Research Article

A Novel Contextual Information Recommendation Model and Its Application in e-Commerce Customer Satisfaction Management

Feipeng Guo[1] and Qibei Lu[2]

[1]*Department of Information Technology, Zhejiang Economic and Trade Polytechnic, Hangzhou 310018, China*
[2]*School of Economics and Trade, Taizhou Vocational and Technical College, Taizhou 318000, China*

Correspondence should be addressed to Qibei Lu; luqibei@hotmail.com

In the current supply chain environment, distributed cognition theory tells us that various types of context information in which a recommendation is provided are important for e-commerce customer satisfaction management. However, traditional recommendation model does not consider the distributed and differentiated impact of different contexts on user needs, and it also lacks adaptive capacity of contextual recommendation service. Thus, a contextual information recommendation model based on distributed cognition theory is proposed. Firstly, the model analyzes the differential impact of various sensitive contexts and specific examples on user interest and designs a user interest extraction algorithm based on distributed cognition theory. Then, the sensitive contexts extracted from user are introduced into the process of collaborative filtering recommendation. The model calculates similarity among user interests. Finally, a novel collaborative filtering algorithm integrating with context and user similarity is designed. The experimental results in e-commerce and benchmark dataset show that this model has a good ability to extract user interest and has higher recommendation accuracy compared with other methods.

1. Introduction

With the development of mobile commerce, Internet of Things, social networking, and ubiquitous computing, the boundaries of supply chain have been greatly expanded. We have entered a business information era called "big data," including user data, resource data, and transaction data [1]. Due to the fact everyone has different age, gender, occupation, education, background, and also interests, toward the same information, what they need from resources and which aspects they are interested in are different [2, 3]. Currently, most of the e-commerce sites can only provide same results for different users, or prediction result in user interest is low. They are facing a problem of "Though people access massive amounts of information, they have the difficulty of obtaining useful information," which is affecting customer satisfaction [4–6].

As to how to enhance the customer satisfaction, there are many approaches such as better quality of products and

appropriate service for customers. Personalized information recommender system is able to meet these needs of individual users. It provides information services based on explicit user request or through the analysis of users' personality and users' habits to reach out to provide information service that may be required by users. Thus, personalized information recommender system gradually becoming an effective means to ease the problem of "information overload," as well as eventually satisfying customers [7, 8]. Currently, recommended system mainly uses content-based algorithm [9], collaborative filtering-based algorithm [10], and hybrid-based algorithm [11]. Among these methods, collaborative filtering recommendation algorithm (CF) has a good effect; meanwhile, it has lower implementation and maintenance costs. Thus, CF gets a lot of research and practical applications [12–14]. But with the deepening of research and applications, people gradually discovered some problems such as poor accuracy and reliability of the acquisition on information needs of individual users and low adaptive of information matching in the

current existing methods [15, 16]. Also, in the current business environment of social networks, users' interest and need will frequently change with some contexts related to time and space [5]. However, in the specific field of recommender systems, some systems tend to focus on the relationship between "users and items," while showing less consideration in contextual environment (such as time, location, surrounding people, mood, activity status, and network conditions). In many application scenarios, relying solely on the "user-item" dual relationship cannot generate accurate recommendations. Some systems only focus on the recommendation generation process, while ignoring the process of user's interest extraction. It means that they assume that they already have plenty and contextual information-included explicit user interest information. In practical applications, this type of contextual user interest is often difficult to obtain or only can obtain a very small amount, resulting in the fact that current system is hard to practice [17, 18].

The issues discussed above show that, in the process of information obtaining, users' personalized information needs have context-sensitivity, which is a major cause of these problems. Context-sensitivity is the context that has a stronger impact on user interest. It is necessary to provide some information services that accurately and efficiently are consistent with the user's current context and meet user preferences; we called it contextual information recommendation [19]. Currently, the context is introduced into the recommender system by scholars [20–23]. These scholars start to study definition, identification, and modeling of the context. But contextual recommender has many problems waiting to solve, such as contextual user preferences extraction and high-dimensional data sparsity [5, 24]. Meanwhile, literatures basically study the influence of single context factors on recommender system [25, 26]; actually, various context factors affect recommender system, considering that a variety of context factors can get a higher recommendation result [27, 28]. On the other hand, obtaining user needs and providing personalized service, in fact, is a process to help people make decisions. At the same time cognitive ability means human brain's processing capacity, storage capacity, and extract information. There is a close link between user needs acquiring and individualized decision-making [14, 29]. Distributed cognition theory tells us that various types of context information affect user decision-making. Therefore, for the problem in contextual recommender, paper proposes a novel contextual information recommendation model based on distributed cognition theory. Firstly, for the problem of user interest extraction under complex contexts, in order to extract more comprehensive, accurate user interest, the theory of distributed cognition is introduced to the modeling process of user interest. The innovative introduction of the distributed cognition theory emphasizes cognitive differences on various environmental factors and how these differentiated contexts affect each level of the recommender, rather than each context factor having equal weight influence on user interest. Secondly, on the basis of the extracting of sensitive contexts proposed above, paper proposes a novel collaborative filtering recommendation algorithm integrating with context similarity. It combines user similarity with context

similarity in context-sensitivity to complete recommended tasks. Experimenting in an e-commerce and benchmark data set, this paper concludes that the new model has superiority over others.

The remainder of the paper is organized as follows. Section 2 discusses theoretical background and related research of recommender systems. The user interest and its information structure modeling within context are described in Section 3. In Section 4, details about the contextual information recommendation model based on distributed cognition theory are proposed and illustrated, including the extraction algorithm of user interest and a new collaborative filtering recommendation algorithm. In Section 5, we discuss the results of our empirical experiments based on the *Moviepilot-mp.mood* data set 1 and the e-commerce data set 2. We then summarize the main characteristics and advantages of the proposed model. Finally, we offer concluding remarks and describe future directions of our research work.

2. Theoretical Background and Related Work

2.1. Collaborative Filtering Recommendation Method. The recommender system based on collaborative filtering is dependent on user rating to build user-item ratings matrix. This method uses statistical techniques to find neighbors users that have the same or similar interest preferences (e.g., similar score and similar purchased goods for different commodities). Using the rating of commodity items from neighbors to predict the rating of target user for new items. Then, it selects the top N items sorted in descending order as recommended set to the target user. (top N recommendation) [2]. It is another advantage that it has no special requirements to the recommended object and can handle unstructured complex objects, such as music and movies. Meanwhile, the relationship between users requires a lot of historical data of study on the behavior of users, which intersects with social networks. Thus, there is a wealth of research foundation and broad prospects. Early research of collaborative filtering includes Tapestry system, GroupLens, and Jester system [2]. Currently the technology for the study is divided into user-based, item-based, and model-based collaborative filtering [5]. Recently, the character where user interests and needs change over context of time and space is more obvious, while traditional collaborative filtering is not considered the effects of context factors on the user's preferences [3].

