

AdNext: A Visit-Pattern-Aware Mobile Advertising System for Urban Commercial Complexes

Byoungjip Kim¹, Jin-Young Ha², SangJeong Lee¹, Seungwoo Kang¹, Youngki Lee¹,
Yunseok Rhee³, Lama Nachman⁴, and June-hwa Song¹

¹Computer Science, KAIST
{bjkim, peterlee, swkang,
youngki, junesong}
@nclab.kaist.ac.kr

²Computer Science and
Engineering, KNU
jyha@kangwon.ac.kr

³Electronics & Information
Engineering, HUFS
rhees@hufs.ac.kr

⁴Intel Corporation
lama.nachman@intel.com

ABSTRACT

As smartphones have become prevalent, mobile advertising is getting significant attention as being not only a killer application in future mobile commerce, but also as an important business model of emerging mobile applications to monetize them. In this paper, we present AdNext, a visit-pattern-aware mobile advertising system for urban commercial complexes. AdNext can provide highly relevant ads to users by predicting places that the users will next visit. AdNext predicts the *next visit place* by learning the sequential visit patterns of commercial complex users in a collective manner. As one of the key enabling techniques for AdNext, we develop a probabilistic prediction model that predicts users' next visit place from their place visit history. To automatically collect the users' place visit history by smartphones, we utilize Wi-Fi-based indoor localization. We demonstrate the feasibility of AdNext by evaluating the accuracy of the prediction model. For the evaluation, we used a dataset collected from COEX Mall, the largest commercial complex in South Korea. Also, we implemented an initial prototype of AdNext with the latest smartphones, and deployed it in COEX Mall.

Categories and Subject Descriptors

C.2.4 [Computer-Communication Networks]: Distributed System; I.2.1 [Artificial Intelligence]: Applications and Expert Systems; K.4.4 [Computer and Society]: Electronic Commerce.

General Terms

Algorithms, Design, Economics, Experimentation, Human Factors.

Keywords

Mobile advertising, Sequential visit patterns, Prediction models, Wi-Fi localization, User survey.

1. INTRODUCTION

Mobile advertising is rapidly growing. According to a report [4], mobile advertising spending worldwide will surpass \$19.1 billion in 2012. The proliferation of smartphones opens up many new opportunities for mobile advertising. Many researchers expect that mobile advertising will be not only a killer application in mobile commerce, but also an important business model for many emerging mobile applications to monetize [14][16]. Accordingly, research on more efficient and effective mobile advertising is

strongly required to meet the requirements of both advertisers and mobile users in the upcoming new mobile era.

Commercial complexes such as Mall of America and COEX Mall have huge potential for mobile advertising. COEX Mall [2], the largest commercial complex in South Korea, has more than 260 stores and attracts more than a hundred thousand visitors per day. From an advertisers' perspective, commercial complexes are strategically important places for advertising, because many people visit a commercial complex for the purpose of purchasing a product or a service. Also, from a customers' perspective, mobile advertising can help them to use a commercial complex in a more convenient way, since there are too many places to know well in detail.

To effectively provide mobile ads in such a commercial complex, customer targeting is important [22]. Through customer targeting, advertisers can identify people who will highly likely purchase a product and a service. Then, they can increase the effectiveness of the advertising by focusing their efforts on those people. Also, customers can avoid spam ads. Especially, for customer targeting at a commercial complex, we should consider the *spatial* and *temporal relevance* of ads to mobile users. If the category of an advertising place is interesting to a user and a place is closely located to a user, the ad will easily attract the user to visit the advertising place (spatial relevance). Also, if an advertised product or service will be highly likely consumed by a user soon, the ad will be able to immediately lead the user to actually purchase the product or service (temporal relevance). However, existing representative mobile advertising, i.e., location-based advertising, is highly limited in effective targeting. This is mainly because it delivers ads of near places just depending on a user's *current location*. For example, ads for nearby restaurant do not attract people who are having dinner or have already had dinner few minutes ago.

