generally recommended data set are all from the 5-point scale, the $r_{i,k}$ is an integer from 0 to 5. If the user u_i does not rate the service s_k, $r_{i,k} = 0$. Considering that each user will have their own habits. For example, some people will habitually rating high on the services they have received. Of course, some people are used to giving low ratings to the service. In order to remove the user's personal rating habits. So we calculate the rating as follows to get the objective rating of the user $w_{i,k}$, and the value is between 0 and 1.

$$w_{i,k} = \frac{r_{i,k}}{\sqrt{\sum_{i=1}^{S} {r_{i,k}}^2}}. \tag{1}$$

The user's rating of a Web service should also be related to the time of the user's evaluation Web service. From a psychological point of view: (1) Similar to memory, people's interest gradually declines with time; (2) The speed of forgetting is gradually slowing down, and the accumulated interest will become more and more stable. The closer the evaluation time is to the current time, the more it reflects the user's preference. The farther the evaluation time is from the current time, the smaller the reference value of the rating is. Considering the time of evaluation Web service and the life cycle of user evaluation behavior, the rating timeliness function TLF is obtained.

$$w_{time}(i,k)=\begin{cases}\exp\{-\ln 2\times time(i,k)/\max\} & time(i,k)\le \max \\ 1/2 & time(i,k)\ge \max\end{cases}. \tag{2}$$

Where $w_{time}(i,k)$ is the timeliness of the user u_i rating the Web service s_k . $time(i,k)$ is a non-negative integer. L is a fresh period of the rating time. When distance between the evaluation time of Web service s_k by user u_i and the last evaluation time is between 0-L, $time(i,k)$ takes a value of 0. When distance between the evaluation time of Web service s_k by user u_i and the last evaluation time is between L-2L, $time(i,k)$ takes a value of 1. The value of L is specifically set in a specific experiment. The closer the time of the evaluation to the latest evaluation date, the smaller the value of $time(i,k)$ is. *max* is a threshold value, which is related to the life cycle of the user u_i evaluation behavior. The longer the life cycle of the user evaluation behavior, the larger the value of *max* is, the slower the user interest decline rate will be, and vice versa. When $time(i,k)\ge \max$, the value of $w_{time}(i,k)$ will remain at 1/2. Therefore, the closer the service evaluation time is, the higher the rating will be. And with the service evaluation time is longer, the time rating is lower.

According to the above calculation and comprehensively considering the individual rating habits and the timeliness of the rating, the total rating of the user u_i to the Web service s_k is obtained as follows:

$$M_{i,k}=\lambda w_{i,k}+(1-\lambda)w_{time}(i,k). \tag{3}$$

Where λ is the weight of $w_{i,k}$, and the value is between 0 and 1. User's objective rating of Web service reflects user's preference very well. The timeliness function can solve user's preference transfer problem very well. So the value of λ can be adjusted according to the importance in specific experiments.

Define a user-service rating matrix to represent the total rating of the user u_i 's preference for the Web service s_k , expressed as:

$$M(U\times S)=\begin{bmatrix} M_{1,1} & M_{1,2} & \cdots & M_{1,I} \\ M_{2,1} & M_{2,2} & \cdots & M_{2,I} \\ \cdots & \cdots & \cdots & \cdots \\ M_{U,1} & M_{U,2} & \cdots & M_{U,I}\end{bmatrix}.$$

The similarity between users is calculated according to the obtained user web service rating matrix. There are many methods for calculating the similarity of users. Pearson correlation and angle cosine are very popular. In this paper, angle cosine is used to calculate user similarity. The formula of similarity calculation is as follows:

$$sim(u_i,u_j)=\frac{\sum_{x\in S_{i,j}} M_{i,x}\times M_{j,x}}{\sqrt{\sum_{x\in S_{i,j}} M_{i,x}^2}\times\sqrt{\sum_{x\in S_{i,j}} M_{j,x}^2}}. \tag{4}$$

The higher the similarity value $sim(u_i,u_j)$ between user u_i and user u_j for Web service rates, the closer their interest will be. However, if there are fewer Web services jointly evaluated by users, the similarity between users will not be well reflected. For example, there is only one Web service that user u_i and user u_2 jointly evaluate. So even if the two users rating the same for this Web service, and the evaluation time is close. It can't be said that the two users have higher similarity. In order to exclude users in the neighbors who have less evaluation Web service with the target users, the neighbors closest to the target users must be selected among a large range of neighbors. Find the weighted similarity value $sim\_w(u_i,u_j)$ between the two users. The specific formula is as follows:

$$sim\_w(u_i,u_j)\frac{\min(num,\varepsilon)}{\varepsilon}\times sim(u_i,u_j). \tag{5}$$

Where $sim(u_i, u_j)$ represents the similarity value of the user u_i and the user u_j based on the Web service rating. $\frac{\min(num, \varepsilon)}{\varepsilon}$ represents a weighting value indicating a similarity between the user u_i and the user u_i. And *num* represents the number of the common evaluation Web services of the user u_i and the user u_j. ε is the threshold and will be reasonably valued in a specific experiment.

