Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8056195, 12 pages
https://doi.org/10.1155/2018/8056195

Research Article

CEPTM: A Cross-Edge Model for Diverse Personalization Service and Topic Migration in MEC

Hongchen Wu,[1] **Huaxiang Zhang**,[1] **Lizhen Cui**,[2,3] **and Xinjun Wang**[3]

[1] *School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China*
[2] *Shandong Provincial Key Laboratory of Software Engineering, Shandong University, Jinan, China*
[3] *School of Computer Science and Technology, Shandong University, Jinan 250101, China*

Correspondence should be addressed to Hongchen Wu; wuhongchen@sdnu.edu.cn

Received 17 April 2018; Accepted 2 August 2018; Published 12 August 2018

Academic Editor: Shangguang Wang

For several reasons, the cloud computing paradigm, e.g., mobile edge computing (MEC), is suffering from the problem of privacy issues. MEC servers provide personalization services to mobile users for better QoE qualities, but the ongoing migrated data from the source edge server to the destination edge server cause users to have privacy concerns and unwillingness of self-disclosure, which further leads to a sparsity problem. As a result, personalization services ignore valuable user profiles across edges where users have accounts in and tend to predict users' potential purchases with insufficient sources, thereby limiting further improvement of QoE through personalization of the contents. This paper proposes a novel model, called CEPTM, which (1) collects mobile user data across multiple MEC edge servers, (2) improves the users' experience in personalization services by loading collected diverse data, and (3) lowers their privacy concern with the improved personalization. This model also reveals that famous topics in one edge server can migrate into several other edge servers with users' favorite content tags and that the diverse types of items could increase the possibility of users accepting the personalization service. In the experiment section, we use exploratory factor analysis to mathematically evaluate the correlations among those factors that influence users' information disclosure in the MEC network, and the results indicate that CEPTM (1) achieves a high rate of personalization acceptance due to the availability of more data as input and highly diverse personalization as output and (2) gains the users' trust because it collects user data while respecting individual privacy concerns and providing better personalization. It outperforms a traditional personalization service that runs on a single-edge server. This paper provides new insights into MEC diverse personalization services and privacy problems, and researchers and personalization providers can apply this model to merge popular users' like trends throughout the MEC edge servers and generate better data management strategies.

1. Introduction

With the proliferation of the Internet of things and the burgeoning of the cloud computing paradigm network, mobile edge computing (MEC) has been widely adopted at a tremendous speed. A large collection of computing resources, such as mobile devices, application servers, and storage units, are utilized to serve users in all types of tenant models. The presentation and wide application of MEC have changed people's daily work and are utilized in the data centers of computing companies. A recent report has demonstrated that extensible infrastructure and processing engine technology have greatly impacted the methods of running applications and websites across multiple edge servers [1]. A large number of data centers are built by venders to serve a very large number of users, and these large-scale, commodity-computer data centers have sufficient resources to provide personalization to a large number of users. As the de facto centralized big data platform, the cloud computing paradigm supports QoE with a fruitful number of benefits—convenience, pay for each use, and ubiquity—which has given birth to a worldwide range of industry companies [2]. The widely applied denominator in MEC is the deployment of cloud computing-like capabilities at the edge of the IoT network. Figure 1 demonstrates the

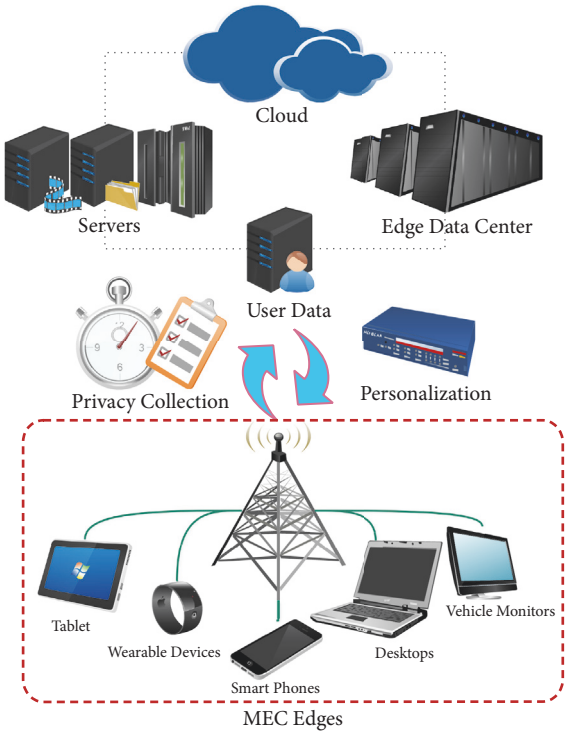

Figure 1: The structure of mobile edge computing in the cloud environment.

structure of the personalization service provided by MEC. Edge data centers, which are owned and deployed by infrastructure providers, implement a multitenant virtualization infrastructure. Edge devices automatically cooperate with one another for service collaboration in the cloud environment, and personalization activity in edge servers collects users' data and enhances end users' experiences by matching the items and their potential buys. These services have forged into various support mechanisms and management platforms, creating an open ecosystem where a multitude of customers can be served.

Despite the achievements, MEC is still far from a panacea. The massive growth of MEC edge servers [3], including websites and applications, has led to an explosive volume of data and has caused the problem of "information overload" [4]. An unprecedented number of information resources are available, but users are often unsure of what to do [4, 5]. Personalization services can apply information-filtering methods to eliminate the items that users would not click on according to their browsing histories and preferences [6, 7]. However, a prerequisite for a personalization service is that it must collect significant amounts of personal data for precise recommendations [8]. This approach introduces three main issues into the research field. First, although a tremendous volume of data is available on the Internet, the valuable data that are useful for analysis constitutes only a small percentage, which is called "data sparsity" [9]. Second, the collection of large amounts of data is in conflict with today's users' high level of privacy concern [10]. In addition, users often have difficulty in deciding on information disclosure and do not possess sufficient knowledge to evaluate the risks and benefits [11]. As a result, users usually deny most information requests. This action leads to a "worse-to-worse" loop, as users do not trust the personalization service and do not share much private information as input for the information-filtering algorithm. The personalization service cannot generate good personalization using such a small amount of data, and the users will, in turn, trust the personalization service even less. Third, users' standards for accepting personalization are not low, and the users may no longer accept personalization in one type of item on an MEC edge server. Instead, users fall prey to various types of "innovative" personalization [12]. This user-tailored personalization approach is difficult to measure, and the trigger that leads a user to accept personalization is unknown. These problems have limited the further development of personalization services in MEC networks, but the available data from users' multiple accounts on various MEC edge servers have provided a solution from a new perspective. The related papers had made some progress on personalization but lack consideration on privacy issues, which indicate the motivation of our work.

