

Mining Personal Context-Aware Preferences for Mobile Users

Hengshu Zhu^{*†} Enhong Chen^{*} Kuifei Yu[†] Huanhuan Cao[†] Hui Xiong[‡] Jilei Tian[†]

^{*}University of Science and Technology of China [†]Nokia Research Center [‡]Rutgers University

^{*}{zhs, cheneh}@ustc.edu.cn [†]{kuifei.yu, happia.cao, jilei.tian}@nokia.com [‡]hxiong@rutgers.edu

Abstract—In this paper, we illustrate how to extract personal context-aware preferences from the context-rich device logs (i.e., *context logs*) for building novel personalized context-aware recommender systems. 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.

Keywords—Personal Context-Aware Preferences, Context-Aware Recommendation, Mobile Users.

I. INTRODUCTION

Recent years have witnessed the rapid growth of smart mobile devices. These devices are usually equipped with some context sensors, such as GPS sensors, and 3D accelerometers which enable them to capture the rich contextual information of mobile users and thus produce a wide range of context-aware services. 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. By considering both the context-aware preferences and the current contexts of users, a personalized context-aware recommender system can be built. Indeed, the personalized context-aware recommender systems can provide better user experiences than traditional context-aware recommender systems which only takes consideration of the contextual information but not different users' preferences under the same context.

In recent years, although many researchers studied the problem of personalized context-aware recommendation [16], [15], [5], [12], [8] 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 to mining personal context-aware preferences from context 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. Therefore, we propose a novel crowd wisdom based approach for mining the personal context-aware preferences for mobile users, which can enable the building of personalized context-aware mobile recommender systems.

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.

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 model 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. 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. Specifically, it 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, 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., [4]). 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. 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 $P(c|C, u)$, which satisfies $P(c|C, u) \propto P(c, C|u)$, thus we have

$$P(c|C, u) \propto \sum_z P(c, C, z|u) \propto \sum_z P(c, C|z)P(z|u),$$

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 converted to learning $P(c, C|z)$ and $P(z|u)$ from many users’ context logs.

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 can 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.

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

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

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 leverage 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 $\langle Tid, C, c \rangle$ where Tid denotes a timestamp, $C = \{p_1, p_2, \dots, p_l\}$ denotes a context and c denotes the category of the used content in the usage record, we can extract l ACP-features, namely, $(c, p_1), (c, p_2), \dots, (c, p_l)$. For 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) [2]. 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 .

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 [6] 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 $P(z|u)$.

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, 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. 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.

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

where C_{max}^r denotes a maximal content-relevant sub-context contained by C , and $|C_{max}^r|$ indicates the number of C_{max}^r .

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 . However, 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 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 “ $C^r \rightarrow c$ ” and $P(C^r|u)$ can be estimated by $\frac{\text{Support}(C^r)}{N_r}$, 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.* [3], 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.

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 classic noise based matrix factorization problem as follows.

$$\Omega_{N \times M} = \Phi_{N \times K} \Theta_{K \times M} + \mathbf{N}_{N \times M},$$

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, Ω denotes the observed matrix of $P(c, C^r|u)$, $\phi_{ik} \in \Phi (1 \leq i \leq N, 1 \leq k \leq K)$ denotes the probability $P(c, C^r|z_k)$, $\theta_{kj} \in \Theta (1 \leq k \leq K, 1 \leq j \leq M)$ 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 Φ and Θ 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 Φ and Θ 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 Φ , Θ and \mathbf{N} . In this paper, we propose to leverage a constraint based Bayesian Matrix Factorization model [13] 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 Φ and Θ

Table I
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}

under a set of model assumptions, which are specified by the likelihood function $P(\Omega|\Phi, \Theta, \mathbf{N})$. The likelihood function denotes the probability of the observed data matrix Ω given priors $P(\Phi, \Theta)$ and $P(\mathbf{N})$. According to [13], 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 $P(n_{ij}) = N(n_{ij}|0, \nu_{ij}) = \frac{1}{\sqrt{2\pi\nu_{ij}}} \exp(-\frac{n_{ij}^2}{2\nu_{ij}})$, 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 Φ and Θ subject to inequality constraints \mathbf{Q} and equality constraints \mathbf{R} as

$$P(\vec{\phi}, \vec{\theta}) \propto \begin{cases} \mathbf{N} \left(\begin{bmatrix} \vec{\phi} \\ \vec{\theta} \end{bmatrix} \middle| \underbrace{\begin{bmatrix} \mu_\phi \\ \mu_\theta \end{bmatrix}}_{\mu}, \underbrace{\begin{bmatrix} \Sigma_\phi & \Sigma_{\phi\theta} \\ \Sigma_{\theta\phi} & \Sigma_\theta \end{bmatrix}}_{\Sigma} \right), & \text{if } \mathbf{Q}(\vec{\phi}, \vec{\theta}) \leq 0, \\ & \mathbf{R}(\vec{\phi}, \vec{\theta}) = 0, \\ 0, & \text{otherwise,} \end{cases}$$

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

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

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 (Context conditional Independence Assumption based Prediction) and CDAP (Context Dependency Assumption based Prediction), with several baseline methods on a real-world data set.

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 I 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 taxonomy of Ovi store (www.oivi.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.

