

## Research Article

# Agatha: Predicting Daily Activities from Place Visit History for Activity-Aware Mobile Services in Smart Cities

Byoungjip Kim,<sup>1</sup> Seungwoo Kang,<sup>2</sup> Jin-Young Ha,<sup>3</sup> and Junchwa Song<sup>4</sup>

<sup>1</sup>Department of Information and Communication Engineering, Sangmyung University, Cheonan 31066, Republic of Korea

<sup>2</sup>School of Computer Science and Engineering, KOREATECH, Cheonan 31253, Republic of Korea

<sup>3</sup>Department of Computer and Communication Engineering, Kangwon National University (KNU), Chuncheon 24341, Republic of Korea

<sup>4</sup>School of Computing, KAIST, Daejeon 34141, Republic of Korea

Correspondence should be addressed to Seungwoo Kang; [swkang@koreatech.ac.kr](mailto:swkang@koreatech.ac.kr)

Received 3 July 2015; Accepted 5 November 2015

Academic Editor: Salvatore Distefano

Copyright © 2015 Byoungjip Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a *place-history-based activity prediction system* called Agatha, in order to enable activity-aware mobile services in smart cities. The system predicts a user's potential subsequent activities that are highly likely to occur given a series of information about activities done before or activity-related contextual information such as visit place and time. To predict the activities, we develop a *causality-based activity prediction model* using Bayesian networks. The basic idea of the prediction is that where a person has been and what he/she has done so far influence what he/she will do next. To show the feasibility, we evaluate the prediction model using the American Time-Use Survey (ATUS) dataset, which includes more than 10,000 people's location and activity history. Our evaluation shows that Agatha can predict users' potential activities with up to 90% accuracy for the top 3 activities, more than 80% for the top 2 activities, and about 65% for the top 1 activity while considering a relatively large number of daily activities defined in the ATUS dataset, that is, 17 activities.

## 1. Introduction

Instrumented with diverse sensors, a smart city can provide a variety of intelligent city services such as city planning, transportation, mobile advertising, and security, enriching people's life experiences in the city. A large volume of sensor data produced in the smart city is an important means to enable such intelligent services. Recently, many researchers have been showing the feasibility of using diverse sensor data to build smart cities. Zheng et al. [1] showed that the GPS trajectories of taxicabs traveling in cities can be used to detect flawed city planning issues such as regions of salient traffic problems. Yuan et al. [2] proposed a recommender system that provides taxi drivers with locations at which to quickly find passengers by analyzing passengers' mobility patterns and taxi drivers' pick-up behaviors from the GPS trajectories of taxicabs. Also, other researchers [3, 4] showed that more relevant mobile ads can be delivered to people by analyzing people's behaviors in a city. Some researchers [5] also showed

that there are spatiotemporal patterns in criminal offense records in a city. With these research efforts, we may imagine predating crimes like *Minority Report* [6], a popular science fiction (SF) movie.

Among many smart city service scenarios, predictive information is quite useful, especially for mobile advertising or recommendation services. Sala et al. [7] proposed a basic concept of activity-based mobile advertising. Through a user study, the authors reported important lessons about how ads related to the current activity are not effective in terms of the relevance and usefulness perceived by users. For example, tennis racket ads for people who are playing tennis are not useful because the people already have rackets and are not interested in buying rackets at that time. In this case, ads or recommendations related to the next activities in which the people are likely to partake after playing tennis might receive more attention. While predicting human behavior is generally considered not easy due to its inherent uncertainty, recent research [8] shows that some human behavior such as

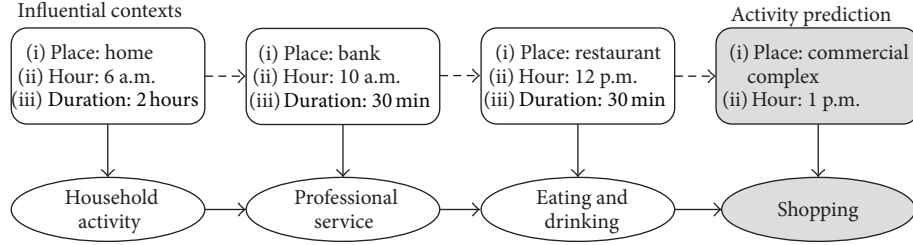


FIGURE 1: Example of causality-based activity prediction.

mobility has repetitive patterns and thus can be reasonably predicted.

In this paper, we present a *place-history-based activity prediction system* called Agatha in order to enable activity-aware mobile services in smart cities. The system predicts a user's potential subsequent activities given a series of information about activities done before or activity-related contextual information such as visit place and time. Key information for the prediction includes people's place visit history, which includes the types of visit place, visit time, and visit duration. While people move around a city, their smartphones collect place in and out events using Wi-Fi fingerprints. Processing those events, the system extracts the place visit history and predicts the users' activities that are likely to succeed the previous ones based on the history information by using a machine learning technique.

To predict the potential subsequent activities, we have developed a *causality-based activity prediction model* using Bayesian networks. The basic idea of predicting activities is that where a person has been or what he/she has done so far influences what he/she will do next (see Figure 1). The idea is based on two main intuitions: *activity causality* and *patterns of activity sequence*. The first one, activity causality, suggests that current activities are influenced by the current contexts, such as time of day and location, as well as the history of the previous contexts, such as visit place sequence. The second one, patterns of activity sequence, implies that people's common daily activities exhibit certain patterns to some extent. For example, housewives and single career women would usually show different activity sequence patterns in daily life. More specifically, housewives are highly likely to complete the following sequence: (a) personal care activities at home and (b) grocery shopping at a commercial complex. On the other hand, single career women are highly likely to do the following: (a) personal care activities at home, (b) work at an office, and (c) socializing and relaxing activities at a commercial complex. Based on these intuitions, the system extracts influential contexts from location streams as a main feature to infer activities. Then, it predicts the next activities based on the history of the influential contexts.

Our contributions are summarized as follows.

First, we present a place-history-based activity prediction system. The system has an advantage in that it can enable predictive mobile services by analyzing the location history

that is usually available in smartphones. It shows the possibility that diverse daily activities such as working, eating, and shopping can be predicted without requiring complex body-worn sensors to recognize daily activities.

Second, we develop a *causality-based activity prediction model* using Bayesian networks. To deal with a large number of random variables in the model, we have developed a hierarchical Bayesian network that reduces the complexity of training and inference.

Third, we present evaluation results using a real-world large dataset, that is, the American Time-Use Survey (ATUS) dataset [9] published by the US government. It includes more than 10,000 people's location and activity history. Our evaluation shows that Agatha can predict users' potential activities with up to 90% accuracy for the top 3 activities, more than 80% for the top 2 activities, and about 65% for the top 1 activity while considering a relatively large number of daily activities defined in the ATUS dataset, that is, 17 activities.

This paper is organized as follows. In Section 2, we discuss related work. We present an overview of the system architecture in Section 3 and an activity prediction model using Bayesian networks in Section 4. We discuss the implementation in Section 5 and evaluation results in Section 6. Finally, we conclude the paper in Section 7.

