Mining Personal Context-Aware Preferences for Mobile Users

Hengshu Zhu[1,2] Kuifei Yu[2] Happia Cao[2] Enhong Chen[1] Hui Xiong[3] Jilei Tian[2]
[1]*University of Science and Technology of China* [2]*Nokia Research Center* [3]*Rutgers University*
[1]{*zhs, cheneh*}*@ustc.edu.cn* [2]{*kuifei.yu, happia.cao, jilei.tian*}*@nokia.com* [3]*hxiong@rutgers.edu*

***Abstract*—Recent advances in mobile devices and their sensing capabilities have enabled the collection of rich contextual information and mobile device usage records through the device logs. These context-rich logs open a venue for mining the personal preferences of users under varying contexts and thus enabling the development of personalized context-aware recommender systems. In this paper, we illustrate how to extract personal context-aware preferences from the context-rich device logs, or *context logs* for short. A critical challenge along this line is that the context log of each individual user may not contain sufficient data for mining his/her context-aware preferences. Therefore, we propose to first learn common context-aware preferences from the context logs of many users. Then, the preference of each user can be represented as a distribution of these common context-aware preferences. Specifically, we develop two approaches for mining common context-aware preferences based on two different assumptions, namely, context independent and context dependent assumptions, which can fit into different application scenarios. Finally, extensive experiments on a real-world data set show that both approaches are effective and outperform baselines with respect to mining personal context-aware preferences for mobile users.**

I. Introduction

Recent years have witnessed the rapid growth of smart mobile devices, such as smart phones and pads. These devices are usually equipped with some context sensors, such as GPS sensors, 3D accelerometers, and optical sensors, which enable them to capture the rich contextual information of mobile users and thus produce a wide range of context-aware services, such as context-aware tour guide [22], location based reminder [20], and context-aware recommendation [3], [11], [13], [23]. In fact, the contextual information and corresponding usage records (e.g., browsing web sites and playing games) can be recorded into context-rich device logs, or *context logs* for short, which can be used for mining the personal context-aware preferences of users, i.e., which category of contents is preferred by a particular user under a certain context. By considering both the context-aware preferences and the current contexts of users, a personalized context-aware recommender system can be built. Indeed, the following example illustrates how personalized context-aware recommender systems can provide better user experiences than traditional context-aware recommender systems which only considers the contextual information but not different users' preferences under the same context.

Example 1 (A motivating example): Suppose that two users, Joy and Kate, would like to play their mobile phones while taking the bus. Through the analysis of the sensing information collected on their smart phones, a context-aware recommender system could discover that these two users are taking a bus (sensed by 3D accelerometers [14]) from the work place to home (sensed by GPS or cell ID combined with data mining on historical trajectories of the users [7], [25]) in a Monday evening (sensed by the system clock) and the light is dim (sensed by optical sensors). Without considering personal context-aware preferences, the recommender system may recommend the same category of contents (e.g., soft music, R&B music, action games, etc) according to most users' preferences under the same context. In contrast, a personalized context-aware recommender system will also consider the personal context-aware preferences of the two users, e.g., Joy often plays action games and Kate often listens to R&B music under the same context, from their historical context-rich logs, and thus recommend some pop action games to Joy and new R&B music to Kate.

In recent years, although many researchers studied the problem of personalized context-aware recommendation [24], [17], [8], [16], [11] and proposed some approaches for mining personal context-aware preferences, most of them did not take into account context-rich information in their approaches. Also, some of these studies are based on item ratings generated by users under different contexts, which are difficult to obtain in practice. In contrast, usage records in context-rich device logs are a rich resource for mining personalized context-aware user preferences. However, it is still under-explored about how to mine context-aware preferences from context-rich logs for developing context-aware recommender systems.

To this end, in this paper, we propose a novel approach for mining personal context-aware preferences from context-rich device logs of mobile users. A critical challenge for mining personal context-aware preferences is that the context log of each individual user usually does not contain sufficient training information. As a result, it can be difficult to learn personal context-aware preferences if we only use the context log of each individual user. Therefore, we propose to first find common context-aware preferences from the context logs of many users and then represent the context-aware preference of each user as a distribution of common context-aware preferences. Moreover, on the basis of two different assumptions about context data dependency, we propose two methods for mining common context-aware preferences. The first one is more efficient but sacrifices a little performance, while the second one needs more training

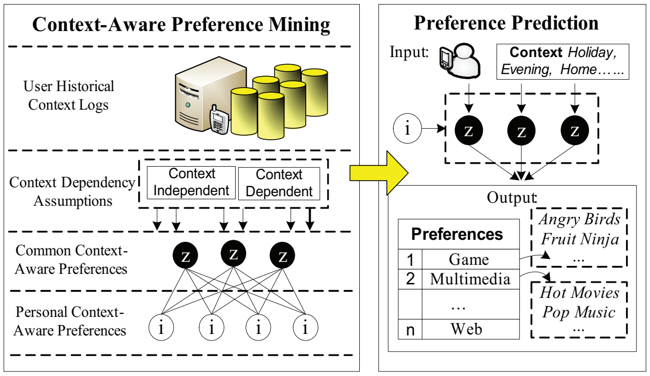


Figure 1. The overview of the proposed approach for mining personal context-aware preferences and how the mined preferences are used for predicting the preferred categories of contents for a given mobile user under a certain context.

time but has better performances. Figure 1 illustrates the overview of the proposed approach and how the mined preferences are used for predicting the preferred categories of contents for a given mobile user under a certain context. The contributions of this paper are summarized as follows.

First, we propose a novel approach for mining the personal context-aware preferences for mobile users through the analysis of context-rich device logs. Specifically, we propose to first mine common context-aware preferences from the context logs of many users and then represent the personal context-aware preference of each user as a distribution of common context-aware preferences. The mined personal context-aware preferences can enable the development of personalized context-aware recommender systems.

Second, we design two effective methods for mining common context-aware preferences based on two different assumptions about context data dependency. If context data are assumed to be conditionally independent, we propose to mine common context-aware preferences through topic models. Otherwise, if context data are assumed to be dependent, we propose to exploit the constraint based Matrix Factorization techniques for mining common context-aware preferences and only consider those contexts which are relevant to content usage for reducing the computation complexity.

Finally, we evaluate the proposed approach using a real-world data set with context logs collected from 443 mobile phone users. In total, there are more than 8.8 million context records. This data set contains much more context-rich information and is much bigger than those reported in previous works on context log mining [4], [6]. The experimental results clearly demonstrate the effectiveness of the proposed approach and indicate some inspiring findings.

II. Mining Personal Context-Aware Preferences From Context Logs

Smart devices can capture the historical context data and the corresponding usage records of users through multiple sensors and record them in context logs. For example, Table I shows a toy context log, which contains several *context records*, and each context record consists of a timestamp, the most detailed context at that time, and the corresponding usage record. A context consists of several *contextual features* (e.g., Day name, Time range, and Location) and their corresponding values (e.g., Saturday, AM8:00-9:00, and Home), which can be annotated as *contextual feature-value pairs*. Moreover, usage records can be empty (denoted as "Null") because a user do not always use the mobile phone.