3.2 User Similarity Calculation Based on Timeliness Tag

To solve the problem of data sparsity, it is necessary to establish a tag set $T = \{t_1, t_2, ..., t_l\}$ according to the tag information. And the set of all Web services marked by the user u_i using the tag t_a is $S_{i,a} = \{s_{i1}, s_{i2}, ..., s_{ib}\}$. Then, based on the value of the user-service rating matrix, the rating value of the user u_i for each tag t_a is calculated. When a user tagging different Web services using the same tag, the user's rating of the tag is the average of all Web service ratings. The relationship between users, Web services, and tags is shown in Fig. 2.

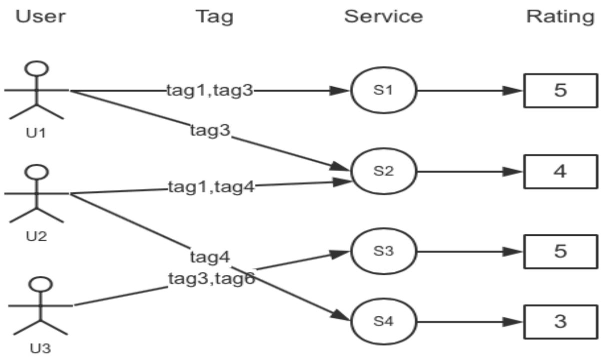

Fig. 2. User-service-tag-rating relationship

When it calculates the user's rating of the tag, it also considers two factors that affect the tag information, namely the timeliness of the tagging and the user's preference for the tag. For the time of tagging, if the target user u_i used tag t_a to tag a service long ago, then the reference value of tag t_a is less for the target user u_i. If the target user used tag t_a to tag a service in the near future, then tag t_a can better reflect the user's preferences; For the degree of preference of the tag, according to the TF-IDF idea, tags with small frequencies are not necessarily less correlated than tags with large frequencies (TF idea); Tagging a service multiple times with the same tag is not necessarily less relevant than a single tag (IDF idea). In other words, the tags frequently used by users may not reflect users' interests very well. However, if a service is marked with the same tag for many times, it indicates that the tag can better reflect the characteristics of the service. Therefore, it is necessary to calculate users' preference for the tag to illustrate the importance of the tag. Therefore, the calculation formula of the user's u_i rating $V_{i,a}$ for the tag t_a is:

$$V_{i,a} = pre(u_i, t_a) w_{time}(u_i, t_a) \frac{\sum_{k=1}^{b} r_{i,k}}{b}. \quad \textbf{(6)}$$

Where $r_{i,k}$ is the rating of the user u_i to the Web service s_k, and the user tags the service s_k using the tag t_a. b is the total number of Web services marked by the user using the tag t_a. $w_{time}(u_i, t_a)$ is a timeliness function TLF of rating the tag. The calculation method is as shown in the formula (2), wherein the value of $time(i, k)$ is based on the time of tagging, and the value is 1, 2, 3... $pre(u_i, t_a)$ is the preference of the user u_i for the tag t_a. It is necessary to consider the frequency of tag use and the importance of the tag. The specific calculation formula is as follows:

$$pre(u_i, t_a) = \begin{cases} 1 & \frac{n_{u_i,t_a}}{n_{u_i,t}} \times \log \frac{U}{n_{t_a,u}} \geq 1 \\ \frac{n_{u_i,t_a}}{n_{u_i,t}} \times \log \frac{U}{n_{t_a,u}} & otherwise \end{cases}. \quad \textbf{(7)}$$

Where n_{u_i,t_a} the number of times the user u_i uses the tag t_a. $n_{u_i,t}$ represents the number of times the user u_i uses the tag used. U represents the number of all users. $n_{t_a,u}$ represents the number of users using t_a. And $\frac{n_{u_i,t_a}}{n_{u_i,t}}$ represents the frequency at which the user u_i uses the tag t_a. $\log \frac{U}{n_{t_a,u}}$ represents the importance of the user u_i at the tag t_a. The higher the frequency of the current user's use of the tag t_a is, the more the current user prefers the tag t_a will be. The fewer users use the tag t_a is, the more the current user prefers the tag t_a will be.

Define a user-tag rating matrix to represent the user u_i's rating on the tag t_a, the matrix representation is:

$$M(U \times T) = \begin{bmatrix} V_{1,1} & V_{1,2} & \cdots & V_{1,I} \\ V_{2,1} & V_{2,2} & \cdots & V_{2,I} \\ \cdots & \cdots & \cdots & \cdots \\ V_{U,1} & V_{U,2} & \cdots & V_{U,T} \end{bmatrix}.$$

Where U is the number of users, T is the number of tags, and $T_{U,T}$ represents the rating of the tag *t* by the user *u* considering the tag labeling time and the user's preference for the tag.