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 (i.e., CASVM). Then, 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 (i.e., CACF and CATF).

- **CASVM** stands for personalized Context-Aware preference prediction by Ranking SVM, which is a learning-to-rank based approach introduced in [9].
- **CACF** stands for Context-Aware preference mining by disjunction CF, which is a memory-based CF approach introduced in [11].
- **CATF** stands for Context-Aware preference mining by Tensor Factorization, which is a model-based CF approach introduced in [10].

C. Evaluation Metrics

In the experiments, we utilize a five-fold cross validation to evaluate each test approach. 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 $AP^{(u)}@K$ 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.

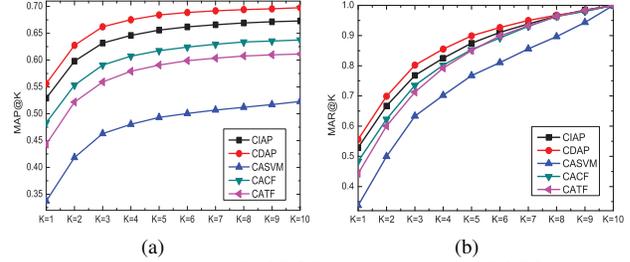


Figure 1. The (a) average $MAP@K$ and (b) average $MAR@K$ of each prediction approach in the five-fold cross validation.

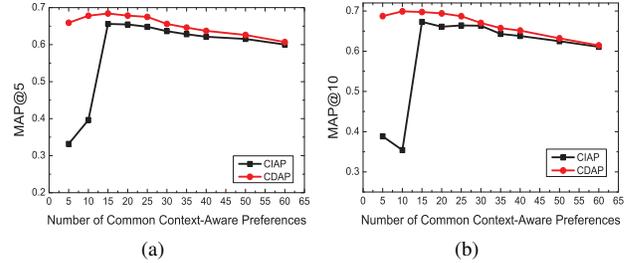


Figure 2. The (a) $MAP@5$ and (b) $MAP@10$ of CIAP and CDAP with respect to varying number of common context-aware preferences.

- **MAR@K** stands for Mean Average Recall at top K prediction results. To be specific, $MAR@K = \frac{\sum AR^{(u)}@K}{|U|}$, where $AR^{(u)}@K$ can be computed by $\frac{1}{N_u} \sum_i \sum_{r=1}^K rel_i(r)$.

D. Overall Results

For selecting a proper number of common context-aware preferences for both CIAP and CDAP, we leverage the perplexity estimation approach introduced in [1] and Chib's method introduced in [14], respectively. Accordingly, 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 [7]. For the Bayesian Matrix Factorization training, according to [13], we use an isotropic noise model and choose a decoupled prior for Φ and Θ with zero mean $\mu = 0$, and a unit diagonal covariance matrix $\Sigma = \mathbf{I}$. 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 [3]. Both our approaches and the baselines are implemented by standard C++ and the experiments are conducted on a 3GHZ×4 quad-core CPU, 3G main memory desktop PC.

From the training process we observe 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.

We first test the $MAP@K$ and $MAR@K$ performance of each test approach with respect to varying K and the average results in the five-fold cross validation are shown in Figure 1 (a) and (b). 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 . From the results we can observe that our approaches consistently outperform other baselines in all five test rounds. 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.

Both CIAP and CDAP need a parameter to determine the number of common context-aware preferences. Figure 2 (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 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.

VI. 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 relatively higher computational cost.

Acknowledgement. This work was supported in part by grants from Natural Science Foundation of China (NSFC, Grant Numbers 61073110, 70890082, 71028002), Research Fund for the Doctoral Program of Higher Education of China (20113402110024), the Key Program of National Natural Science Foundation of China (Grant No. 60933013), the NSFC Major Program (Grant No. 71090401/71090400), and Nokia. It was also partially supported by grants from National Science Foundation (NSF) via grant numbers CCF-1018151, IIS-1256016, DUE-1241315.

REFERENCES

- [1] 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.
- [2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. In *Journal of Machine Learning Research*, pages 993–1022, 2003.
- [3] 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.
- [4] 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.
- [5] Y. Ge, Q. Liu, H. Xiong, A. Tuzhilin, and J. Chen. Cost-aware travel tour recommendation. In *KDD'11*, 2011.
- [6] T. L. Griffiths and M. Steyvers. Finding scientific topics. In *Proceedings of National Academy of Science of the USA*, pages 5228–5235, 2004.
- [7] G. Heinrich. Parameter estimation for text analysis. *Technical report, University of Leipzig*, 2009.
- [8] 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.
- [9] M. Kahng, S. Lee, and S.-g. Lee. Ranking in context-aware recommender systems. In *WWW'11*, pages 65–66, 2011.
- [10] 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.
- [11] 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.
- [12] Q. Liu, Y. Ge, Z. Li, E. Chen, and H. Xiong. Personalized travel package recommendation. In *ICDM'11*, 2011.
- [13] M. N. Schmidt. Linearly constrained bayesian matrix factorization for blind source separation. In *NIPS'09*, pages 1624–1632. 2009.
- [14] 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.
- [15] K. Yu, B. Zhang, H. Zhu, H. Cao, and J. Tian. Towards personalized context-aware recommendation by mining context logs through topic models. In *PAKDD'12*, 2012.
- [16] 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.