2.2. Distributed Cognition Theory. Distributed cognition theory is related to not only the perception in the distribution of the individual, but also more emphasis on cognitive distributed in a variety of external factors (environment, media, culture, society, time, etc.). This theory describes the distributed and difference cognition of various external elements, which are important to the whole cognition activities. In the broad sense, cognitive psychology includes structural psychology, psychologism, and information processing [28]. In 1990s, Hutchins proposed the concept of "distributed cognition" where the nature of cognition is distributed. Cognitive phenomena include not only cognitive activities that occur in the minds of individuals, but also the external environment,

media, cultural, social, and time, and so forth. In the follow-up study, Hutchins team is committed to the application of distributed cognition theory. Unremittingly, they took the distributed cognition as research of framework or theoretical basis to analyze human-computer interaction (HCI) and computer supported cooperative work (CSCW) [30, 31].

From the conclusion of related work, we can see that (1) existing user interest extraction technique is mainly based on traditional machine learning, data mining, with little regard to the impact on the user context information needs. That means ignoring the inherent relevance between demand level of user information content and user interest under different contexts. (2) Some techniques for extraction of user interest had considered context, but they only consider the same kind of context types (such as time) for different context instances (morning, afternoon, etc.) on the users' interests, while ignoring the various context types for distributed and differentiated impact of user needs in recommendation service. It also means that little research had detected and quantified the sensitive contexts and priority order of impact to user interest. (3) Existing recommendation algorithm did not well consider the impact of contextual factors on recommendation process. Although some did it, most of them failed to consider the effective integration of existing methods perfectly. They mostly have the problems of low predicting accuracy or bad coverage. Therefore, it is difficult for company to put it in practice and the contribution of this paper is to solve this problem.

3. Description of User Interest and Information Structure Modeling

3.1. Description of Contextual User Interest. Interest characteristics of e-commerce users are mainly influenced by internal factors and external factors related to user preferences. We called these factors context. Internal factors are user's individual characteristics such as gender, age, occupation, personality, education, income, emotions, intentions and other aspects of human cognition. External factors include cultural, social and environment, region, device, time, and so forth. Internal and external contexts lead to the emergence of many different behaviors of Internet users. Because of this reason, it makes a difference in many aspects of each user and makes a difference in level of interest for commodities and commodity interest bias.

Part of user interest contexts is described as following.

3.1.1. User Context Information

(1) Regional culture: in addition to the various regions having different networks of total retail sales, per capita GDP, per capita consumption expenditure, network penetration, and so forth, they also have different cultural backgrounds among users, and they have different lifestyles; what is more is that some consumer behaviors have local characteristics. For instance, "Shanghai culture," "Beijing culture," and "Guangzhou culture," all of them affect individuals in consumer behavior. To facilitate quantitative analysis,

this paper characterizes the culture as a geographical feature points. Different regions may have similar or different shopping interests due to local customs or cultural theme.

(2) Gender: there is a great difference in interest between men and women. Women generally have more interests in cosmetics, garment, and accessories, while men are interested in digital electronics products.

(3) Age: people of different ages have different interests in commodity groups. Young people are generally focused on fashion, the old tend to be more focused on practicality, and, with age, they have a gradual shift of interest. For example, after students enter the society from campus, their interest in things will change.

(4) Marriage: whether a person is married or not will also make the user's interest drift. People who are not married will decide their own interests by individual needs. The married person shows their points of interest in entire family, and after a person gets married, he or she shows more focus on children's goods.

(5) Education: people of different educational background have different requirements for commodities. People of higher educational background have the more comprehensive understanding of online shopping and requirements.

(6) Incomes: incomes directly affect the quality requirements of users' interested in the product. People who have high income always have high quality requirements. With the increase in personal income, he or she pays more attention to the quality of goods.

3.1.2. User Behavior Information. Often the user's interest can be reflected in their behavior and there will be a certain tendency when they are interested in something. The user's needs and interests are recorded in their behavior information. Therefore, we can deduce the true interest of the user by analyzing searching behavior, browsing behavior, purchase records, and other information. In this paper, user behavior information includes the following aspects.

(1) *User's Search Keywords.* To a certain extent, search keywords reflect the exact user's current interest, which illustrates that users have a strong desire to obtain the concept. But it is difficult for users to describe their interest completely and clearly, which they need to be adjusted to match the concept in knowledge base, in order to describe the user's interest more accurately.

(2) *User's Purchase Records.* Commodities that people have purchased could reflect user's previous interest; this interest may be long in term or just temporary. According to knowledge rules we can identify the user's long-term interest theme or impact users' interest in the topic content in the near term.

(3) *User's History Browsing Behavior.* User's browsing behavior also reflects to some extent the user's current

interest that may include the subject of browsing page, the number of visiting the product category, add to favorites, the residence time of the page, drag the scroll bar, bookmarked, copy, click on the link, and so on.

3.2. Modeling User Interest Information Structure. For description of the users and their interest discussed above in e-commerce site, this paper models the structure of user interest information, making the definition of relevant data as follows.

Definition 1. The user (u) has the following features in the internet: register on the site and have unique access account; provide background information about the individual; access the website; browse, purchase, and other operations.

Then the set consisting of all users is defined as $U = \{u_1, u_2, \ldots, u_N\}$ and N represents the number of users. Any user ($u_i = \{u_{ik} \mid k \in [1, X]\}$) is defined as the set of user attributes; X represents the number of types of user attributes.

Definition 2. Recommended services set (S) is defined as $S = \{s_i \mid j \in [1, M]\}$, M represents the number of services, and $s_j = \{s_{jp} \mid p \in [1, Y]\}$ is defined as a collection of service attributes, where Y represents the number of types of service attributes.

Definition 3.

User Context. Context object is any associated object in recommendation process, including user attributes and external factors. A nonempty set of properties $C = C_1 \times C_2 \cdots C_L$ is used to describe the characteristic of object. This set is expressed as Cartesian product of a single-dimension context type c_i. $C_i = \{c_{i_1}, c_{i_2}, \ldots, c_{i_m}\}$, $i \in [1, L]$ and M is the number of specific instance of a context type.