Based on the established user-tag rating matrix, the similarity based on the tag rating between users is calculated. As with the calculation method based on the user-Web service rating matrix, the cosine of the angle is used to calculate the similarity between users. The formula is as follows:

$$sim\_t(u_i, u_j) = \frac{\sum_{a \in T_{i,j}} \times V_{j,a}}{\sqrt{\sum_{a \in T_{i,j}} V_{j,a}^2} \times \sqrt{\sum_{a \in T_{i,j}} V_{j,a}^2}}. \quad \textbf{(8)}$$

Where $sim\_t(u_i, u_j)$ represents the similarity between the user u_i and the user u_j based on the tag rating, $V_{i,a}$ represents the rating of the user u_i for the tag t_a, and $V_{j,a}$ represents the rating of the user u_j for the tag t_a.

3.3 User Trust Calculation Based on Social Network

Social networking is a small society that exists in virtual networks. Users in these small societies have real social activities between each other. In order to achieve one or more purposes, a community established by groups with common interests exists in the form of a network, that is, a social network. In social networks, people tend to trust friends with more familiar relationships and are willing to accept their recommendations. Therefore, the calculation of trust between users has become a major research content. The user trust relationship in the social network refers to the degree of intimacy between users in the social network. Each trust involves two users, and there may be a direct trust relationship between them, as shown in A of Fig. 3. The users may also be indirect trust relationships, as shown in B of Fig. 3. This paper combines direct trust and indirect trust between users to calculate the final trust.

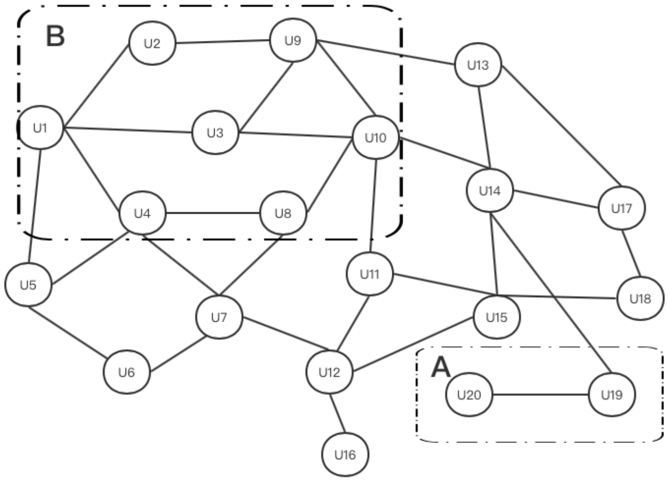

Fig. 3. User social network

Definition 1. Social relationship familiarity refers to the degree of familiarity between users according to the frequency of interaction between users. That is to say, users are more familiar with friends who often interact with each other, and usually trust friends who often interact with each other. The familiarity of social relationships is calculated as follows:

$$Fam(u_i, u_j) = \frac{I_{i,j} - I_{i\min}}{I_{i\max} - I_{i\min}}. \quad \textbf{(9)}$$

Among them, $Fam(u_i, u_j)$ indicates the familiarity of the social relationship between the user u_i and the user u_j. $I_{i,j}$ represents the number of interactions between user u_i and user u_j. $I_{i\min}$ represents the minimum number of times user u_i interacts with other users. $I_{i\max}$ represents the maximum number of times user u_i interacts with other users. Therefore, the more frequently users interact with each other, the more familiar they become.

Definition 2. The core degree of social relationship refers to the importance of a user in the social network in which they are located. Generally speaking, in a small social network, some people will be the core people in social networks and have a greater impact on other members. Therefore, the services recommended by core members are often believed by more people. The calculation of the core degree of social relationships is as follows:

$$Core(u_j) = \begin{cases} 1, & d_j = d_{\max} \\ lb(1 + \frac{d_j}{d_{\max}}), & 0 < d_j < d_{\max} \\ 0, & d_j = 0 \end{cases}. \quad \textbf{(10)}$$

Where $Core(u_j)$ represents the user u_j's core degree of the social relationship. d_j represents the number of user u_j's friends, and $d_{\max}$ represents the maximum number of friends in the social relationship. When the number of friends is more, the user is considered to have a higher influence.

Definition 3. Direct trust of social relationship refers to a trust relationship in which the recommended user and the target user are friends and can directly communicate in the social network. As shown in B of Fig. 3. In the social network, according to the social relationship familiarity $Fam(u_i, u_j)$ between the target user u_i and the recommended user u_j and the social relationship core degree $Core(u_j)$ of the recommended user, the obtained recommended user u_j is relative to the target user u_i direct trust in social relationships. The formula for direct trust calculation is as follows:

$$DT(u_i, u_j) = \alpha Fam(u_i, u_j) + \beta Core(u_j). \quad \textbf{(11)}$$

α, β respectively represent the weighting factors of familiarity and core degree in social relationships, α + β = 1.

Definition 4. Indirect trust of social relationship refers to the indirect relationship established through several common friends rather than the direct friendship between the recommended user and the target user. The trust relationship is a kind of trust relationship generated through the transmission of several direct trust relationships between common friends. For example, user u_a and user u_b are not friends and there is no direct contact. But they have a common friend u_c, through the user u_c can generate contact. The relationship model is an indirect relationship. There is a direct trust between the user u_a , u_b and the user u_c. The trust relationship between the user u_a and the user u_b through the user u_c is an indirect trust relationship. The strength of the indirect relationship between users is considered to be the indirect trust of the service recommendation, abbreviated as IDT. As shown in B of Fig. 3.

