When And Where Next: Individual Mobility Prediction*

Győző Gidófalvi†
Geodesy and Geoinformatics
KTH Royal Institute of Technology
Stockholm, Sweden
gyozo.gidofalvi@abe.kth.se

Fang Dong
Geodesy and Geoinformatics
KTH Royal Institute of Technology
Stockholm, Sweden
fdong@kth.se

ABSTRACT

The ability to predict *when* an individual mobile user will leave his current location and *where* we will move next enables a myriad of qualitatively different Location-Based Services (LBSes) and applications. To this extent, the present paper proposes a statistical method that explicitly performs these related temporal and spatial prediction tasks in three continuous, sequential phases. In the first phase, the method continuously extracts grid-based staytime statistics from the GPS coordinate stream of the location-aware mobile device of the user. In the second phase, from the grid-based staytime statistics, the method periodically extracts and manages regions that the user frequently visits. Finally, in the third phase, from the stream of region-visits, the method continuously estimates parameters for an inhomogeneous continuous-time Markov model and in a continuous fashion predicts *when* the user will leave his current region and *where* he will move *next*. Empirical evaluations, using a number of long, real world trajectories from the *GeoLife* data set, show that the proposed method outperforms a state-of-the-art, rule-based trajectory predictor both in terms of temporal and spatial prediction accuracy.

Categories and Subject Descriptors

H.2.8 [**Database Applications**]: [Data mining, Spatial Databases and GIS]

General Terms

Algorithms

Keywords

spatio-temporal data mining, mobility patterns, location prediction, inhomogeneous continuous-time Markov model

*This work was partially supported by TRENoP, the strategic research program in Transport at KTH.

†Corresponding author.

ACM SIGSPATIAL MobiGIS'12, November 6, 2012. Redondo Beach, CA, USA

1. INTRODUCTION

The movements of an individual contain a high degree of regularity [3,13]. An individual regularly visits a small set of location / regions and regularly moves between those locations. The movement regularities can be *temporal*, *periodic* and *sequential*. With the increased availability and adoption of mobile positioning, computing and communication technologies, researchers have quickly identified the utility in extracting these regularities and using them to predict future movement of objects for a broad domain of applications including transport and mobility studies for urban planning, mobile communication network optimization and prefetching for Location Based Services.

Two popular extraction / prediction approaches emerged: discrete-time Markov model based [1–3, 7] and sequential rule / trajectory pattern based [4,5,8,11,14–18]. The methods can also be classified according to what movement information they use to model the movement of objects into methods using a general model for all objects [3–5,7,8,11,14], methods using a type-base model for similar (type of) objects [1, 18] and methods using a specific model for each individual object [2,16,17]. The methods can also be classified according to their definition of Regions Of Interest (ROIs) for prediction and consequently their spatial and temporal scale and granularity into methods using application-specific ROIs (road segment, network cell, sensors, etc.) [1, 3, 5, 11, 16], density-based ROIs [2, 4, 8, 14, 17, 18] and grid-based ROIs [4, 7, 11, 14]. Finally and most importantly, the methods can also be classified according to their prediction provision into methods that provide only sequential spatial predictions (location of next ROI) [1–3,7,16,18] and methods that provide spatio-temporal predictions [4, 5, 8, 11, 14, 17].

A major shortcoming of the methods that can provide spatio-temporal predictions –which are all sequential rule / trajectory pattern based– is that it is expensive and difficult to construct and combine individual prediction models with them that capture temporal and period regularities at different scales. Therefore, the present paper proposes the use of a dynamically weighted ensemble of Inhomogeneous Continuous-Time Markov (ICTM) models to simply but effectively capture the temporal-, periodic- and sequential regularities in movements of an object to predict *when* and *where* the object will move *next*.

The rest of this paper is organized as follows. Section 2 defines the prediction problems and gives preliminaries. Section 3 details individual parts of the proposed ICTM-based prediction model. Section 4 empirically evaluates and compares the proposed method. Finally, Section 5 concludes.

2. DEFINITIONS AND PRELIMINARIES

The following section first, in Section 2.1, formalizes the temporal and spatial mobility prediction tasks, then, in Section 2.2 presents basic theory related to the ICTM model.

2.1 Definitions of Mobility Prediction Tasks

Let the time domain be denoted by $\mathbb{T}$ and be modeled as the totally ordered set of non-negative natural numbers $\mathbb{N}^+$. Let the trajectory of a moving object o in the 2-dimensional (2D) space be modeled and defined as a sequence of tuples $S = \langle (l_1, t_1), \ldots, (l_n, t_n) \rangle$, where $l_i \in \mathbb{R}^2$ $(i = 1, \ldots, n)$ describe locations, and $t_1 < \ldots < t_n \in \mathbb{T}$ are irregularly spaced but temporally ordered time instances, i.e., gaps are allowed.

According to the typical linear movement model the continuous movement of the object is assumed to be linear and at constant speed between two consecutive locations. Assuming that the object's position is sampled relatively frequently with respect to the displacement, according to the *discretized movement model*, the object is assumed to be at location l_i during the period $[t_i, t_{i+1})$ $(1 \leq i < n)$.

Then, given the trajectory $S = \langle (l_1, t_1), \ldots, (l_n, t_n) \rangle$ of an object o, according to the discretized movement model the *staytime* of o in any given region of space $R \subset \mathbb{R}^2$ is defined as: $t_{st}^R = \sum_{1 \leq i < n : l_i \in R} (t_{i+1} - t_i)$.