Note that, in Table I, raw locations in context data, such as GPS coordinates or cell IDs, have been transformed into semantic locations such as "Home" and "Work Place" by some location mining approaches (e.g., [7]). The basic idea of these approaches is to find the clusters of user locations and recognize their semantic meaning by a time pattern analysis. Moreover, we also map the raw usage records to the usage records of particular categories of contents. For example, we can map two raw usage records "*Play Angry Birds*" and "*Play Fruit Ninja*" to the usage records of content category"*Action games*". In this way, the context data and usage records in context logs are normalized and the data sparseness problem is somewhat alleviated. The above helps the task of personal context-aware preference mining.

Intuitively, context logs contain rich information about content usage given particular contexts and can be used for mining the personal context-aware preferences of users. However, the context log of each individual user is usually too sparse for this task. This is also demonstrated by the experiments on a real-world data set in the experimental section. The main reason is that, while the context logs of individual users may contain many context records, only a small proportion of them have non-empty usage records which can be used as meaningful mining source. To that end, we propose a novel approach for mining personal context-aware preferences as follows.

The basic idea is first mining common context-aware preferences from the context logs of many users and then represent each user's context-aware preference by a distribution of common context-aware preferences. Let us denote the variable of common context-aware preference as z, the conditional probability that a user u prefers the content category c given a context C can be represented as

$$P(c|C,u) = \frac{P(c,C|u)P(u)}{P(C,u)} \propto P(c,C|u)$$
$$\propto \sum_z P(c,C,z|u) \propto \sum_z P(c,C|z)P(z|u),$$

where we assume a user's preference given a context only relies on the common context-aware preferences followed by many users, i.e., $P(c,C|z)$, and his (her) personal context-aware preference expressed by a distribution of common context-aware preferences, i.e., $P(z|u)$. Then the task is

Table I
A TOY CONTEXT LOG.

Timestamp	Context	Usage records
t_1	{(Day name: Monday),(Time range: AM8:00-9:00),(Profile: General),(Battery: 5),(Location: Home)}	Null
t_2	{(Day name: Monday),(Time range: AM8:00-9:00),(Profile: General),(Battery: 5),(Location: On the way)}	Play action games
t_3	{(Day name: Monday),(Time range: AM8:00-9:00),(Profile: General),(Battery: 5),(Location: On the way)}	Null
	......	
t_{359}	{(Day name: Monday),(Time range: AM10:00-11:00),(Profile: Meeting),(Battery: 4),(Location: Work Place)}	Null
t_{360}	{(Day name: Monday),(Time range: AM10:00-11:00),(Profile: Meeting),(Battery: 4),(Location: Work Place)}	Browse sports web sites
	......	
t_{448}	{(Day name: Monday),(Time range: AM11:00-12:00),(Profile: General),(Battery: 4),(Location: Work Place)}	Play with SNS
t_{449}	{(Day name: Monday),(Time range: AM11:00-12:00),(Profile: General),(Battery: 4),(Location: Work Place)}	Null

converted to learning $P(c, C|z)$ and $P(z|u)$ from many users' context logs. Specifically, given a user u and context C, both $P(u)$ and $P(C, u)$ are constant, thus we have $P(c|C, u) \propto P(c, C|u)$ in the above equation.

After mining the personal context-aware preference of each mobile user, we predict which category of contents will be preferred for a given user according to the corresponding context. Specially, we first rank content categories according to the probability $P(c|C, u)$ of each content category c, then we infer user preferred content category c^* by $c^* = \arg\max_c P(c|C, u)$ and recommend corresponding contents. For example, if we infer the user would like "action games", we will recommend some popular action games to the user.

We observe that modeling and mining common context-aware preferences rely on the assumption about context data dependency. Basically, we can have two different assumptions about context data dependency as follows. The first assumption is that different types of context data are conditionally independent given a particular common context-aware preference, which is relatively strong but simplifies the problem. For example, under such an assumption, given a context "{(*Time range: PM10:00-11:00*), (*Location: Home*)}" and a user u, if we can infer the latent common context-aware preference distribution of u, we only need to consider which content category u may prefer under the context (*Time range: PM10:00-11:00*) and the context (*Location: Home*) given each common context-aware preference, but not need to consider which content category u may prefer given the co-occurrence of (*Time range: PM10:00-11:00*) and (*Location: Home*) given each common context-aware preference.

The second assumption is that different types of context data are mutually dependent, which is relatively weak and may be more proper in practice. However, such an assumption makes it more difficult for modeling context-aware preferences. For example, under such an assumption, given the above context, we have to consider the co-occurrence of (*Time range: PM10:00-11:00*) and (*Location: Home*) when making a preference prediction. Obviously, the corresponding models may be more complex than the ones based on the first assumption. In this paper, we propose two approaches based on the above two assumptions and conduct extensive experiments to evaluate them. The details of two approaches are presented in the following two sections, respectively.

III. Context-Aware Preference Mining based on Context Conditional Independency Assumption

We first propose a method based on the assumption that different types of context data are conditional independent given a particular common context-aware preference. Under such an assumption, given a context $C = \{p\}$ where p denotes an *atomic context*, i.e., a contextual feature-value pair, the probability that a user u prefers content category c can be represented as

$$P(c|C, u) \propto \sum_z P(c, C|z)P(z|u)$$
$$\propto \sum_z \prod_{p \in C} P(c, p|z)P(z|u)$$

Therefore, the problem is further converted to learn $P(c, p|z)$ and $P(z|u)$ from many users' context logs, which can be solved by widely used topic models. In this section, we present how to utilize topic models for mining common context-aware preferences by estimating $P(c, p|z)$ and $P(z|u)$. For simplicity, we refer to the co-occurrence of a usage of a content in category c and the corresponding contextual feature-value pair p, i.e., (c, p), as an *Atomic Context-aware Preference feature*, and *ACP-feature* for short.

A. Mining Common Context-Aware Preferences through Topic Models

Topic models are generative models that are successfully used for document modeling. They assume that there exist several topics for a corpus D and a document d_i in D can be taken as a bag of words $\{w_{i,j}\}$ which are generated by these topics. Intuitively, if we take ACP-features as words, take context logs as bags of ACP-features to correspond documents, and take common context-aware preferences as topics, we can take advantage of topic models to learn common context-aware preferences from many users' context logs.