## 2. Related Work

**2.1. Activity Recognition.** In a ubiquitous computing domain, activity recognition has been receiving much attention. In general, there are three different approaches for activity recognition: (1) kinematic-based activity recognition [10, 11], (2) object-use-based activity recognition [12, 13], and (3) location-based activity recognition [14–16]. Kinematic-based recognition and object-use-based recognition are usually achieved with inertial sensors such as accelerometers attached to either people or objects. Body-worn or object-attached inertial sensors enable fine-grained activity recognition mainly related to the movement of a body, for example, walking, running, standing still, and bicycling, or the movement of an object, for example, drinking a cup of water and cooking. However, the semantics of the recognizable activities are limited for daily activities performed in diverse places in a city, for example, eating, shopping, and socializing. As an alternative, location-based activity recognition is considered

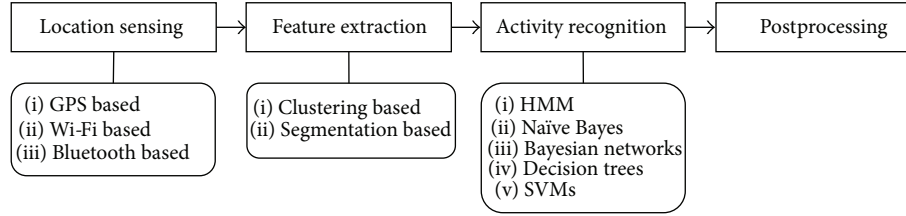


FIGURE 2: General processing stages of location-based activity recognition.

simple and practical in order to recognize daily activities related to diverse city places.

**2.2. Location-Based Activity Recognition.** Focusing on the prevalence of localization technologies such as GPS and Wi-Fi, there has been much effort to propose location-based activity recognition approaches [14, 15, 17]. Generally, location-based activity recognition consists of three processing stages: (1) location sensing, (2) feature extraction, and (3) activity inference (see Figure 2). For location sensing, GPS- and Wi-Fi-based localization are commonly used. Feature extraction and inference are different for the specific goal of each proposed approach. Broadly speaking, there are two types of machine learning techniques for activity inference models: (1) stochastic models such as Naïve Bayes models, hidden Markov models (HMMs), and Bayesian networks (BNs) and (2) nonstochastic models such as decision trees and Support Vector Machines (SVMs). In many research projects [14–16], stochastic models have been preferred due to the uncertainty of and the variability in human activities.

Ashbrook and Starner [14] proposed a HMM-based scheme to learn significant locations from GPS traces and predict transitions between locations. Different from our work to predict potential activities from previous activities or related contextual information, the previous work focused on extracting meaningful locations from raw GPS traces. In addition, while a HMM-based approach can model highly variable sequence data, it is not adequate for modeling causality relations. Also, it is hard to accept  $N$ -step previous states due to the computational complexity and high memory consumption, although the 2nd- and the 3rd-order Markov assumptions are possible.

The work by Liao et al. [15] used a relational Markov network (RMN), which is an extension of conditional random fields (CRFs). An RMN is able to model diverse relations of activities and their contexts, addressing the limitations of the HMM. However, it is still hard to accept  $N$ -step previous states because it assumes a Markov process on an activity sequence. Also, it is still limited in that it simply maps an activity to a place, for example, “home activity” to “home,” which is not practical in real situations. Moreover, the evaluation was performed under a limited condition with a small number of people and activity types, that is, 5 people and 6 activity types.

A Bayesian network (BN) is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph.

For example, a BN can represent the probabilistic relationship between diseases and symptoms. The most important advantage of the BN is that it is appropriate for modeling causality. Also, it is a simple and efficient approach compared to CRFs and HMMs, especially when considering problems in which the number of random variables is large and the available memory is limited. However, there have been only a few applications of BNs for activity inference. Patterson et al. [17] used a dynamic Bayesian network (DBN) for a quite different goal, that is, inferring transportation modes. The transportation mode could be more strictly classified than the activity. Also, the number of its states, for example, 4 transportation modes, was far less than that of daily activities. We also adopted the BN-based approach to benefit from its advantage to predict potential subsequent activities, given a series of previous states. Our BN model is trained and used to predict a number of activities, that is, 17 activities defined in the ATUS dataset.

**2.3. Localization.** For location-based activity recognition, choosing the underlying localization technology is important. The localization techniques can be roughly classified into (1) GPS-based [14, 15, 17], (2) Wi-Fi-based [18, 19], and (3) Bluetooth-based approaches [20]. The traditional location-based activity recognition systems mainly relied on GPS-based localization. However, GPS-based localization is inherently not appropriate for place-level (or store-level) localization, because the accuracy of GPS significantly decreases in downtown building valley and does not work indoors. Thus, Wi-Fi and Bluetooth localization techniques are more appropriate for place-level localization. Since Bluetooth-based localization requires intensive deployment of Bluetooth beacon devices over a large city, we adopt Wi-Fi-based localization for detecting place in/out events. We utilize an open-source Wi-Fi fingerprinting tool, Elekspot [21].

**2.4. Location Prediction.** Location prediction has been investigated by many researchers, since it can enable interesting and promising applications such as targeted mobile advertising and efficient wireless communication. BreadCrumbs [22] uses location prediction to facilitate the wireless communication of mobile users. WhereNext [23] predicts the next location of moving objects by matching a new GPS trajectory to a set of similar GPS trajectories. NextPlace [24] presents a prediction model based on nonlinear time series analysis in order to predict not only the next location but also the arrival and residence time of a mobile user. Exploiting frequent

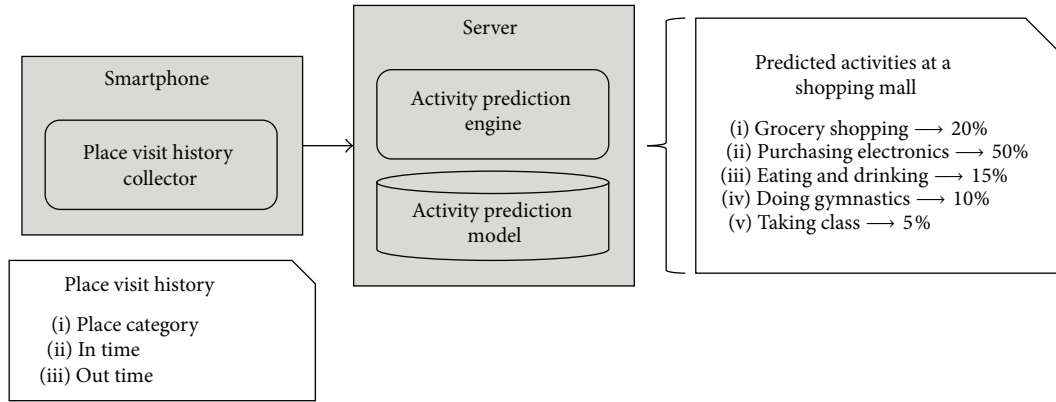


FIGURE 3: Overall system architecture.

sequential place visit patterns in a large shopping mall, VisitSense [25] provides a probabilistic prediction model based on Bayesian networks. Different from location prediction, we investigate activity prediction. Due to the different characteristics of people's daily activity patterns, for accurate prediction, we need to carefully devise a prediction algorithm that is a bit different. We briefly present the related work of activity prediction in the following subsection.