Let $\mathcal{R} = \{R_i, \ldots, R_k\}$ denote a set of spatially contiguous, mutually exclusive regions of space, i.e., $R_i, R_j \subset \mathbb{R}^2$, $R_i \cap R_j = \emptyset$, $i, j \in \{1, \ldots, k\}$, $i \neq j$, such that:

(i) the regions are *prevalent* - the sum of the relative staytimes of the object in the regions of $\mathcal{R}$ is above a user-defined *minimum relative prevalence* parameter, *min_rp*, i.e., $\sum_{R_i \in \mathcal{R}} t_{st}^{R_i} / t_{st}^{\mathbb{R}^2} \geq min\_rp$, and

(ii) the regions are *maximally discriminative* - the total area of the regions in $\mathcal{R}$ is *minimal*.

The set of regions $\mathcal{R} = \{R_i, \ldots, R_k\}$ fulfilling these conditions is termed as *prevalent* and *maximally discriminative regions*. A member of this set $R_i \in \mathcal{R}$ is called a *prevalent* and *maximally discriminative region*, or *pmd-region*.

Given the set of regions $\mathcal{R}$, the object's continuous movement can be further approximated by the object's *region-based trajectory* as a sequence of tuples $S^{\mathcal{R}} = \langle (R_1, t_1^s, t_1^e), \ldots, (R_m, t_m^s, t_m^e) \rangle$, where $R_i \in \mathcal{R}$ $(i = 1, \ldots, m)$ describe regions, and $t_1^s < t_1^e < \ldots < t_m^s < t_m^e \in \mathbb{T}$ are irregularly spaced but temporally ordered time instances, i.e., gaps are allowed.

Given a time instance $t \in \mathbb{T}$ at which the object is in one of the regions $R_j \in \mathcal{R}$, let $\mathcal{H}(t)$ denote the object's *region-based trajectory history* up to time instance t, i.e., $\mathcal{H}(t) = \langle (R_1, t_1^s, t_1^e), \ldots, (R_j, t_j^s, ?) \rangle$, where $t_i^s < t_i^e < t_j^s \leq t$, $i = 1, \ldots, j-1$. Furthermore, for the ease of exposition let $h(S)$ represent the head of an ordered tuple sequence S.

Then, the *temporal mobility prediction task* is informally defined as upon the object's arrival at a region R_j predicting the time of stay at R_j. Formally:

Definition 1. Temporal Mobility Prediction: Given an object o and its *region-based trajectory history* up to time instance t, i.e., $\mathcal{H}(t)$ such that $h(\mathcal{H}(t)) = (R_j, t_j^s = t, ?)$, a user-defined minimum staytime likelihood threshold, *min_stl*, and a time-parameterized discrete random variable $X(t)$ taking on values from the set of regions $\mathcal{R}$, predict the *staytime*, s^*, or equivalently the *departure time*, $t + s^*$, of o as:

$$s^* = \underset{s}{\operatorname{argmin}} \Pr(X(t+s) = R_j | \mathcal{H}(t)) \leq min\_stl.$$

Subsequently, the *spatial mobility prediction task* is informally defined as upon the object's arrival at a region R_j predicting the next region that the object will enter after leaving R_j. Formally:

Definition 2. Spatial Mobility Prediction: Given an object o and its *region-based trajectory history* up to time instance t, i.e., $\mathcal{H}(t)$ such that $h(\mathcal{H}(t)) = (R_j, t_j^s = t, ?)$, and a time-parameterized discrete random variable $X(t)$ taking on values from the set of regions $\mathcal{R}$ predict the *next region*, R_{j+1}^*, that o will start moving to at the afore predicted departure time, $t + s^*$, as:

$$R_{j+1}^* = \underset{R_{j+1} \in \{\mathcal{R} - R_j\}}{\operatorname{argmax}} \Pr(X(t+s^*) = R_{j+1} | \mathcal{H}(t)).$$

2.2 ICTM Model Related Statistical Theory

Adapted from Chapters 4-6 of [12], the following subsections describe basic statistical theory that is relevant for the proposed ICTM modeling and prediction.

2.2.1 Discrete-Time Markov Chain

A discrete-time, finite state stochastic process $\{X_n : n = 0, 1, 2, \ldots\}$ takes on a finite number of possible values, e.g., $X_n \in \{0, 1, 2, \ldots\}$. If $X_n = i$, then the process is said to be in state i at time instance n. Suppose that whenever the state is i, there is a fixed probability $P_{ij} \geq 0$ that the next state will be j. That is, suppose that:

$$\Pr(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P_{ij} \tag{1}$$

for *all* states $i_0, \ldots, i_{n-1}, i, j$ and *all* $n \geq 0$. Such a stochastic process is known as a *discrete-time Markov chain*. Alternatively, the distribution of any future state X_{n+1} is conditionally independent of the past states $X_0, \ldots, X_{n-1}$ given the current state X_n,i.e.:

$$\begin{aligned} \Pr(X_{n+1} &= j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ &= \Pr(X_{n+1} = j | X_n = i) = P_{ij}. \end{aligned} \tag{2}$$

This conditional independence is referred to as the *Markovian property* or *memorylessness*. A Markov chain has *memory* m or *order* m if the future state depends on the past m states only.

2.2.2 Exponential Distribution

A continuous random variable X is said to have an *exponential distribution* with rate parameter $\lambda > 0$ if its probability density function (PDF) is $f(x) = \lambda e^{-\lambda x}$ when $x \geq 0$ and $f(x) = 0$ otherwise, or, equivalently, its cumulative distribution function (CDF) is $F(x) = \Pr(X \leq x) = 1 - e^{-\lambda x}$ when $x \geq 0$ and $f(x) = 0$ otherwise. The following are the relevant properties of the exponential distribution:

1. The expectation or mean of an exponentially distributed random variable X is $1/\lambda$.
2. The exponentially distributed random variable X is said to be without memory, or *memoryless* because for *all* $t, s \geq 0$: $\Pr(X > t + s | X > t) = \Pr(X > s)$.
3. The exponential distribution is the *only* distribution that possesses the memoryless property.