In this paper, we present AdNext, a visit-pattern-aware mobile advertising system for urban commercial complexes. To provide spatially and temporally relevant ads to users, AdNext predicts users' *next visit place*. Because the next visit place of a user implies a product or a service for the user to potentially purchase in the near future, AdNext can send spatially and temporally relevant ads based on such information. To effectively predict users' next visit place, we propose an approach that exploits behavioral patterns of commercial complex users. More specifically, we predict the next visit place by learning *sequential visit patterns* of commercial complex users in a collective manner. For example, at COEX mall, we might discover that many people who have sequentially visited a cinema and a restaurant are highly likely to visit either a café or a fashion shop next. Based on this fact, AdNext provides ads related to cafés and fashion shops to people who have sequentially visited a cinema and a restaurant.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HotMobile'11, March 1–2, 2011, Phoenix, Arizona, USA.

Copyright 2011 ACM 978-1-4503-0649-2...\$10.00.

There are many challenges in developing a visit-pattern-aware mobile advertising system. Important challenges include user contexts collection (i.e., location, time, profile, etc.), user behavior modeling, ad presentation design, user feedback collection, user privacy preservation, ad pricing modeling, etc. In this paper, we focus on effective prediction of the next visit places of commercial complex users. Accordingly, we explore the relevant challenges and discuss potential solutions to them. The first challenge is how to automatically collect the place visit history of commercial complex users. It is critical to enable store-level indoor localization in a cost-effective way. Furthermore, it is required to detect *place-in* and *place-out* events necessary for accurately capturing sequential visit patterns. We develop a place visit detection method that traces user’s place-in/out events and corresponding visit places based on Wi-Fi indoor localization. More importantly, the second challenge is how to predict the next visit place of users. It is difficult to predict the next visit place with certainty, because there is inherent uncertainty in human behaviors. Therefore, we use a probabilistic prediction model to learn the sequential visit patterns of commercial complex users. AdNext builds a next visit place prediction model using Bayesian networks, and makes a prediction from a user’s partial place visit history.

The contributions of this paper are as follows. First, we introduce a visit-pattern-aware mobile advertising system that exploits sequential visit patterns of commercial complex users to provide highly targeted mobile ads to users. Also, we design and implement an initial prototype that actually works in a real commercial complex, i.e. COEX Mall. Second, we develop a next visit prediction model using Bayesian networks (BN) as an enabling technique of AdNext. To confirm the validity of using BN, we compare our BN-based model with other possible machine learning techniques such as decision trees and conditional random fields (CRF). Third, we show the feasibility of AdNext by evaluating the prediction model. For the evaluation, we collected a datasets by surveying about 150 people in COEX Mall.

This paper is organized as follows. Section 2 presents related work. AdNext design is presented in Section 3. In Section 4, we show preliminary evaluation results. Section 5 explains the initial implementation and deployment of AdNext. In Section 6, we briefly discuss related issues of AdNext. Finally, Section 7 concludes this paper.

2. RELATED WORK

Mobile Advertising. In general, mobile advertising broadly refers to diverse forms of advertising via mobile devices. The types of mobile advertising include SMS or MMS advertising, mobile Web banner advertising, in-app banner advertising, etc. For instance, AdMob [1], one of the rapidly growing mobile advertising companies, provides mobile Web and in-app banner advertising for many smartphone platforms. As a more experimental approach to mobile advertising, activity-based mobile advertising has been introduced. The basic idea is to send ads according to a user’s current activity. However, Sala, et al. reported very important lessons that the ads related to users’ *current* activity are less effective in terms of relevance and usefulness than they expected [20].

Location-Based Advertising. Location-based advertising (LBA) provides target customers with location-specific ads on their mobile devices. They can be categorized according to

proactiveness: pull-type and push-type. In pull-type LBAs, users explicitly request location-specific ads from the system. However, in push-type LBAs, advertising systems proactively send location-specific ads to target users.

As an experimental LBA system, Alto et al. proposed a push-type LBA system called B-MAD (for Bluetooth Mobile Advertising) that proactively sends ads to mobile phones when a user passes by a certain store [9]. The proximity between a user and a store is detected by using Bluetooth localization. Recently, with the advances of wireless networks and smartphones, many commercial LBAs such as WHERE [7], Shopkick [5], etc. have been emerging. For instance, Shopkick sends discount coupons to a user visiting a shopping mall: when the user explicitly requests coupons by “checking-in” (pull-type) or when the user is “walking into” a shopping mall (push-type). Similar to the activity-based mobile advertising mentioned above, existing LBAs are also limited in providing highly effective ads, because they exploit only the *current* location to provide ads.