As can be seen from B in Fig. 3. There may be multiple paths between the recommended user u_{10} and the target user u_1, and the intermediate user may be one or more. Therefore, the indirect trust of the user is obtained by calculating the relationship path length between the two users, the number of relationship paths, the edge weight of the relationship path, and the direct trust degree of the intermediate user. It describes the degree of intimacy between a recommended user and one or more intermediate users and target users. And uses $IDT(u_i, u_j)$ to indicate the indirect trust between the target user u_i and the recommended user u_j :

$$IDT(u_i, u_j) = \frac{\sum_{k=1}^{n} e^{-\lambda p_k} \prod_{l=1}^{p_k} DT_l(x, y)}{\sum_{k=1}^{n} e^{-\lambda p_k}}. \tag{12}$$

Where $e^{-\lambda p_k}$ is an attenuation function whose value varies continuously between 0 and 1, and is a representation of the weight coefficient of the kth relationship path in all paths. λ represents the attenuation coefficient of the relationship path length. p_k indicates the length of the kth relationship path. It is assumed that the user u_i and the user u_j have *n* relational paths, and $p_{ij} = \{p_1, p_2, \dots p_n\}$ is used to represent the set of relationship paths between the two users in question. The weights of the relationship paths are $\{e^{-\lambda p_1}, e^{-\lambda p_2}, \dots e^{-\lambda p_n}\}$. $DT_i(x, y)$ represents the direct trust between the user u_x and u_y in the kth path. Let the set D be a set of all users connecting the recommended user u_j and the target user u_i in one path. Including users u_i and u_j , then u_i, $u_j \in D$. There may be no direct trust relationship between any two users. Therefore, adding indirect trust when calculating user trust can avoid the problem of data sparsity and increase the accuracy of recommendation.

Definition 5. Social relationship trust is a comprehensive measure of trust relationship between target users and recommended users. The direct trust DT of service recommendation and indirect trust IDT of service recommendation are synthesized, which are abbreviated as $Tru(u_i, u_j)$. The measurement method is as shown in formula (13):

$$Tru(u_i, u_j) = \mu DT(u_i, u_j) + \omega IDT(u_i, u_j). \tag{13}$$

Where μ denotes the weighting factor of direct trust and ω denotes the weighting factor of indirect trust. Both satisfy μ+ω=1, μ, ω∈[0,1]. From the perspective of social relations, in the process of user interaction, the degree of trust obtained by direct interaction is higher than that obtained by indirect interaction. Therefore, as the frequency of direct interaction between users increases, the direct trust between users will the more intense. That is, the value of μ, ω will change dynamically with factors such as the number of direct interactions between users. When μ is larger, ω is smaller, indicating that as the number of direct interactions increases, the proportion of direct trust between users will become larger and larger, and the proportion of indirect trust will become smaller and smaller. When μ =1, ω =0, the trust between users comes from the direct trust relationship, and there is no indirect trust between users. When μ =0, ω =1, it means that there is no direct connection between the two users, and the trust is all derived from the indirect trust relationship.

3.4 Choice of Trusted Neighbor Users

3.4.1 Direct Neighbor Users Based on Service Ratings and Trust

The neighbor user of the target user is selected in combination with the user similarity $sim\_w(u_i, u_j)$ and the social relationship trust degree $Tru(u_i, u_j)$ calculated above. The main choice is to select users with relatively high similarity and trust to the target users. According to the similarity and trust degree, the combined influence $affect\_w(u_i, u_j)$ of the recommended user u_j on the target user u_i is calculated. Then the neighbor user of the target user is searched. Through analysis, it is found that when the similarity value $sim\_w(u_i, u_j)$ and the recommendation trust value $Tru(u_i, u_j)$ are higher, the comprehensive influence of the recommended users should be greater. When the similarity or trust value between users is 0, the comprehensive influence value of recommended users should be defined as 0. Therefore, the formula for calculating the combined influence of user u_j on target user u_i is given as follows: $affect\_w(u_i, u_j)$ is as follows:

$$affect\_w(u_i, u_j) = \frac{2sim\_w(u_i, u_j) \times Tur(u_i, u_j)}{sim\_w(u_i, u_j) + Tur(u_i, u_j)}. \quad \mathbf{(14)}$$

Among them, $sim\_w(u_i, u_j)$ is the user similarity calculated by formula (5) considering the timeliness of the rating. $Tru(u_i, u_j)$ is the trust degree of social relationship between users calculated by formula (13). When $sim\_w(u_i, u_j) = 1$, $Tru(u_i, u_j) = 1$, the recommended user u_j has the highest influence on the target user u_i, is 1. When $sim\_w(u_i, u_j) = 0$, $Tru(u_i, u_j) = 0$, the recommended user u_j has the lowest influence on the target user u_i , is 0. When $sim\_w(u_i, u_j) \in (0,1)$, $Tru(u_i, u_j) \in (0,1)$ the higher their value, the recommended user u_j has the higher comprehensive influence on the target user u_i . By considering the trust relationship between users based on the similarity between users. it can effectively avoid the adverse effects of some malicious users on service recommendation and improve the credibility of neighbor users.