2.2.3 Continuous-Time Markov Process

The continuous time analogue of a discrete time Markov chain is a *Continuous-Time Markov (CTM) process*. Namely, a continuous-time stochastic process $\{X(t) : t \geq 0\}$ takes on

values from set of possible values, e.g., $X_n \in \{0, 1, 2, \ldots\}$, which in general is referred to the *state space*. The process $\{X(t) : t \geq 0\}$ is a CTM process if for all $s, t \geq 0$ and nonnegative integers $i, j, x(u), 0 \leq u < t$:

$$\begin{aligned} \Pr(X(t+s) &= j | X(t) = i, X(u) = x(u), 0 \leq u < t) \\ &= \Pr(X(t+s) = j | X(t) = i). \end{aligned} \quad (3)$$

That is, a CTM process is a stochastic process that has the Markovian property, i.e., the distribution of the future state $X(t+s)$ is conditionally independent of the past states $X(u), 0 \leq u < t$ given the current state $X(t)$. If the conditional probability $\Pr(X(t+s) = j | X(t) = i)$ is independent of t, then the CTM process is said to have *stationary / homogeneous* transition probabilities or it is said to be stationary / homogeneous. Inversely, if the conditional probability depends on t, then the CTM process is said to have *non-stationary / inhomogeneous* transition probabilities or it is said to be non-stationary / inhomogeneous. The latter type of process is explicitly referred to as an *Inhomogeneous Continuous-Time Markov (ICTM) process*. Analogous to the discrete-time case, a CTM process has memory m or order m if the future state depends on the past m unique consecutive states only.

Let T_i, referred to as the *holding time*, denote the amount of time that a CTM process stays in state i before making a transition to some other state. Due to the Markovian property, assuming the process has entered state i at some time, say, time 0, the probability that the process will remain in state i for at least s more time units given that the process has already been in state i up to the current time $t > 0$ is just the unconditional probability that the process will remain in state i for at least s time units. That is:

$$\Pr(T_i > t + s | T_i > t) = \Pr(T_i > s) \quad (4)$$

for *all* $s, t \geq 0$. Hence, T_i must be memoryless and must thus be exponentially distributed. This gives rise to an alternative definition of a CTM process, i.e., a stochastic process is a CTM process *if and only if* the holding times in each state are exponentially distributed with some mean.

3. METHOD

The proposed method essentially can be divided into two sequentially dependent parts: the preprocessing part, where the trajectory is transformed (Sections 3.1-3.4), and the prediction part, where the model parameters are estimated from historical part of the transformed trajectory and the model is applied to the most recent part of the transformed trajectory (Section 3.5).

3.1 Grid-Aggregation of Mobility Statistics

In order to deal with the noise in GPS measurements and make online processing efficient, the proposed method employs the simple but effective grid-based spatial aggregation of temporal mobility statistics as follows. Let $\mathcal{G}$ denote a uniform *grid* and $g_1, g_2, \ldots$ its grid cells that have side length *glen* and that uniformly partition the 2D Euclidean space. Given $\mathcal{G}$, the proposed method uses straight forward coordinate arithmetic to map the individual locations of the object's trajectory $S = \langle (l_1, t_1), \ldots, (l_n, t_n) \rangle$ to a grid-based trajectory $S^{\mathcal{G}} = \langle (g_1, t_1), \ldots, (g_n, t_n) \rangle$ and incrementally maintains for every grid cell $g_i \in \mathcal{G}$ the *staytime* of the object in g_i, which, according to the discretized movement model, is defined as: $t_{st}^{g_i} = \sum_{1 \leq j < n : g_j = g_i} (t_{j+1} - t_j)$.

3.2 Grid-based Detection of pmd-regions

A pmd-region can span several spatially contiguous grid cells. After observing that, like many other natural geospatial processes [10, 13], the staytimes of individual objects in grid cells also exhibit the powerlaw distribution, for which the elements in the "head" part of the distribution tend to be qualitatively different and have distinct semantic meaning, the proposed method iteratively, always starting with the grid cell with the highest staytime g^* among the so far not yet grouped grid cells, grows and extracts groups of contiguous dense grid cells (i.e., $t_{st}^{g_i} > \overline{t_{st}}$) within the *maxR*-neighborhood of the current g^* until the prevalence of the so extracted pmd-regions reaches the *min_rp*-requirement.

3.3 Tracking the Evolution of pmd-regions

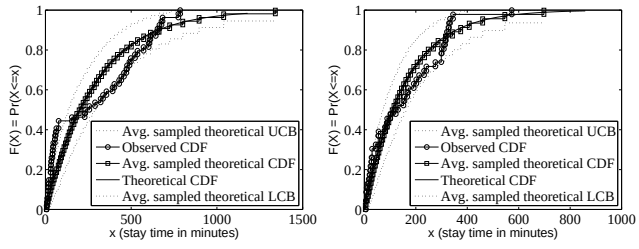
The method presented in Section 3.2, at a given time, derives pmd-regions based on the grid-based temporal mobility statistics (Section 3.1). However, as the object continues to move, the grid-based statistics naturally change over time and consequently the pmd-regions also evolve over time. In particular, pmd-regions can shift in space, grow or shrink in size, disappear (become non-pmd), reappear (become pmd again), or emerge (become pmd for the first time). Although the method in Section 3.2 can be used to extract pmd-regions periodically, due to this evolution, it is not guaranteed that the same pmd-regions will be detected in two consecutive executions of the method. Thus, to perform the prediction tasks effectively based on relevant staytime information about pmd-regions and transition information between pmd-regions, it is essential to correctly track and maintain the evolution of the pmd-regions as they are discovered periodically by the pmd-region detection method. The proposed method for the correct tracking and maintenance of pmd-regions is as follows:

1. Detect the current pmd-regions.
2. Spatially intersect the pmd-regions that have been discovered at any time in the past with the currently detected pmd-regions.
3. For each current pmd-region that intersects with a previously detected pmd-region assign the same ID to the newly detected pmd-region as the ID of the intersecting previously detected pmd-region. Update the spatial information associated with the pmd-region ID with the spatial information of the currently detected pmd-region.
4. For any remaining currently detected pmd-region that does not intersect with any previous pmd-region, assign a new unique pmd-region ID that is one higher than the currently highest pmd-region ID.

To facilitate this tracking / maintenance of pmd-regions, pmd-region IDs and spatial information about pmd-regions in terms of constituent grid cells are stored in a relational format in the table `reg: <reg_ID, grid_ID>`.

3.4 From grid- to region-based trajectory

In order to predict the departure time from or staytime at pmd-regions and the transitions between them, relevant staytime information about pmd-regions and transition information between pmd-regions needs to be gathered. To achieve this the grid-based trajectory $S^{\mathcal{G}} = \langle (g_1, t_1), \ldots, (g_n, t_n) \rangle$ needs to be transformed into a region-based trajectory $S^{\mathcal{R}} = \langle (R_1, t_1^a, t_1^d), \ldots, (R_m, t_m^a, t_m^d) \rangle$, i.e., for each



(a) R_1 (b) R_2

Figure 1: Observed, sampled and theoretical CDFs of an object's staytime in its top two pmd-regions.

pmd-region R_i that the object enters and later on leaves, an *arrival time* $,t_m^a$, and a *departure time* $,t_m^d$, needs to captured and stored. As signal failure can cause a decrease in positioning accuracy, during an actual visit to a pmd-region, the object's position may inaccurately indicate the presence of the object in a cell outside of the pmd-region for a brief period. Such interruptions of visits must be filtered out to accurately capture the arrival time and departure time for each pmd-region visit. To do this, based on the grid-based trajectory the events of *valid arrival* and *valid departure* are defined as follows:

Definition 3. Valid Arrival (Departure): Given an object o's grid-based trajectory $S^{\mathcal{G}} = \langle (g_1, t_1), \ldots, (g_n, t_n) \rangle$, a set of pmd-regions $\mathcal{R} = \{R_1, \ldots, R_k\}$, and a *minimum staytime threshold min_t$_{st}$* (*maximum interruption time threshold max_t$_{int}$*), o is defined to *arrive* at (*depart* from) a pmd-region $R_j \in \mathcal{R}$ at time instance t_i if for the consecutive subsequence $S^{\mathcal{G}}_{\sqsubseteq} = \langle (g_{i-1}, t_{i-1}), (g_i, t_i), \ldots, (g_{i+m}, t_{i+m}) \rangle \sqsubseteq S^{\mathcal{G}}$ it is true that for all grid cells g_l, $(g_l, t_l) \in S^{\mathcal{G}}_{\sqsubseteq}$ s.t. $i \leq l \leq i+m$ the grid cell $g_l \subseteq R_j$ ($g_l \not\subset R_j$), $g_{i-1} \not\subset R_j$ ($g_{i-1} \subseteq R_j$), and $\sum_{i \leq l < i+m} (t_{l+1} - t_l) \geq min\_t_{st}$ ($\geq max\_t_{int}$).

Given the two definitions, using a finite set of temporary variables, arrival and departure times are easily identified from the grid-based trajectory stream in a continuous fashion. This information and the transitions between two consecutive regions in the so-transformed region-based trajectory stream are stored in the table `reg_vis_trans: <reg_id, arr_time, dep_time, prv_reg, nxt_reg, date, day_of_week, isweekend>`.

3.5 Prediction Method

The proposed methods to predict the departure time and the next region is facilitated by modeling the object's pmd-region visit and transition sequence as an ICTM process.

3.5.1 Applicability of the ICTM Model

For this modeling approach to be appropriate, according to the alternative definition of a CTM process, the holding time (staytime), in each state (pmd-region) has to be exponentially distributed. Figure 1 verifies the correctness of this assumption. Figure 1 compares the observed cumulative distribution functions (CDFs) of staytimes in the top two pmd-regions of an object to 1) the CDFs of (theoretical) exponentially distributions that have the same mean values as the respective observations and 2) the average CDFs of $k = 100$ sets of samples that are randomly drawn from the respective theoretical distributions and have the same sample size as the respective observations. Observed and sampled CDFs are calculated via the Kaplan-Meier estimate [9]. Upper and lower confidence bounds (UCB, LCB) are calculated at $\alpha = 0.05$ for 95% confidence levels using Greenwood's formula [6]. As Figure 1 shows the observed CDFs for most of the staytime values fall between the 95% confidence interval of the averaged sampled distribution, which is what one would expect in 95% of the cases for samples drawn from the respective theoretical distributions. Therefore, the proposed modeling aproach is appropriate.