Since raw context logs are not naturally in the form of bags of ACP-features, we need to extract bags of ACP-features from them as training data. Specially, we first remove all context records without any usage record and then extract ACP-feature from the remaining ones. Given a context record $< Tid, C, c >$ where Tid denotes a timestamp, $C = \{p_1, p_2, ..., p_l\}$ denotes a context and c denotes the category of the used content in the usage record, we can extract l ACP-features, namely, (c, p_1),(c, p_2),...,(c, p_l). For

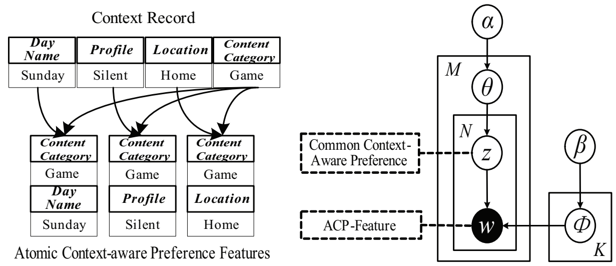


Figure 2. The graphic representation of modeling ACP-feature bags by LDA.

simplicity, we refer the bag of ACP-features extracted from user u's context log as the *ACP-feature bag* of u.

Among several existing topic models, in this paper, we leverage the widely used Latent Dirichlet Allocation model (LDA) [5]. According to LDA, the ACP-feature bag of user u_i denoted as d_i is generated as follows. First, before generating any ACP-feature bag, K prior ACP-feature conditional distributions given context-aware preferences $\{\phi_z\}$ are generated from a prior Dirichlet distribution β. Secondly, a prior common context-aware preference distribution θ_i is generated from a prior Dirichlet distribution α for each user u_i. Then, for generating the j-th ACP-feature in d_i denoted as $w_{i,j}$, the model firstly generates a common context-aware preference z from θ_i and then generates $w_{i,j}$ from ϕ_z. Figure 2 shows the graphic representation of modeling ACP-feature bags by LDA.

The process of LDA model training is to learn the proper latent variables θ and ϕ to maximize the posterior distribution of the observed ACP-feature bags, i.e., $P(u|\alpha, \beta, \theta, \phi)$. In this paper, we take advantage of a Markov chain Monte Carlo method named Gibbs sampling [9] for training LDA models. This method begins with a random assignment of common context-aware preferences to ACP-features for initializing the state of Markov chain. In each of the following iterations, the method will re-estimate the conditional probability of assigning a common context-aware preference to each ACP-feature, which is conditional on the assignment of all other ACP-features. Then a new assignment of common context-aware preferences to ACP-features according to those latest calculated conditional probabilities will be scored as a new state of Markov chain. Finally, after rounds of iterations, the assignment will converge, which means each ACP-feature is assigned a stable and final common context-aware preference and we can obtain the estimation of $P(c, p|z)$ and $P(z|u)$.

B. Selecting the Number of Common Context-Aware Preferences

The LDA model needs a predefined parameter K to determine the number of common context-aware preferences. In this paper we utilize the method proposed in [4] to estimate K. To be specific, we first empirically define a topic number range $[K_{min}, K_{max}]$ and then select two groups of ACP-feature bags as the training set S_1 and the test set S_2, respectively. Then we can determine K by the corresponding perplexity [2], [5] of the test set S_2, which is defined as follows.

$$Perplexity(S_2) = Exp\left\{-\frac{\sum_{d_i \in S_2} log\{P(d_i|S_1)\}}{\sum_{d_i \in S_2} N_{d_i}}\right\},$$

where N_{d_i} indicates the number of ACP-features in d_i and $P(d_i|S_1) = \prod_{(c,p)\in d_i} P(c, p|S_1) = \prod_{(c,p)\in d_i} \sum_{j=1}^{K} P(c, p|z_j)P(z_j|S_1)$. Herein, $P(c, p|z_j)$ can be obtained after model training, and $P(z_j|S_1)$ can be estimated by $\frac{n_{d_i,j}+\alpha}{\sum_{k=1}^{K} n_{d_i,k}+\alpha}$, where $n_{d_i,j}$ indicates the number of ACP-features labeled with z_j in d_i.

The perplexity is widely used to index the modeling performance. The smaller the perplexity, the better modeling performance it implies. However, the perplexity may consistently drop with the increase of K in practical data sets. To avoid the over-fitting problem, we cannot only utilize the minimum perplexity as the metric for determining K [2], [5]. A complementary method is to define a decline rate ζ of perplexity and stop seeking better K if the decline rate of perplexity is less than ζ. In our experiments, we set ζ to be 10% according to [4].

IV. Context-Aware Preference Mining based on Context Dependency Assumption

Since it may be relatively strong to assume that different types of context data are conditionally independent, we also propose a method for mining common context-aware preferences based on the assumption that different types of context data are mutually dependent. Under such an assumption, we cannot decompose contexts into atomic contexts and need to learn $P(c, C|z)$ directly from user context logs. A major challenge is that we cannot learn all conditional distributions $P(c, C|z)$ for all C simply because the number of unique C is exponential to the number of unique contextual feature-value pairs and we will suffer the assemble explosion problem if we learn all of them. Fortunately, we observe that usually not all parts of a context are relevant to content usage and thus the corresponding preferences of content categories. For example, given a context "$\{$(*Day name: Sunday*), (*Day period: Evening*), (*Location: Home*), (*Battery level: Low*)$\}$", we may not be able to demonstrate the whole context is relevant to a particular content category through the analysis of context logs. Instead, we may find context logs show that some parts of them such as "$\{$(*Day name: Sunday*), (*Day period: Evening*)" and "$\{$(*Day period: Evening*), (*Location: Home*)$\}$" are indeed relevant to a particular content category. To this end, an intuitive idea is to only consider the content-relevant parts of contexts for predicting personalized context-aware preferences of content categories. These content-relevant parts can be referred to as *content-relevant contexts* for simplicity. Along this line, we only need to learn the conditional distributions $P(c, C^r|z)$

and $P(z|u)$, where C^r denotes a content-relevant context. Moreover, given a context C, it can be divided into two situations to calculate $P(c, C|z)$.

First, if C contains some content-relevant contexts, we can calculate $P(c, C|z)$ directly by its maximal sub-contexts which are also content-relevant contexts as follows.

$$P(c, C|z) = \frac{1}{|C^r_{max}|} \sum_{C^r_{max}} P(c, C^r_{max}|z),$$

where C^r_{max} denotes a maximal content-relevant sub-context contained by C, and $|C^r_{max}|$ indicates of the number of C^r_{max}. Take the above context C={(*Day name: Sunday*), (*Day period: Evening*), (*Location: Home*), (*Battery level: Low*)} for example, suppose that there are two maximum content-relevant sub-contexts C^r_1={(*Day name: Sunday*), (*Day period: Evening*) and C^r_2={(*Day period: Evening*), (*Location: Home*)} in it, we will recommend corresponding contents to the given user by considering his/her preference under both C^r_1 and C^r_2 equally.

Second, if C does not contain any content-relevant context, we can estimate $P(c, C|z)$ by normalizing the probabilistic space of the joint distribution of c and C conditional on z. Especially, we let

$$P(c, C^{\phi}|z) = \frac{1}{N_{\phi} \cdot N_c} \times (1 - \sum_{a} \sum_{C^{\Delta}} P(c, C^{\Delta}|z)),$$

where C^{ϕ} denotes a context without any content-relevant sub-context, N_{ϕ} indicates the total number of C^{ϕ}, N_c indicates the total number of unique content categories, and C^{Δ} denotes a context with at least one content-relevant sub-context. Actually, we do not need to calculate $P(c, C|z)$ in this case because it is the same with varying c and cannot help to make a recommendation decision. In practice, we do not recommend any content in this case.