According to the calculation, the influence value of recommended users on target users is obtained. *N* users with the highest influence are selected as the direct neighbor users based on service rating and trust of target users, denoted as G1.

3.4.2 Expand Neighboring Users Based on Tag Rating and Trust.

Combined with the tag rating similarity $sim\_t(u_i, u_j)$ and social relationship trust $Tru(u_i, u_j)$ calculated above. The combined influence of recommended users on target users based on tag similarity and social trust is $affect\_t(u_i, u_j)$, which is calculated as follows:

$$affect\_t(u_i, u_j) = \frac{2sim\_t(u_i, u_j) \times Tur(u_i, u_j)}{sim\_t(u_i, u_j) + Tur(u_i, u_j)}. \quad \mathbf{(15)}$$

According to the calculated influence value, the first *N* users with high influence were selected as the target users' extended neighbor user set, denoted as G2.

3.4.3 Determination of Trusted Neighbor Users.

When recommending to the target user, the target project may not be rated by the direct neighbor user. In this case, because the data is sparse, the accuracy of the recommendation may be affected. Therefore, taking the union of the target user's direct neighbor user set and the extended neighbor user set as the final trusted neighbor user set to solve the problem of data sparseness, denoted as G, G=G1 ∪ G2.

3.5 A Trusted Hybrid Recommendation Method for Web Services Considering Timeliness

After the calculation above, a group of users considering timeliness and social trust relationships are screened out and recorded as G, which is the trusted neighbor user of the target user. According to the set of trusted neighbor user, the prediction of service rating and service recommendation behavior of target users are realized. According to the collaborative filtering theory, the predicted value $p_{i,k}$ of service rating can be calculated according to the following formula (16):

$$p_{i,k} = \overline{r_i} + \frac{\sum_{j \in G} affect(i, j)(r_{j,k} - \overline{r_j})}{\sum_{j \in G} affect(i, j)}. \quad \textbf{(16)}$$

Where $p_{i,k}$ represents the predicted evaluation of the service s_k by the target user u_i . $r_{j,k}$ is the evaluation of the service s_k by the user u_j . $\overline{r_i}$ and $\overline{r_j}$ are the average of the service ratings of all user used. For each service s_k , the predicted rating value $p_{i,k}$ of the target user u_i to service s_k is calculated according to the formula (16). The services are sorted according to the predicted rating value, wherein the first n services are used as the service recommendation set of the user u_i , and the service recommendation ends.

The specific steps of the recommended method are as follows:

Input: target user u_i , user rating information for the Web service, user tag information, user socialization relationship information

Output: target user's service recommendation set

1. Consider the user's rating habits and timeliness of the ratings, and calculate the user's comprehensive rating $M_{i,k}$ for the Web service.
2. A user-Web service rating matrix is established based on the user's overall rating for the Web service.
3. Similarity $sim\_t(u_i, u_j)$ between users based on Web service rating is calculated by formula (5) according to user-Web service rating matrix.
4. Considering the timeliness of tags and user preference for tags, the user's rating $V_{i,a}$ for tags is calculated.
5. Establish a user-tag rating matrix based on user's rating of tags.
6. According to the user-tag rating matrix, the similarity $sim\_t(u_i, u_j)$ based on the tag rating is calculated by formula (8).
7. Considering the familiarity among users and the core of recommending users, the direct trust $DT(u_i, u_j)$ between users is calculated according to formula (11).
8. Considering the relationship path length, the number of relationship paths and the edge weight of relationship paths, the formula (12) calculates the indirect trust $IDT(u_i, u_j)$ between users.
9. The user's trust degree $Tru(u_i, u_j)$ is obtained according to the formula (13) in combination with the direct trust degree and the indirect trust degree between the users.
10. Considering the similarity and trust between users, formula (14) can be used to get the first N1 users with high influence as the direct neighbor user set G1 of the target users. And formula (15) can be used to get the first N2 users with high influence as the expanded neighbor user set G2 of the target users, and the final neighbor users can be recorded as G, G=G1 ∪ G2
11. For each Web service s_k , the predicted rating value $P_{i,k}$ of service s_k for target user u_i is calculated according to formula (16).
12. The Web services are sorted according to the size of the predicted rating. The top N Web services are regarded as the Web service recommendation set of the target user u_i , and the service recommendation ends.

4 Experiment and Result Analysis

This section mainly carries out the simulation experiment of effectiveness verification. The proposed method in this paper is compared with several classic service recommendation methods. And the differences of several service recommendation methods are analyzed.

