Research Article

Cultural Distance-Aware Service Recommendation Approach in Mobile Edge Computing

Yan Li¹ and Yan Guo²

¹School of Business and Management, Shanghai International Studies University, Shanghai, China
²State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Correspondence should be addressed to Yan Guo; guoyan@bupt.edu.cn

Received 13 October 2017; Accepted 26 December 2017; Published 14 February 2018

Academic Editor: Youngjae Kim

Copyright © 2018 Yan Li and Yan Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the era of big data, traditional computing systems and paradigms are not efficient and even difficult to use. For high performance big data processing, mobile edge computing is emerging as a complement framework of cloud computing. In this new computing architecture, services are provided within a close proximity of mobile users by servers at the edge of network. Traditional collaborative filtering recommendation approach only focuses on the similarity extracted from the rating data, which may lead to an inaccuracy expression of user preference. In this paper, we propose a cultural distance-aware service recommendation approach which focuses on not only the similarity but also the local characteristics and preference of users. Our approach employs the cultural distance to express the user preference and combines it with similarity to predict the user ratings and recommend the services with higher rating. In addition, considering the extreme sparsity of the rating data, missing rating prediction based on collaboration filtering is introduced in our approach. The experimental results based on real-world datasets show that our approach outperforms the traditional recommendation approaches in terms of the reliability of recommendation.

1. Introduction

With the rapid development of Internet of Things and 5G network, a large number of distributed data and computation tasks are generated, which makes traditional centralized computing paradigm suffer large processing pressure. Mobile edge computing is emerging to process the data at the edge of network for high performance big data processing [1].

In mobile edge computing, services are provided within a close proximity of mobile users by servers at the edge of network [2]. With the increasing number of edge services, there are too much choices for users to meet their requirements [3, 4]; therefore, the service recommendation technology is needed to help people find the optimal edge services from a huge mass of services. Because of the ubiquitous services, service recommendation is playing an increasingly significant role in our diary life [5, 6]; for example, Amazon has deployed its recommending system to help recommending books and other products to its users [7].

Although there is a significant research in mobile edge computing, there has been little attention paid to services recommendation in mobile edge computing. Traditional service recommendation system is to dig up service users’ preferences based on user’s history records and scores and then recommend similar users’ services which have never been used by the service user [8, 9]. It can be reduced to a problem of estimating ratings for the services that have not been used by a user. Intuitively, this estimation is usually based on the ratings given by the user to other services. Once the ratings for the yet unrated services can be estimated, we can recommend to the service user the service with the highest estimated rating.

Much research over the past decade has focused on developing service recommendation systems, and they are usually classified based on recommendation approach as content-based recommendation [10], collaborative filtering-based recommendation [11], and model-based collaborative filtering recommendation [12].

Collaborative filtering is one of the most successful recommendation technologies in many recommended systems [13]. The basic idea is to predict the ratings of the services that have not been used by the service user according to the users
who have the common experience and the same interest or
the services which are similar to the services that have been
rated by the service user [14]. However, the efficiency and
accuracy of collaborative filtering are not high in practical
applications [15].

Traditional collaborative filtering recommendation ap-
proach only focuses on the similarity extracted from the
rating data in the calculation of similarity, the selection
of the nearest neighbor, and recommendation. Specifically,
the recommendation is according to the weighted average
of evaluation from the nearest neighbor, and the weight is
the similarity between the user and the nearest neighbors.
However, the similarity is not the only decision factor;
there are many other factors having an important role. The
preference of users only depends on the similarity, which may
lead to an inaccuracy expression of user preference and low
reliability of recommendation results.

In addition, traditional service recommendation ap-
proaches perform poorly when facing sparse user rating
data. There are some approaches that have been proposed
to solve the sparsity of ratings data, and it can be divided
into three categories. The first is attribution and dimension
reduction; however, attribution may lead to the missing
of user’s personal characteristics data, and dimension reduction
may delete useful data. The second effective way to solve
the sparse data problem is to introduce the trust into the
recommendation system [16, 17]. The third way is to predict
the missing ratings by adopting the approach in the similarity
calculation [18]. This approach can exploit the effective
data while also enlarging the influence of the erroneous
data.