However, staytimes in- and transitions between pmd-regions are inhomogeneous. First, they are *temporally* inhomogeneous, i.e., $\Pr(X(t+s) = j | X(t) = i)$ depends on t. Second, because human activity and hence movement is governed by periodic natural events (changes of days, seasons, etc.) and because time is thus referenced using a multi-level periodic reference system (60 minutes in one hour, 24 hours in one day, etc.), the transition probabilities are likely to be *periodically* inhomogeneous. Finally, last but not least, sequential regularities in staytimes at- and/or transitions between pmd-regions are likely to exist, making the transition probabilities *sequentially* inhomogeneous. For example the semantic location sequence *daycare*→*work*→*daycare* may be frequently observed in the pmd-region visit- and transition sequence, while the sequence *home*→*work*→*daycare* might not be observed as frequently. Thus, the conditional probability that the next region will be *daycare* given that the previous region has been *daycare* and the current region is *work* is higher than the same probability given that the previous region is *home* and the current region is *work*. An approach to estimate the temporally-, periodically- and sequentially inhomogeneous transition probabilities of an ICTM model is present in Section 3.5.2.

3.5.2 Prediction Using the ICTM Model

Combining the theories presented in Section 2.2, the object's visits to and transitions between pmd-regions is modeled by a CTM process $\{X(t), t \geq 0\}$ as follows. Let the state space S of the process be the incrementally labeled set of pmd-regions $\mathcal{R}$. Similar to the P_{ij} transition probabilities for the discrete-time analogue, the *transition rate* q_{ij} from state i, $i \in S$ to state i, $i \in S, j \neq i$ is defined as the number of times the process transitions from state i to j during the unit time interval. For each state i associate for each other state j a random alarm clock a_{ij} that has an alarm time that is exponentially distributed with rate parameter q_{ij}. Assume that as soon as the process enters a state i the set of alarm clocks that are associated with state i, i.e., $\{a_{ij}, j \in S, j \neq i\}$, are activated. The process remains in state i until one of the alarm clocks goes off, at which time the process transitions to the state that is associated with the first alarm clock then went off. It can be shown that the time until first alarm clock goes off in state i, i.e., the holding time in state i, is exponentially distributed with rate parameter $v_i = \sum_{j \in S, j \neq i} q_{ij}$, where $q_{ij} = 0$ if the process cannot enter state j from state i. When the process leaves state i it transitions to state j with probability $p_{ij} = q_{ij}/v_i$, where p_{ij} is the transition probability of the embedded discrete-time Markov chain from state i to j. It can be shown that the above described stochastic process is a CTM process.

Consequently, according to the properties of the exponential distribution, given that at the current time t the process is in state i, the probability that the process will remain in state i during the interval $(t, t+s]$ is:

$$\Pr(X(t+s) = j | X(t) = i) = e^{-v_i s} \tag{5}$$

Given the user-defined minimum staytime likelihood threshold *min_stl* and that the object is in pmd-region $R_i \in \mathcal{R}$ at time t the object's departure time $(t + s^*)$ from pmd-region R_i to some other pmd-region $R_j \in \mathcal{R}, R_j \neq R_i$ is predicted by equating the expression in Equation 5 to *min_stl* and solving for s as follows:

$$s^* = s = \frac{\ln(min\_stl)}{-v_i} \tag{6}$$

Independently of when the process makes the transition from state i, the probability that the process will transition from state i to state j, i.e., the transitions probability of the embedded discrete-time Markov chain, is $p_{ij} = q_{ij}/v_i$. Therefore, the next pmd-region is predicted to be the pmd-region R^* that has the maximum transition probability from pmd-region R_i, i.e:

$$R^* = \underset{j \in \{S-i\}}{\operatorname{argmax}} p_{ij} \tag{7}$$

Provided the information stored in the `reg_vis_trans` table, given that the current pmd-region of the object is `R_c` the transition rates to all other states are calculated by the following SQL query:

```
SELECT r.nxt_reg AS R_j, sum(*)/t.durr AS q_cj
FROM reg_vis_trans AS r,
     (SELECT sum(dep_time-arr_time) AS durr
      WHERE reg_id = R_c) AS t
WHERE r.reg_id = R_c;
```

As the object moves between the pmd-regions the entries in the table `reg_vis_trans` increase. Consequently, the *evidence* for the calculation of transition rates increases and changes. However, as the change is likely to degrease over time, the departure time and next pmd-region predictions from any given pmd-region are likely to become deterministic over time. This is because according to the above presented transition rate estimations, the modeled continuous-time Markov process is assumed to have homogeneous transitions rates. As it is argued in Section 3.5.1, this assumption clearly does not hold in the case of the modeled pmd-region visit- and transition sequence. However, given the additional information that the object has arrived at the current pmd-region `R_c` at time `t_a` on date `d_a` and that the previous pmd-regions visited by the object was `R_p` the temporal-, periodic- and sequential inhomogeneity of the transitions rates can be modeled by generally imposing additional sequential (spatial) and temporal constraints on the SQL query used in the estimations as follows:

TI-Q: Temporally Inhomogeneous Transition Rates

```
SELECT r.nxt_reg AS R_j, sum(*)/t.durr AS q_cj
FROM reg_vis_trans AS r,
     (SELECT sum(dep_time-arr_time) AS durr
      WHERE reg_id = R_c) AS t
WHERE r.reg_id = R_c
  AND t_a BETWEEN r.arr_time AND r.dep_time;
```

PI-Q: Periodically Inhomogeneous Transition Rates

```
SELECT r.nxt_reg AS R_j, sum(*)/t.durr AS q_cj
FROM reg_vis_trans AS r,
     (SELECT sum(dep_time-arr_time) AS durr
      WHERE reg_id = R_c) AS t
WHERE r.reg_id = R_c
  AND dow(d_a) = r.day_of_week;
```

SI-Q: Sequentially Inhomogeneous Transition Rates

```
SELECT r.nxt_reg AS R_j, sum(*)/t.durr AS q_cj
FROM reg_vis_trans AS r,
     (SELECT sum(dep_time-arr_time) AS durr
      WHERE reg_id = R_c) AS t
WHERE r.reg_id = R_c
  AND R_p = r.prv_reg;
```

3.5.3 Weighted Ensemble of ICTM Models

It is clear that each of the queries presented in Section 3.5.2 constructs an *evidence* set E for the calculation of the transition rates. Based on a single evidence set E, one can construct an ICTM model $\mathcal{M}$ and perform departure time and next pmd-region predictions as it is described in Section 3.5.2.