Definition 4. User historical behavioral information set (UHB) is defined as UHB = $\{$uhb$_{ij} \mid i \in [1, N]$, $j \in [1, M]$, uhb$_{ij} \in B_1 \times B_2 \times \cdots B_w\}$. The uhb$_{ij}$ is the behavior variable expressed by a single nonnegative real number or multiple, such as search keywords, user interest page topic, time of browsing the page, add bookmarks, save the page, drag the scroll bar, and click on the link, which composes the behavior vector UBH = {Keywords, Topics, Times, Book, Save page, Scroll, Links}.

Definition 5. Context information set of user history behavior is defined as

$$\text{UHBC} = \left\{ \text{uhbc}_{ij\overline{r}} \mid i \in [1, N], j \in [1, M], \right.$$

$$\text{uhbc}_{ij\overline{r}} \in B_1 \times B_2 \times \cdots B_w, c_{\overline{r}} \in C, \quad (1)$$

$$\left. r \in \left[1, \prod_{}^{L} m\right] \right\}.$$

The UHBC means user history behavior information collection in the contexts, where $c_{\overline{r}} \in C$ is a multidimensional

example of user context. uhbc$_{ij\overline{r}}$ is the behavior variable expressed by a single nonnegative real number or multiple, such as user purchase goods information at a certain time (Product), buy time (BTime), and purchase intentions (Intention), which components the behavior vector UHBC = {Product, BTime, Intention}.

Definition 6. User interest set (UI), UI = $\{$ui$_{ijt'} \mid i \in [1, N]$, $j \in [1, M], c_{\overline{t'}} \in C_{\overline{T}} \in C\}$, ui$_{ijt'}$ is a real number. UI is defined as an interest set of different services in different contexts. A larger value of ui$_{ijt'}$ indicates a higher preference, wherein $c_{\overline{t'}}$ represents an instance of user combination context.

4. Contextual Information Recommendation Model Based on Distributed Cognition Theory

This paper proposes to construct a contextual information recommendation model based on distributed cognition theory. The basic idea is to solve existing problems and is inspired by "distributed cognition theory"; this paper researches the hierarchical user needs, distributed cognition of context-sensitive information, and differential calculation and gradually extracted the user preferences precisely. Firstly, to some extent, for a user's preference affected by internal and external context, this model makes the context factors distributed cognition and analyzes the differential influence for user preferences caused by various effective context type and their instances. Then, model extracts the user preferences based on cognition of context-sensitivity. Secondly, model introduces user context information into collaborative filtering recommendation process and calculates the similarity among context in the context-sensitive set. Then, it constructs a similar context set of target user with current contexts and designs a novel collaborative filtering recommendation algorithm that integrated contextual similarity. Detailed model is shown in Figure 1.

4.1. Extraction Algorithm of User Interest Based on Distributed Cognition Theory. After analyzing distributed cognition theory and user interest model discussed perviously, this section firstly defines the concept set of context-sensitive cognition, detects the context-sensitivity, and quantifies the contexts by using distributed cognition theory. Secondly, the algorithm sets user historical behavior data of single-dimension context as data source in order to extract single-dimension contextual user interest. Finally, the algorithm integrates the distributed influence factors to extract multidimensional contextual user interest based on context-sensitive cognition.

Definition 7. uhbc$_{ijk_q}$ represents a single-dimension context of user historical behavior and context-sensitive set $(C_{\text{sensitivity}})_{ij}$ represents the set of context types which have a significant impact on users' interests. $\delta_{\text{threshold}}$ represents the threshold of context-sensitivity. The impact factor ε_{ijk} represents influence degree of one context-sensitive type C_k. It also means users (u_i) interest for service s_j is affected by C_k. (ui$_c$)$_{ijk_q}$ represents the interest of u_i for s_j under the single
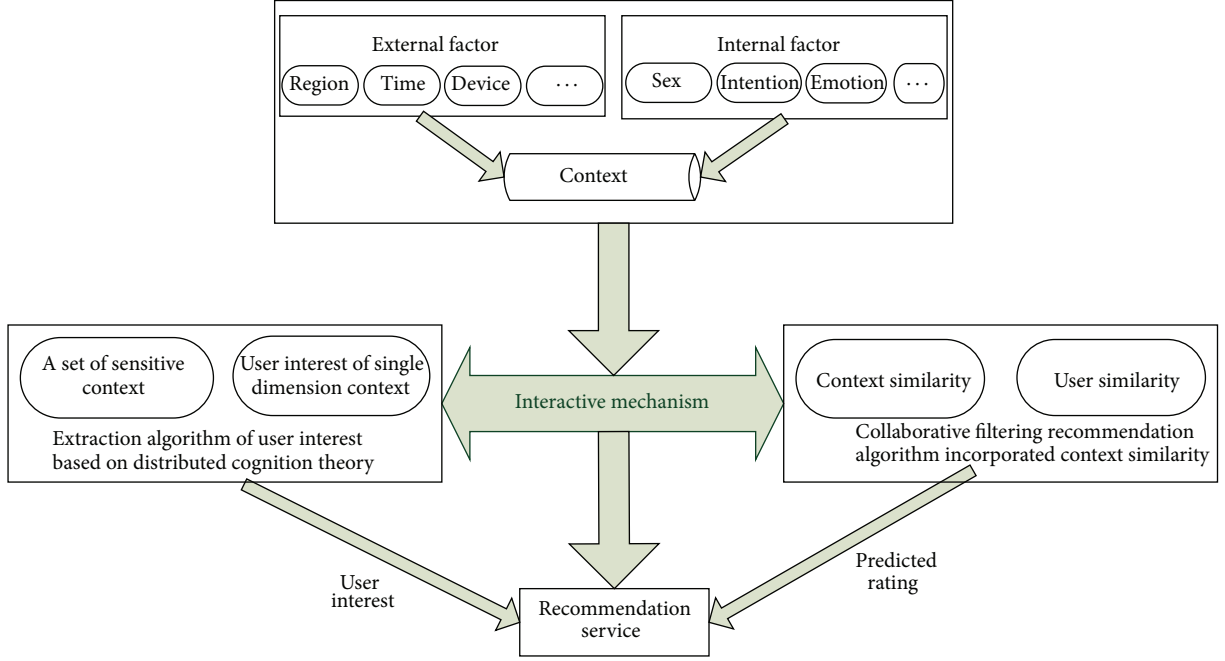
FIGURE 1: Contextual information recommendation model based on distributed cognition theory.

dimension context instance of c_{k_q}. $(\mathrm{UI}_C)_{ij\overline{t'}}$ represents the interest based on context-sensitive cognition, namely, represents the interest of u_i for s_j under multidimension context instance of $c_{\overline{t'}}$.