**2.5. Activity Prediction.** Activity prediction has not been as heavily investigated as activity recognition. The goal of activity prediction is to predict the future activity occurrence. Similar to location prediction, the basic approach is to analyze sequential patterns. However, activity prediction requires more sophisticated prediction models, since activity patterns are more complex. For example, diverse activities can happen in the same location. More specifically, diverse activities such as eating, sleeping, and working can happen at home. This means that the uncertainty of activity prediction is higher than that of location prediction. Recently, there have been some research efforts [26–28] in robotics. Their main goal has been to predict activities to facilitate interaction between humans and robots. Consequently, they try to predict fine-grained physical activities such as taking an object and moving an object and, thus, mainly rely on computer vision techniques. Instead, we attempt to predict daily routine activities that span a relatively longer term, for example, shopping in a mall, by using unobtrusive smartphone sensors such as Wi-Fi or GPS. With this approach, we can enable activity-aware applications in an unobtrusive way.

### 3. System Overview

**3.1. Usage Scenario.** As a usage scenario of activity prediction system in a smart city, we can imagine activity-aware mobile advertising at a commercial complex. On a Friday evening, a newly married couple, Tom and Jane, plan to have dinner at the Mall of America (or Macy's) downtown. After finishing his work, Tom drives to nearby Starbucks to pick up Jane, who is waiting for him. Tom is worried about where to go because they have not decided yet. After parking his car at a parking

lot, and entering the Mall of America, Tom is alerted from his mobile phone. Tom's mobile phone shows a map of the Mall of America that displays some attractive places for him. The place advertising is based on which activities Tom will be highly likely to do at the Mall of America. Therefore, Tom can see highly relevant and attractive place advertisements. After looking over some advertisements, Tom decides to go to an Italian restaurant Giovanni's that specially offers a bottle of luxurious Italian wine to customers who have a course menu. At the same time, Jane's mobile phone shows a map that displays some advertisements for shopping places, such as jewelry shops attracting her.

**3.2. Overall System Architecture.** We present a *place-history-based activity prediction system*, *Agatha*, as an enabler of the activity-aware mobile services in a smart city. Figure 3 shows the overall system architecture. The system consists of *place visit history collectors* running on users' smartphones and an *activity prediction engine* running on a remote cloud server. The place visit history collectors are responsible for detecting a user's place in/out events and reporting the place visit history to the prediction engine. While people go around a city, their smartphones collect place in/out events based on Wi-Fi-based localization. The prediction engine performs feature extraction and activity prediction using the place-related events collected from mobile users. Assuming the previous usage scenario, when Tom and Jane come into the Mall of America, their mobile phones automatically upload the logged place visit history to the prediction engine. Then, the engine predicts their highly likely activities at the commercial complex. Based on such prediction, mobile advertising application providers offer target ads or coupons to them.

**3.3. Place Visit History Collector.** Detecting place in and out events is an essential component for place-history-based activity prediction. We develop a place visit history collector for the place in/out event detection on users' smartphones. It basically employs a Wi-Fi-based localization approach for place-level localization as mentioned in Related Work. For Wi-Fi fingerprinting, we use an open-source tool called



Elekspot [21]. It can uniquely identify a place (or room) by using a set of accessible Wi-Fi access points (APs) and the corresponding received signal strength indicator (RSSI) values or fingerprints. One of the important features of the Elekspot is that it exploits a crowdsourcing approach, in which users of Wi-Fi-enabled mobile devices contribute fingerprints and help to build a radio map in a collective manner.

Although the Elekspot is a very useful tool for place-level localization, it is still not enough to detect place in/out events. For place in/out event detection, we further develop a simple and efficient algorithm based on the Elekspot. The Elekspot periodically performs Wi-Fi scanning and resolves the current place. From the Elekspot, the place visit history collector periodically receives the Elekspot's place update events that consist of place name, place category, place geographic coordinates, and confidence value. It then determines whether a user moves in or out of a place by comparing the current and the previous place name. To detect false positives and negatives, it uses the confidence value that indicates how well the Wi-Fi fingerprints match. By monitoring a threshold confidence value  $\alpha$  ( $0 < \alpha < 1$ ), it finally detects place in/out events.

For place visit history collection, we define a *place visit event*,  $V(u_i, p_j) = \langle \text{uid}, \text{pid}, t_{\text{in}}, t_{\text{out}} \rangle$ , where the uid is an identifier of the user  $u_i$ , the pid is an identifier of the place  $p_j$ , the  $t_{\text{in}}$  is the time when the user  $u_i$  moves into the place  $p_j$ , and the  $t_{\text{out}}$  is the time when the user  $u_i$  moves out of the place  $p_j$ . The place visit event  $V(u_i, p_j)$  is generated by composing the primitive events, that is, *place in* and *place out* events detected by the place visit history collector. They can be represented as follows:  $I(u_i, p_j) = \langle \text{uid}, \text{pid}, t_{\text{in}}, \text{null} \rangle$  for a place in event and  $O(u_i, p_j) = \langle \text{uid}, \text{pid}, \text{null}, t_{\text{out}} \rangle$  for a place out event.

**3.4. Activity Prediction Engine.** The prediction engine builds and updates a causality-based activity prediction model that is effectively and flexibly learned by using Bayesian networks. It operates in two modes: offline learning and online prediction. In offline learning mode, the engine trains the activity model by using training data collected from a large number of people. In online prediction mode, the prediction engine receives a large number of place visit histories from many users' smartphones. It extracts and estimates features previously defined for an activity prediction model. It then predicts activities based on the trained activity model. Finally, it performs postprocessing by using the profile knowledge previously extracted at the offline learning mode.

## 4. Causality-Based Activity Prediction Model

To develop the activity prediction engine, we use Bayesian networks (BNs) that can capture activity causality and activity sequence patterns effectively:

- (i) *Activity causality*: activities are influenced not only by the current contexts but also by the history of the previous contexts.
- (ii) *Activity sequence pattern*: people's daily activities imply some patterns. For example, it is highly likely

that average housewives do personal care at home and afterward go grocery shopping at a shopping mall. Meanwhile, single career women would do personal activities at home, work at their offices, and then go for socializing and relaxing at a commercial complex.

Our activity prediction model learns the activity sequence patterns by using BNs and predicts potential subsequent activities based on the history of *influential contexts*. We take visit place, visit time, visit duration, and transportation mode from ATUS data and use them as the influential contexts.

The challenges in building a BN come from scalability and efficiency. The model has a number of random variables corresponding to the influential contexts. Also, each random variable has a large number of states. For example, there are 17 activity types in our activity category. This increases the complexity significantly in terms of memory consumption and learning time. To address these challenges, we consider conditional independence and develop a *hierarchical Bayesian network*.