This paper establishes a cross-edge model, called CEPTM, which integrates valuable resources for analyzing users' "like" behaviors from multiple MEC edge servers for diverse personalization and for merging real-time popular buys from a source edge server into all destination edge servers. This model is equipped with a personalization engine for producing user-tailored suggestions of what items to view. CEPTM is applied in the experiment section and achieves higher prediction accuracy than a baseline personalization service that is running on a single-edge server. This finding is attributed to it determining the factors that influence users' decisions on accepting personalization. Our study also focuses on reconciling the tension between user acceptance of a personalization and various other factors, e.g., personalization diversity and personalization quality, by using exploratory factor analysis (EFA), which is a common modeling method for explaining the correlations among the variables in terms of fundamental factors. From both the user and managerial perspectives, the results of the model on real-life datasets from the MEC edge servers enabled us to answer the following question: what should personalization providers do to cause mobile users to be more likely to accept the personalization and more willing to disclose their privacy? Our suggestions are as follows:

(1) Collect as many user profiles as possible. Taking this action gives the personalization provider access to more available information for generating high-accuracy personalization, which could further optimize user QoE and increase user satisfaction with the personalization service.

(2) The scope of personalization should be diverse and, especially, not restricted to the types of products with which users are very familiar, as this could generate repeated items. Repeated items are those that users have already liked, and showing them will cause users to lose interest in the personalization service and MEC edge server. Compared with a traditional personalization service method, our model uses common tags from various types of items by similarity computing.

(3) Respect users' privacy concerns. An effective user-tailored personalization approach should not only produce a personalization that the user will like according to his or her preferences but also collect the user's privacy in a volume that will not raise his or her privacy concerns. It is noted that a user's tolerance of information disclosure differs from person to person. Once users feel that their privacy has been compromised by data collection, their trust in the personalization service will decrease. Thus, they will decrease their amount of disclosure. However, we observed that most MEC managers still seek an agreement with a data collection policy from whoever signs into their edge servers.

2. Related Work

This study builds on existing research on personalization services and privacy in MEC networks and addresses several theoretical and practical gaps in prior studies. Here, we briefly introduce some related work and describe the manner in which the problems of personalization service and privacy were solved. Gaps were identified in these studies, which we subsequently discuss in this work.

2.1. Personalization Service versus Privacy Concern. MEC aims to offer computing resources and information technology services at the network edge, where mobile users face a bewildering array of options in an MEC edge server, such as which movie to watch [13], which news to read [14], and which people to add as friends [15]. Personalization services have been applied to help users find the items that best fit their interests by learning the past item-user relationships and performing information-filtering methods. Common methods of personalization services are collaborative filtering methods [16], content-based methods [17], and hybrid methods [18]. Collaborative filtering methods plot the user-item relationship as a metric and use similarity computing, such as Pearson correlations, to filter the items that have large gaps with users' preferences in the distance function. Content-based methods integrate the information that best describes the items, such as manually tagged annotations from experts and other users, and rank the items according to how well they match the users' preferences. A hybrid method is a mixture of the previous two categories.

However, there remain unsolved problems in this body of work. All of the methods necessitate the availability of users' rudimentary data in such a way that the personalization can know the users' needs at a basic level. One general rule is that the quality of the personalization improves with the available volume of personal data [19], but users usually do not agree to this method of data collection [20], which could lead to unwanted exposure of information. Conversely, users could disclose their long-term privacy for temporary benefits, such as filling in questionnaires in exchange for coupons and discounts [21]. This issue is called the "personalization-privacy paradox" [22], and it has attracted the interest of several researchers. Some researchers have made progress by extending this paradox, and they believe that the amount of privacy that is disclosed to personalization services depends on how much users trust the system. Meyffret argued that user anonymity allows users to feel less concerned about their privacy [23]. Letting users decide what privacy information to share with the personalization service is an option [24], but users often rely on default settings when considering their privacy controls [25]. Most users lack sufficient knowledge with regard to personalization services and privacy and are not certain about privacy control [26]. As a result, recent studies have proposed the idea of "usable privacy" [27], which states that users shall be assisted with decision support for information disclosure. This topic aims at kindly informing users that a proper method of data collection would not harm their personal information. Personalization providers should make use of personal data and convince users that good data collection generates more benefits than potential risks.

2.2. Single-Edge versus Cross-Edge. Personalization service techniques aim at predicting missing ratings for an item based on previous ratings, which are collected from user-item connections. However, the sparsity problem in single-edge personalization services has become a major bottleneck for prediction algorithms. There are various reasons for this problem. For example, users' privacy concerns may cause anxiety, thereby limiting the amount of shared information; each user may only view a small number of items, which yields limited useful data for similarity computing in personalization services; and data for new users are very limited, which leaves the personalization service with a "cold start" problem [28]. To alleviate data sparsity, one solution is to integrate the data and knowledge using cross-edge rather than single-edge solutions. The joint method has several advantages. First, there is no need to learn new users' behaviors and rules from the very beginning. Although globalization could lead to communication and coordination delays, empirical studies have suggested that the representation of user-item relationships can be efficiently accomplished [29]. Second, cloud computing make it feasible to provide cross-edge personalization services, thereby increasing the number of sites that allow users to sign into multiple accounts. A joint approach that combines user data from various sites will become a major trend in personalization services and MEC. Providing heterogeneous cross-edge personalization could shed light on brand new and promising models, which could tremendously increase the interaction between users and items using both single-edge and cross-edge methods. Users' interests in items could also be discovered in a wider scope, thereby leading to diverse personalization from which users can choose [30]. As a result, the likelihood that users would accept the personalization would increase. Finally, users' privacy concerns could be alleviated. Data collection conflicts with users' high level of privacy protection, so they may share very little information with each edge server. Cross-edge methods could integrate users' privacy to better predict their potential buys [31]. By establishing relatedness across multiple rating matrices, Li constructed a user-item joint mixture model and the corresponding ratings and showed that their proposed algorithm for effective cross-domain collaborative filtering personalization service can outperform individual models that are trained on a single domain [32]. Pan observed that, in different domains, the

user feedback could be heterogeneous, such as ratings versus clicks. To determine a solution using this feedback, both user and item knowledge from auxiliary data sources are integrated through a principled matrix-based transfer learning framework, and the discovered principle can be used to coordinate both users and items in the auxiliary data matrices and transfer them to the target domain to reduce the effect of data sparsity [33]. The sparsity problem can be addressed by jointly considering multiple heterogeneous link prediction tasks, such as the links between users and different types of items, including books, movies, and songs, and the learned intertask similarity was applied to correct the biases and skewness of the distribution from the single domain.

In summary, the current shortcomings of single-edge methods and single-domain algorithms have limited the further development of personalization services in MEC due to limited data integration, and new problems are encountered when incorporating a cross-edge personalization service. The following question was posed: Is it possible to migrate useful knowledge by establishing bridges between the resource edge server and the destination edge servers, to optimize the performance? This question still remains unsolved in recent personalization service studies. Before setting up a personalization service from multiple MEC edge servers, we must determine what mobile users truly want from personalization, what factors influence the likelihood that they will accept personalization from a cross-edge personalization service, and what we should do to make them more willing to disclose their privacy in the MEC network. This paper is aimed at these three targets and designs a cross-edge model prototype, which is focused on personalization service and data management for users from a cross-edge MEC environment. This model, CEPTM, provides an informing function of what users' friends are doing on various MEC edge servers and is equipped with a personalization engine to produce user-tailored suggestions of what items to view. By integrating the resources of user-item relationships from multiple MEC edge servers, we also discover the factors that affect users' willingness to accept personalization with exploratory factor analysis.