AdNext is different from existing LBAs. AdNext aims to predict the most effective ads for a user by exploiting the user’s behavioral history (e.g., sequential visit patterns), not only her current location. Based on many people’s behavioral patterns, AdNext also predicts a highly probable next place from the user’s behaviors. In a way, it looks similar to the behavioral targeting of online advertising whose main idea is that people with similar online behaviors such as page view and keyword search will be highly likely to click the same ads on a page [23]. However, AdNext raises challenges such as offline user behavior modeling, and consequently requires very different design and approaches such as sensing users’ place visits with smartphones.

Probabilistic Reasoning. Probabilistic reasoning has become the dominant approach to human behavior modeling because of the uncertainty and the variability of human behaviors. Much research work has developed probabilistic models and techniques to infer a variety of human behaviors such as meaningful locations visited, activities, and transportation modes based on GPS location traces.

Ashbrook et al. proposed a system that automatically clusters GPS data taken over an extended period of time into meaningful locations at multiple scales [8]. In addition to this, they suggested building a Markov model to model a user’s transition across the meaningful locations.

Lin Liao et al. proposed a relational Markov network (RMN) which is an extension of conditional random fields (CRF) for inferring human activities from GPS traces [15]. They suggest that human activities can be recognized by location and time. However, the approach can only recognize course-grained activities such as home activity and work activity.

Patterson et al. proposed using Bayesian networks for inferring transportation modes from GPS traces [19]. Their main idea was that transportation modes such as bus, foot, and car can be inferred from a user’s average speed and variance. They also demonstrate that prediction accuracy can be improved by adding more external knowledge about bus routes and bus stops.

Unlike this prior work, we aim at enabling mobile advertising based on user behaviors in commercial complexes. For the purpose, we develop a prediction model to infer next visit places of mobile users in commercial complexes. To build the model, we adopt Bayesian networks which are effective in modeling the causality between users’ visit places.

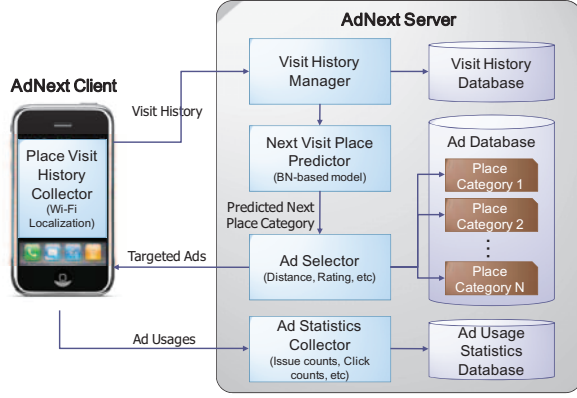


Figure 1. AdNext System Architecture.

3. AdNext DESIGN

AdNext consists of *mobile phones* (i.e., AdNext clients) and an *advertising server* (i.e., AdNext server). Figure 1 shows the overall system architecture. The AdNext client is responsible for collecting a user’s place visit history and reporting the history to the advertising server. While people move around a commercial complex, their mobile phones collect place-in and place-out events by using Wi-Fi fingerprints. Also, the AdNext client is responsible for automatically notifying mobile users when relevant ads are available. The AdNext server is responsible for (1) building a prediction model based on sequential visit patterns of commercial complex users, (2) sending spatially and temporally relevant ads based on the prediction model, and (3) collecting the statistics of ads usages such as issue counts, click counts, actual purchase counts, etc.

The advertising server operates in two modes: offline learning and online prediction mode. In offline learning mode, the advertising server builds the prediction model by using training data collected from a large number of commercial complex users. The process of model build can be periodically conducted to capture the change in the sequential visit patterns of a commercial complex. In online prediction mode, the advertising server receives place visit histories from users’ mobile phones, predicts the next visit place by using a trained probabilistic prediction model, and sends the most relevant ads based on the prediction. At the same time, the advertising server collects users’ place visit histories to train the prediction model in offline mode. Each user’s place visit history is encrypted and anonymized for preserving privacy.