Definition 8. Detection of context-sensitive cognition has a distributed cognition to the various contexts which may influence the extraction of user interest. It sets user historical behavior data under the effect of a single-dimension context C_k as data source in order to analyze the fluctuation characteristics of user interest, which is influenced by this context. If D_{ijk} is not smaller than the threshold value $\delta_{\mathrm{threshold}}$, then the users (u_i) interest towards service s_j is affected by C_k and, otherwise, is not significantly affected by C_k.

The description of distributed cognitive theory-based algorithm to user interest elicitation (DCUIE) is as follows:

input: u_i, s_j, C_k;

output: $(\mathrm{UI}_C)_{ij\overline{t'}}$, rule set.

Step 1. Calculate the average value through a specific instance of a type of context and call it the single-dimension context of user historical behavior (uhbc_{ijk_q}), wherein the number of context of user historical behavior is d_{ij} and contains the context c_{k_q}. Consider

$$\mathrm{uhbc}_{ijk_q} = \frac{1}{d_{ij}} \sum_{r=1}^{d_{ij}} \mathrm{uhbc}_{ij\overline{r}} \mid C_{\overline{r}} = c_1 \cdots c_{k-1} c_{k_q} c_{k+1} \cdots c_N. \quad (2)$$

Step 2. Calculate user behavior variable variance (D_{ijk}) of u_i for s_j under the single dimension context and obtain concept

set of sensitive contextual cognition $((C_{\mathrm{sensitivity}})_{ij} = \{C_k \mid D_{ijk} \geq \delta_{\mathrm{threshold}}\})$. Consider

$$D_{ijk} = \sum_{q=1}^{Z_k} \left\{ \left[\left(\mathrm{uhbc}_{ijk_q} \right) - E_{ijk} \right]^2 \times p_{kq} \right\},$$

$$E_{ijk} = \left\{ \sum_{q=1}^{Z_k} \left[\left(\mathrm{uhbc}_{ijk_q} \right) \times p_{kq} \right] \right\}. \quad (3)$$

E_{ijk} means the value of user behavior expectations of u_i for s_j under the context C_k. p_{kq} means the probability of interactions between u_i and s_j under the instance of context c_{k_q} (comparing to other specific instances of C_k).

Step 3. Calculate the impact extent of certain sensitive context type C_k to detect the distributed cognition of sensitive context. Consder

$$\varepsilon_{ijk} = \begin{cases} \dfrac{D_{ijk}}{\sum_{k=1}^{N'} D_{ijk}}, & C_k \in \left(C_{\mathrm{sensitivity}} \right)_{ij} \\ 0, & C_k \notin \left(C_{\mathrm{sensitivity}} \right)_{ij}, \end{cases} \quad (4)$$

where $N' \leq N$ is the number of type of sensitive contexts.

Step 4. Calculate the user interest of single dimension. Algorithm sets single-dimension context of user historical behavior uhbc_{ijk_q} as data source and computes the user interest

of single-dimension $(\text{ui}_c)_{ijk_q}$ by K-means algorithm and its value range is $[a, b]$. Consider

$$(\text{ui}_c)_{ijk_q} = f_{K\text{-Means}}\left(\text{uhbc}_{ijk_q}\right) \mid \text{uhbc}_{ijk_q} \neq 0. \atop {\substack{1 \leq j \leq M \\ 1 \leq k \leq N' \\ 1 \leq q \leq Z_k}}} \tag{5}$$

Step 5. Calculate the multidimensional user interest based on sensitive context cognition. Algorithm calculates the multidimensional user interest by using $(\text{ui}_c)_{ijk_q}$ with the action of ε_{ijk}. Consider

$$(\text{UI}_C)_{ij\overline{t}'} = \sum_{k=1}^{N'} \left(\varepsilon_{ijk} \times (\text{ui}_c)_{ijk_q}\right). \tag{6}$$

Step 6. According to the set of context-sensitive $(C_{\text{sensitivity}})_{ij}$ and $(\text{UI}_C)_{ij\overline{t}'}$, algorithm uses C5.0 for rule extraction and interest analysis by combining the key attributes of the selected user characteristics, user behavior data.

4.2. Collaborative Filtering Recommendation Algorithm Incorporated Context Similarity. In this section, after extraction of users' interest, this paper gives the description on inputs, outputs, and detailed process of collaborative filtering recommendation algorithm incorporated context similarity (CFRCS):

> input: three-dimensional model of "User-Service-Context"; target user (u) and current context (C); recommendation service set (R);
>
> output: user's maximum top-N service preference in service set (R) under current context.

Specific process of algorithm is divided into the following steps.

Step 1. Using DCUIE algorithm to extract context-sensitive set $(C_{\text{sensitivity}})_{ij}$ and then obtaining the sheets of target user (u) preference data $(\text{UI}_C)_{ij\overline{t}'}$ by using the selected sheet of key attributes, user behaviors, and user ratings, then, CFRCS constructs the 2-dimensional preference matrix of context-service.

Step 2. Using the method in literature 5 to calculate the context similarity among contexts $(C_{\text{sensitivity}})_{ij}$,

$$\text{Sim}_u\left(u, c_i, c_j\right)$$

$$= \frac{\sum_{s \in S} \left(r_{(u,s,c_i)} - \overline{r}_{u,c_i}\right)\left(r_{(u,s,c_j)} - \overline{r}_{u,c_j}\right)}{\sqrt{\sum_{s \in S} \left(r_{(u,s,c_i)} - \overline{r}_{u,c_i}\right)^2 \times \sum_{s \in S} \left(r_{(u,s,c_j)} - \overline{r}_{u,c_j}\right)^2}}. \tag{7}$$

S is the service set that user (u) has scored under context c_i and context c_j ($c_i, c_j \in (C_{\text{sensitivity}})_{ij}$). The rated value of u for s under context c_i is expressed as $r_{(u,s,c_i)}$. The average rated value of u under context c_i is expressed as $\overline{r}_{u,c_i}$.

Step 3. It calculates the similarity among users under various contexts for the construction of user similarity matrix. This similarity matrix is 3-dimensional, and each element $\text{Sim}(c, i, j)$ represents cosine similarity between user i and user j under context c:

$$\text{Sim}\left(c, i, j\right) = \cos\left(\vec{V}_{c,i}, \vec{V}_{c,j}\right) = \frac{\vec{V}_{c,i} \times \vec{V}_{c,j}}{\left\|\vec{V}_{c,i}\right\| \left\|\vec{V}_{c,j}\right\|}. \tag{8}$$

$\vec{V}_{c,i}, \vec{V}_{c,j}$ is the eigenvector of attribute and behavior under the same context c, which is, respectively, for user i and user j.