We use the term *place* as below:

- (i) A *place* is a structured space in a city where people perform some activities. It may be either an open or a closed space.
- (ii) A *simple place* is where only one or two typical activities occur, for example, restaurant, and grocery store.
- (iii) A *complex place* is where multiple simple places are colocated so that diverse activities can occur. For example, in a shopping mall, diverse activities such as shopping, socializing, eating, and entertaining occur [29, 30].

It is also important to select the categories of places and activities properly in designing and evaluating an activity prediction system. We use the classification of tier 1 activity in the ATUS dataset [9]. The categories of places and activities are summarized below.

### 4.1. Features for Activity Prediction

- (i) *Activity* is denoted by  $A$  and the domain of  $A$  by  $\text{dom}(A)$ :

$\text{dom}(A) = \{\text{personal care, household activities, caring for household members, caring for non-household members, work, education, consumer purchases, personal care service, household services, government services, eating and drinking, socializing, sports, religious activities, volunteer activities, telephone calls, others}\}.$

- (ii) *Visit place* is denoted by  $P$ :

$\text{dom}(P) = \{\text{home, workplace, other's home, restaurant, worship place, grocery store, other store, school, outdoors away from home, library, bank, gym, post office, other place}\}.$

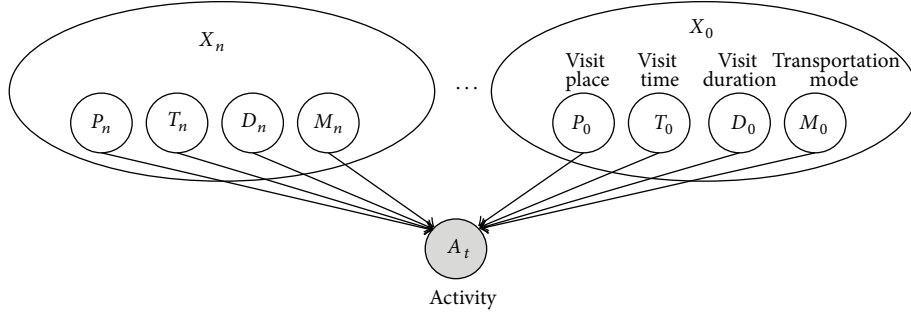


FIGURE 4: A flat Bayesian network for activity prediction.

(iii) *Visit time* is denoted by  $T$ :

$\text{dom}(T) = \{t_i \mid t_i \text{ is a discretized state. } 0 \text{ hour} \leq R_{\min}(t_i) < 24 \text{ hours, where } R_{\min}(t_i) \text{ is the minimum value of the range of state } t_i\}.$

(iv) *Visit duration* is denoted by  $D$ :

$\text{dom}(D) = \{d_i \mid d_i \text{ is a discretized state. } 0 \text{ min} \leq R_{\min}(d_i), \text{ where } R_{\min}(d_i) \text{ is the value of the minimum value of the range of state } d_i\}.$

(v) *Transportation mode* is denoted by  $M$ :

$\text{dom}(M) = \{\text{car, walk, bus, bicycle, other mode}\}.$

We use  $X$  to denote the influential contexts collectively; that is,  $X = \{P, T, D, M\}$ . The states of  $A$  and  $P$  are determined from ATUS data. The two continuous random variables,  $T$  and  $D$ , are discretized. We further denote the current activity and influential contexts by  $A_0$  and  $X_0 = \{P_0, T_0, D_0, M_0\}$  and the  $i$ th previous activities and contexts by  $A_i$  and  $X_i = \{P_i, T_i, D_i, M_i\}$ ,  $1 \leq i \leq n$ .

**4.2. Hierarchical Bayesian Network.** Causality plays an important role in the process of constructing a BN [31]. Proper consideration of the causality relations reduces the complexity of the BN especially when building a large-scale network consisting of a large number of random variables.

**Challenges.** To show a snapshot of the problem, we consider a straightforward BN shown in Figure 4, where  $A_0$  is connected from each influential context,  $X_i$ ,  $0 \leq i \leq n$ , and each  $X_i$  is in turn connected by  $X_{i+1}$  in a sequence. The model is similar to a naïve Bayes model. We call this simple network a flat BN for comparison with our hierarchical model. The model can be expressed in the following form under the conditional independence assumption of BN:

$$P(A_0, X_0, X_1, \dots, X_n) = P(A_0 \mid X_0, X_1, \dots, X_n) \cdot \left[ \prod_{i=0}^{n-1} P(X_i \mid X_{i+1}) \right] P(X_n). \quad (1)$$

The flat BN is mathematically valid. However, in practice, it is highly impractical due to the limitations in resources, for

example, limited memory, and learning time. More specifically, the model would require a table with  $s^{n+1}$  entries. Here,  $s$  is the number of states of the individual random variables, and  $n$  is the number of random variables  $\{X_0, X_1, \dots, X_n\}$ . Also, the learning time would increase exponentially according to  $n$ . For instance, to compute the joint probability  $P(A_0, X_0, X_1, \dots, X_n)$ , we should compute the conditional probability  $P(A_0 \mid X_0, X_1, \dots, X_n)$ , where  $n+1$  variables of  $s$  dimension should be considered in the conditional part, which requires  $s^{n+1}$  dimensional computation.

**Proposed Activity Prediction Model.** Figure 5 shows the graphical representation of the hierarchical BN. It significantly reduces the complexity in dealing with a large number of random variables for the influential contexts. It is also very efficient in terms of memory consumption and learning time.

To design the hierarchical BN, we carefully consider causality relations and probabilistic independence among variables, avoiding excessive links between the nodes for contexts and activities.

An important feature of the BN is that previous activities are inferred from previous influential contexts. The influential contexts are observable nodes, and the activities are latent nodes that are not known in real situations. Then, the current activity is inferred from the history of previous influential contexts.

For the subdomain  $X_i = \{P_i, T_i, D_i, M_i\}$ , we build the following model by using the causality between current activity and current influential contexts. That is, the current activity is inferred from the current place, time of day, activity duration, and transportation mode:

$$P(A_i, X_i) = P(A_i, P_i, T_i, D_i, M_i) = P(A_i \mid P_i) \cdot P(A_i \mid T_i) P(A_i \mid D_i) P(A_i \mid M_i) P(P_i \mid T_i) \cdot P(D_i \mid T_i, P_i) P(T_i) P(M_i). \quad (2)$$

Then, the proposed hierarchical BN is described as follows:

$$P(A_0, A_1, \dots, A_n, X_0, X_1, \dots, X_n) = \prod_{i=0}^n P(A_i \mid A_{i+1}, \dots, A_n, X_0, X_1, \dots, X_n) \cdot P(X_0, X_1, \dots, X_n) \quad (3)$$