3. Methods and Materials

3.1. Model for Cross-Edge MEC. In this paper, we propose a prototype of the model, called CEPTM, which consists of the interactions between users and diverse items from multiple MEC edge servers, personalization for new items, and privacy collection strategies. CEPTM has two basic functions: (a) informing users about what their friends are doing in all MEC servers and (b) generating personalization for items by calculating the possibility that one user may like the personalization, while focusing on examining the factors that affect users' decisions on accepting the personalization. The prototype of the model is depicted in Figure 2, where Johnson has friends across different MEC devices; e.g., Mike is his friend on photo and music sites, and Nancy is his friend on a music and movie sites. Friends share favorite items of diverse types but usually have common tags, such as fiction and entertainment, which make personalization of various types of items feasible.

3.1.1. Basic Data Structure. The data structures of users and items that are loaded to the model are as follows.

User (Id, Profile, Friends $[\boldsymbol{id_1}, \boldsymbol{id_2}, \ldots, \boldsymbol{id_x}]$**, ViewHistory, Servers** $[\boldsymbol{s_1}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_y}]$**)**: a user may have multiple names on different MEC servers to which (s)he has registered. However, the model should use only one **Id** to identify a user. This data structure also includes private user information, e.g., age, gender, phone number, postage code, and affiliation, which can be collected from all of the servers under the users' permissions and should be stored in the user **Profile**. A user can add other users as his/her friends in the list of **Friends**, while **ViewHistory** records users' browsing histories for all types of items in the following format:

Item (Id, Category, Tags $[\boldsymbol{t_1}, \boldsymbol{t_2}, \ldots, \boldsymbol{t_i}]$**, Weight, Servers** $[\boldsymbol{s_1}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_y}]$**)**: an item must be named with an **Id** that is not equal to any existing items, while **Category** records the groups to which this item belongs, e.g., books, flowers, and stories. When an item is uploaded by a user to a **Server** for the first time, the user can add descriptions about this item with several **Tags**, e.g., romantic, comedy, and tragedy. Users can rate an item with a five-scale range, and the average rate can be calculated as **Weight**.

3.1.2. Personalization Engine. CEPTM is equipped with a personalization algorithm that calculates the recommended items in three steps. First, the algorithm determines the friends and computes their similarities. Specifically, the similarities among friends are given by the common "ratings", as defined by Pearson's correlation coefficient [34–36], which could indicate the linear correlation between two variables:

$$S_{friends} = \frac{Weight_{x,y}}{\sqrt{Weight_x^2 \times Weight_y^2}} \tag{1}$$

where $Weight_{x,y}$ is the total **Weight** of the rated items that users x and y have in common, while $Weight_x$ is the total weight of the items that are rated by user x. The **Weight** of each item is initialized to 1. Second, personalization is generated by assigning **Weight**s to items that are rated by friends:

$$\textbf{Weight} = \sum_{friends\ rate\ items} S_{friends} \tag{2}$$

where the **Weight** of a recommended item is the sum of the rates given by friends. All of the recommended items are presented to users in decreasing order of **Weight**. Finally, items that are rated by the users are removed, and an item is recommended in the first priority that has common **Tag**(s) with the removed items.

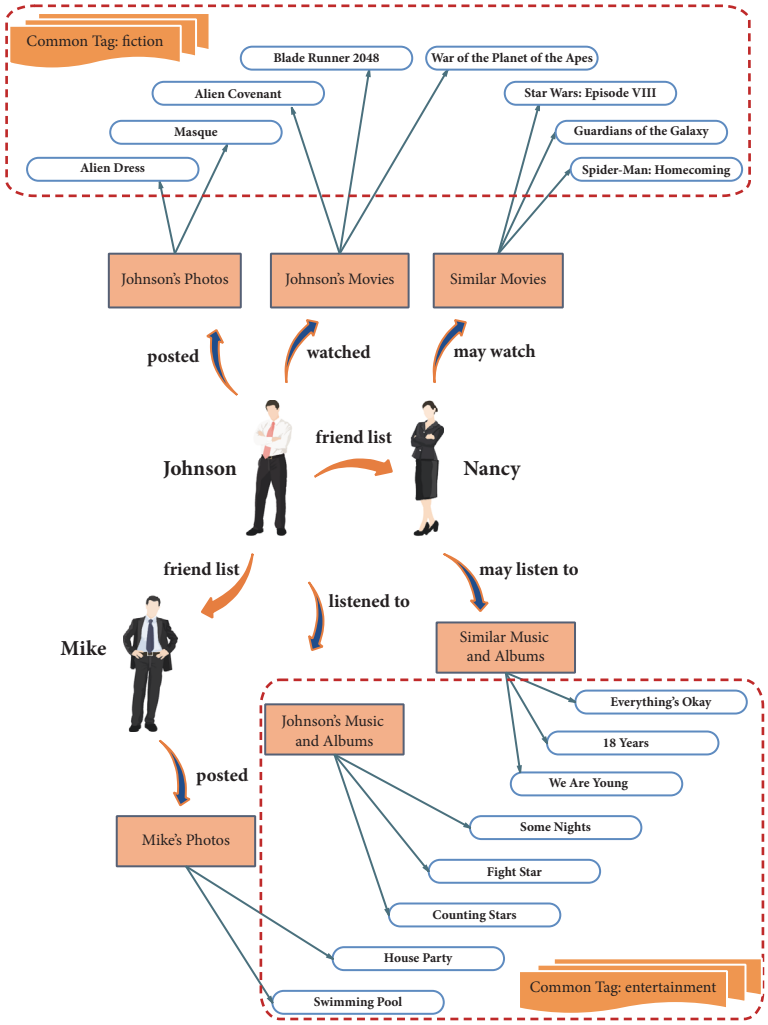

Figure 2: Prototype of the cross-edge model CEPTM.

3.1.3. Informing Function. The informing function can inform users about their friends' activities on all MEC edge servers, such as new posts and relevant personalization. The following is an example of a message that could be provided by CEPTM:

"Dear **Mike** (**User Id**$_{Mike}$), your friend **Nancy** (**User Id**$_{Nancy}$) has posted 'Surf all the way to the North with my paddle boat!' on **Sina Weibo** at 7:21 p.m. on January. 21, 2018."