3.1 Collecting Place Visit History

The first problem we address in designing and implementing AdNext is how to collect place visit histories of commercial complex users. The problem is challenging because of the following reasons. (1) AdNext requires store-level localization accuracy. Although there are many localization techniques, only a few techniques support store-level localization accuracy in a cost-effective way. (2) AdNext needs not only to identify users’ location, but also to accurately detect the time when a user moves in and out of a certain place to learn users’ sequential visit patterns.

To collect a user’s place visit history, we develop a place visit detection method in the AdNext client. It traces a user’s place-in/out events and corresponding visit places based on a Wi-Fi localization technique. Wi-Fi localization is appropriate for

AdNext, compared to other in-door localization such as Bluetooth-based or IR-based localization. This is mainly because a number of Wi-Fi access points (APs) are widely and densely deployed in commercial complexes such as COEX Mall. Accordingly, we can build a visit-pattern-aware mobile advertising system without installing additional infrastructure, and achieve reasonable performance for store-level indoor localization. For Wi-Fi fingerprinting and localization, we utilize Elekspot [17], a crowd-sourced Wi-Fi localization system.

The place visit detection involves two major steps: current location detection and location change validation. First, it performs periodic Wi-Fi scanning (e.g., every 10 seconds) and obtains information about the current location (i.e., Elekspot ID and its confidence value) from the Elekspot server. Receiving information about accessible Wi-Fi APs and received signal strength indicator (RSSI) values from the AdNext client, the Elekspot server resolves the current location by searching for the most similar Wi-Fi fingerprint stored in its database. It also provides the corresponding confidence value, which indicates how well the currently captured Wi-Fi fingerprints match the current location.

Second, the location change validation checks if a user enters or leaves a store. A location change is validated with an Elekspot ID change, confidence value threshold, and stable period. Specifically, it checks if the acquired Elekspot ID has changed. If so, it checks if the confidence value of location inference is larger than the predefined threshold. Location change with a low confidence value indicates situations such as “just passing by a place”. Finally, if it is stable for a certain amount of time, e.g., a period of three consecutive scan, the new Elekspot ID is determined as a visit place.

We further consider utilizing an accelerometer embedded in a smartphone to enhance the accuracy of place visit detection. The location change validation could exploit user’s activity such as sitting, walking for a certain amount of time, and the number of steps along with the distance between stores.

3.2 Predicting the Next Visit Place

3.2.1 Basic Idea

One of the most important enabling techniques in realizing AdNext is how to predict mobile users’ next visit place. The problem is challenging because of the following reasons. First, it is highly difficult to predict with certainty the next visit place of a certain user, since there is inherent uncertainty in people’s behavior. Second, some people may not want to allow personal behavior such as place visit history to be identified by others. In this paper, we argue that we can predict a certain user’s next visit place from *common sequential visit patterns* learned from many people in a *collective* manner. The basic idea enabling our approach is as follows.

Visit Causality. People’s current visit is influenced by the previous visit history. The causality exists between sequentially visited places. For example, in COEX Mall, if a user visits a restaurant, the user will highly likely not visit other restaurants consecutively, but may visit a café to have a dessert.

Common Visit Pattern. There exist frequent and similar sequential visit patterns among many people. For example, in COEX Mall, many people may show a sequential visit pattern such as cinema→restaurant→café. Thus, a next visit place of a

user could be predicted based on the collective place visit behaviors of other users, rather than learning individual user's sequential visit patterns. Using the common sequential visit patterns can significantly reduce privacy concerns, since the prediction is based on the collective information where individual user's place visit history is anonymized.

3.2.2 Next Visit Prediction Model

To predict the next visit place from place visit history, we develop a probabilistic prediction model by using Bayesian Networks. First of all, probabilistic models are more appropriate for learning customer behavior that inherently includes a degree of the uncertainty. Among many models proposed for probabilistic reasoning, we carefully select Bayesian networks.