Therefore, the original problem is divided into two sub-problems, namely, *how to discover those content-relevant contexts?* and *how to learn common context-aware preferences and user personal distributions of common context-aware preferences, i.e., $P(c, C^r|z)$ and $P(z|u)$?* The solutions for the two sub-problems are presented in the following sections in detail, respectively.

A. Discovering the Content-Relevant Context

An intuitive way of discovering the context relevant to some content categories is mining association rules [1] between them with predefined minimum supports and minimum confidences. Therefore, given a content-relevant context C^r and a content category c, $P(c|C^r, u) \propto P(c, C^r|u)$ can be calculated as $P(c, C^r|u) = P(c|C^r, u)P(C^r|u)$, where $P(c|C^r, u)$ can be estimated by the corresponding confidence of the association "$C^r \longrightarrow c$" and $P(C^r|u)$ can be estimated by $\frac{Support(C^r)}{N_r}$, where N_r indicates the total number of context records in the context log of user u.

However, as pointed out by Cao *et al.* [6], the amounts of context data and user usage records are usually extremely unbalanced, which makes it difficult to mine such association rules by traditional association rule mining approaches. To that end, they proposed a novel algorithm called GCPM (Generating Candidates for behavior Pattern Mining) for mining such association rules, which are referred as behavior patterns in their work, by utilizing different ways of calculating supports and confidences.

In this paper, we take advantage of GCPM for mining association rules between contexts and content categories, and then take the contexts which appear in any of such associations as content-relevant contexts. It is worth noting that the mining is performed on individual users' context logs because merging all context logs may normalize the associations between contexts and content categories. For example, given several users who usually play action games in bus and several users usually play other games in bus. If we try to mine the associations between contexts and content category by merging all users' context logs, we may falsely conclude that "In bus" has no significant relevance with any content category. In contrast, we can discover "In bus" is both relevant to action games and other games according to different people by taking into account each user's context log separately.

B. Mining Common Context-Aware Preferences through Constraint based Bayesian Matrix Factorization

After finding content-relevant contexts, the remaining task is to learn common context-aware preferences and user personal distributions of common context-aware preferences, i.e., $P(c, C^r|z)$ and $P(z|u)$. By building a matrix of $P(c, C^r|u)$, where each column denotes a probabilistic distribution of different (c, C^r) pairs for a given user u, we can convert this task into a matrix factorization problem as follows.

$$\mathbf{\Omega}_{N \times M} = \mathbf{\Phi}_{N \times K} \mathbf{\Theta}_{K \times M} + \mathbf{N}_{N \times M},$$

where N indicates the number of unique (c, C^r) pairs, M indicates the number of users and K indicates the number of common context-aware preferences. To be specific, $\mathbf{\Omega}$ denotes the observed matrix of $P(c, C^r|u)$, $\phi_{ik} \in \mathbf{\Phi}(1 \leq i \leq N, 1 \leq k \leq K)$ denotes the probability $P(c, C^r|z_k)$, $\theta_{kj} \in \mathbf{\Theta}(1 \leq k \leq K, 1 \leq j \leq M)$ denotes the probability $P(z_k|u_j)$, and the matrix $\mathbf{N}$ denotes the residual noise information. Moreover, the matrix factorization task has two additional constraints for possible solutions as follows: 1) all elements in matrix $\mathbf{\Phi}$ and $\mathbf{\Theta}$ should be non-negative values, 2) $\forall_{j:1 \leq j \leq M} \sum_{k=1}^{K} \theta_{kj} = 1$ and $\forall_{k:1 \leq k \leq K} \sum_{i=1}^{N} \phi_{ik} = 1$, which are both obvious since each column of $\mathbf{\Phi}$ and $\mathbf{\Theta}$ denotes a probabilistic distribution.

According to the problem statement and constraints above, the objective of our matrix factorization task is to find a possible solution for matrix $\mathbf{\Phi}$, $\mathbf{\Theta}$ and $\mathbf{N}$. In this paper, we propose to leverage a constraint based Bayesian Matrix Factorization model [18] for resolving this problem. In this model, we can perform matrix factorization with multiple inequality and equality constraints. Specifically, we aim to infer the posterior probabilistic distributions of $\mathbf{\Phi}$ and $\mathbf{\Theta}$ under a set of model assumptions, which are specified by the likelihood function $P(\mathbf{\Omega}|\mathbf{\Phi},\mathbf{\Theta},\mathbf{N})$. The likelihood function denotes the probability of the observed data matrix $\mathbf{\Omega}$ given priors $P(\mathbf{\Phi},\mathbf{\Theta})$ and $P(\mathbf{N})$. According to [18], to perform efficient inference based on Gibbs sampling, we select priors as follows. First, we select an i.i.d. zero mean Gaussian noise model as follows.

$$P(n_{ij}) = N(n_{ij}|0,\nu_{ij}) = \frac{1}{\sqrt{2\pi\nu_{ij}}}exp(-\frac{n_{ij}^2}{2\nu_{ij}}),$$

where parameter ν_{ij} satisfies conjugate inverse-gamma prior that $P(\nu_{ij}) = IG(\nu_{ij}|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\nu_{ij}^{-(\alpha+1)}exp(\frac{-\beta}{\nu_{ij}})$. Then, we select a Gaussian prior over $\mathbf{\Phi}$ and $\mathbf{\Theta}$ subject to inequality constraints $\mathbf{Q}$ and equality constraints $\mathbf{R}$ as

$$P(\overrightarrow{\phi},\overrightarrow{\theta}) \propto$$
$$\begin{cases} \mathrm{N}\left(\begin{bmatrix}\overrightarrow{\phi}\\ \overrightarrow{\theta}\end{bmatrix} | \underbrace{\begin{bmatrix}\mu_\phi\\ \mu_\theta\end{bmatrix}}_{\mu}, \underbrace{\begin{bmatrix}\mathbf{\Sigma}_\phi & \mathbf{\Sigma}_{\phi\theta}\\ \mathbf{\Sigma}_{\phi\theta}^{\mathbf{T}} & \mathbf{\Sigma}_s\end{bmatrix}}_{\mathbf{\Sigma}}\right), & if\ \mathbf{Q}(\overrightarrow{\phi},\overrightarrow{\theta}) \le 0, \\ & \mathbf{R}(\overrightarrow{\phi},\overrightarrow{\theta}) = 0, \\ 0, & otherwise, \end{cases}$$

where $\overrightarrow{\phi} = (\phi_{11},\phi_{12},...,\phi_{NK})^{\mathbf{T}}$ and $\overrightarrow{\theta} = (\theta_{11},\theta_{12},...,\theta_{KM})^{\mathbf{T}}$.