Considering that there is a close relationship between the
user preference and their cultural background, we introduce
cultural distance which can express the user preference to
reduce the decision-making power of similarity. For example,
in mobile edge computing, users in different edge clouds may
have different preference, and users in the same edge clouds
may have the same preference.

Responding to the above problems, we propose a new
collaborative filtering recommendation approach based on
cultural distance in which we combine the cultural dis-
tance and similarity to represent the user preference. In the
approach, we combine the user-based collaborative filtering
and the service-based collaborative filtering approach to take
full advantage of the information of the similar users and
similar services, and then we recommend service to users
based on a cultural distance-aware collaborative filtering rec-
ommendation approach. In addition, we predict the missing
ratings based on the cultural distance-aware collaborative
filtering approach to improve the sparsity of user rating
data.

The remainder of the paper is organized as follows:
Section 2 introduces the background of service recommenda-
tion; Section 3 describes our cultural distance-aware hybrid
collaborative filtering recommendation approach; Section 4
presents the experiment results and Section 5 concludes this
paper.

2. Background

2.1. Related Concept. The main idea of our approach is the
introduction of cultural distance; therefore, we introduce a
related concept: cultural distance before the introduction of
our approach.

Cultural distance is an important concept to study cul-
tural differences, which is widely used in cross-culture study
and cross-culture practice, such as enterprise international-
ization [19]. Culture directly impacts on people’s beliefs,
thoughts, and social behaviors. While some people may not
know this impact, culture and values have been rooted in
people’s minds and are the decisive factor in people’s social
behavior and their preferences. All computations of cultural
distance are based on cultural dimensions.

The most famous cultural dimension research is Hof-
steded’s cultural dimension theory, which first realizes the
quantitative description of the abstract complicated concept
of culture [20]. This cultural dimension theory enables
researchers to compute cultures as data in a more intuitive
way to compare differences between different cultures and
various behaviors.

The Hofstede’s cultural dimension theory consists of six
dimensions. The six dimensions are as follows: Power Dis-
tance Index, Individualism versus Collectivism, Masculinity
versus Femininity, Uncertainty Avoidance Index, Long Term
Orientation versus Short Term Normative Orientation, and
Indulgence versus Restraint. Based on Hofstede’s cultural
dimension theory, [21] first uses a simple mathematical
formula to define the cultural distance and describes the
cultural differences more simply. The formula is as follows:

$$CD_j = \frac{\sum_{i=1}^{m} \left( (I_{ij} - I_{ik})^2 / V_i \right)}{m},$$

where $CD_j$ is the cultural distance for the $j$th country, $I_{ij}$ is
Hofstede’s $i$th cultural dimension score of the $j$th country, the
$k$th country is the host country, $V_i$ is the variance of the $i$th
cultural dimension scores, and $m$ is the number of cultural
dimensions.

2.2. Related Work. A large number of service recommenda-
tion approaches have been proposed in the literature, and we
only review some notable approaches here.

Among the numerous service recommendation ap-
proaches, collaborative filtering is one of the most successful
recommendation technologies. And it includes user-based
collaborative filtering [14, 22, 23], service-based collaborative
filtering [24–28], and its fusion [18]. User-based collaborative
filtering is to predict the rating of users based on the ratings of
their similar users, and service-based collaborative filtering is
to predict the rating of users based on the ratings of services
which are similar to the services chosen by the users. The
basic steps of collaborative filtering are rating collection, simi-
liarity computation, neighbor selection, rating prediction, and
recommendation, and the similarity computation is always
based on Person Correlation Coefficient [29]. Besides, [18]
proposes a fusion approach which combines user-based and
service-based collaborative filtering approaches. Different
from [18], our approach considers the cultural factor in rating prediction, and the similarity is not the only weigh factor of neighbors.

However, there are some problems in collaborative filtering approaches, such as the sparse rating data, cold start, malicious attack, and bad system scalability. To solve these problems, many trust-based recommendation approaches are proposed [16, 17, 30, 31]. The paper [30] proposes a TPCF model which joins trust propagation into the collaborative recommendation system and combines the trust in trust propagation model and the similarity in traditional collaborative filtering algorithm by a mixing index to obtain the final user similarity. The paper [17] put forward a random walk algorithm which combines collaborative recommendation based on the item and the recommendation algorithm based on trust to predict the rating of a single item, and the principle is that the ratings of similar projects by highly trusted neighbors are more reliable than the ratings of the target project by lowly trusted neighbors. Besides, some recommendation approach is based on missing rating prediction [18] to improve the sparsity of rating data.