The message contains the user's name and **Id**, his friend's name and **Id**, the post status and relevant item **paddle boat**, and the name of the **Sina Weibo**. The hyperlinks are also provided, as users may click them for message details, e.g., **Mike** (**User Id**$_{Mike}$) will direct the user "Mike" to his homepage, "Surf all the way to the North with my paddle boat!" will direct the user to the real poster on his friend Nancy's homepage, and **Weibo** will direct the user to Weibo. As a result, all of the messages that are presented by the model will contain the following information: the users' **Ids**, friends' **Ids**, friends' posts, the item **Id**, and the **Server**.

A nearby stranger could be suggested as an addable friend to the users. A user must disclose relevant private information to CEPTM to receive messages from this stranger and may add him/her as a friend later. For example, a user could provide his current location to see a message from a nearby stranger, or the model could display an upcoming motor show that will be held at the university at which a user studies if the user discloses the campus on which he is studying. A concert tour event could be displayed by CEPTM if the user is willing to reveal which style of music he likes the most. Furthermore, some information will be grayed out in the message if the person is not a friend of the user. This approach is a privacy protection strategy, but it also encourages users to add more friends. The differences between the formats of the messages from a friend and a stranger can be identified from the following:

> Friend (User name & Id, **Friend name & Id**, poster & personalization)
>
> Stranger (User name & Id, **Stranger Id**, poster & personalization)

where the stranger's hyperlink will direct the clicker to a page that asks whether the user would like to add the stranger

as a friend with the poster's information as the reference. The stranger and the user may have no overlapping MEC server, so the model may suggest that the user register first by clicking **Server**. Before the stranger accepts the user as a friend, the details of the stranger's homepage and the comments and likes will not be available to the user. All of the descriptions of recommended items, persons (users, friends, and strangers), and posters are available when clicking a corresponding link to the homepage. The user can adjust his threshold for eliminating strangers who have low similarity with him and avoiding recommended items with low scores.

3.2. Exploratory Factor Analysis. CEPTM can not only provide personalization and informing services but also analyze the latent variables that could affect a user's willingness to accept personalization. Here, we introduce the exploratory factor analysis (EFA) that will be applied in the experiment section. When the latent variables that lead to an observed item (e.g., users could accept personalization) are unknown, one solution for determining the variables is to apply an EFA to primitively identify the number of factors and their associated correlations. An EFA is used to determine the minimum volume of factors that explain the correlations among the latent variables. The dimensions are reduced when possible. The factors include common factors, unique factors, and errors. There is no correlation among the latent variables when the common factors are extracted. Factor analysis mainly focuses on the common factors. As a result, the error is usually combined with a unique factor.

$$B_{mn} = f_{n1} \times w_{n1} + f_{n2} \times w_{n2} + \cdots + f_{nx} \times w_{nx} + d_n \times u_{mn} \tag{3}$$

where B_{mn} is the standard deviation of sample m on observed item n, f_{nx} is a common factor with corresponding **Weight** w_{nx}, and d_n is the unique factor with corresponding u_{mn}. An EFA has to cover as many aspects as possible to find the factors that affect the observed item. Our method was to invite EMC mobile users to participate in an online survey, which required each person to consider "on what condition I will accept the personalization" for as much personalization as possible. The natural language processing technique was applied to the dataset of their answers to determine which content was mentioned. The code was written in Python and selected nouns as keywords in such a way that we could identify the primary concerns when a personalization was accepted. Keywords were ranked based on the number of times they were mentioned.

$$|Factor| = |keyword_1| + |keyword_2| + \cdots + |keyword_x| \tag{4}$$

For example, the respondents mentioned "satisfaction" 47 times and "interesting" 83 times, and the technique categorized both of these keywords under "user experience". Then, the score for "user experience" was calculated as 47 + 83 = 130. The participants also mentioned "security" 52 times, "fraud" 73 times, and "illegal usage" 20 times. Therefore, these three keywords were categorized under "data collection aspect", and its score was calculated as 52 + 73 + 20 = 145. If "data collection aspect" was not identified as a factor in EFA, then "satisfaction" would not be a factor either. The load factor, namely, Cronbach's Alpha, should be more than .70 for a good fit. The top-ranking (e.g., i_1-i_6 and j_1-j_4) keywords would be regarded as the indicators of the factor; see step 1 in Figure 2. Kaiser's Rule is applied in EFA to determine the number of factors, where the eigenvalues of each factor were visualized according to its **Weight** contribution to the observed item. Kaiser's rule is to drop all of the necessary components with eigenvalues that are less than 1.0 or, in other words, to consider information that is accounted for by an average single item while recommending the use of a sole cut-off criterion for estimating the number of factors as overextracted factors. The calculation is stopped when the confidence interval for each eigenvalue is greater than 1.0. Once the number of factors has been confirmed, GEOMIN, which is an oblique rotation method in EFA, is used to determine the factor structure, which should look similar to the model formation in step 2 and step 3 of Figure 3. For example, in step 2, i_3 is not defined as a key indicator of factor f_1, j_2 is not defined as a key indicator of factor f_2, *and* i_6 is not defined as a key indicator of factor f_3. Once all factors have found their determining indicators, we should know that the effects from f_3 on f_2 and f_2 on f_1 are all positive, shown in step 3.

The final structure consists of the latent variables, including the top-ranking factors and indicators, combined with the analysis methods, e.g., maximum likelihood parameter estimation and GEOMIN rotation. The output should consist of Cronbach's Alpha of each factor with the indicators and the model fitting information, e.g., values of log likelihood, Chi-Squared Test with the p value, RMSEA (Root Mean Square Error of Approximation), and GEOMIN rotated loadings and correlations.

4. Experiments

4.1. Prestudies before Setting Up CEPTM. This survey addresses the users' behaviors when managing their accounts. We regard users who have accounts in more than three MEC edge servers equal to users who are involved in more than three servers with a joint account. The total number of users who completed our survey was 860, and 89.77% of them (N = 772) indicated that they have multiple accounts in at least three MEC servers. We further asked these users for the reasons that they have so many accounts, and their answers cover five aspects: to keep in touch with friends (N = 353, 41.05%), e.g., "My best friend Nancy had registered a new account in another app, and she posts more updates than in the site to which we both registered"; to meet new people (N = 246, 28.60%), e.g., "Tim Cook mostly releases his posts on Weibo" and "I want to know my classmates better from Weibo"; to be well known (N = 394, 45.81%), e.g., "I want to find a good internship opportunities"; to try something new (N = 95, 11.05%), e.g., "The interface of the old website is so hard to use" and "So many ads in my old account"; to gain access to available resources