We consider visit place, visit time, visit duration, gender and age as the main features for the prediction model. Visit place is denoted by $\mathbf{P}$. For example, $\mathbf{P}$ has four states, i.e., {fashion shop, café, restaurant, entertainment store}. Visit time is denoted by $\mathbf{T}$. For example, $\mathbf{T}$ has 24 states which are divided by hour. Visit duration is denoted by $\mathbf{D}$. For example, $\mathbf{D}$ has 8 states whose max value is limited to 4 hours and each state is divided into 30 minutes chunks. Age is denoted by $\mathbf{A}$. For example, $\mathbf{A}$ has 4 states, i.e., {under 20's, 20's, 30's, over 40's}. Gender is denoted by $\mathbf{G}$, and has two states, i.e., {male, female}. We further denote the current features by $\mathbf{P}_0$, $\mathbf{T}_0$ and $\mathbf{D}_0$, and the i -th previous features by $\mathbf{P}_i$, $\mathbf{T}_i$, and $\mathbf{D}_i$. Note that age $\mathbf{A}$ and gender $\mathbf{G}$ are static features.

We propose a next visit prediction model using Bayesian networks. A Bayesian network (BN) [13] is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. A BN models the joint probability $P(\mathbf{X}, \mathbf{Y})$, where $\mathbf{X}$ represents features (or observations) and $\mathbf{Y}$ represents labels (or classifications). To make inference tractable, BN assumes conditional independences among $\mathbf{X}$ and $\mathbf{Y}$, and represents the joint probability $P(\mathbf{X}, \mathbf{Y})$ as products of conditional probabilities. It is represented by the following formula.

$$P(\mathbf{V}) = \prod_{v \in \mathbf{V}} P(v | \text{parents}(v))$$

where, $\mathbf{V}$ represents nodes (or vortexes) in a Bayesian network (i.e., including both $\mathbf{X}$ and $\mathbf{Y}$), and $\text{parents}(v)$ represents parent nodes that are linked with a node v by edges. To make an inference, it calculates the posterior probability $P(\mathbf{Y} | \mathbf{X})$ from the conditional probability $P(\mathbf{X} | \mathbf{Y})$ by using the Bayes theorem.

Figure 2 shows the proposed model. Conceptually, in the model, an edge represents the causality between linked nodes (or random variables). More specifically, the model implies that the current visit place $\mathbf{P}_0$ is influenced by *special influential contexts* such as the previous visit place $\mathbf{P}_1$ and $\mathbf{P}_2$ and *temporal influential contexts* such as the previous visit duration $\mathbf{D}_1$ and current time $\mathbf{T}_0$. Also, it is influenced by *profile information* such as gender $\mathbf{G}$ and age $\mathbf{A}$. Formally, it is represented as follows.

$$\begin{aligned} P(\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2, \mathbf{D}_1, \mathbf{D}_2, \mathbf{G}, \mathbf{A}) \\ = P(\mathbf{P}_0 | \mathbf{P}_1, \mathbf{P}_2, \mathbf{T}_0, \mathbf{D}_1, \mathbf{G}, \mathbf{A}) \\ \cdot P(\mathbf{P}_1 | \mathbf{P}_2, \mathbf{T}_1, \mathbf{D}_2) \cdot P(\mathbf{P}_2 | \mathbf{T}_2) \end{aligned}$$

Here, we only consider two previous visit places as features. There is a trade-off between the number of the previous visit places and the complexity (i.e., memory and computation). According to our experiment, if the number of visit places is more than three, the improvement in the prediction accuracy is marginal.

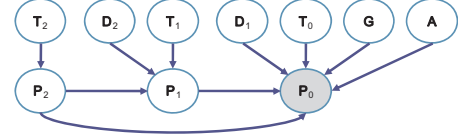


Figure 2. Next visit prediction model using Bayesian networks.

We use the E-M algorithm [10] for training the model. In the E-M algorithm, E-step estimates missing values in the current model and M-step maximizes likelihood. The E-M algorithm allows us to train the model with latent variables. Once the model is trained, the model can evaluate the probability distribution of $\mathbf{P}_0$ when some of the other features are given as evidence.

For the visit place $\mathbf{P}$ of the prediction model, we can consider a different level of place category. Place category level is one of the important parameters that affect the effectiveness of AdNext. As place category is more fine-grained, advertisers will be able to distinguish customers more specifically. However, the prediction accuracy may decrease, because there are many similar and confusing choices to predict. Therefore, we should adjust the place category level to achieve the required effectiveness of the mobile advertising.