3. Our Approach

Aiming at the more reliable recommendation, we propose a cultural distance-aware service recommendation approach based on missing ratings prediction. Actually, the missing ratings prediction of training data is the same as the ratings prediction of active user in service collaborative filtering recommendation; therefore, we use the same cultural distance-aware collaborative filtering approach to solve them. The first step of our approach is the missing ratings prediction of training data which contains the similarity computation based on initial ratings, neighbor selection, and missing ratings prediction based on cultural distance; the second step is service recommendation which contains the similarity computation based on prediction rating data, neighbor selection, and ratings prediction based on cultural distance and recommendation.

3.1. Ratings Prediction. The foundation of the collaborative filtering recommendation approach is the user-service ratings matrix; however, these matrices are always sparse in practical applications. In our approach, we predict the missing rating values before recommendation to obtain a denser user-service ratings matrix. The main idea of prediction is similar to the collaborative filtering, and it can be divided into three phases: similarity computation, neighbor selection, and ratings prediction.

3.1.1. Similarity Computation. Assume that the original user-service ratings matrix \( R \) which expresses users preferences contains \( n \) users and \( m \) services; the element in this two-dimensional matrix \( r_{ui} \) is a vector of QoS values which is obtained by the user \( u \) on the service \( i \). In general, a user only invokes a small part of the services and rates less services; as a result, many elements \( r_{ui} \) in \( R \) are missing.

To predict the missing ratings, we adopt two collaborative filtering prediction approaches: the user-based prediction approach and the service-based prediction approach. The former is to first match the user’s other ratings against other users, compute the similarity between them, and find users with the most similar preferences. The latter is to first match the service’s other ratings against other services, compute the similarity between them, and find the most similar services. No matter in user-based or service-based prediction approaches, the similarity computation is based on Pearson correlation coefficient.

The Pearson correlation coefficient is a measure of the linear correlation between two variables. In user-based prediction approach, the similarities between user \( u \) and other users can be represented as follows:

\[
\text{Sim}_u = \{ \text{sim}_{ua} | a \in C, a \neq u \},
\]

where \( C \) represents the set of users, \( \text{sim}_{ua} \) represents the similarity between user \( u \) and user \( a \), and it is computed as follows:

\[
\text{sim}_{ua} = \frac{\sum_{i \in I} (r_{ai} - \bar{r}_a)(r_{ui} - \bar{r}_u)}{\sqrt{\sum_{i \in I} (r_{ai} - \bar{r}_a)^2} \sqrt{\sum_{i \in I} (r_{ui} - \bar{r}_u)^2}},
\]

where \( I \) represents the subset of services which are rated by both user \( a \) and user \( u \), \( \bar{r}_a \) represents the average of ratings obtained by the user \( a \) on services in \( I \), and \( \bar{r}_u \) represents the average of ratings obtained by the user \( u \) on the services in \( I \). The value of \( \text{sim}_{ua} \) is between \(-1 \) and \( 1 \), where \(-1 \) represents total negative linear correlation, \( 0 \) represents no correlation, and \( 1 \) represents total positive linear correlation. The larger value of \( \text{sim}_{ua} \) indicates the higher similarity between user \( a \) and user \( u \). Note that when \( I \) is an empty set, \( \text{sim}_{ua} \) is \(-1 \).

In service-based prediction approach, the similarities between service \( i \) and other services can be represented as follows:

\[
\text{Sim}_i = \{ \text{sim}_{ij} | j \in S, j \neq i \},
\]

where \( S \) represents the set of services, \( \text{sim}_{ij} \) represents the similarity between service \( i \) and service \( j \), and it is computed as follows:

\[
\text{sim}_{ij} = \frac{\sum_{u \in U} (r_{ui} - \bar{r}_i)(r_{uj} - \bar{r}_j)}{\sqrt{\sum_{u \in U} (r_{ui} - \bar{r}_i)^2} \sqrt{\sum_{u \in U} (r_{uj} - \bar{r}_j)^2}},
\]

where \( U \) is the subset of users who rate both service \( i \) and service \( j \), \( \bar{r}_i \) represents the average of ratings of service \( i \) obtained by the users in \( U \), \( \bar{r}_j \) represents the average of ratings of service \( j \) obtained by the users in \( U \). Similar to \( \text{sim}_{ua} \), the value of \( \text{sim}_{ij} \) is between \(-1 \) and \( 1 \), and the larger value of \( \text{sim}_{ij} \) indicates the higher similarity between service \( i \) and service \( j \). Note that when \( U \) is an empty set, \( \text{sim}_{ij} \) is \(-1 \).