Step 4. Create a nearest neighbor set of users. Every user in each context builds a set of nearest neighbors. The nearest neighbors set of user j is $N_j = \{N_{j,c_1}, N_{j,c_2}, \ldots, N_{j,c_k}\}$, $1 \leq j \leq \text{Num}(U)$, N_{j,c_i} is the set of nearest neighbor of user j under the context c_i, where $\text{Num}(U)$ is the total number of users and $c_1, c_2, \ldots, c_k$ means the number of context is k. The users who are obtained according to similarity calculation of context and users have similarity in context c_i also incorporated into the set of nearest neighbors. The number of users of nearest neighbor set under each context may be the same or may be different, such as the number of neighbors depending on the scale of context.

Step 5. According to the nearest neighbor set of target user (N_j) and multidimensional user interest in sensitive context $(\text{UI}_C)_{ij\overline{t}'}$ for generating recommendation, then, it brings the top-N services in maximum preference set for target user u.

User u has a prediction score $S_{j,s}$ for service s. $S_{j,s}$ can be obtained by calculating a weighted average value of user's (u) scores for service in N_j. Specific formula is as follows:

$$S_{j,s} = \frac{\sum_{i \in N_j} \left\lfloor (\text{UI}_C)_{ij\overline{t}'} \times \text{Sim}\left(c, i, j\right)\right\rfloor}{\sum_{i \in N_j} \left|\text{Sim}\left(c, i, j\right)\right|}. \tag{9}$$

5. Experiment Results and Discussion

Experimental hardware and software environment: computers have Intel Dual Core CPU 2.8 GHz, 4 GB RAM, and Windows 7 operating system. Development language environment is JDK1.7, Eclipse3.7, Matlab7, database MySQL5.5. Data uses benchmark set and real-world e-commerce data set.

5.1. Description of Data Set and Evaluation Criteria. Benchmark data set uses *Moviepilot-mp.mood*, which is specifically used for emotional-based recommendation task, mainly including time, location, target groups, and emotional status. Training data set of *Moviepilot-mp.mood* contains 105,137 users who have scored 4,544,409 records on 25,058 movies; testing data set contains 160 users who have scored 19,506 records on 3,396 movies. Score is an integer between 0 and 100 and increments at intervals of 5, 100 he said he likes. Since, before calculating the similarity between two users, the user needs to find a common movie rating, the experimental

set minimum number of corated movies is $|M_{cc'}|_{\min} = \min\_corated\_num = 10$. Found through the analysis of the training set and testing set under usually situation, user rating of at least 100 movies, target user, and other parts of users have common rating of not less than 10 films. Therefore, we select no less than 100 films from the training set, which are scored by 10,051 users, and then randomly select one of the 3,000 users as the ultimate training set users. After data preprocessing, we get 3,000 users who score 1,211,291 records on 21,333 movies.

e-commerce data set is got from a B2C platform and is described as follows:

(1) user data set: 500 users, whose attributes contain user identity, gender, age, occupation, personality, education, income, and so forth;

(2) network services data set: 100 kinds of services, whose attributes contain service ID, service price, quality of service, type of service, and so forth;

(3) context data set: 5 kinds of context: time, equipment, geography, emotion, and purpose; time is divided into quarter, month, week, and day and distinguished between weekdays and weekends; area is divided into north, east, south, northeast, central, northwest, and southwest; device is divided into PDA, mobile phones, and computers; emotion is divided into negative emotions, positive emotions, and other emotions; purpose is divided into self-use and gifts;

(4) context data set of user historical behavior: 8600 records (statistics within 30 days, user behavior variables take staying time on the same interest service as the key factor);

(5) user behavior data set: user-service behavior matrix (500×100) is composed by number 0 or 1.

User has used the service behavior marked as 1 and user has not used the service behavior marked as 0.

Evaluation criteria adopted $P@R$ and the mean absolute error (MAE). $P@R$ means ranking by users interests to recommend top-R kinds of service. These are compared in accuracy with the top-R kinds of service in test data set:

$$P@R = \frac{N_{\text{Top-}R}}{R}. \tag{10}$$

$N_{\text{Top-}R}$ is the number of top-R service both in recommend service set and test date set.

MAE measures the prediction accuracy by calculating the deviation between predicted value and actual value. It has a higher accuracy with a less bias.

Let us suppose that a collection of prediction score is expressed as $(p_1, p_2, \ldots, p_N)$, the corresponding set of actual rating is $(q_1, q_2, \ldots, q_N)$, and then

$$\text{MAE} = \frac{\left(\sum_{i=1}^{N} |p_i - q_i| \right)}{N}. \tag{11}$$

Furthermore, the evaluation standard also uses DOA (degree of agreement).

$\text{NW}_{U_j} = I - (L_{u_j} \cup T_{u_j})$ means the item does not appear in both training set of U_j and testing set of U_j. L_{u_j} is the item scored by U_j in training set and T_{u_j} is the item scored by U_j in testing set. Consider

$$\text{DOA}_{u_j} = \frac{\sum_{(i \in T_{U_j}, k \in \text{NW}_{U_j})} \text{check\_order}_{U_j}(I_i, I_k)}{|T_{U_i}| * |\text{NW}_{U_j}|}.$$

$$\text{check\_order}_{U_j}(I_i, I_k) \tag{12}$$

$$= \begin{cases} 1, & \text{predict\_rank}_{I_i} \geq \text{predict\_rank}_{I_k} \\ 0, & \text{otherwise.} \end{cases}$$

predict_rank$_{I_i}$ means predicted position of I_i in a recommended list. The DOA value of random prediction is about 50%, while DOA is 100% when all of the predictions are sorted correctly. This paper adopts averaged DOA among all users as the overall effect evaluation.

5.2. Experimental Results Analysis of User Interest Extraction. DCUIE validation experiment uses the e-commerce data set. Data is divided into two parts: the data of 24 days before (80%) is set as training set to extract user interest; the data of 6 days after (20%) is set as testing set to evaluate user interest extraction accuracy. Steps are as follows.

(1) Paper uses training set as input data and utilizes DCUIE and MCUIE (multidimensional contextual user interest elicitation) to extract context-related user interest. Then it, respectively, recommends top-R services under various combined context examples for users. Contrast number one: baseline method adopts random service selection. This means obtaining service list used by users and then randomly selecting R services. Contrast number two: MCUIE method calculates a variety of single-dimensional contexts of user interest in accordance with DCUIE. MCUIE regards user interest affected by various types of contexts as being of equal weight (namely, ε_{ijk} is the same). Then MCUIE extracts user interest under multidimensional context according to DCUIE.