4.1 Data Source and Processing

In this paper, the trusted hybrid recommendation method is proposed to help users select the appropriate Web service. But so as to make the verification results of the proposed method to be believed, we choose the public dataset as the training set and the test set. At present, the public dataset based on social network recommendation mainly includes the following: Epinions dataset, MovieLens dataset and Hetrec2011 dataset. This paper selects Hetrec2011-Last.fm dataset under Hetrec2011 dataset. The dataset includes user information, artist information, tag information, user's listening to the artist, friend relationship, tag timestamp information, etc. The sematic mapping relationships between the concepts related to Web service recommendation and the concepts in the data set is given in the Table 1. It is used to explain the rationality of that Hetrec2011-Last.fm dataset is used to verify the proposed method.

Table 1. The sematic mapping relationships between the two types of concepts

The concepts related to Web service recommendation	The concepts in Hetrec2011-Last.fm dataset
Users	User provided by the Web service
Artists	Web service
Number of times users listen to artists	User ratings for Web services
Friendship	Relationship between users
Tags	Tags for Web services
Timestamp	Time to tag the Web service

The dataset Hetrec2011-Last.fm includes 1892 users, 17632 artists, 12727 two-way user friend relationships, ie 25434 pairs, 11946 tags and tagged timestamps. Pre-process the data and select the top 200 users with the most friends to experiment. The 200 users listen to a total of 2783 artist. In this data set, the user's rating of the artist is determined by the number of times the artist listens. Since the ratings in the general data set are all from the 5-point scale. This article also quantifies the rating to 1 to 5 points for the interaction between users. The interaction between users is determined by the number of times the same work is common commented. We divided the data set into two groups, the training set and the test set, in which 80% of the data was placed in the training set and 20% of the data was placed in the test set.

4.2 Evaluation Index

Whether the recommendation algorithm is reasonable depends on whether the prediction of the rating is accurate. Mean Absolute Error (MAE) is a commonly used method for comparing and measuring accuracy. It can avoid the problem that the errors cancel each other accurately to reflect the actual prediction error, so as to accurately reflect the quality of the recommendation. Root Mean Square Error (RMSE) can be used to measure the deviation between the predicted value and the actual value to better measure the accuracy of the predicted data. MAE and RMSE are commonly used as experimental comparison criteria. The evaluation criteria used in literature [17] and [18] in this paper are also MAE and RMSE. Therefore, MAE and RMSE are used for verification in this paper. The calculation formulas for MAE and RMSE are given below:

$$MAE = \frac{\sum_{i=1}^{n} | r_{i,k} - p_{i,k} |}{n}. \tag{17}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(r_{i,k} - p_{i,k})^2}{n}}. \quad \textbf{(18)}$$

Where $r_{i,k}$ represents the actual rating of the target user u_i for the artist s_k, $p_{i,k}$ represents the predicted rating of the target user u_i for the artist s_k, and n represents the number of samples. The smaller the values of MAE and RMSE, the higher the accuracy of the recommendation.

4.3 Experimental Results and Analysis

This paper selects the collaborative filtering service recommendation based on a novel similarity calculation method (RACF) [17] and the model-based social network recommendation method (SocialMF) [18] as the representative, and compares it with the proposed method.

4.3.1 Effectiveness of Recommendation Method

Experiment 1: Based on the calculation of the similarity of the traditional user service rating, consider the habit of user rating and the timeliness of rating. The top k users with the highest similarity are obtained as the nearest neighbor users for recommendation. The effectiveness of the recommendation is verified by experiments. The experimental results are shown in the Fig. 4 and Fig. 5.

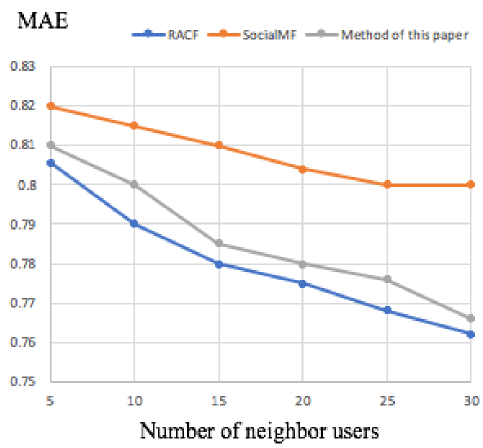

Fig. 4. MAE comparison of different recommendation methods

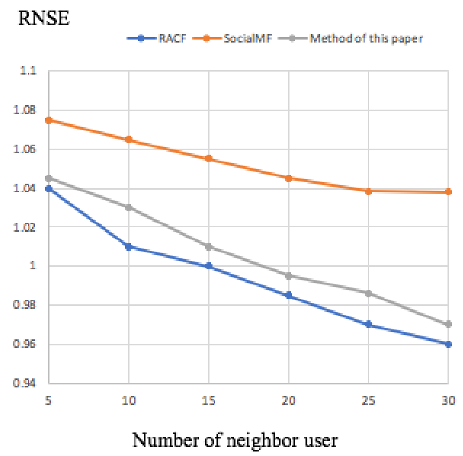

Fig. 5. RMAE comparison of different recommendation methods

It can be seen from the comparison of the MAE and RMSE of the Fig. 4 and Fig. 5. The values of the recommended MAE and RMSE are decreasing as the value of the neighbor user k increases. The more neighbors are, the more accurate the recommendation is. The SocialMF method considers trust delivery and uses matrix decomposition to perform service recommendation. But it does not consider the similarity between users, so MAE and RMSE are relatively high. The RACF method proposes a new method of similarity calculation, which uses a ratio-based method to calculate similarity. So the recommended MAE and RMSE are relatively low. The method of this paper considers user rating habits and rating time factors when calculating similarity, so the values of MAE and RMSE are lower than other methods.