3.1.2. Neighbor Selection. The main idea of collaborative filtering is to harness the collective intelligence and learn from the opinions of correlative populations; therefore, it is important to select the neighbors which are similar users or similar services. Before ratings prediction, neighbor selection
according to similarities is needed to find users with the most similar preferences or the most similar services.

There are two common ways to identify neighbors in traditional collaborative filtering approaches: Top-K approach and threshold-based approach. Top-K approach is to select the most similar K users or services as neighbors, no matter how similar the neighbors are. As shown in Figure 1(a), to select the 5 neighbors of Point 1, we take the recent five points: Points 2, 3, 4, 7, and 5 according to the distance between the points. However, we can see clearly that this approach has a bad performance on the calculation of isolated points, because the number of neighbors is fixed, and we have to take some points which are not very similar as neighbors when there are not enough similar points nearby. For example, in some points which are not very similar as neighbors when because the number of neighbors is fixed, and we have to take some points which are not very similar as neighbors when there are not enough similar points nearby. For example, in Figure 1(a), Point 5 is selected as a neighbor of Point 1, but it is not very similar to Point 1. Different from the principle of Top-K approach, threshold-based approach is to select the users or services of which similarity is smaller than threshold as the neighbors. As shown in Figure 1(b) where ts represents the value of threshold, to select the neighbors of Point 1 based on threshold, we consider the circular region with center Point 1, and a radius of ts, and take the points within the circular region as neighbors. In this example, Points 2, 3, 4, and 7 are selected as neighbors. This approach can improve the performance on the calculation of isolated points; however, this approach is very sensitive to the value of threshold; the neighbors selected are few when ts is too small, and when ts is too large. Therefore, how to set the value of threshold is very difficult.

To ensure the similarity of neighbors, we propose a novel neighbor selection approach which combines the Top-K approach and the threshold-based approach. We first set a loose threshold to filter the users or services with low similarity and then select the remaining users or services which are the top K most similar users or services as neighbors.

In user-based prediction approach, the selected neighbors of user $u$ can be described as follows:

$$N_u = \{ a \mid \text{sim}_{au} < ts, \ r_{au} < K \},$$

(6)

where $N_u$ represents the neighbors of user $u$ and $r_{au}$ represents the ranking of sim$_{au}$ among Sim$_u$. The formulation means that user $a$ is a neighbor of user $u$ if user $a$ belongs to the top $K$ similar users of user $u$, and sim$_{au}$ is less than the threshold.

In service-based prediction approach, the selected neighbors can be described as follows:

$$N_i = \{ j \mid \text{sim}_{ij} < ts, \ r_{ij} < K \},$$

(7)

where $N_i$ represents the neighbors of service $i$, and $r_{ij}$ represents the ranking of sim$_{ij}$ among Sim. The formulation means that service $j$ is a neighbor of service $i$ if service $j$ belongs to the top $K$ similar users of service $i$, and sim$_{ij}$ is less than the threshold.

3.1.3. Ratings Prediction. In traditional collaborative filtering approaches, the similarity between users or services is the only factor to determine the accuracy of prediction results. However, there are other factors having an important role in prediction. Therefore, we adopt the cultural distance in our collaborative filtering approach to improve the accuracy of prediction.

In this paper, we design a cultural factor to represent the preference of users to their neighbors. The cultural factor between user $u$ and its neighbor $a$ is as follows:

$$c_{au} = \left( 1 - \frac{|c_{d_{u}} - c_{d_{a}}|}{\max_{cd_{u}} - \min_{cd_{u}}} \right),$$

(8)

where $c_{au}$ represents the cultural factor of user $u$ to user $a$, $c_{d_{u}}$ represents the cultural distance value of user $u$, $c_{d_{a}}$ represents the cultural distance value of user $a$, $\max_{cd_{u}}$ represents the maximum value among the cultural distance values of neighbors of user $u$, and $\min_{cd_{u}}$ represents the minimum value among the cultural distance values of neighbors of user $u$.