(2) Paper uses testing set as input data. Contrast number one: baseline method adopts the users' used services which are sorted by the stay total time on the same interest service. Then this method obtains top-R services as users often use. Contrast number two: paper sorts the order of used service under various combined context examples for user thorough counting total stay time on the same interest service. Then this method obtains top-R services as users often use. After that, this contrast utilizes $P@R$ to compute the average recommendation accuracy of MCUIE, DCUIE under various combined context example, and comparative analysis of the experimental results.

Figure 2 shows that when $\delta_{\text{threshold}} = 0.6$, the overall accuracy of DCUIE is better than random service selection
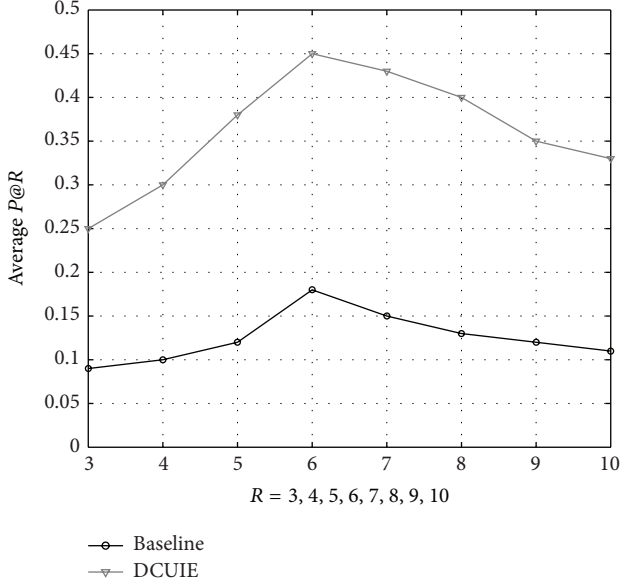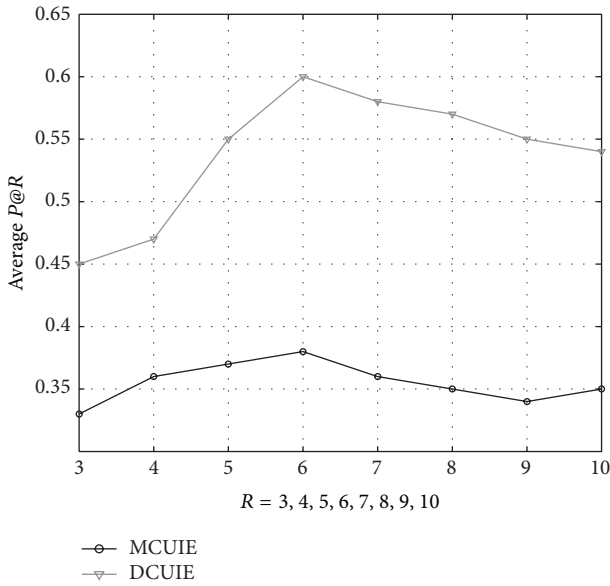
FIGURE 2: Validation of DCUIE ($\delta_{\text{threshold}} = 0.6$).

FIGURE 3: Comparison of experimental result between MCUIE and DCUIE.

method without personalized services. It also shows the introduction of user interest extraction method based on the fact that a distributed cognitive model is meaningful.

Figure 3 shows the $P@R$ comparative results of MCUIE and DCUIE when $\delta_{\text{threshold}} = 1.2$. As can be seen, set $P@R$ as the evaluation criteria and DCUIE method is better than MCUIE method. It states that paper discriminated and quantified impact of the user preference with various contexts and its examples can help improve the extraction accuracy of contextual user preferences. Meanwhile, the value of $\delta_{\text{threshold}}$ will affect the accuracy of DCUIE. To some sense, this is due to a user preference for certain services which is more affected

by context of location and activities, less of time, device, and personnel around, and vice versa. Namely, it can not take all contexts as the sensitive ones, and also can not consider them as being of equal weight. The analysis results have verified the practicality and validity of theories and methods established in the paper.

Experiments of DCUIE in e-commerce data set show user's interest is affected by contexts in a certain extent. Different context factors can lead to differences in user interest topics, such as the fact users who own different cultural backgrounds have different lifestyles; even some consumer behaviors have local characteristics. Someone may "like" a certain entertainment, but his interest value in this service may be low when he stays under the instance of combination context of "office" and "work." The results also find that contexts such as age, gender, income, geography, time, and purpose are more important factors that influence consumer choice of interest. For example, the rules indicate that the one who focuses on "brand" is most high-income young man, but the proportion of online shopping is not high. The one who is most concerned about the "cost" is highly educated people in middle-income level, especially among married women. The regional contexts affect consumer interest, and shopping behavior, such as per capita GDP, network penetration rate (including mobile, Internet), per capita consumption expenditures. Interest rule examples are shown in Table 1, where some of the rules have been merged.

5.3. Experimental Results Analysis of Contextual Recommendation. Paper selects the optimal context-sensitive test threshold $\delta_{\text{threshold}}$ for testing context based on distributed cognition theory in order to improve accuracy of contextual recommendation by filter of noise context. Therefore, this paper divides e-commerce datasets into many data pairs and compares DOA with different $\delta_{\text{threshold}}$ value, as shown in Table 2. The results show that different $\delta_{\text{threshold}}$ affects sorting accuracy of CFRCS on different testing datasets. When $\delta_{\text{threshold}} = 0$, since all the contexts are considered as sensitivity, CFRCS had more poor sorting results. When $\delta_{\text{threshold}} = 1.2$, CFRCS has best sorting results on DOA. However, the $\delta_{\text{threshold}}$ increases beyond 1.2 resulting in a reduced number of context-sensitive sets and DOA downward again. On the other hand, different data pair also has a greater effect on the DOA. On the whole, CFRCS divides data into 80%–20% which is the best pair.

Paper adjusts the number of neighbors k in e-commerce data set and analyzes the association between CFRCS and context influence. Table 2 shows that the accuracy of CFRCS achieving the best when the test threshold of context-sensitive $\delta_{\text{threshold}} = 1.2$. Therefore, paper sets $\delta_{\text{threshold}} = 1.2$ to compare the MAE of CFRCS, collaborative recommendation algorithm based on the prefiltering context (CCF), and traditional collaborative filtering algorithm without considering the context (TCF). These MAE are even value obtained from three training set, which are showed on the ordinate corresponding in Figure 4. The results find that CFRCS has a better recommendation quality than CCF and TCF with different number of neighbors, and this algorithm has the best recommendation quality when $k = 30$. It can be seen, by

TABLE 1: Users' contexts and interest rules table (part of C5.0 rules).