With above definitions, we can utilize Gibss sampling methods to estimate the posterior distributions as follows. In the first round of sampling, we randomly assign values for $\overrightarrow{\phi}$ and $\overrightarrow{\theta}$ according to the two constraints to initialize the state of Markov chain. Then, we calculate the density of noise variance $P(\nu_{ij}|\overrightarrow{\phi},\overrightarrow{\theta})$ by inverse-gamma distribution due to the choice of conjugate prior. Next, we can estimate $P(\overrightarrow{\phi}|\mathbf{\Omega},\overrightarrow{\theta},\mathbf{N})$ and $P(\overrightarrow{\theta}|\mathbf{\Omega},\overrightarrow{\phi},\mathbf{N})$ from the constraint Gaussian density. Finally, we re-generate values for $\overrightarrow{\phi}$ and $\overrightarrow{\theta}$ according to the new posterior probabilities to score a new state of Markov chain. After many rounds of iterations, the results of matrixes $\mathbf{\Phi}$ and $\mathbf{\Theta}$ will converge.

C. Selecting the Number of Common Context-Aware Preferences

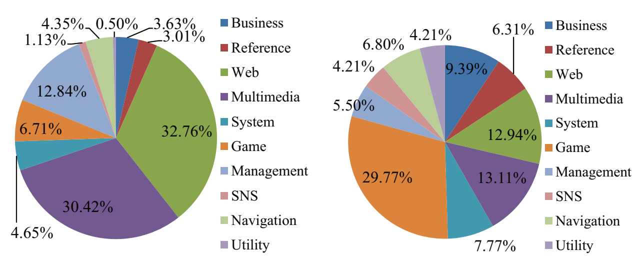
As mentioned in Section 3, an important problem for mining common context-aware preferences is to select a proper number of common context-aware preferences. For the Matrix Factorization based approach for mining common context-aware preferences, we utilize the Chib's method introduced in [19] to infer the proper number of common context-aware preferences. To be specific, in the Bayesian framework, model selection can be performed by evaluating the marginal likelihood $P(\mathbf{\Omega})$. The Chib's method is based on the Bayes relation $P(\mathbf{\Omega}) = \frac{P(\mathbf{\Omega}|\mathbf{\Theta})P(\mathbf{\Theta})}{P(\mathbf{\Theta}|\mathbf{\Omega})}$, where the numerator can be easily estimated by the trained model and the key problem is to estimate $p(\mathbf{\Theta}|\mathbf{\Omega})$. Denoting each row of $\mathbf{\Theta}$ as $\mathbf{\Theta}_i$, we can calculate the denominator as $P(\mathbf{\Theta}|\mathbf{\Omega}) = P(\mathbf{\Theta}_1|\mathbf{\Omega}) \times P(\mathbf{\Theta}_2|\mathbf{\Theta}_1,\mathbf{\Omega}) \times ... \times P(\mathbf{\Theta}_K|\mathbf{\Theta}_1,...,\mathbf{\Theta}_{K-1},\mathbf{\Omega})$. After R rounds of Gibbs sampling in our approach, we can estimate each term by averaging over the conditional density $P(\mathbf{\Theta}_K|\mathbf{\Theta}_1,...,\mathbf{\Theta}_{K-1},\mathbf{\Omega}) \approx \frac{1}{R}\sum_{r=1}^{R} P(\mathbf{\Theta}_i|\mathbf{\Theta}_1,...,\mathbf{\Theta}_{i-1},\mathbf{\Theta}_{i+1}^{(r)},...,\mathbf{\Theta}_K^{(r)},\mathbf{U})$, where $\mathbf{\Theta}_{i+1}^{(r)},...,\mathbf{\Theta}_K^{(r)}$ are Gibbs samples from $P(\mathbf{\Theta}_{i+1},...,\mathbf{\Theta}_K|\mathbf{\Theta}_1,...,\mathbf{\Theta}_{i-1},\mathbf{\Omega})$. Thus, given a range $[K_{min},K_{max}]$ for the number of common context-aware preferences, we can select the $K \in [K_{min},K_{max}]$ which maximizes the likelihood.

Table II
THE TYPES OF CONTEXTUAL INFORMATION IN OUR DATA SET.

Context	Value range
Week	{Monday, Tuesday,... , Sunday}
Is a holiday?	{Yes, No}
Day period	{Morning(7:00-11:00), Noon(11:00-14:00), Afternoon(14:00-18:00), Evening(18:00-21:00), Night(21:00-Next day 7:00)}
Time range	{0:00-1:00, 1:00-2:00, ... , 23:00-24:00}
Profile type	{General, Silent, Meeting, Outdoor, Pager, Offline}
Battery Level	{Level 1, Level 2, ... , Level 7}
Charging State	{Charging, Complete, Not Connected}
Social location	{Home, Work Place, On the way}.

V. EXPERIMENTAL RESULTS

In this section, we evaluate the performances of the two implementations of the proposed approach for predicting user preferences of content categories, namely, CIAP (**C**ontext conditional **I**ndependency **A**ssumption based **P**rediction) and CDAP (**C**ontext **D**ependency **A**ssumption based **P**rediction), with several baseline methods on a real-world data set.

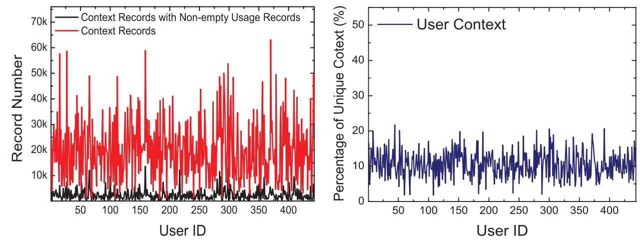


(a) (b)

Figure 3. (a) The distribution of 618 unique contents w.r.t. the corresponding content categories and (b) the distribution of context records w.r.t. the content categories their corresponding usage records belong to.

A. Experimental Data

The data set used in the experiments is collected from many volunteers by a major manufacturer of smart mobile devices. The data set consists of 8,852,187 context records which contain rich contextual information and usage records of 443 smart phone users spanning for from several weeks to several months. Table II shows the concrete types of context data the data set contains. In the experiments, we manually classified the 665 unique contents appearing in raw usage records into 12 content categories based on the



(a) (b)

Figure 4. (a) The distributions of all context records and the context records with non-empty usage records for all users. (b) The coverage of unique contexts in each user's context log compared with all unique contexts.

taxonomy of Ovi store (www.ovi.com), which are *Call*, *Web*, *Multimedia*, *Management*, *Game*, *System*, *Navigation*, *Business*, *Reference*, *Social Network Service (SNS)*, *Utility* and *Others*. In our experiments, we do not utilize the categories *Call* and *others* because they is not useful for generating corresponding recommendations. Instead, we only utilize other 10 content categories which contain 618 unique contents appearing in total 408,299 usage records. Figure 3 (a) and (b) show the distribution of 618 unique contents in raw usage records with respect to the corresponding content categories and the distribution of context records with respect to the content categories their corresponding usage records belong to, respectively. The context records with empty usage records are not taken into account.