3.3 Selecting Relevant Ads

Predicting the next place is not enough for providing ads. We should select a small number of ads among many candidate ads classified in the same place category that AdNext predicts. There can be many candidate ads classified in the same place category. For example, in COEX Mall, there are about 60 places classified into fashion shops, and about 50 places classified into restaurants. However, smartphones have small screens that can display only a few ads without scrolling. Therefore, too many ads displayed on a small screen would not attract users' attentions, and would decrease the effectiveness of the mobile advertising.

To select ads that are relevant to each user, we propose an ad selection method that evaluates a score for each candidate ads classified into a certain place category that have been predicted for the next visit place. Basically, AdNext gives a high score to an ad, as the place of the ad is closer to the current location of a user and the user-given rating of the place is higher. AdNext uses a score function as follows.

$$\text{Score}(c, a) = \alpha \cdot 1/\text{distance}(c, p_a) + \beta \cdot \text{rating}(p_a)$$

where c is a user, a is an ad, p_a is a place of ad a , and α, β are scaling factors. Once each ad is scored, AdNext selects ads with high scores.

To effectively display ads on smartphones, we adjust the number of ads to display according to the probability that the next visit prediction model calculates. That is, AdNext provides more ads for the place category that is predicted with higher probability. For example, consider that AdNext predicts a user's next visit place as café 60% and fashion shop 40%, and the number of ads to display is limited to five. Then, AdNext provides three ads related to cafés and two ads related to fashion shops.

4. PRELIMINARY EVALUATION

Data Collection. To perform an evaluation that can show the feasibility of AdNext, we have collected a real dataset from COEX Mall. We have collected a sequential visit history of COEX Mall users. More specifically, we requested COEX Mall

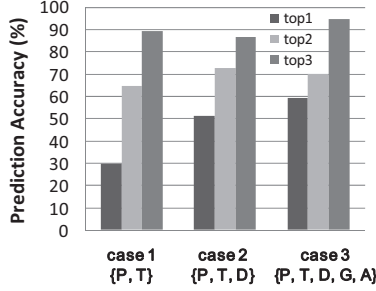


Figure 3. Prediction accuracy of the proposed model.

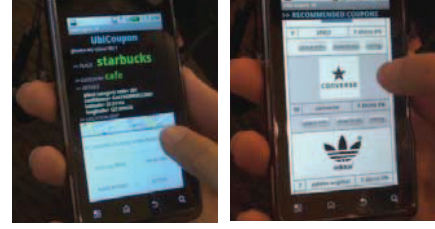
Table 1. Prediction accuracy comparison

	top1	top2	top3
Decision Tree	40.54%	n/a	n/a
CRF	29.72%	n/a	n/a
Bayesian Networks	59.45%	70.27%	94.59%

users to list all the places they have visited in a sequential manner. Also, we requested they write down the visit time, visit duration, and their rating to the visit places. We surveyed about 130 people. From the data, we filtered out some users who reported too few visit places (i.e., less than three places), or whose data duplicate to others (i.e., possibly companion group members of couples or families). We thus obtained the clean dataset of 76 people. The total number of visits is 351. We used 80% of the entire dataset as the training dataset, and 20% as test dataset.

Prediction Accuracy. We show the prediction accuracy of the next visit place prediction model using Bayesian networks. For evaluating the models, we used GeNIe and SMILE [10], a Bayesian network toolkit. We examine how much the features of the model affect the prediction accuracy. We used four place categories, i.e., {fashion shop, café, restaurant, entertainment store}. Figure 3 shows the evaluation result. As the features of the model increase, the prediction accuracy increases. Case 1 only considers visit place **P** and visit time **T** as features. It is derived from the model in Figure 2 by eliminating gender **G**, age **A**, and visit duration **D**. Case 2 includes visit place **P**, visit time **T**, and visit duration **D**. Case 3 includes all the features mentioned in Section 3.2.2. In the figure, top 1 means the percentage accuracy that the next visit place is correctly predicted as the first rank (i.e., the highest probability), and top 2 means the percentage accuracy that the next visit place is correctly predicted within the second rank (i.e., the second highest probability). Case 3 shows the highest prediction accuracy, i.e., 59.45%. If top2 is considered, the prediction accuracy of the model increases up to 70.27%. In the current implementation of AdNext, the system provides ads related to places which it predicts within the second highest probability, i.e., top2. This means that we can have about 70% confidence that the provided ads are spatially and temporally relevant to COEX Mall users. We believe that the prediction accuracy will increase as the number of training data increases.