Similarly, the cultural factor between service $i$ and its neighbor $j$ is as follows:

$$c_{ij} = \left( 1 - \frac{|c_{d_{i}} - c_{d_{j}}|}{\max_{cd_{i}} - \min_{cd_{i}}} \right),$$

(9)

where $c_{ij}$ represents the cultural factor of service $i$ to service $j$, $c_{d_{i}}$ represents the cultural distance value of service $i$, $c_{d_{j}}$ represents the cultural distance value of service $j$, $\max_{cd_{i}}$
represents the maximum value among the cultural distance values of neighbors of service \( i \), \( \min_{\text{cd}} \) represents the minimum value among the cultural distance values of neighbors of service \( i \).

In our missing ratings prediction approach, we design a hybrid weight which can consider not only the similarity of neighbors but also the cultural factor. In our user-based prediction approach, the compound weight between user \( u \) and its neighbor \( a \) is

\[
\text{\( w_{u,a} = \beta \text{sim}_{u,a} + (1 - \beta) c_{u,a}, \)}
\]

where \( \beta \) represents the degree of the emphasis on the similarity. And then the prediction value of missing ratings can be computed based on the ratings of neighbors as follows:

\[
\text{\( p_u (r_{u,i}) = \bar{r}_u + \frac{\sum_{a \in N_u} w_{u,a} (r_{a,i} - \bar{r}_a)}{\sum_{a \in N_u} w_{u,a}}, \)}
\]

where \( p_u (r_{u,i}) \) represents the prediction value of the missing rating \( r_{u,i} \) based on the neighbor user, \( \bar{r}_u \) represents the basic prediction value of \( r_{a,i} \), and it is the average of ratings obtained by the user \( u \); similarly, \( \bar{r}_a \) is the average of ratings obtained by the user \( a \) on the services besides service \( i \).

Similarly, in our service-based prediction approach, the compound weight between service \( i \) and its neighbor \( j \) is

\[
\text{\( w_{i,j} = \beta \text{sim}_{i,j} + (1 - \beta) c_{i,j}, \)}
\]

And then the prediction value of missing ratings can be computed based on the ratings of similar services as follows:

\[
\text{\( p_s (r_{u,i}) = \bar{r}_j + \frac{\sum_{j \in N_i} w_{i,j} (r_{u,j} - \bar{r}_j)}{\sum_{j \in N_i} w_{i,j}}, \)}
\]

where \( p_s (r_{u,i}) \) represents the prediction value of the missing rating \( r_{u,i} \) based on the similar neighbor services, \( \bar{r}_i \) represents the average of ratings of service \( i \), and \( \bar{r}_j \) represents the average of ratings of service \( j \) besides the rating obtained by user \( u \).

In order to take full advantage of the information of the similar users and similar services and ensure the diversity of the recommendation results, our prediction approach combines the two collaborative filtering prediction approaches. Therefore, the prediction value of the missing rating \( r_{a,i} \) in our prediction approach is as follows:

\[
\text{\( p (r_{a,i}) = \lambda p_u (r_{a,i}) + (1 - \lambda) p_s (r_{u,i}), \)}
\]

where \( \lambda \) represents the weight of the user-based prediction approach.

3.2. Service Recommendation. After the missing ratings prediction, we can recommend services for the active users who require service recommendations based on the prediction rating data of the training users whose ratings have been obtained. Similar to our missing ratings prediction approach, our service recommendation approach can be divided into three phases: similarity computation, neighbor selection, ratings prediction, and recommendation.

The first phase is similar to the last section. The similarities between the active user and the training users and the similarities between the candidate services and other services are computed by Formula (3) and Formula (5). The second phase is to select neighbors for the active users and the candidate services by Formula (6) and Formula (7). The third phase is to predict the ratings of the candidate services from the active users by Formula (14), and the final phase is to recommend service according to the prediction ratings.