It is likely that different, mutually-dependent aspects of inhomogeneity of the process have difference importance in the overall behavior of the process. Hence, the proposed method combines the predictions of a weighted ensemble of ICTM models $\mathcal{M}_1, \ldots, \mathcal{M}_d$ as follows. For brevity, let $\Pr_{\mathcal{M}}(i(s)|i)$ denote the probability according to model $\mathcal{M}$ that the process will remain in the current state i within the next s time units. Similarly, let $\Pr_{\mathcal{M}}(j|i)$ denote the probability according to model $\mathcal{M}$ that the process will transition from the current state i to the next state j. Then, the departure time prediction based on an ensemble of models $\mathcal{M}_1, \ldots, \mathcal{M}_d$ associated with weights $w_1, \ldots, w_d$ is performed by solving the following equation:

$$s^* = \underset{s}{\operatorname{argmin}} \frac{\sum_{k=1}^{d} w_k * \Pr_{\mathcal{M}_k}(i(s)|i)}{\sum_{k=1}^{d} w_k} \leq min\_stl. \tag{8}$$

The solution for Equation 8 is obtained by performing a binary search in the possible range of s, i.e., $[\min_{k=1}^{d}(s^*_{\mathcal{M}_k}), \max_{k=1}^{d}(s^*_{\mathcal{M}_k})]$, where $s^*_{\mathcal{M}_k}$ represents the solution to Equation 6 by model $\mathcal{M}_k$, and testing for the equality condition.

Similarly, the next state prediction based on an ensemble of models $\mathcal{M}_1, \ldots, \mathcal{M}_d$ associated with weights $w_1, \ldots, w_d$ is performed by solving the following equation:

$$R^* = \underset{j \in \{S-i\}}{\operatorname{argmax}} \frac{\sum_{k=1}^{d} w_k * \Pr_{\mathcal{M}_k}(j|i)}{\sum_{k=1}^{d} w_k}. \tag{9}$$

3.5.4 Model Weights in the Ensemble

Individual models in an ensemble capture different aspects of the inhomogeneity of transition rates. Individual model weights indicate the relative importance of each of the models and the corresponding inhomogeneity aspects. Model weights can be set in a *static* fashion based on intuition or expert knowledge or based on a parameter optimization procedure on the basis of a general training set of several users. However, the relative importance of the models and the corresponding inhomogeneity aspects are expected to vary from individual to individual as well as over time. To capture this variation the calculation of *dynamic* weights is proposed.

As it is described Sections3.5.2 and 3.5.3, the transition rates of an ICTM model $\mathcal{M}$ are estimated on the basis of a query condition QC and a resulting evidence set E^{QC}. In case of the TI-Q, PI-Q, and SI-Q queries for the query parameters time of arrival `t_a`, date of arrival `d_a`, and previous pmd-regions visited `R_p`, these query conditions are

- `t_a BETWEEN r.arr_time AND r.dep_time`,
- `dow(d_a) = r.day_of_week`, and

- `R_p = r.prv_reg`

and capture the temporal-, periodic- and sequential inhomogeneity of the transitions rates, respectively. Each of these query conditions puts a constraint on the evidence set in either the projected temporal- or the discretized spatial dimensions each with finite domain size. Intuitively, the relative importance of a model $\mathcal{M}$ with query condition QC over a finite-domain dimension D and evidence set E^{QC} should be directly proportional to the *relative size of the evidence set*, $|E^{QC}|/|E^{\emptyset}|$, and should be inversely proportional to the *relative expected domain selectivity of the query condition*, $S_D(QC)/S_D(\emptyset)$, where $\emptyset$ represents the empty constraint and $S_D(.)$ returns the size of its argument w.r.t. the domain D. Since for query TI-Q the size of a time instance `t_a` is not well defined w.r.t. to the projected temporal dimension, $S_D(\texttt{t\_a})$ is approximated with the average length of the projected time intervals that overlap with `t_a`, i.e., $S_D(\texttt{t\_a}) = \sum_{\texttt{e} \in E^{QC}} |\texttt{e.dep\_time} - \texttt{e.arr\_time}|/|E^{QC}|$. Consequently, the weights of a model $\mathcal{M}$ with query condition QC over a finite-domain dimension D and evidence set E^{QC} is calculated as:

$$w_{\mathcal{M}} = \frac{|E^{QC}|}{|E^{\emptyset}|} \times \frac{S_D(\emptyset)}{S_D(QC)}. \quad (10)$$

The SQL queries to calculate the weights of models from relevant evidence sets are straight forward and therefore are omitted to preserve clarity and save space.

4. EMPIRICAL EVALUATION

The following section describes the process and results of the experimental evaluation of the proposed mobility prediction method.

4.1 Test Environments

To realize the evaluations the proposed method has been implemented in Java both as a desktop and a mobile application. The prediction performance experiments have been conducted in the desktop environment on a PC with Intel core i7 2630QM processor with 8 GB of main memory running a 64-bit Windows 7 OS. The execution time and resource consumption experiments have been conducted in the mobile environment on a HTC G7 smartphone with 1 GHz CPU, 512M memory and Android 2.3.7 OS.