Comparison. We compare the BN-based prediction model with other prediction models. We used a decision tree for a non-probabilistic model, and Conditional Random Fields (CRF) for a competitive probabilistic model. For evaluating the decision tree, we used Weka [6], a data mining and machine learning toolkit. More specifically, we used the C4.5 algorithm. For evaluating CRF, we used a Java CRF package [3]. We used default settings



(a) A main screen showing current place “Starbucks”. (left)
(b) An ad screen showing pushed mobile ads. (right)

Figure 4. Screenshot of AdNext client.



Figure 5. Screenshot of AdNext server.

provided by the package. Table 1 shows the evaluation result. The decision tree shows about 40.54% accuracy. CRF shows about 29.72% accuracy. The evaluation shows that the models using BN outperforms other models. Generative models such as BN can converge relatively faster during training and have less variance. Therefore, when the independence assumption holds, or when only a small amount of training data is available, a generative approach could outperform a discriminative approach such as CRF [18].

5. IMPLEMENTATION AND DEPLOYMENT

We have implemented an initial prototype of AdNext, deployed in COEX Mall. The implementation details are as follows.

We implemented the AdNext client on Android smartphones (i.e., Android version 2.1). The AdNext client consists of continuously running services such as place visit history collector, visit history reporter, ads agent, etc. To implement such software design, we fully exploit the multi-thread programming feature that Android successfully supports. Figure 4 shows screenshots of the AdNext client. For automatically collecting a user’s place visit history, AdNext uses the APIs provided by Elekspot that runs on the same smartphones. Collected place visit history is stored in SQLite, a light-weight relational database, so that users can browse their visit history. AdNext client uses a Wi-Fi interface for localization and a 3G interface for communication to the advertising server. When there are available ads for users, the AdNext client alerts users by using the vibration functionality of the smartphones.

We implemented the AdNext server based on Java Web technology. The main server components such as Place History Manager, Next Place Predictor, Ads Selector, and Ads Statistics Collector have been implemented as Servlets. We used Apache Tomcat 5.5.28 as a Servlet container. For the Next Place Predictor, we used Java APIs provided by SMILE [10], a Bayesian network library. For storing ads, we used MySQL 5.0.88 as a relational database. Figure 5 shows a screenshot of the Web-based administration interface of AdNext server. It shows the ad usage statistics such as issue counts, click counts, and user rating of each advertisement.

We have deployed AdNext in COEX Mall. AdNext can identify almost all the stores (about 200 stores) in COEX Mall by using Wi-Fi localization. Currently, AdNext provides mobile ads of several selected places for the purpose of experiments. However, the prototype has been designed to be extended to provide ads for all the places.

6. DISCUSSION

Privacy concerns. AdNext can raise some privacy concerns, since it collects users' place visit history in a centralized server. Accordingly, we need to provide a certain level of privacy control to end users, and a set of privacy-preserving mechanisms to protect an adversary's potential attacks. As one of an architectural alternative, we may consider adding a privacy proxy in our system architecture. A privacy proxy would be responsible for (1) authenticating each user, (2) encrypting each user's update message, and (3) anonymizing visit history. For anonymizing a user's place visit, we can apply existing anonymity techniques such as *k-anonymity* [21] that provides a form of plausible deniability by ensuring that the user cannot be individually identified from a group of k users. In the case of AdNext, the privacy proxy can only accept a user's update when there are at least k people visiting a certain place during a given time period. Also, we can apply other anonymity techniques such as *path confusion* [12]. For preserving privacy more strictly, it is highly required to avoid updating place visit history to a central server. To achieve this goal, we can think of moving the place prediction module from the server to each client. Then, each client predicts the next visit place by using locally collected visit history, and internally requests relevant ads from the server based on the prediction.

Energy consumption. AdNext can raise energy consumption issues, since it performs continuous location sensing in resource-scarce smartphones. To increase energy-efficiency, periodic Wi-Fi scanning for collecting place visit history should be avoided. One effective solution to do that would be to use multi-axis accelerometers in smartphones. The accelerometers allow us to infer a user's physical activity (i.e., walking, running, sitting up/down, etc.). If a user visits a restaurant and "sitting down" is detected by the accelerometers, then we can avoid periodic Wi-Fi scanning until "sitting up" is detected.