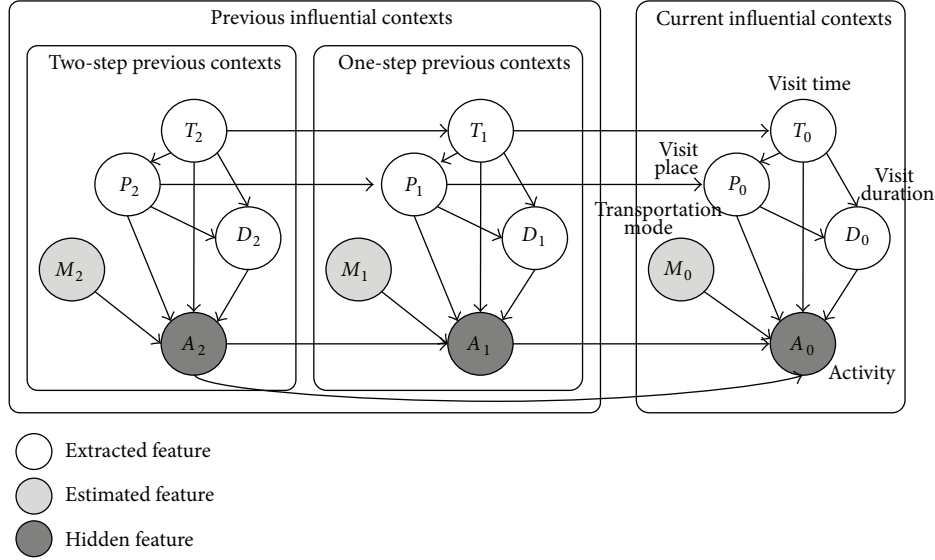


FIGURE 5: A hierarchical Bayesian network.

$$= \left[ \prod_{i=0}^{n-1} P(A_i | A_{i+1}, \dots, A_n, X_i) P(X_i | X_{i+1}) \right] \quad (4)$$

$$\begin{aligned} & \cdot P(A_n | X_n) P(X_n) \\ &= P(A_0 | A_1, A_2, X_0) P(A_1 | A_2, X_1) P(X_0 | X_1) \\ & \cdot P(X_1 | X_2) P(X_2). \end{aligned} \quad (5)$$

The model is constructed with important independence assumptions. First, we assume the following independence among contexts,  $X_i$ , and between activity,  $A_i$ , and contexts,  $X_i$  (i.e., intercontext independence):

- (i)  $X_i$  is independent of  $X_j$ , where  $j \neq i + 1$ , for  $i$  such that  $1 \leq i \leq n$ .
- (ii)  $A_i$  is conditionally independent of  $X_{i+1}, X_{i+2}, \dots, X_n$ , given  $A_{i+1}$  and  $A_{i+2}$ , for  $i$  such that  $1 \leq i \leq n$ .

Unlike the flat BN, with independence, we do not need to make links from the previous influential contexts to the current activity. Instead, the previous contexts,  $X_i = \{P_i, T_i, D_i, M_i\}$ , at each stage are first connected to the activity  $A_i$  at the stage and serially connected to the current activity  $A_0$ . (The assumption is made by deriving (4) from (3) in the above equations.)

We also assume the following independence among the activities,  $A_i$ :

- (iii)  $A_0$  is independent of  $A_i$ , for  $i$  such that  $3 \leq i \leq n$ .

This independence allows us to consider only recent previous influential contexts. (Note that the recent previous activities  $A_1$  and  $A_2$  are latent variables.) It might be good to compromise between accuracy and complexity.

**Learning and Inference Algorithm.** We use the EM algorithm [32] for learning the proposed hierarchical BN. In the EM algorithm, the  $E$ -step estimates missing values in the current model and the  $M$ -step maximizes likelihood. The EM

algorithm allows us to train the model with latent variables, for example,  $A_1$  and  $A_2$ .

**Continuous Feature Discretization.** Continuous features such as visit time and visit duration should be appropriately discretized. For the discretization of the visit time, we simply divide a day into 24 states (hours). For the visit duration, a more sophisticated method is required, because the distribution of the duration can be highly skewed. We use a simple and useful method based on a logarithm function:  $\log_x d = s$ . Here,  $d$  is a continuous value of duration that ranges from 0 to the maximum duration, and  $s$  is a discrete state of the duration. We use 24 states for the duration and a day ( $24 \times 60$ ) as the maximum duration. Then,  $x$  is determined by  $\log_x(24 \times 60) = 24$ .

**High-Level Feature Estimation.** Transportation mode is not directly extracted but rather estimated from the place visit events. For the estimation, we develop a simple heuristic that discriminates the transportation mode based on average speed between consecutive places. If the speed exceeds the average of 30 km/h, people's transportation mode is determined as by car. Note that the average speed can be easily estimated from place visit events by using the following formula:

$$v = \frac{|R(p_j) - R(p_{j-1})|}{V(o_i, p_i) t_{in} - V(o_i, p_{i-1}) t_{out}}. \quad (6)$$

Therein, place  $p_{j-1}$  is the immediately previous place of the place  $p_j$ ,  $R(p_j)$  is the minimum bounded rectangle of place  $p_j$ , and  $V(o_i, p_j)$  is people's visit event into place  $p_j$ .

Alternatively, the transportation mode can be inferred from separate machine learning modules such as HMM and Bayesian networks. For transportation mode inference, Patterson et al. [17] proposed a dynamic Bayesian network, and Zheng et al. [33] proposed a decision-tree-based inference

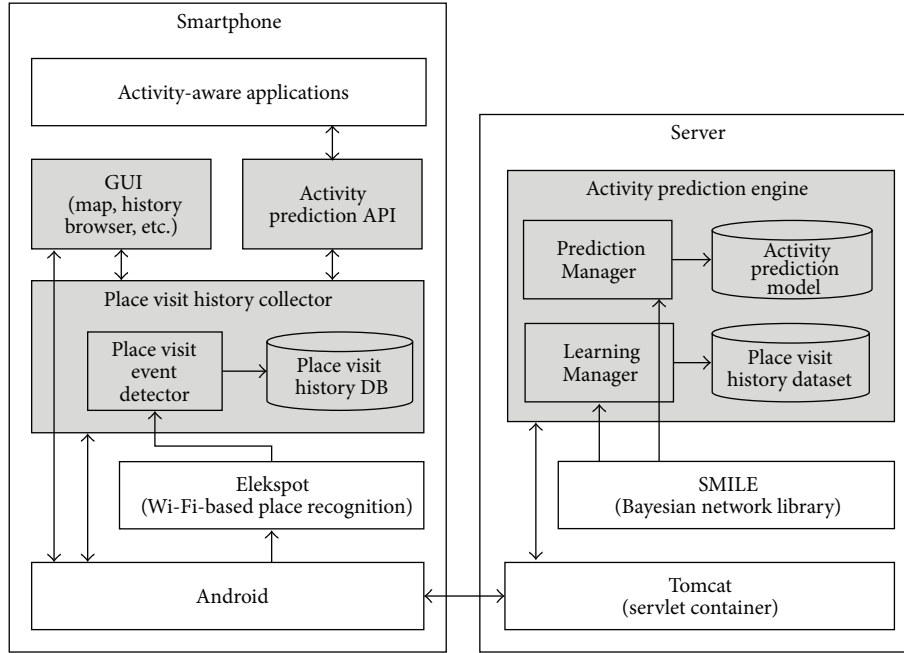


FIGURE 6: Implementation overview of Agatha.

model. Note that our activity model can easily incorporate such method because of the flexibility of Bayesian networks.