Number	Sex	Monthly income	Age	Region	Time	Purpose	Subject interest 1	Subject interest 2	Subject interest 3
1	Female	8000	32	Eastern China	Weekend	Self-use	Clothing	Hairdressing	Entertainment
2	Male	5400	28	Northeast	Workday	Gift	Digit	Recreation and sports	Food
3	Female	14381	27	Eastern China	Workday	Self-use	Accessories	Hairdressing	Food
4	Female	5930	33	Eastern China	Weekend	Self-use	Infant & mom	Home furnishing	Clothing
5	Female	24103	31	South China	Weekend	Gift	Hairdressing	Entertainment	Clothing
6	Male	4590	38	South China	Weekend	Gift	Digit	Home furnishing	Recreation and sports
7	Female	5089	40	Eastern China	Workday	Self-use	Home furnishing	Food	Hairdressing
8	Female	14093	46	Northeast	Workday	Gift	Clothing	Hairdressing	Food
9	Male	14903	31	Southwest	Weekend	Self-use	Food	Digit	Entertainment
10	Male	6250	42	North China	Weekend	Self-use	Home furnishing	Food	Recreation and sports

TABLE 2: Comparison results of CFRCS with different $\delta_{\text{threshold}}$ (DOA %).

$\delta_{\text{threshold}}$\split	80%–20%	90%–10%	70%–30%	50%-50%
0	86.4	85.5	84.9	83.9
0.4	87.1	86.3	85.5	85.1
0.8	88.5	87.3	86.3	85.9
1.2 (Cut-off point)	89.2	88.4	87.3	86.2
1.6	88.8	87.1	86.5	85.6
1.8	87.5	86.8	85.7	85.1
2.2	86.3	85.1	84.8	83.2

TABLE 3: Comparison results of different methods (DOA %).

Alg.\split	80%–20%	90%–10%	70%–30%	50%-50%
KNN	72.4	70.5	68.9	65.9
W-SlopeOne	75.1	73.3	71.5	68.1
CCF	80.5	79.3	75.3	70.9
CFRCS	88.2	86.4	83.3	81.2

FIGURE 4: MAE of three algorithms in an e-commerce data set.

selecting the appropriate test threshold of context-sensitiviy, the CFRCS has a better recommendation quality. This method makes the means of measure of context similarity more flexible in order to adapt to dynamic context change which affects the service needs change.

Overall, the experiment validation shows that DCUIE makes contextual recommendation more favorable to enhance the accuracy of the recommendation than traditional methods. Therefore, based on context-sensitivity, it can integrate some more sophistication and focus on user similarity measure into collaborative filtering recommendation algorithm.

To further verify the score prediction effect of CFRCS, this paper adopts *Moviepilot-mp.mood* movie rating data sets to do numerical experiments and selects kNN-100, W-SlopeOne, and CCF for contrastive analysis. Firstly, CFRCS extracts context-sensitive set, such as age, occupation, mood, time, and location by DCUIE and then calls CFRCS algorithm to recommend. It could be found that CFRCS gets the maximum DOA value on three different data sets by comparing with the other three recommendation algorithms, as shown in Table 3. It also finds that CFRCS gets the minimum
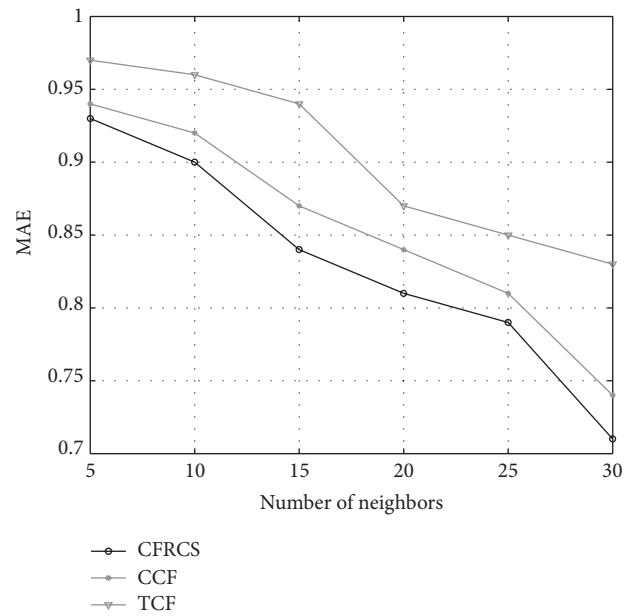
MAE value by comparing with the other two algorithms. Specifically in $K = 600$, CFRCS compared with traditional user-based collaborative filtering (TCF) and context pretreatment-based collaborative filtering (CCF) increases by 10.57% and 18.62%, respectively, as shown in Figure 5. Therefore, CFRCS can significantly improve the quality of recommender systems and be more suitable for personalized recommendations problem under future mobile business environment.

6. Conclusion

Personalized information recommendation service is an effective means to ease the "Internet information overload" problem and improve the customer satisfaction. It has important significance and practical value. This paper builds a novel contextual information recommendation model based on distributed cognition theory after modeling user interest information structure. Firstly, it proposes a user interest
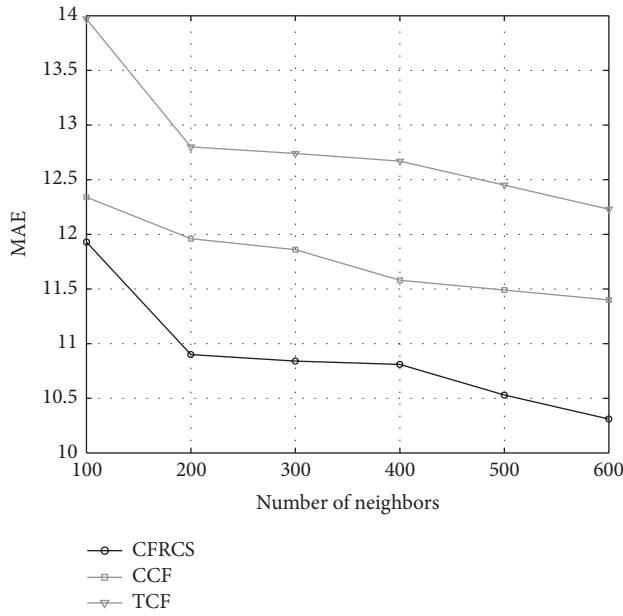
Figure 5: MAE of three algorithms on the standard data set.