4.2 CPU and Battery Consumption

The prototype mobile application can in an online fashion collect and project GPS coordinates onto the grid using 14% of the CPU when the average sampling frequency is 4.7 seconds. One run of pmd-region detection and and tracking, which is executed infrequently, e.g., daily, takes on average 4.8 seconds and one prediction, which is executed on average 5-10 times daily, takes on average 1.4 seconds. After running the prototype mobile application for 1 hour 15 minutes, the total battery consumption is $214431\mu Ah$, so the transient battery consumption is $47\mu Ah/sec$. This is really low battery consumption, with one $1300mAh$ battery can enable this application to run 7.68 hours. Since the overall resource consumption is dominated by the location sampling and grid-projection, the prototype application is likely to be able to run up to 10-12 times longer than 7.68 hours if the sampling frequency is set to one minute, which arguably will not significantly affect the predictive performance of the application. Overall, the execution time and resource consumption experiments show that proposed method can be executed on a conventional smartphone meeting real-time application requirements while using minimal resources.

4.3 Real Word Data Set

The trajectory data set used in the experiments is a subset of the *GeoLife* dataset [15]. It is a GPS trajectory data set that has been collected in *GeoLife* project by Microsoft Research Asia. It contains the GPS trajectories of 178 users sampled between the period April 2007 and August 2009. A GPS trajectory in the data set is represented by a sequence of time-stamped points of geographical coordinates (longitude, latitude, height) an additional information about the movement (speed and heading direction). These trajectories were recorded by different GPS loggers or GPS-phones using a variety of sampling rates. 95% of the trajectories are sampled frequently, i.e., every 2-5 seconds or 5-10 meters. The data set records a broad range of users' outdoor movements, including not only life routines like go home and go to work but also some entertainment and sports activities, such as shopping, dining, hiking, and cycling. Because the sampling rate and period are not constant for all users and because some user trajectories contain siggnificant sampling gaps, the trajectories of the top users with the higest avarage sampling rate, longest continuous sampling period, and least amount of sampling gaps were selected for the exeriments. The number of samples and the number of observation days in the individual trajectories of the selected users ranges between approximatelly 210000 and 640000 samples and between 19 and 61 observation days.

4.4 Prediction Performance Measures

Given an object's *observed* region-based trajectory, $S^{\mathcal{R}} = \langle (R_1, t_1^s, t_1^e), \ldots, (R_m, t_m^s, t_m^e) \rangle$, and the region-based trajectory $S^{\mathcal{R}*} = \langle (R_1^*, t_1^{s*}, t_1^{e*}), \ldots, (R_m^*, t_m^{s*}, t_m^{e*}) \rangle$ that is *predicted* by a model $\mathcal{M}$, for the two prediction tasks five evaluation criteria are defined as follows:

- **Absolute Temporal Prediction Error** (ATPE) is equal to $\sum_{1 \leq i \leq m} abs(t_i^{e*} - t_i^e)/m$.
- **Relative Temporal Prediction Error** (RTPE) is equal to $\sum_{1 \leq i \leq m} abs(\frac{t_i^{e*} - t_i^e}{t_i^e - t_i^s})/m$.
- **Overall Spatial Prediction Accuracy** (OSPA) is equal to $\sum_{1 \leq i \leq m} \mathrm{I}\,(R_i^* = R_i)/m$, where $\mathrm{I}\,()$ is the Boolean indicator function that returns 1 if its argument evaluates to TRUE and returns 0 otherwise.
- **True Spatial Prediction Confidence** (TSPC) is equal to $\sum_{1 \leq i \leq m : \mathrm{I}\,(R_i^* = R_i)} \Pr_{\mathcal{M}}(X(t_i^s) = R_i | \mathcal{H}(t_i^s))$.
- **False Spatial Prediction "Confusion"** (FSPC) is equal to $\sum_{1 \leq i \leq m : \mathrm{I}\,(R_i^* \neq R_i)} \Pr_{\mathcal{M}}(X(t_i^s) = R_i | \mathcal{H}(t_i^s))$.

The meaning, possible ranges, and desired values for the measures ATPE, RTPE and OSPA are self-evident from the definitions. The measures TSPC and FSPC try to evaluate the probabilistic predictions of the model in a cost-sensitive setting where correct predictions with higher confidence are valued higher than correct predictions with lower confidence (higher TSPC is better) and where incorrect predictions with lower confusion are valued higher than incorrect predictions with higher confusion (lower FSPC is better).

4.5 Rule-Based Baseline Prediction Method

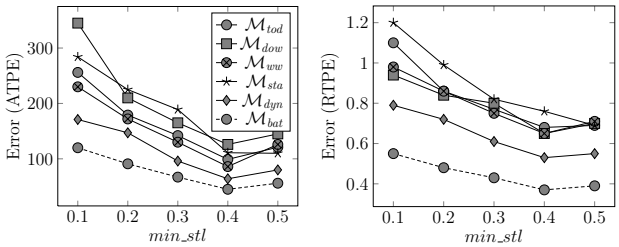
As Section 1 states, no previously proposed method was designed to perform the same mobility prediction tasks, in particular the temporal mobility prediction task, under the given application setting. Thus, to form a baseline for comparison, a representative, state-of-the-art, rule-based trajectory mining and prediction method [5] has been modified and applied to the given application setting as follows.