## 5. Implementation

We have implemented a prototype of the proposed activity prediction system including the place visit history collector on the mobile side and the activity prediction engine on the server side. The implementation overview is shown in Figure 6. The shaded boxes are developed modules. The place visit history collector has been implemented on an Android smartphone. For collecting place visit history of a user, it uses the APIs provided by Elekspot [21] running on the same smartphone. As mentioned in Section 3 (i.e., System Overview), Elekspot is an accurate room-level localization system based on a Wi-Fi fingerprint matching technique. Because it exploits crowdsourcing for building a Wi-Fi radio map, we can expand our system to a city scale. Elekspot provides APIs that give an access to current place ID, name, and localization confidence. Using Elekspot APIs, the place visit event detector detects place visit events (i.e., enter and departure events) by periodically monitoring the current place ID. Then, the place visit event is collected and stored on SQLite, a lightweight relational database provided by Android platform. Also, we have implemented a GUI that shows the current place on a map, lists place visit history, and provides user options. The GUI is shown in Figure 7. We provide activity prediction APIs that developers can use to develop interesting activity-aware applications on smartphones.

The prediction engine has been implemented on a server. It mainly consists of the Learning Manager and Prediction Manager. The Learning Manager trains the BN-based activity

prediction model by using a place visit history dataset collected from many users. The Prediction Manager predicts the next activity by using the trained activity prediction model. The prediction engine has been implemented based on Java Web technology. Main components such as the Learning Manager and Prediction Manager have been implemented as servlets. We used Apache Tomcat 5.x as a servlet container. For the activity prediction model, we used Java APIs provided by SMILE [34]. SMILE (Structural Modeling, Inference, and Learning Engine) is a platform-independent library that implements graphical probabilistic models such as Bayesian networks. Although SMILE is implanted in C++, we used jSMILE, a Java interface. The prediction engine runs on a Linux machine equipped with Intel Core Duo 2.4 Hz CPU and 4 GB memory.

## 6. Evaluation

**6.1. Dataset.** ATUS provides meaningful data for learning and evaluation of causality-based activity prediction models because they include a large number of real activity sequences recorded in respondents' daily lives. For example, 2008 ATUS data contains about 260,000 activity episodes collected from about 12,000 households. Partridge and Golle [16] also pointed out that time-use surveys such as ATUS can be used as useful data for activity inference in ubiquitous computing. To compare the accuracy of our model with that of the previous approach, we used the 2008 ATUS dataset that is used in the previous work [16]. We used 150,000 activity episodes for learning and 30,000 episodes for evaluation. We exclude travelling activities such as moving in vehicles from the data.



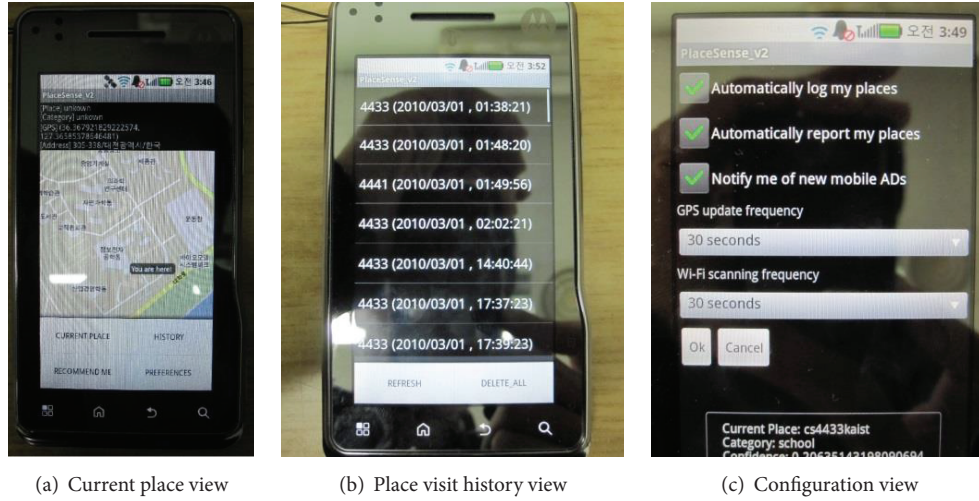


FIGURE 7: (a) Current place view that shows user's current geographical location on a map and shows place name identified by Wi-Fi localization, (b) place visit history log view, and (c) configuration view.

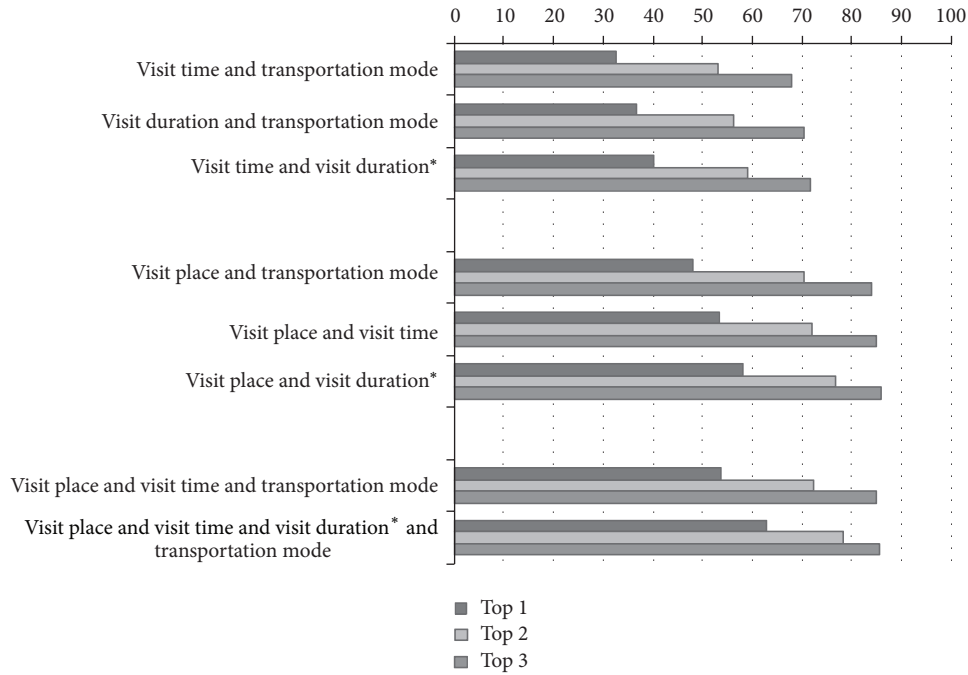


FIGURE 8: Prediction accuracy of the combination of each feature.

**6.2. Effect of Influential Contexts.** We evaluate the effect of the current influential contexts, that is, visit place, visit time, visit duration, transportation mode, and their combinations, on the activity prediction. Figure 8 shows that the prediction accuracy varies depending on different combinations of influential contexts. When all four contexts are given, the prediction model achieves the highest accuracy, that is, up to 86% for top 3 activities, 78% for top 2 activities, and 63% for top 1 activity. In practice, however, the duration is not available at the time of prediction; it can be obtained only after an activity is finished. With the combination of the remaining three contexts, that is, place, time, and transportation,

the prediction accuracy slightly decreases, that is, about 85% for top 3 activities, 73% for top 2 activities, and 54% for top 1 activity. Note that use of additional information such as previous influential contexts can improve the accuracy achieved by current influential contexts.