extraction method based on distributed cognition theory. After researching the hierarchical user needs, the distributed and differentiated contexts, paper analyzes the impact of various factors in the process of interest extraction and gradually obtain precise interest. Secondly, it proposes contexts and contextual users' similarity calculation method to deal with user interest based on differentiated context-sensitive value. After that, it improves collaborative filtering algorithm. The results show that the model can effectively mine user behavior data in context. It also finds the differences in the degree of impact on users' interests with different contexts and its examples and reduces the negative impact caused by noise contexts, thereby improving accurate of contextual information recommendation service. Mainly a next step includes dependency relationship and its impact on user interest among different types of context. Study of how to alleviate the sparse problem of multidimensional contextual user interest and study of contextual recommendation problem based on user interest drift.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] B. Vinayak, J. C. Michael, and L. Chen, "Inside Big data management: ogres, onions, or parfaits?" in *Proceedings of the 15th International Conference on Extending Database Technology*, pp. 3–14, ACM, March 2012.

[2] F. Ricci, L. Rokach, B. Shapira et al., *Recommender Systems Handbook*, Springer, Berlin, Germany, 2011.

[3] J. Han, H. R. Schmidtke, X. Xie, and W. Woo, "Adaptive content recommendation for mobile users: ordering recommendations using a hierarchical context model with granularity," *Pervasive and Mobile Computing*, vol. 13, pp. 85–98, 2014.

[4] M. Terpstra and F. H. M. Verbeeten, "Customer satisfaction: cost driver or value driver? Empirical evidence from the financial services industry," *European Management Journal*, vol. 32, no. 3, pp. 499–508, 2014.

[5] G. Adomavicius and A. Tuzhilin, "Personalization technologies: a process-oriented perspective," *Communications of the ACM*, vol. 48, no. 10, pp. 83–90, 2005.

[6] Z. Lin, "An empirical investigation of user and system recommendations in e-commerce," *Decision Support Systems*, 2014.

[7] A. Finn, L. Wang, and T. Frank, "Attribute perceptions, customer satisfaction and intention to recommend E-services," *Journal of Interactive Marketing*, vol. 23, no. 3, pp. 209–220, 2009.

[8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[9] J. W. Kim, B. H. Lee, M. J. Shaw, H.-L. Chang, and M. Nelson, "Application of decision-tree induction techniques to personalized advertisements on internet storefronts," *International Journal of Electronic Commerce*, vol. 5, no. 3, pp. 45–62, 2001.

[10] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, 2003.

[11] S. Debnath, N. Ganguly, and P. Mitra, "Feature weighting in content based recommendation system using social network analysis," in *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*, pp. 1041–1042, Beijing, China, April 2008.

[12] Y. Ding and X. Li, "Time weight collaborative filtering," in *Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM '05)*, pp. 485–492, New York, NY, USA, 2005.

[13] H. Yildirim and M. S. Krishnamoorthy, "A random walk method for alleviating the sparsity problem in collaborative filtering," in *Proceedings of the ACM Conference on Recommender Systems*, pp. 131–138, New York, NY, USA, October 2008.

[14] S. Hosany and G. Prayag, "Patterns of tourists' emotional responses, satisfaction, and intention to recommend," *Journal of Business Research*, vol. 66, no. 6, pp. 730–737, 2013.

[15] G. Adomavicius and F. Ricci, "RecSys'09 workshop 3: workshop on context-aware recommender systems (CARS-2009)," in *Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys '09)*, pp. 423–424, New York, NY, USA, October 2009.

[16] W.-P. Lee and K.-H. Lee, "Making smartphone service recommendations by predicting users' intentions: a context-aware approach," *Information Sciences*, vol. 277, pp. 21–35, 2014.

[17] Z. D. Champiri, S. R. Shahamiri, and S. S. Binti Salim, "A systematic review of scholar context-aware recommender systems," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1743–1758, 2014.

[18] A. Hawalah and M. Fasli, "Utilizing contextual ontological user profiles for personalized recommendations," *Expert Systems with Applications*, vol. 41, no. 10, pp. 4777–4797, 2014.

[19] H. Jonghyun, R. S. Hedda, X. Xing, and W. Woontack, "Adaptive content recommendation for mobile users: ordering recommendations using a hierarchical context model with granularity," *Pervasive and Mobile Computing*, vol. 13, pp. 85–98, 2014.

[20] B. N. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Proceedings of the Workshop on Mobile Computing Systems and Applications (WMCSA '94)*, pp. 89–101, 1994.

[21] A. K. Dey, "Understanding and using context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, 2001.

[22] T. Strang, C. Linllhof, and K. Frank, "CoOL: a context ontology language to enable contextual interoperability," in *Proceedings of the 4th IFIP International Conference on Distributed Applications and Interoperable Systems*, pp. 236–247, Paris, France, 2008.

[23] M. Kaminskas and F. Ricci, "Contextual music information retrieval and recommendation: state of the art and challenges," *Computer Science Review*, vol. 6, no. 2-3, pp. 89–119, 2012.

[24] G. Adomavicius and A. Tuzhilin, "Context-aware recommender system," in *Recommender Systems Handbook*, pp. 217–253, Springer, 2011.

[25] Y. Shi, L. Martha, and H. Alan, "Mining mood-specific movie similarity with matrix factorization for context-aware recommendation," in *Proceedings of the Workshop on Context-Aware Movie Recommendation*, pp. 34–40, ACM, 2010.

[26] S. Rendle, Z. Gantner, C. Freudenthaler, and L. Schmidt-Thieme, "Fast context-aware recommendations with factorization machines," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*, pp. 635–644, New York, NY, USA, July 2011.

[27] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[28] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition: toward a new foundation for human-computer interaction research," *ACM Transactions on Computer-Human Interaction*, vol. 7, no. 20, pp. 174–196, 2000.

[29] D.-S. Liu and S.-J. Fan, "Tourist behavior pattern mining model based on context," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 108062, 12 pages, 2013.

[30] D. Y. Pavlov and D. M. Pennock, "A maximum entropy approach to collaborative filtering in dynamic, sparse, high-dimensional domains," in *Proceedings of the 16th Annual Neural Information Processing Systems Conference (NIPS '02)*, December 2002.

[31] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin, "Incorporating contextual information in recommender systems using a multidimensional approach," *ACM Transactions on Information Systems*, vol. 23, no. 1, pp. 103–145, 2005.