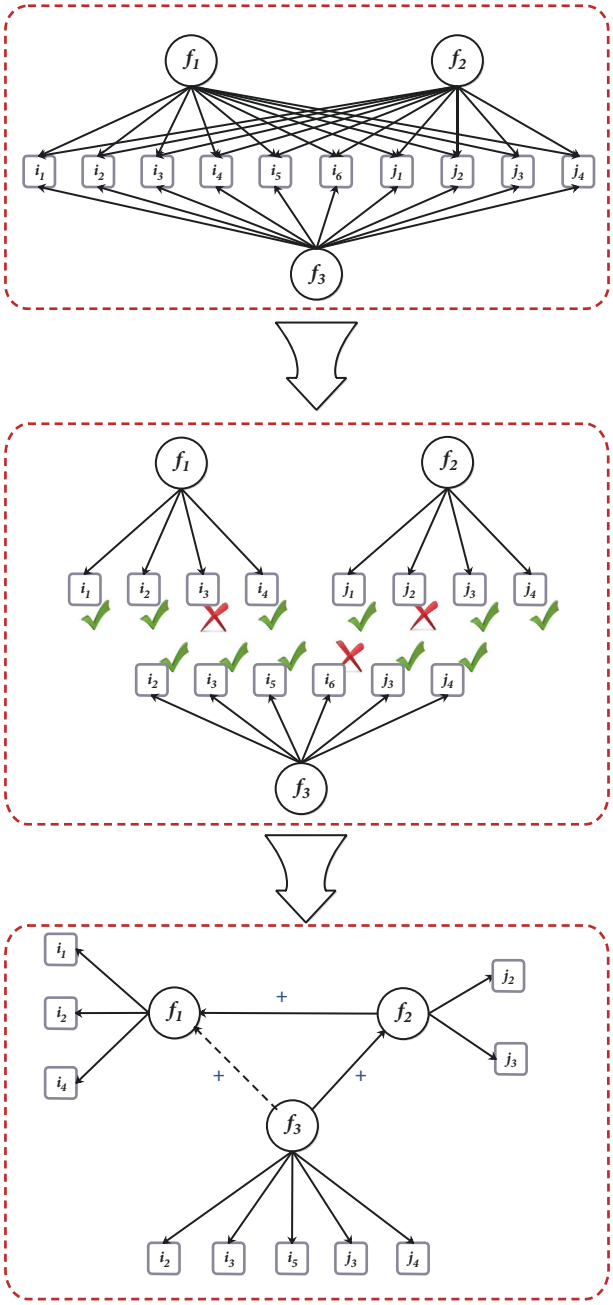

Figure 3: Structure formation and factor determination using EFA.

(N = 503, 58.49%), e.g., "Weibo provides a tremendous amount of knowledge and news for reference, and I can simply sign in with my Tencent account"; and for other reasons (N = 85, 9.88%), e.g., "For fun." From these aspects, the general reason for having multiple MEC server accounts was summarized as "being connected as much as possible". In the circumstances of cloud computing, users require several types of information for supporting daily decisions in work, study, etc. For example, to fix a car, users can browse an autocare site for suggestions. Visiting a new campus requires users to visit its college website beforehand. However, the users also gave some feedback on the flaws of having so many accounts (N = 612, 71.16%). They mainly stated that they did not have enough time to browse so many sites, e.g., "The site that I often visit is different from the one where my friends visit and post mostly", "I will not post my updates repeatedly in all of my accounts", and "I post my updates on one site only, and I sometimes miss the updates that my friend post on the other site." We conclude that users have difficulty in managing multiple accounts. Designing the CEPTM that could distribute and post users' information from all of their MEC server accounts in a cross-edge manner is truly necessary.

From the managerial perspective, a personalization service would increase its prediction accuracy if it could collect enough user information, especially when the collected information has a common **Tag**. This rule is also applicable for multiple edge servers. Compared with collecting user data from a single-edge server, collecting user data from cross-edge would increase the likelihood of locating popular topics through MEC networks, thereby guaranteeing higher personalization accuracy. For example, suppose that a user had viewed cake on a picture site, listened to romantic and peaceful songs on a music site, viewed roses and lilies on a flower app, watched romance and fiction films on a movie website, added her family members and bridesmaids as friends on Weibo, and read love and friendship tales on a book app. All of the viewed items share the common **Tag** "romantic". The personalization engine could generate an ad for wedding services accordingly, and this user would be more likely to accept this suggestion. In contrast, if the personalization engine only collected this user's actions on the flower and book sites, where the common **Tags** are "friendship" and "romantic", the prediction could miss the appropriate ad personalization. Another advantage of integrating the users' data from cross-edge is collecting more information while keeping the users' privacy concerns low. Users are likely to disclose their information only on a specific MEC site, e.g., a patient may only post his hypertension status on a medical website, rather than sharing this information with his friends on Weibo, or a person may only share her monthly income with Wells Fargo bank for a loan, while choosing not to disclose this information elsewhere. Table 1 shows that users' volumes of information disclosure in response to daily requests depend on how much they trust the information recipient. Users mostly disclose the sensitive privacy to their trusted people as expected, e.g., family members. However, the interesting thing is that users mostly disclose low sensitivity information to the people they work alongside or see only occasionally, e.g., friends and companies, regardless of whether the information was collected from a phone, a computer, or a car. A good model collects the users' data across MEC edge servers to maximize the available information for generating precise personalization. All of the items were collected under the users' permission. Another 137 MEC mobile users were invited online to register for accounts, including WeChat (Photos, http://www.wechat.com/), Kuwo Music (Music and Albums, http://www.kuwo.cn/), Youku (Movies, http://www.youku.com/), and Weibo (People, http://www.weibo.com/). Each user must view the items on each site for at least 15 minutes. The goal was to predict which

Table 1: Disclosure rates of users privacy with different recipient.

ITEMS\OPTIONS	Family	Friend	Stranger	None
Sensitive Information on portable device	153(52.76%)	115(39.66%)	13(4.48%)	106(36.55%)
Low-Sensitive Information on portable device	121(41.72%)	133(45.86%)	35(12.07%)	102(35.17%)
Low-Sensitive Information on computer	77(26.55%)	88(30.34%)	15(5.17%)	156(53.79%)
Sensitive Information on computer	88(30.34%)	62(21.38%)	18(6.21%)	162(55.86%)
Sensitive Information on connected vehicles	165(56.9%)	85(29.31%)	14(4.83%)	88(30.34%)
Low Sensitive Information on portable device	149(51.38%)	160(55.17%)	51(17.59%)	66(22.76%)

ad generated by the personalization engine users would like to view. Once a user had clicked the recommended ad, the prediction accuracy is calculated as the number of clicks multiplied by the total number of personalization. The statistics in Table 1 also show that the more servers from which user profiles were loaded as training data, the higher the probability that users accepted the generated ad personalization. This survey revealed that a model that could help them manage multiple MEC accounts is truly needed. The prestudy used approved datasets from many MEC edge servers. Thus, it could guarantee higher prediction accuracy by learning users' preferences through their browsing histories. Establishing a cross-edge model could provide mutual benefits from both the users' perspective and the managerial perspective.