Figure 8 also shows that a moderate level of prediction accuracy can be achieved with combination of only two contexts. If duration is included, the combination of place and duration results in the highest prediction accuracy, 58% accuracy for top 1 activity. If duration is excluded, the combination of place and time shows the highest accuracy. Note that place is the most important context; the combinations

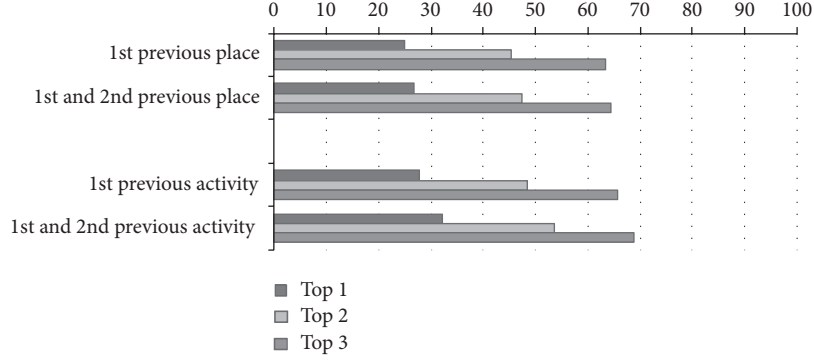


FIGURE 9: Prediction accuracy of the previous influential contexts.

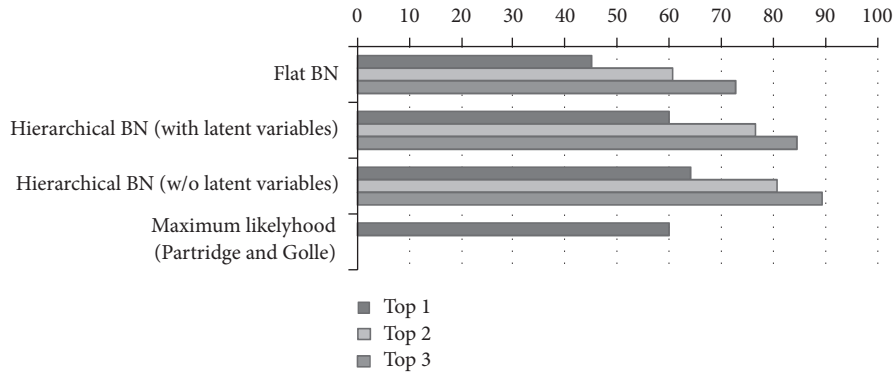


FIGURE 10: Prediction accuracy of different BNs.

including place show higher accuracy than the combinations without place information.

We also evaluate the effect of the previous contexts on the current activity. Figure 9 shows that it is possible to predict activities to some extent only using the previous contexts; for top 3 activities, about 64% of accuracy could be achieved with the previous place information only. Among them, the combination of the immediate previous activity and the second previous activity is most influential to the current activity. In practice, however, it is not straightforward to know the previous activities with certainty; they also should be inferred from relevant contexts. The proposed hierarchical BN infers the previous activities from the previous influential contexts to better predict the current activity together with current influential contexts.

**6.3. Prediction Accuracy of Hierarchical BN.** We compare the prediction accuracy of the hierarchical BN with that of the flat BN as well as the maximum likelihood estimator [16]. Figure 10 shows that the accuracy of the hierarchical BN is better than the other prediction models. We evaluate the prediction accuracy of the hierarchical BN in two different cases: (1) with and (2) without latent variables. The case of “with latent variables” means that previous activities are not given when performing prediction. Instead, the previous

activities are inferred from the previous influential contexts such as visit place, visit time, and visit duration. With latent variables, the hierarchical BN shows about 85% of prediction accuracy for top 3 activities, about 75% for top 2 activities, and about 60% for top 1 activity, respectively. If the previous activities are given when performing prediction, the accuracy increases more than 5%. Without latent variables, the hierarchical BN shows about 90% of accuracy for top 3 activities, about 80% for top 2, and about 65% for top 1, respectively.

The accuracy of the flat BN is about 73% for top 3, 61% for top 2, and 45% for top 1, respectively. These are much lower than those of the hierarchical BN. This is mainly because the flat BN is very inefficient in terms of memory use. In the given experiment setting (i.e., 4 GB memory), we could not include all the random variables when training the model due to memory limitation. More specifically, we could include only the immediate previous influential contexts in the flat BN. We failed to train the flat BN with SMILE/GeNIe [34] because the tool reached maximum memory use. The flat BN consumes memory heavily because the conditional probability table of the current activity node increases exponentially proportional to the number of directly linked nodes. Note that the hierarchical BN consumes only about 116 MB to deal with immediate and second previous influential contexts. This result shows that the hierarchical BN achieves much higher

prediction accuracy than the flat BN, while using much less memory (i.e., almost 10 times less memory use).

The maximum likelihood estimator shows relatively high accuracy, that is, 60%, compared to the flat BN. Note that the accuracy has been achieved under the assumption that the previous activities are given. In other words, the activity prediction is done from the previous activity history, not from the previous location history. In the same situation (i.e., without latent variables), the hierarchical BN achieves a top 1 prediction accuracy of about 65%, which is 5% higher than that of the maximum likelihood estimator.

Even though the prediction accuracy might not be significantly improved over that of the maximum likelihood estimator, the proposed hierarchical BN is more efficient than the maximum likelihood estimator due to its conditional independence assumptions, so it is adequate to build a city scale system. Generally, the maximum likelihood estimator requires more computational resources such as memory for training and inference, as it maintains a global joint probability distribution table. Also, the proposed hierarchical BN is more robust than the maximum likelihood estimator, because the proposed model handles the latent variables well by using the EM algorithm.

## 7. Conclusion

In this paper, we propose a place visit history-based activity prediction system to enable activity-aware mobile services in smart cities. For this purpose, the system adopts a causality-based activity prediction model using Bayesian networks. It uses location-based activity causality, that is, the idea that where you have been so far influences what you will do next. The causalities are inherent in common daily activities that usually disclose an activity sequence pattern. The model effectively captures such activity causalities as well as activity sequence patterns from so-called influential contexts such as visit place, visit time, duration, and transportation mode. We develop a hierarchical BN to efficiently tackle the complexity of learning activity causalities between many diverse influential contexts. We evaluate the prediction model using the American Time-Use Survey (ATUS) dataset that includes more than 10,000 people's location and activity history. Evaluation results show that it can predict users' potential activities with up to 90% of accuracy for top 3 activities, more than 80% for top 2 activities, and about 65% for top 1 activity while considering a relatively large number of daily activities defined in the ATUS dataset.

## Conflict of Interests

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

## Acknowledgment

This paper was partially supported by the Education and Research Promotion Program of KOREATECH.