7. CONCLUSION AND FUTURE WORK

This paper presents AdNext, a visit-pattern-aware mobile advertising system that provides highly relevant mobile ads to commercial complex users by predicting users' next visit place. AdNext predicts users' next visit place by learning sequential visit patterns of commercial complex users. To effectively predict users' next visit place, we develop a probabilistic prediction model using Bayesian networks. We demonstrate the feasibility of AdNext by showing the initial evaluation results. For the evaluation, we use a dataset collected from COEX Mall. According to the evaluation, we can predict the next visit place of COEX Mall users with about 60% accuracy. We have designed and implemented an initial prototype of AdNext.

We plan to further investigate diverse issues of AdNext based on the initial prototype deployed in COEX Mall. The issues may include comparison with existing location-based advertising, privacy preserving mechanisms and architecture, and energy efficiency in collecting place visit history.

8. ACKNOWLEDGEMENTS

This work was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0004). We would like to thank the reviewers and the shepherd, Prof. David Johnson, for their helpful feedback.

9. REFERENCES

- [1] AdMob. <http://www.admob.com/>
- [2] COEX Mall. <http://www.coex.co.kr/eng/index.asp>
- [3] CRF software package. <http://crf.sourceforge.net/>
- [4] Mobile Advertising Spending. <http://www.emarketer.com>
- [5] Shopkick. <http://www.shopkick.com>
- [6] Weka 3. <http://www.cs.waikato.ac.nz/ml/weka/>
- [7] WHERE Ads. <http://site.where.com/where-ads/>
- [8] D. Ashbrook, T. Starner, "Using GPS to learn significant locations and predict movement across multiple users", *Personal Ubiquitous Computing*, vol.7, p275-286, 2003.
- [9] L. Alto, N. Gothlin, J. Korhonen, T. Ojala, "Bluetooth and WAP Push Based Location-Aware Advertising System", *ACM MobiSys* 2004.
- [10] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum-likelihood from incomplete data via the EM Algorithm", *Journal of the Royal Statistical Society*, 39, 1977.
- [11] M. J. Druzdzel, "SMILE: Structural Modeling, Inference, and Learning Engine and GeNIe: A Development Environment for Graphical Decision-Theoretic Models", *AAAI* 1999.
- [12] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Preserving Privacy in GPS Traces via Uncertainty-Aware Path Cloaking", *ACM CCS* 2007.
- [13] U. B. Kjaerulff, A. L. Madsen, "Probabilistic Networks: An Introduction to Bayesian Networks and Influence Diagrams", Springer 2005.
- [14] J. Krumm, "Ubiquitous Advertising: The Killer Applications for the 21st Century", *IEEE Pervasive* 2010.
- [15] L. Liao, D. Fox, H. Kautz, "Location-based Activity Recognition using Relational Markov Networks", *IJCAI* 2005.
- [16] C. Narayanaswami, D. Coffman, M. Lee, Y. Moon, J. Han, H. Jang, S. McFaddin, Y. Paik, J. Kim, J. Lee, J. Park, D. Soroaker, "Pervasive Symbiotic Advertising", *ACM HotMobile* 2008.
- [17] M. Lee, H. Yang, D. Han, C. Yu, "Crowdsourced Radiomap for Room-Level Recognition in Urban Environment", *IEEE PerCom workshop* 2010.
- [18] A.Y. Ng, M.I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and Naïve Bayes", *NIPS* 2002.
- [19] D. J. Patterson, L. Liao, D. Fox, H. Kautz, "Inferring High-Level Behavior from Low-Level Sensors", *ACM UbiComp* 2003.
- [20] M. C. Sala, K. Partridge, L. Jacobson, J. B. Begole, "An Exploration into Activity-Informed Physical Advertising Using PEST", *Pervasive* 2007.
- [21] L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression", *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5): 571-588, 2002.
- [22] W. D. Wells, S. Moriarty, J. Burnett, "Advertising: Principles and Practice", 7th ed. Pearson, 2005.
- [23] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, Z. Chen, "How much can Behavioral Targeting Help Online Advertising?", *WWW* 2009.