4. Experiment

In this section, we implement our cultural distance-based collaborative filtering recommendation approach (named CDCF) to verify the performance and comparatively evaluate CDCF against several service recommendation approaches in terms of the prediction failure rate under the training rating data with different density and the active rating data with different density. The experimental evaluation results show that CDCF can improve the accuracy of recommendation.

4.1. Experiment Setup. To evaluate the performance of CDCF, we implement it by adopting two real-world datasets: WS-Dream dataset [32] and Hofstede’s cultural dimension dataset [33].

The WS-Dream dataset contains the real-world QoS evaluation results from 339 users on 5,825 Web services and the location information (e.g., IP address, country, region, latitude, and longitude) of these users and services. Specifically, the QoS evaluation results are the user-service matrices on response time and throughout which can be used as the rating matrix in our approach.

Hofstede’s cultural dimension dataset has been published in [33] and it contains the scores for the six cultural dimensions. These scores are obtained according to the items in the IBM database plus extensions and the items in the World Values Survey.

Based on the two datasets, we compare CDCF with other recommendation approaches.

(i) User-Based Collaborative Filtering Recommendation Approach (UCF). This approach is to predict the missing ratings using the traditional user-based prediction approach without adopting the cultural distance firstly and then recommend services for active users by employing the traditional user-based collaborative filtering approach based on the prediction ratings data.

(ii) Service-Based Collaborative Filtering Recommendation Approach (SCF). This approach is to predict the missing ratings using the prediction approach without adopting the cultural distance firstly and then recommend services for active users by employing the traditional service-based collaborative filtering approach based on the prediction ratings data.

(iii) Traditional Hybrid Collaborative Filtering Recommendation Approach (TCF). This approach is to predict the missing ratings by combining the traditional user-based
prediction approach and traditional service-based prediction approach and then recommend services for active users by combining the traditional user-based approach and service-based prediction approach. Compared with CDCF, both of the prediction and recommendation approaches only depend on the similarity of neighbors.

We select 150 users from the WS-Dream dataset randomly and then divide them into training users and the active users; the rate of training users is 0.75 in our experiments. To evaluate the performance of CDCF in data with different density, we randomly remove some ratings in training rating matrix or active rating matrix and employ different approaches to verify the reliability of CDCF.

The evaluation index used in this paper is Root Mean Square Error (RMSE) to measure the recommendation reliability of CDCF in comparison with other collaborative filtering approaches. RMSE is defined as follows:

\[
\text{RMSE} = \sqrt{\frac{\sum (r_{u,i} - \tilde{r}_{u,i})^2}{N}},
\]

where \(r_{u,i}\) represents the removed rating of active user \(u\), \(\tilde{r}_{u,i}\) represents the prediction value, and \(N\) represents the number of the removed ratings.

4.2. Performance Comparison. Figure 2 shows the comparison results in terms of RMSE with different density of training data ranging from 0.1 to 0.3, while the density of active data is 0.3. As shown in Figure 2, our approach CDCF is much better than other approaches, and its RMSE is 22.2%, 7.5%, and 13.4% lower than SCF, UCF, and TCF on average. In addition, there is no obvious relationship between the recommendation reliability and the density of training data in SCF, UCF, TCF, and CDCF.

In a word, as shown in Figures 2 and 3, although TCF takes full advantage of the information of the similar users and similar services and ensures the diversity of the recommendation results, its accuracy is lower than SCF with the low accuracy of UCF. Our approach CDCF can ensure not only the diversity of the recommendation results, but also the high recommendation reliability.

5. Conclusion

In this paper, we propose a cultural distance-aware service recommendation approach based on missing rating prediction. The main idea is to adopt the cultural distance to represent the user preference and combine the cultural factor with similarity to improve the reliability of recommendation. In addition, we predict the missing ratings by cultural distance-aware collaborative filtering approach to improve the sparsity of rating data. Experimental results show that compared to previous approaches, our approach can significantly improve the accuracy of the recommendation results. However, the cultural factor is at the regional/national level in this paper. A narrower grouping is required to improve the reliability of recommendation. Consequently, the next study in this area could focus on the presentation of the user cultural factor in a narrower group.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The work presented in this study was supported by the National Natural Science Foundation of China (Grant no.
References