Figure 4 (a) compares the distributions of all context records and the context records with non-empty usage records for all users. From the figure we can see that usually though many context records of individual mobile users are collected, only a small proportion of them have non-empty usage records and can be used as training data, which implies the limit of mining personal context-aware preferences only from individual users' context logs. Moreover, Figure 4 (b) shows the coverage ratio of unique contexts in each user's context log compared with all unique contexts, from which we can see that the unique contexts in each individual user's context log is relatively limited, which motivates learning from many users' context logs as well.

B. Benchmark Methods

First, we select a state-of-the-arts personalized context-aware recommendation approach based on individual users' context logs as a baseline.

CASVM stands for personalized **C**ontext-**A**ware preference prediction by Ranking **SVM**, which is introduced in [12]. To be specific, in this approach, given a user's u and a context C, we calculate five types of features $P(c)$, $P(c|u)$,$P(c|p)$,$P(c|u,p)$ and $P(c|u,p_1,...,p_n)$, where p denotes a contextual feature-value pair in C, according to the context log of u and then leverage Ranking SVM for ranking content category c.

Moreover, to validate the performance of leveraging many users' context logs for mining personal context-aware preferences, we also select two state-of-the-arts collaborative filtering (CF) based approaches as baselines.

CACF stands for **C**ontext-**A**ware preference mining by disjunction **CF**, which is a memory-based CF approach introduced in [15]. In this approach, the preference score $s_{c,u,C}$ of the user u on content category c in the context C is calculated by the disjunctive aggregation of the estimated preferences of $s_{c,u,p}$, where p is a feature-value pair in C and $s_{c,u,p}$ is measured by counting how many context records of u contain p and contents in category c.

CATF stands for **C**ontext-**A**ware preference mining by **T**ensor **F**actorization, which is a model-based CF approach introduced in [13]. In this approach, all users' context logs can be represented by a high dimensional tensor T, and the objective is to complete the missing values for ranking content categories. Specifically, each value $t_{u,c,v_1,...,v_n}$ in original tensor is measured by counting how many context records of user u contain contents in category c and feature value pairs $\{p = (f_i, v_i)\}$.

C. Evaluation Metrics

In the experiments, we utilize a five-fold cross validation to evaluate each test approach. To be specific, we first randomly divide each user's context log into five equal parts, and then use each part as the test data while using other four parts as the training data in five test rounds. Finally, we report the average performance of the five runs. In the test process, we only take into account the context records with non-empty usage records, and use the contexts and the category of the content indicated by the usage record as context inputs and ground truth, respectively. Moreover, to evaluate the ranking of content categories generated by each approach, we leverage two metrics as follows.

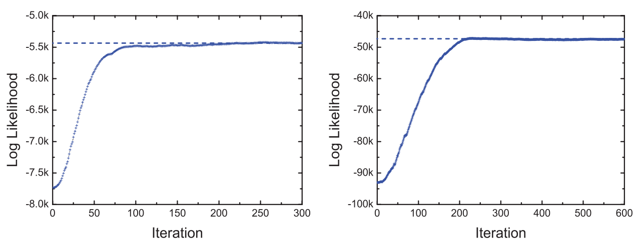
MAP@K stands for Mean Average Precision at top K recommendation results. To be specific, $MAP@K = \frac{\sum AP^{(u)}@K}{|U|}$, where $AP^{(u)}@K$ denotes the average precision at top k prediction results on the test cases of user u, and $|U|$ indicates the number of test users. $AP^{(u)}@K$ can be computed by $\frac{1}{N_u}\sum_i \sum_{r=1}^{K}(P_i(r) \times rel_i(r))$, where N_u denotes the number of test cases for user u, r denotes a given cut-off rank, $P_i(r)$ denotes the precision on the i-th test case of u at a given cut-off rank r, and $rel_i()$ is the binary function on the relevance of a given rank.

MAR@K stands for Mean Average Recall at top K prediction results. To be specific, $MAR@K = \frac{\sum AR^{(u)}@K}{|U|}$, where $AR^{(u)}@K$ denotes the average recall at top k prediction results on the test cases of user u, and $|U|$ indicates the number of test users. $AR^{(u)}@K$ can be computed by $\frac{1}{N_u}\sum_i \sum_{r=1}^{K} rel_i(r)$.

D. Overall Results

According to the parameter estimation approaches introduced in Section III and Section IV, the number of common context-aware preferences for both LDA and NMF training denoted as K is empirically set to be 15. For the LDA training, the two parameters α and β are empirically set to be $50/K$ and 0.2 according to [10]. For the Bayesian Matrix Factorization training, according to [18], we use an isotropic noise model and choose a decoupled prior for $\mathbf{\Phi}$ and $\mathbf{\Theta}$ with zero mean $\mu = 0$, and an unit diagonal covariance matrix $\mathbf{\Sigma} = \mathbf{I}$. The maximum iterations of Gibbs sampling are set to be 2000 in our experiments. Moreover, the behavior patterns are mined by GCPM algorithms introduced in [6]. Both our approaches and the baselines are implemented by C++ and the experiments are conducted on a 3GHZ×4 quad-core CPU, 3G main memory PC.

Figure 5 (a) and (b) show the convergence curves of Gibbs sampling for our two implementations of the proposed approach by measuring their log likelihood for the training data set in one of the five test rounds. From the figure we can see that the Gibbs sampling of both implementations converge quickly. The convergence curves for other test rounds follow the similar trend. Moreover, each iteration of Gibbs sampling averagely costs 89 milliseconds for CIAP and 423 milliseconds for CDAP, respectively. It is because that the NMF training is more complex than LDA and the number of associations between context and content category for matrix factorization is greater than the ACP-features in LDA model.



(a) (b)

Figure 5. The log likelihood of training sets in each iteration of Gibbs sampling for (a) CIAP and (b) CDAP.

It is worth noting that both our approach and baselines except CATF may be not able to generate preference predictions for all contexts. To be specific, CIAP, CASVM and CACF can only generate preference predictions for the contexts which contain the ACP-features appeared in training sets, CDAP can only generate preference predictions for the contexts which contain content-relevant sub-contexts. However, in our test data sets, we observe that CIAP, CDAP, CASVM and CACF can all cover 100% test contexts.

We first test the $MAP@K$ performance of each test approach with respect to varying K and the average results in the five-fold cross validation are shown in Figure 6. From the results we can observe that CDAP and CIAP consistently outperforms other baselines with varying K. We also find that two CF based approaches CACF and CATF outperform CASVM, which indicates leveraging many users' context logs other than individual users' context logs can improve the recommendation performance. Moreover, we can see that CDAP outperforms CIAP slightly with varying K.