4.3.2 Fraud Resistance of Recommended Methods

Experiment 2: Based on the similarity of service ratings among users, this paper adds the trust degree in social relationships and considers two aspects to make recommendations. In the experiment, the proportion of malicious referral users was continuously increased to verify the anti-fraud ability of several recommended methods. The experimental results are shown in Fig. 6 and Fig. 7.

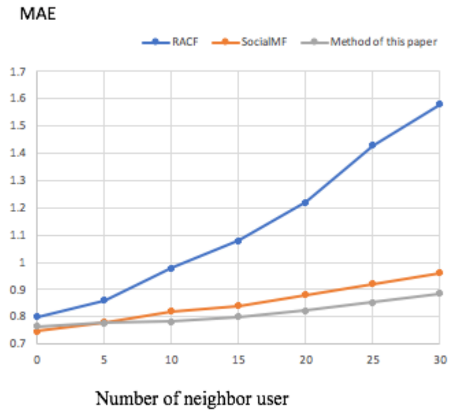

Fig. 6. MAE comparison of different recommendation methods under malicious users

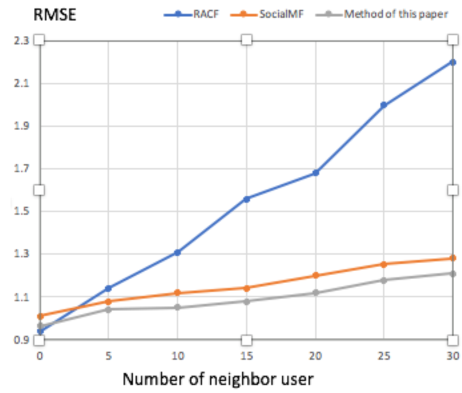

Fig. 7. RMAE comparison of different recommendation methods under malicious users

As can be seen from the Fig. 6 and Fig. 7, with the increasing proportion of malicious referral users, the values of MAE and RMSE of several recommended methods have increased to some extent. Among them, the RACF method mainly performs rating prediction and recommendation based on the similarity of users. Therefore, when the number of malicious recommendation users increases, the MAE and RMSE of the method increases significantly. That is, the anti-fraud ability is poor. The SocialMF and the proposed method are less interfered and have less growth. Because both methods consider the trust relationship between users. The SocialMF method only considers the trust transfer between users by matrix decomposition. The method of this paper also considers the similarity between users. At the same time, when calculating the trust degree, it adds the interactivity between users, the core degree of recommending users and the indirect trust between users, so it can better resist fraud.

4.3.3 Alleviating the Sparsity of Recommendation Data

Experiment 3: In this paper, based on the direct neighbor users, the extended neighbor users that integrate the label information are added to the neighbor users to increase the number of neighbor users. Several methods have been used to verify the problem of mitigating data sparseness. The experimental results are shown in the Fig. 8 and Fig. 9.

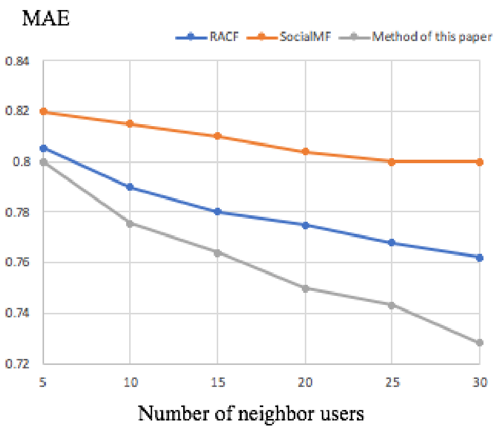

Fig. 8. MAE comparison of different recommended methods for data sparseness

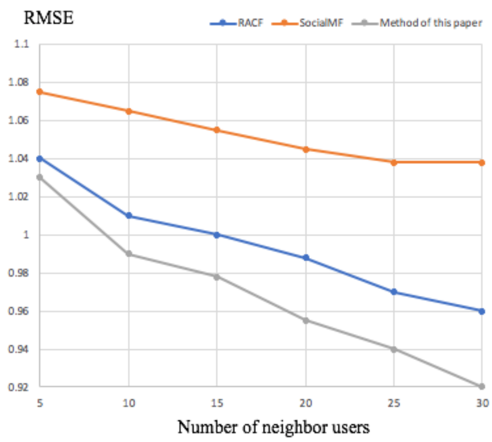

Fig. 9. RMAE comparison of different recommended methods for data sparseness

As can be seen from the Fig. 8 and Fig. 9, as the neighbor user k grows, the values of MAE and RMSE are reduced to different degrees. The RACF method and the SocialMF method are relatively slow to reduce. The method of this paper is more obvious. Because the method in this paper considers that the neighbor user may not score the target item, the number of valid recommendation users among the neighbor users is relatively small. there is a problem that the data is sparse. In this paper, the extended neighbor user based on the tag information is used to expand the final neighbor user of the target user, which better solves the problem of data sparseness.