4.2. Participants and Procedure in Main Experiment. We recruited 860 people from Sojump, which is a website that provides online survey services that connect more than 2 million members throughout China and enables individuals and businesses to coordinate the use of human intelligence to perform tasks that computers are currently unable to complete. For quality purposes, only adult online users who had signed into at least three of the four sites, e.g., WeChat, Kuwo, Youku, and Weibo, and spent more than 2 hours per day on these websites were allowed to participate in the main experiment. Since there is little difference among which tools users applied in browsing sites in Table 1, users can utilize any MEC device through the procedure. At least 1/3 of their friends must have multiple accounts on these websites. Of the MEC mobile users, 561 were male, 329 were between the ages of 18 and 26, 413 were between the ages of 26 and 45, and 118 were older than 45. Additionally, 278 users were working in or studying computer science or used their mobile device and computer very often, e.g., programmers and network engineers; 539 participants had received a B.S. degree or less education, 253 had received an M.A. degree, and 68 had received a Ph.D. degree or above. Each participant was required to have at least 10 common ratings with his friends, and otherwise it was suggested that he perform more ratings and attempt to join the study again to avoid the cold start problem.

The experiment was conducted for 4 days, and the performance of CEPTM demonstrated that hot topics that were raised on one site could spread to the other 3 sites over time in the MEC network. The keywords that had been clicked the most were considered to be hot top-10 topics. As shown in Figure 4, topics were segmented into 4 styles, e.g., romance, comedy, adventure, and thriller. Each style came from only one site, e.g., romance in WeChat in Figure 4(a), comedy in Kuwo Music in Figure 4(b), adventure in Youku in Figure 4(c), and thriller in Weibo in Figure 4(d). Hot topics in the other 3 sections in each site were removed in the first 24 hours. Users can update their status and forward friends' statuses within the site or across the sites. During the second 24 hours, the number of clicks on the hot topics in all 4 sites started rising. On the third day, hot topics started to migrate to the other 3 sites. During the last 24 hours, the number of clicks remained high in each section of each site and stayed equally high as the clicks on the same topics in the other 3 sites. A baseline personalization service (BLPS) was also applied in each site, which only allows forwarding friends' statuses within the site.

During the 4 days, we also compared the number of clicks between CEPTM running on the four sites (red squares and lines) and BLPS running on each site (black squares and lines). The BLPS utilizes users' profiles from a single site and recommends items by similarity computing, and items from friends who share the most common interests in hot topics are provided. The results indicate that the clicks increased much faster in CEPTM than the clicks in each BLPS, and the average number of clicks on personalization provided by CEPTM was 87.17% higher than 58.09% provided by BLPS. Therefore, CEPTM outperforms BLPS.

4.3. Results and Discussion. We examine the possible reasons behind the experimental results. First, all qualified users were asked about their personal traits, including item expertise: how much they are familiar with the item category, trust propensity: how much they trust the BLPS or CEPTM, and familiarity with online privacy and personalization: how well they understand the privacy collection. Second, they were briefly introduced to the functions of CEPTM and informed of how the personalization service works. Afterward, we administered a quiz to ensure that they understood our research target and had a basic understanding of what we were trying to accomplish. Finally, we requested 30 privacy items that belonged to one type in Table 1 from the users, and they could choose to agree to or deny the disclosures. The personalization service produced an item by similarity computing, and when a user finished the 30th request, we asked them 8 questions regarding their experience. These questions involved the users' understanding of CEPTM, their satisfaction with CEPTM, the quality of the recommended items, privacy concerns, etc. The responses to the 30 requests and eight questions on personal traits, privacy collection

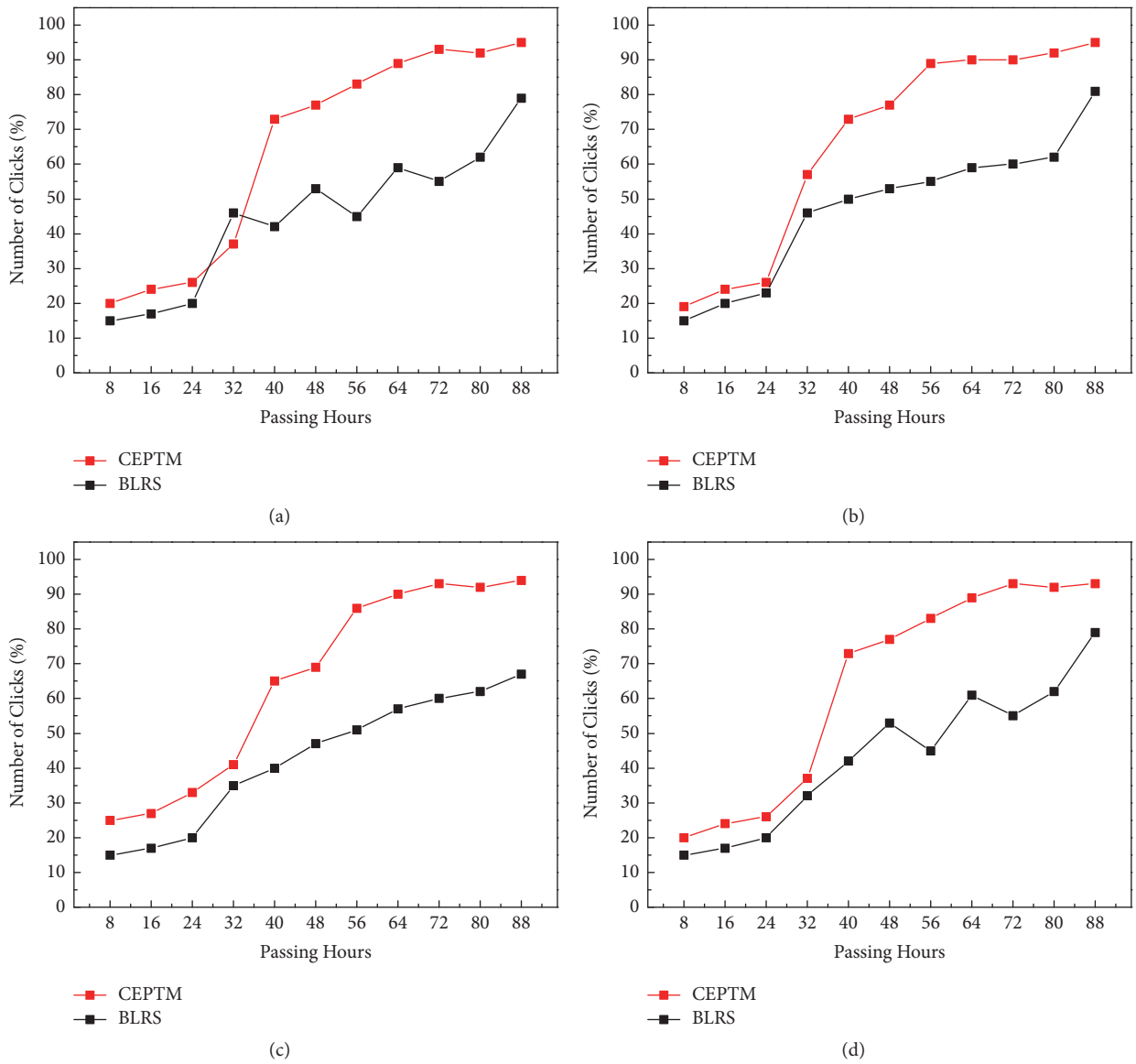

Figure 4: Four comparisons between CEPTM and BLRS running on each MEC.

aspects, user experience, and personalization quality were inputted into an exploratory factor analysis. Table 2 provides details on the results.