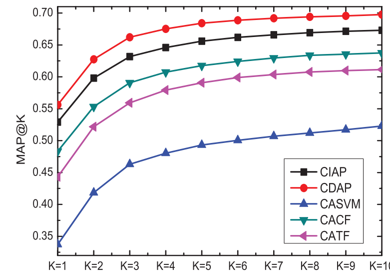
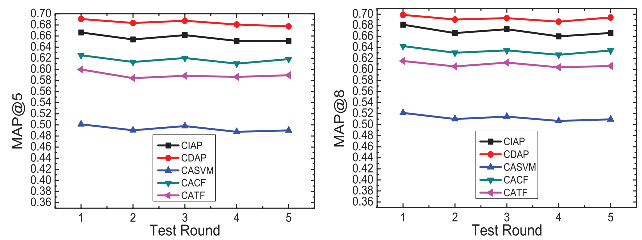


Figure 6. The average $MAP@K$ of each prediction approach in the five-fold cross validation.



(a) (b)

Figure 7. The $MAP@K$ of each prediction approach when (a) $K = 5$, (b) $K = 8$ in each test round.

Figure 7 (a) and (b) further show the $MAP@K$ of each recommendation approach in each test round when $K = 5$ and $K = 8$, respectively. From the results we can observe that our approaches consistently outperform other baselines in all five test rounds.

Figure 8 shows the average $MAR@K$ of each approach in five-fold cross validation. From the results, we can observe that CIAP and CDAP outperform other baselines and CDAP outperforms CIAP slightly in terms of $MAR@K$. Figure 9 (a) and (b) further show the $MAR@5$ and $MAR@8$ of each recommendation approach in each test round. The results also validate the effectiveness of CIAP and CDAP in terms of $MAR@K$. Specially, we conduct a series of paired T-tests with 0.95 confidence level in each K. The test results show that the improvements of CIAP and CDAP on $MAP@K$ and $MAR@K$ compared with other baselines are statistically significant.

From the above experimental results we can clearly see that both CDAP and CIAP outperform other baselines under different metrics and experimental settings, which demonstrates the effectiveness of our framework for personalized context-aware recommendation. Moreover, CDAP outperforms CIAP slightly though its training cost is much higher than the latter.

E. Robustness Analysis

Both CIAP and CDAP need a parameter to determine the number of common context-aware preferences. Figure 10 (a) and (b) show the $MAP@5$ and $MAP@10$ of CIAP and CDAP with respect to varying settings of the number. From these figures we can observe that both $MAP@5$ and $MAP@10$ of CDAP are relatively not sensitive to the parameter. In contrast, the robustness of CIAP is not good

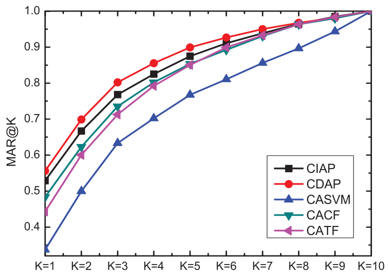
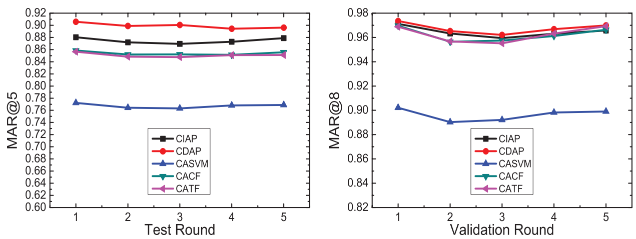


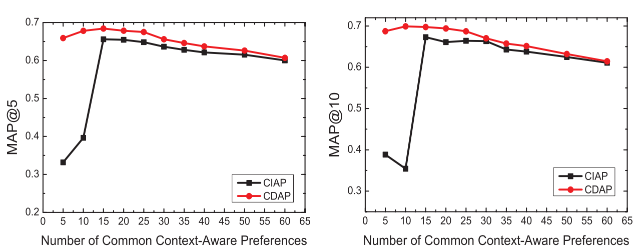
Figure 8. The average $MAR@K$ of each prediction approach in the five-fold cross validation.



(a) (b)

Figure 9. The $MAR@K$ of each prediction approach when (a) $K = 5$, (b) $K = 8$ in each test round.

with small numbers of common context-aware preferences but becomes stable when the setting of the number increases. It may be because that CDAP leverages associations between contexts and user content categories for extracting common context-aware preferences and such associations have been filtered from noisy data. Thus, the quality of mined common context-aware preferences is always relatively good with different parameters since the mining are on the basis of pruned training data. In contrast, CIAP leverages ACP-features for extracting common context-aware preferences, where ACP-features usually contain more noisy information and thus make the mining results more sensitive to parameters.



(a) (b)

Figure 10. The (a) $MAP@5$ and (b) $MAP@10$ of CIAP and CDAP with respect to varying number of common context-aware preferences.

F. Case Study

In addition to the studies on the overall performance of our approach, we also study the cases in which CIAP and CDAP outperform the baselines.

For example, Table III shows an example of the top 3 predicted preferences of each approach for two different users given the same context {(*Time range: 21:00-22:00*), (*Profile: Silent*), (*Is holiday: Yes*), (*Day name: Sunday*), (*Day period: Night*), (*Location: Home*)}, which may imply the leisure time at home. Under that context, the ground truth preferred content categories of user #162 and user #423 are *Game* and *Web*, respectively. From the results, we can observe that both CIAP and CDAP can predict correct content categories for user #162 and user #423 in the top one position. In contrast, CASVM, CACF and CATF predicted correct content categories in lower positions for both users.

Table III
PREDICTION RESULTS FOR USER #162 AND #423.

Context	{(*Time range: 21:00-22:00*), (*Profile: Silent*) (*Is holiday: Yes*), (*Day name: Sunday*), (*Day period: Night*), (*Location: Home*)}
Top 3 predicted preferences for user *#162*	
Ground truth	*Game*
CIAP	**Game** ($\surd$), Multimedia, Web
CDAP	**Game** ($\surd$), Web, Multimedia
CASVM	Multimedia, Web, **Game** ($\surd$)
CACF	Multimedia, **Game** ($\surd$), System
CATF	Multimedia, **Game** ($\surd$), Web
Top 3 predicted preferences for user *#423*	
Ground truth	*Web*
CIAP	**Web** ($\surd$), Multimedia, SNS
CDAP	**Web** ($\surd$), Reference, Game
CASVM	Multimedia, **Web**($\surd$), Reference
CACF	Game, **Web** ($\surd$),Multimedia
CATF	Game, **Web** ($\surd$), System

VI. Related Work

Many previous works about personalized context-aware recommendation for mobile users have been reported. For example, Tung *et al.* [21] have proposed a prototype design for building a personalized recommender system to recommend travel related information according to users' contextual information. Park *et al.* [17] proposed a location-based personalized recommender system, which can reflect users' personal preferences by modeling user contextual information through Bayesian Networks. Bader *et al.* [3] have proposed a novel context-aware approach to recommending points-of-interest (POI) for users in an automotive scenario. Specifically, they studied the scenario of recommending gas stations for car drivers by leveraging a Multi-Criteria Decision Making (MCDM) based methods to modeling context and different routes. However, most of these works only leverage individual users' historical context data for modeling personal context-aware preferences, and do not take into account the problem of insufficient personal training data.