5 Conclusion

Aiming at the problem of recommending appropriate Web services for constructing high-quality Mashup applications, the trusted hybrid recommendation method of Web services is proposed in this paper. Firstly, the timeliness function TLF of Web service ratings is used to solve the problem of user preference transfer. Secondly, the calculation of the user trust based on social relations is used to resist malicious recommendation, which is considering the interaction frequency between users and the core degree of users in social network as well as the relationship path between users. Finally, the calculation of the user similarity based on tag timeliness is used to extend the neighbor users, and then alleviate the issues of data sparseness. Compared with other recommendation methods, our method has better recommendation accuracy and resistance to malicious recommendation. In the follow work, we can further improve the accuracy of recommendation by adding location information to recommendation.

Acknowledgements

Foundation items: Project supported by the National Natural Science Foundation, China (No. 61601146), the Natural Science Foundation of Heilongjiang Province, China (No. QC2018081), and the Training Program for Yong Innovated Talent of Universities in Heilongjiang Province, China (No. UNPYSCT-2016032).

References

[1] B. Xia, Y. Fan, W. Tan, K. Huang, J. Zhang, C. Wu, Category-aware API clustering and distributed recommendation for automatic mashup creation, IEEE Transactions on Services Computing 8(5)(2015) 674-687.

[2] X. Xu, Z. Wang, Z. Feng, Big service: applications and influence in the big data environment, Communications of the CCF 13(2)(2017) 24-30.

[3] L. Liu, F. Lecue, N. Mehandjiev, Semantic content-based recommendation of software services using context, ACM Transactions on the Web 7(3)(2013) 17-20.

[4] H.-Y. Wang, W.-B. Yang, S.-C. Wang, S.-R. Li, A service recommendation method based on trustworthy community, Chinese Journal of Computers 37(2)(2014) 301-311.

[5] L. Yao, Q. Sheng, H. Ngu, Y. Yu, A. Yu, Unified collaborative and content-based web service recommendation, IEEE Transactions on Services Computing 8(3)(2015) 453-466.

[6] F. Ricci, L. Rokach, B. Shapira, P.B. Kantor, Recommender Systems Handbook, Springer, 2010.

[7] R. Jaschke, L.B. Marinho, A. Hotho, L. Schmidt-Thieme, G. Stumme, Tag recommendations in folksonomies, in: Proc. 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, 2007.

[8] H.-N. Kim, A.-T. Ji, I. Ha, G.-S. Jo, Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation, Electronic Commerce Research and Applications 9(1)(2010) 73-83.

[9] X. Ding, F. Peng, Q. Wang, J. Lu, Tensor factorization recommendation algorithm combined with social network and tag information, Journal of Computer Applications 35(7)(2015) 1979-1983.

[10] Z. Zhang, X. Zheng, D.-D. Zeng, A framework for diversifying recommendation lists by user interest expansion, Knowledge-Based Systems 105(2016) 83-95.

[11] J.-L. Zhang, M.-X. Huang, Y. Zhang, Collaborative filtering recommendation algorithm based on tag optimization, Application Research of Computers 35(10)(2018) 2916-2919.

[12] Z. Wang, L.-F. Sun, W.-W. Zhu, S.-Q. Yang, H.-Z. Li, D.-P. Wu, Joint social and content recommendation for user-generated videos in online social network, IEEE Transactions on Multimedia 5(3)(2013) 698-710.

[13] Y. Yu, G.-H. Qiu, Research on collaborative filtering recommendation algorithm by fusing social network, New Technology of Library and Information Service 28(6)(2012) 54-59.

[14] R. Forsati, M. Mahdavi, M. Shamsfard, M. Sarwat, Matrix factorization with explicit trust and distrust side information for improved social recommendation, ACM Transactions on Information Systems 32(4)(2014) 17-24.

[15] X. Lin, M. Zhang, Y. Zhang, Y. Liu, S. Ma, Learning and transferring social and item visibilities for personalized recommendation, in: Proc. 2017 ACM on Conference on Information and Knowledge Management, 2017.

[16] Y.-C. Li, L. Chen, C.-C. Shi, X.-Y. Lan, Enhanced collaborative filtering algorithm adopting trust network, Application Research of Computers 35(1)(2018) 16-24.

[17] X. Wu, B. Cheng, J. Chen, Collaborative filtering service recommendation based on a novel similarity computation method, IEEE Transactions on Services Computing 10(3)(2017) 352-365.

[18] M. Jamali, M. Ester, A matrix factorization technique with trust propagation for recommendation in social networks, in: Proc. Fourth ACM Conference on Recommender Systems, 2010.