Table 2 lists the eight factors in the experiment that were loaded to fit the factor measurement model and the structural relations between the factors and variables in the exploratory factor analysis. The model achieved a good fit: $p < 0.01$, CFI = 0.985, TLI = 0.99, RMSEA = 0.035, SRMR = 0.631, and it revealed interesting findings among the factors, e.g., **item expertise** and **personalization diversity** had a positive effect on the users' **understandability** (+0.471, +0.273), which in turn raised the **personalization quality** (+0.371) and lowered the users' privacy concerns (-0.239). Furthermore, users' **trust propensity** had a negative effect on **satisfaction** (-0.351), while **satisfaction** increased the acceptance rate of a recommended item or stranger (+0.291). Users' **item expertise** and **familiarity with privacy and PS** both positively affected users' understanding of the cross-edge model (+0.418, +0.227). According to the above results, we describe why those factors influenced one another and briefly discuss how to manage users' information disclosure and the quality of the personalization engine. We strongly argue that users' knowledge that supported their decisions on information disclosure played an important role in this study. **Item expertise** and **familiarity with privacy and PS** determined users' **understandability**, which in turn increased the probability of accepting recommended items. This aspect further lowered users' privacy concerns in such a way that the users decided to disclose more information, and they increased their satisfaction with the personalization of the model. Users could benefit from the personalization service and the informing function, so they disclose more privacy for increased personalization quality. In this study, if users should know the underlying strategy of the privacy collection and

Table 2: Eight factors in the users' feedback on personalization services (PS) and the related parameters.

Factors	Questions	User Feedback Sample	Alpha	AVE
item expertise	4	I know the artists very well, and my friends always accept my music suggestions.	0.83	0.832
trusting propensity	4	I seldom accept a personalization from a stranger.	0.73	0.681
familiarity with privacy & PS	3	I understand that the personalized needs people's private information to generate better personalization.	0.89	0.893
understandability	3	The model provides better management of friends' posts across multiple sites.	0.75	0.709
satisfaction	5	The informing function of the model saved me time in viewing my friends' status updates on all sites.	0.90	0.738
PS quality	4	I like the recommended item.	0.78	0.712
privacy concern	5	I don't like requests that are too sensitive.	0.75	0.709
PS diversity	3	I like the personalization of a new type of item.	0.81	0.731

personalization, they might decrease their privacy concerns and increase their propensity for information disclosure. Users chased after the benefits of using our model, such as receiving high-quality personalization for unknown items, as well as being informed of their friends' real-time activities on other edge servers in MEC network. This trend was reflected in their satisfaction being directly determined by the personalization quality.

This study reveals that personalization should collect users' data while respecting their privacy concerns. This approach would allow users to share their information without experiencing worry or unnecessary anxiety. However, current privacy policy has not been designed well enough to manage and collect the users' information based on their personal traits or background knowledge. For example, a site in the MEC network will post the same set of terms and data policy to whoever signs in, even though most users do not read it. From a managerial perspective, why not design a personalized data collection policy in such a way that the users' satisfaction could be maximized? That should be an important consideration when adjusting the interactions between MEC edge servers and their users.

5. Conclusion

In this paper, we proposed a novel cross-edge model, called CEPTM, for better personalization service, and we reveal how famous topics in one resource edge server can emerge on several other destination edge servers in MEC. The prototype of the model consists of the interaction between the users and diverse items, the personalization of new items, and privacy collection and distribution strategies. Our study indicates that CEPTM (1) achieves a high rate of personalization acceptance due to the availability of a large amount of data as inputs and highly diverse personalization as outputs and (2) gains users' trust because it collects user data while respecting an individual's privacy concerns. It outperforms a baseline personalization service that is running on a single MEC edge server. The experiment results, which were performed on several real-life datasets, indicated that the trends of hot topics on a resource edge server could migrate to other edge servers in an MEC environment. An exploratory factor analysis revealed that personalization of diverse types is more likely to be accepted than single-type personalization and that helping users gain more knowledge and understanding can increase their disclosure of personal information, which enables the personalization service to predict items better, with higher accuracy. Additionally, CEPTM has provided new insight into handling the cold start problem in cloud computing, to which scholars and personalized providers can refer. In future work, we will continue to focus on user-tailored personalization, which covers both providing items that fit users' preferences and requesting users' disclosures according to their individual privacy concerns in MEC networks. Hopefully, additional interesting results will be discovered.

Data Availability

Users' data for the study are collected from our previous work [37, 38], and users who have at least 3 server accounts in EMC can join the study in Section 5.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61702312), the Natural Science Foundation of Shandong Province of China (No. ZR2017BF019), and the project of Shandong Province Higher Educational Science and Technology Program (No. J17KB178); and Hongchen Wu acknowledges the financial support from the China Scholarship Council (CSC, File No. 201306220132).

References

[1] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, vol. 78, part 2, pp. 680–698, 2016.

[2] Z. Wang, Q. Zhao, F. Xu, H. Dai, and Y. Zhang, "Detection performance of packet arrival under downclocking for mobile edge

computing," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 9641712, 7 pages, 2018.

[3] Z. Ruobin, Z. Weihua, and H. Jiang, "Scalable multiple description coding and distributed video streaming in 3G mobile communications," *Wireless Communications and Mobile Computing*, vol. 5, no. 1, pp. 95–111, 2005.

[4] T. Verhagen, J. van Nes, F. Feldberg, and W. van Dolen, "Virtual customer service agents: Using social presence and personalization to shape online service encounters," *Journal of Computer-Mediated Communication*, vol. 19, no. 3, pp. 529–545, 2014.

[5] P. M. Herceg, T. B. Allison, R. S. Belvin, and E. Tzoukermann, "Collaborative exploratory search for information filtering and large-scale information triage," *Journal of the Association for Information Science and Technology*, vol. 69, no. 3, pp. 395–409, 2018.

[6] J. Teevan, S. T. Dumais, and E. Horvitz, "Personalizing Search via Automated Analysis of Interests and Activities," *ACM SIGIR Forum*, vol. 51, no. 3, pp. 10–17, 2018.

[7] B. Loepp, K. Herrmanny, and J. Ziegler, "Blended recommending: Integrating interactive information filtering and algorithmic recommender techniques," in *Proceedings of the 33rd Annual CHI Conference on Human Factors in Computing Systems, CHI 2015*, pp. 975–984, Seoul, Republic of Korea, April 2015.

[8] Y. Li, C. Zhai, and Y. Chen, "Exploiting rich user information for one-class collaborative filtering," *Knowledge and Information Systems*, vol. 38, no. 2, pp. 277–301, 2014.