Actually, the problem of insufficient personal training data is common in practice and many researchers have studied how to address this problem. For example, Woerndl *et al.* [23] proposed a hybrid framework named "play.tools" for recommending mobile applications by leveraging users' context information. This recommendation framework are based on what other users have installed in similar context will be liked by a given user. Kim *et al.* [11] investigated several Collaborative Filtering (CF) based approaches for recommendation and developed a memory based CF approach to providing context-aware advertisement recommendation. Specially, the proposed approach can leverage a classification rule of decision tree to understand users' personal preference. Zheng *et al.* [24] have studied a model based CF approach to recommending user locations and activities according to users' GPS trajectories. The approach can model user, location and activity as a 3-dimensional matrix, namely tensor, and perform tensor factorization with several constraints to capture users' preferences. Alexandros *et al* [13] proposed a model based CF approach for making recommendation with respect to rich contextual information, namely multiverse recommendation. Specifically, they modeled the rich contextual information with item by N-dimensional tensor, and proposed a novel algorithm to make tensor factorization. In a word, most of these approaches are based on rating logs of mobile users and the objective is to predict accurate ratings for the unobserved items under different contexts. However, we usually cannot obtain such rating data in user mobile devices. In contrast, it is relatively easier to collect context logs which contain the users' historical context information and their usage records, which can be used for mining context-aware user preferences.

Recently, some researchers studied how to mine event logs for personalized context-aware recommendation. For example, Lee *et al* [15] investigated to convert the event logs into implicit ratings and tested various memory based CF approaches to make personalized context-aware recommendation. Kahng *et al* [12] proposed a novel approach for ranking item in event logs for personalized context-aware recommendation. Compared with context logs, event logs only record the context records with non-empty usage records and lose thus some discriminative information to capture the relevance between contexts and content categories [6]. But these event log based approaches can still be used for mining context logs towards personalized context-aware recommendation. Compared with these works, our framework explicitly models common context-aware preferences and represent individual users' personal context-aware preferences by distributions of common context-aware preferences. Moreover, one proposed approach in our framework (CDAP) utilizes the contexts with only empty usage records.

VII. Concluding Remarks

In this paper, we proposed to exploit user context logs for mining the personal context-aware preferences of mobile users. First, we identified common context-aware preferences from the context logs of many users. Then, the personal context-aware preference of an individual user can be represented as a distribution of common context-aware preferences. Moreover, we designed two methods for mining common context-aware preferences based on two different assumptions about context data dependency. Finally, the experimental results on a real-world data set clearly showed that the proposed approach could achieve better performances than benchmark methods for mining personal context-aware preferences, and the one implementation based on the independent assumption of context data slightly outperforms another one but has higher computational cost.

References

[1] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. In *VLDB'94*, pages 487–499, 1994.

[2] L. Azzopardi, M. Girolami, and K. V. Risjbergen. Investigating the relationship between language model perplexity and ir precision-recall measures. *In SIGIR'03*, pages 369–370, 2003.

[3] R. Bader, E. Neufeld, W. Woerndl, and V. Prinz. Context-aware poi recommendations in an automotive scenario using multi-criteria decision making methods. In *CaRR '11*, pages 23–30, 2011.

[4] T. Bao, H. Cao, E. Chen, J. Tian, and H. Xiong. An unsupervised approach to modeling personalized contexts of mobile users. In *ICDM'10*, pages 38–47, 2010.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Lantent dirichlet allocation. In *Journal of Machine Learning Research*, pages 993–1022, 2003.

[6] H. Cao, T. Bao, Q. Yang, E. Chen, and J. Tian. An effective approach for mining mobile user habits. In *CIKM'10*, pages 1677–1680, 2010.

[7] N. Eagle, A. Clauset, and J. A. Quinn. Location segmentation, inference and prediction for anticipatory computing. In *AAAI Spring Symposium on Technosocial Predictive Analytics*, 2009.

[8] Y. Ge, Q. Liu, H. Xiong, A. Tuzhilin, and J. Chen. Cost-aware travel tour recommendation. In *KDD'11*, 2011.

[9] T. L. Griffiths and M. Steyvers. Finding scientific topics. *In Proceedings of National Academy of Science of the USA*, pages 5228–5235, 2004.

[10] G. Heinrich. Paramter stimaion for text analysis. *Technical report, University of Lipzig*, 2009.

[11] K. jae Kim, H. Ahn, and S. Jeong. Context-aware recommender systems using data mining techniques. *In Journal of World Academy of Science, Engineering and Technology*, 64:357–362, 2010.

[12] M. Kahng, S. Lee, and S.-g. Lee. Ranking in context-aware recommender systems. In *WWW'11*, pages 65–66, 2011.

[13] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *RecSys '10*, pages 79–86, 2010.

[14] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl.*, 12:74–82, 2011.

[15] D. Lee, S. Park, M. Kahng, S. Lee, and S.-g. Lee. Exploiting contextual information from event logs for personalized recommendation. In *ICIS'10*, pages 121–139, 2010.

[16] Q. Liu, Y. Ge, Z. Li, E. Chen, and H. Xiong. Personalized travel package recommendation. In *ICDM'11*, 2011.

[17] M.-H. Park, J.-H. Hong, and S.-B. Cho. Location-based recommendation system using bayesian user's preference model in mobile devices. In *UIC'07*, pages 1130–1139, 2007.

[18] M. N. Schmidt. Linearly constrained bayesian matrix factorization for blind source separation. In *NIPS'09*, pages 1624–1632. 2009.

[19] M. N. Schmidt, O. Winther, and L. K. Hansen. Bayesian non-negative matrix factorization. In *International Conference on Independent Component Analysis and Signal Separation*, pages 540–547, 2009.

[20] T. Sohn, K. A. Li, G. Lee, I. Smith, J. Scott, and W. G. Griswold. Place-its: A study of location-based reminders on mobile phones. In *Ubicomp'05*, pages 232–250, 2005.

[21] H.-W. Tung and V.-W. Soo. A personalized restaurant recommender agent for mobile e-service. *In EEE'04*, pages 259–262, 2004.

[22] M. van Setten, S. Pokraev, and J. Koolwaaij. Context-aware recommendations in the mobile tourist application compass. In *AH'2004*, pages 235–244, 2004.

[23] W. Woerndl, C. Schueller, and R. Wojtech. A hybrid recommender system for context-aware recommendations of mobile applications. In *ICDE'07*, pages 871–878, 2007.

[24] V. W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang. Collaborative filtering meets mobile recommendation: A user-centered approach. In *AAAI'10*, pages 236–241, 2010.

[25] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW'09*, pages 791–800, 2009.