## References

- [1] Y. Zheng, Y. Liu, J. Yuan, and X. Xie, "Urban computing with taxicabs," in *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*, pp. 89–98, Beijing, China, September 2011.
- [2] N. J. Yuan, Y. Zheng, L. Zhang, and X. Xie, "T-Finder: a recommender system for finding passengers and vacant taxis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2390–2403, 2013.
- [3] J. Krumm, "Ubiquitous advertising: the killer application for the 21st century," *IEEE Pervasive Computing*, vol. 10, no. 1, pp. 66–73, 2010.
- [4] B. Kim, J.-Y. Ha, S. J. Lee et al., "AdNext: a visit-pattern-aware mobile advertising system for urban commercial complexes," in *Proceedings of the 12th International Workshop on Mobile Computing Systems and Applications (HotMobile '11)*, pp. 7–12, Phoenix, Ariz, USA, March 2011.
- [5] J. L. Toole, N. Eagle, and J. B. Plotkin, "Spatiotemporal correlations in criminal offense records," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 4, article 38, 2011.
- [6] Minority Report, April 2015, <http://www.imdb.com/title/tt0181689/>.
- [7] M. C. Sala, K. Patridge, L. Jacobson, and J. B. Begole, "An exploration into activity-informed physical advertising using PEST," in *Pervasive Computing: 5th International Conference, PERVASIVE 2007, Toronto, Canada, May 13–16, 2007. Proceedings*, vol. 4480 of *Lecture Notes in Computer Science*, pp. 73–90, Springer, Berlin, Germany, 2007.
- [8] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [9] American Time Use Survey, <http://www.bls.gov/tus/>.
- [10] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21–23, 2004. Proceedings*, vol. 3001 of *Lecture Notes in Computer Science*, pp. 1–17, Springer, Berlin, Germany, 2004.
- [11] P. Lukowicz, J. A. Ward, H. Junker et al., "Recognizing workshop activity using body worn microphones and accelerometers," in *Proceedings of the 2nd International Conference on Pervasive Computing (Pervasive '04)*, pp. 18–32, Vienna Austria, April 2004.
- [12] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall, "Recognizing daily activities with RFID-based sensors," in *Proceedings of the 11th ACM International Conference on Ubiquitous Computing (UbiComp '09)*, pp. 51–60, ACM, Orlando, Fla, USA, October 2009.
- [13] M. Philipose, K. P. Fishkin, M. Perkowitz et al., "Inferring activities from interactions with objects," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 50–57, 2004.
- [14] D. Ashbrook and T. Starner, "Using GPS to learn significant locations and predict movement across multiple users," *Personal and Ubiquitous Computing*, vol. 7, no. 5, pp. 275–286, 2003.
- [15] L. Liao, D. Fox, and H. Kautz, "Location-based activity recognition using relational Markov networks," in *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI '05)*, pp. 773–778, Edinburgh, Scotland, July 2005.
- [16] K. Partridge and P. Golle, "On using existing time-use study data for ubiquitous computing applications," in *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*, pp. 144–153, Seoul, South Korea, September 2008.

- [17] D. J. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring high-level behavior from low-level sensors," in *Proceedings of the 5th International Conference on Ubiquitous Computing (UbiComp '03)*, pp. 73–89, Seattle, Wash, USA, October 2003.
- [18] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proceedings of the 9th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, vol. 2, pp. 775–784, IEEE, Tel Aviv, Israel, 2000.
- [19] A. LaMarca, Y. Chawathe, S. Consolvo et al., "Place lab: device positioning using radio beacons in the wild," in *Proceedings the 3rd International Conference on Pervasive Computing (Pervasive '05)*, pp. 116–133, Munich, Germany, May 2005.
- [20] E. O'Neill, V. Kostakos, T. Kindberg et al., "Instrumenting the city: developing methods for observing and understanding the digital cityscape," in *UbiComp 2006: Ubiquitous Computing: 8th International Conference, UbiComp 2006 Orange County, CA, USA, September 17–21, 2006 Proceedings*, vol. 4206 of *Lecture Notes in Computer Science*, pp. 315–332, Springer, Berlin, Germany, 2006.
- [21] M. Lee, H. Yang, D. Han, and C. Yu, "Crowdsourced radiomap for room-level place recognition in urban environment," in *Proceedings of the 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops '10)*, pp. 648–653, Mannheim, Germany, April 2010.
- [22] A. J. Nicholson and B. D. Noble, "BreadCrumbs: forecasting mobile connectivity," in *Proceedings of the 14th ACM Annual International Conference on Mobile Computing and Networking (MobiCom '08)*, pp. 46–57, ACM, San Francisco, Calif, USA, September 2008.
- [23] A. Monreale, F. Pinelli, R. Trasarti, and F. Giannotti, "WhereN-ext: a location predictor on trajectory pattern mining," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD '09)*, pp. 637–646, Paris, France, June 2009.
- [24] S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell, "NextPlace: a spatio-temporal prediction framework for pervasive systems," in *Pervasive Computing: 9th International Conference, Pervasive 2011, San Francisco, USA, June 12–15, 2011. Proceedings*, vol. 6696 of *Lecture Notes in Computer Science*, pp. 152–169, 2011.
- [25] B. Kim, S. Kang, J. Ha, and J. Song, "VisitSense: sensing place visit patterns from ambient radio on smartphones for targeted mobile ads in shopping malls," *Sensors*, vol. 15, no. 7, pp. 17274–17299, 2015.
- [26] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *Proceedings of the European Conference on Computer Vision (ECCV '12)*, pp. 1–14, Florence, Italy, October 2012.
- [27] K. P. Hawkins, N. Vo, S. Bansal, and A. F. Bobick, "Probabilistic human action prediction and wait-sensitive planning for responsive human-robot collaboration," in *Proceedings of the 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids '13)*, pp. 499–506, Atlanta, Ga, USA, October 2013.
- [28] H. S. Koppula and A. Saxena, "Anticipating human activities for reactive robotic response," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '13)*, p. 2071, IEEE, Tokyo, Japan, November 2013.
- [29] S. Wesley, M. LeHew, and A. G. Woodside, "Consumer decision-making styles and mall shopping behavior: building theory using exploratory data analysis and the comparative method," *Journal of Business Research*, vol. 59, no. 5, pp. 535–548, 2006.
- [30] K. J. Shelley, "Developing the American time use survey activity classification system," *Monthly Labor Review*, vol. 128, no. 6, pp. 3–15, 2005.
- [31] U. B. Kjærulff and A. L. Madsen, *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*, Springer, 2008.
- [32] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.
- [33] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on GPS data," in *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*, pp. 312–321, ACM, Seoul, Republic of Korea, September 2008.
- [34] M. J. Druzdzel, "SMILE: structural modeling, inference, and learning engine and GeNIe: a development environment for graphical decision-theoretic models," in *Proceedings of the 16th National Conference on Artificial Intelligence and the 11th Conference on Innovative Applications of Artificial Intelligence (AAAI '99/IAAI '99)*, pp. 902–903, Orlando, Fla, USA, July 1999.