[9] G. Guo, J. Zhang, and D. Thalmann, "Merging trust in collaborative filtering to alleviate data sparsity and cold start," *Knowledge-Based Systems*, vol. 57, pp. 57–68, 2014.

[10] S. Guizani, "Relay attacks concerns in wireless ad hoc, sensors, and RFID networks," *Wireless Communications and Mobile Computing*, vol. 16, no. 11, pp. 1431–1435, 2016.

[11] G. Xu, Q. Wu, M. Daneshmand, Y. Liu, and M. Wang, "A data privacy protective mechanism for wireless body area networks," *Wireless Communications and Mobile Computing*, vol. 16, no. 13, pp. 1746–1758, 2016.

[12] B. Paudel, F. Christoffel, C. Newell, and A. Bernstein, "Updatable, accurate, diverse, and scalable recommendations for interactive applications," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 7, no. 1, 2016.

[13] D. Margaris, C. Vassilakis, and P. Georgiadis, "Query personalization using social network information and collaborative filtering techniques," *Future Generation Computer Systems*, vol. 78, pp. 440–450, 2018.

[14] X. Bai, B. Barla Cambazoglu, F. Gullo, A. Mantrach, and F. Silvestri, "Exploiting search history of users for news personalization," *Information Sciences*, pp. 125–137, 2017.

[15] Z. Yu, C. Wang, J. Bu, X. Wang, Y. Wu, and C. Chen, "Friend recommendation with content spread enhancement in social networks," *Information Sciences*, vol. 309, pp. 102–118, 2015.

[16] S. Wang, Y. Zhao, L. Huang, J. Xu, and C.-H. Hsu, "QoS prediction for service recommendations in mobile edge computing," *Journal of Parallel and Distributed Computing*, 2017, In press.

[17] T. Spyropoulos and P. Sermpezis, "Soft cache hits and the impact of alternative content recommendations on mobile edge caching," in *Proceedings of the the Eleventh ACM Workshop*, pp. 51–56, New York City, NY, USA, October 2016.

[18] D. Kim, C. Park, J. Oh, and H. Yu, "Deep hybrid recommender systems via exploiting document context and statistics of items," *Information Sciences*, vol. 417, pp. 72–87, 2017.

[19] L. Yang and X. Amatriain, "Recommending the world's knowledge: Application of recommender systems at Quora," in *Proceedings of the 10th ACM Conference on Recommender Systems, RecSys 2016*, p. 389, Boston, Mass, USA, September 2016.

[20] E. Aguirre, D. Mahr, D. Grewal, K. de Ruyter, and M. Wetzels, "Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness," *Journal of Retailing*, vol. 91, no. 1, pp. 34–49, 2015.

[21] Y. Sun, N. Wang, X.-L. Shen, and J. X. Zhang, "Location information disclosure in location-based social network services: Privacy calculus, benefit structure, and gender differences," *Computers in Human Behavior*, vol. 52, pp. 278–292, 2015.

[22] J.-M. Lee and J.-Y. Rha, "Personalization-privacy paradox and consumer conflict with the use of location-based mobile commerce," *Computers in Human Behavior*, vol. 63, pp. 453–462, 2016.

[23] S. Meyffret, L. Médini, and F. Laforest, "Trust-based local and social recommendation," in *Proceedings of the 4th ACM RecSys Workshop on Recommender Systems and the Social Web, RSWeb 2012*, pp. 53–60, Dublin, Ireland, September 2012.

[24] G. Liu, Y. Wang, M. A. Orgun, and E.-P. Lim, "Finding the optimal social trust path for the selection of trustworthy service providers in complex social networks," *IEEE Transactions on Services Computing*, vol. 6, no. 2, pp. 152–167, 2013.

[25] E. Toch, N. M. Sadeh, and J. Hong, "Generating default privacy policies for online social networks," in *Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems, CHI 2010*, pp. 4243–4248, Atlanta, Ga, USA, April 2010.

[26] T. De Pessemier, K. Vanhecke, and L. Martens, "A hybrid strategy for privacy-preserving personalizations for mobile shopping," in *Proceedings of the the 8th ACM Conference on Personalized System*, pp. 22–25, 2014.

[27] C. Brodie, C.-M. Karat, J. Karat, and J. Feng, "Usable security and privacy: A case study of developing privacy management tools," in *Proceedings of the Symposium on Usable Privacy and Security, SOUPS 2005*, pp. 35–43, Pittsburgh, Pa, USA, July 2005.

[28] Y. Li and Y. Guo, "Cultural distance-aware service recommendation approach in mobile edge computing," *Scientific Programming*, vol. 2018, Article ID 2181974, 8 pages, 2018.

[29] J. Yang, H. Wang, Z. Lv, and H. Song, "Multimedia personalization and transmission system based on cloud platform," *Future Generation Computer Systems*, vol. 70, 2016.

[30] Y. Zhang, "Browser-oriented universal cross-site recommendation and explanation based on user browsing logs," in *Proceedings of the the 8th ACM Conference*, pp. 433–436, Foster City, Silicon Valley, Calif, USA, October 2014.

[31] A. Barth, C. Jackson, and J. C. Mitchell, "Robust defenses for cross-site request forgery," in *Proceedings of the 15th ACM conference on Computer and Communications Security, CCS'08*, pp. 75–87, Alexandria, Va, USA, October 2008.

[32] B. Li, "Cross-domain collaborative filtering: A brief survey," in *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2011*, pp. 1085-1086, Boca Raton, Fla, USA, November 2011.

[33] W. Pan, E. W. Xiang, N. N. Liu, and Q. Yang, "Transfer learning in collaborative filtering for sparsity reduction," in *Proceedings of the 24th AAAI Conference on Artificial Intelligence and the 22nd Innovative Applications of Artificial Intelligence Conference, AAAI-10 / IAAI-10*, pp. 230–235, Atlanta, Ga, USA, July 2010.

[34] H. Zhou, Z. Deng, Y. Xia, and M. Fu, "A new sampling method in particle filter based on Pearson correlation coefficient," *Neurocomputing*, vol. 216, pp. 208–215, 2016.

[35] Q. Zou, J. Zeng, L. Cao, and R. Ji, "A novel features ranking metric with application to scalable visual and bioinformatics data classification," *Neurocomputing*, vol. 173, part 2, pp. 346–354, 2016.

[36] N. Altman and M. Krzywinski, "Points of Significance: Association, correlation and causation," *Nature Methods*, vol. 12, no. 10, pp. 899-900, 2015.

[37] H. Wu, B. P. Knijnenburg, and A. Kobsa, "Improving the prediction of users' disclosure behavior… by making them disclose more predictably?" in *Proceedings of the SOUPS2014 Workshop on Privacy Personas and Segmentation*, Menlo Park, Calif, USA, 2014.

[38] H. Wu, H. Zhang, L. Cui, and X. Wang, "A heuristic model for supporting users decision-making in privacy disclosure for recommendation," *Security & Communication Networks*, vol. 4, pp. 1–13, 2018.