The object's *observed* region-based trajectory $S^{\mathcal{R}} = \langle e_1, \ldots, e_m \rangle$, where $e_i = (R_i, t_i^s, t_i^e)$, $1 \leq i \leq n$, is partitioned into a set of $m' = \lfloor m/n \rfloor$, n-long, consecutive, overlapping subsequences using a sliding window over $S^{\mathcal{R}}$. The so obtained set of subsequences is $DS^{\mathcal{R}} = \{\langle e_1, \ldots, e_n \rangle, \langle e_2, \ldots, e_{n+1} \rangle, \ldots, \langle e_{n(m'-1)+1}, \ldots, e_{nm'} \rangle\}$. Using the pmd-region visit staytimes, i.e., $(t_i^e - t_i^s)$, as temporal annotations, the mining algorithm in [5] is applied in an offline fashion to extract *all* ($min\_supp = 0$) *Closed Contiguous Temporally Annotated Pattern Sequences* (CCTAPSes) from $DS^{\mathcal{R}}$, where the temporal annotation Δt_i of a pattern element $pe_i = (R_i, \Delta t_i)$, $1 \leq i \leq l \leq n$ in the CCTAPS, $ps = \langle (R_1, \Delta t_1), \ldots, (R_l, \Delta t_l) \rangle$, is the average staytime at the pmd-region R_i among the set of sequences in $DS^{\mathcal{R}}$ in which the the pmd-region R_i is immediately preceded by the consecutive sequence of pmd-regions $\langle R_1, \ldots, R_{i-1} \rangle$.

Subsequently, in the prediction phase, the extracted CCTAPSes are inserted in a Frequent Pattern tree (FP-tree), i.e., a prefix tree that is augmented with a header table and vertical links through identical pmd-regions. Then, for all sequences $s_i = \langle (R_1, t_1^s, t_1^e), \ldots, (R_n, t_n^s, t_n^e) \rangle \in DS^{\mathcal{R}}$ the subsequence of the first $n-1$ elements of the sequence s_i is matched against the CCTAPSes in the FP-tree. Subsequently, the longest matching branch $\langle (R_1, \Delta t_1), \ldots, (R_{n-1}, \Delta t_{n-1}) \rangle$ and its children $(R_j, \Delta t_j)$, $1 \leq j \leq k$ in the FP-tree are used to predict the departure time and the next region as follows. The departure time from pmd-region R_{n-1} is predicted as $t_{n-1}^s + \Delta t_{n-1}$. The next pmd-region is predicted as the pmd-region of the child of the matched branch that has the highest support amongst the children. Note that due to the unique application of the rule-based trajectory mining and prediction, the subsequence of the first $n-1$ elements of the sequence s_i can always be fully matched to a unique path in the FR-tree that starts from the root node and the matched path always has at least one child. Indeed, the baseline predictor is presented all the data before it has to make any predictions. The effects of this "batch learning" advantage are considered in the prediction performance evaluations.

4.6 Prediction Performance Results

The prediction performance of the proposed method has been evaluated for a number of different second-order ICTM models each capturing in different ways, different aspects of the inhomogeneity of the transitions rates using different temporal domain projections. In particular, three single ICTM models $\mathcal{M}_{tod}$, $\mathcal{M}_{dow}$ and $\mathcal{M}_{ww}$ have been evaluated based on the time-of-day, day-of-week and workday-weekend temporal domain projections, respectively. In addition, the single ICTM models have been combined into two weighted ensemble ICTM models: $\mathcal{M}_{sta}$ and $\mathcal{M}_{dyn}$. In the case of $\mathcal{M}_{sta}$ model the weights of the submodels in the ensemble are statically set to the same constant value c. In the case of $\mathcal{M}_{dyn}$they are set dynamically according to Equation 10 in Section 3.5.4. Finally, noting the "batch learning" advantage



(a) ATPE (b) RTPE

Figure 2: Absolute 2(a) and relative 2(b) temporal prediction errors of different prediction models for varying *min_stl* values in comparison rule-based baseline predictor's performance (ATPE($\mathcal{M}_{rule}$) = 82 minutes and RTPE($\mathcal{M}_{rule}$) = 0.57).

of the rule-based baseline predictor in Section 4.5, a dynamic ensemble model, $\mathcal{M}_{bat}$ that exploits the same advantage, has also been evaluated and compared to the rule-based baseline predictor $\mathcal{M}_{rule}$. For all the experiments equally affecting the prediction performance of all models, including the baseline, the preprocessing parameters of the proposed method have been set as follows: $glen = 50$ meters, $maxR = 150$ meters, $min\_rp = 0.8$ and $min\_t_{st} = max\_t_{int} = 30$ seconds. The effects of these preprocessing parameters on the prediction performance are investigated in Section 4.6.3.

4.6.1 Evaluation of Temporal Prediction Error

The absolute (ATPE) and the relative (RTPE) temporal prediction errors of the different models for varying *min_stl* values are shown on Figures 2(a) and 2(b), respectively. The measurements represent the uniform average of the temporal prediction errors of a given model for the 10 selected users. Almost all the models achieve optimal temporal prediction accuracy at $min\_stl = 0.4$, which is expected according to Equation 6. Considering only the single ICTM models, the $\mathcal{M}_{ww}$ model based on the workday-weekend temporal domain projection achieves the best results in the absolute and relative sense, i.e., ATPE($\mathcal{M}_{ww}$) = 86 minutes and RTPE($\mathcal{M}_{ww}$) = 0.65. The results also show that the single ICTM models cannot be effectively combined into an ensemble using the static weighting scheme to decrease the prediction error. In comparison, the ensemble based on the dynamic weighting scheme, $\mathcal{M}_{dyn}$, can successfully benefit from the specialization of its submodels and archive optimal results at ATPE($\mathcal{M}_{dyn}$) = 64 and RTPE($\mathcal{M}_{dyn}$) = 0.53. Finally, when presented with the "batch learning" advantage, the $\mathcal{M}_{bat}$ model can improve on the prediction error of the $\mathcal{M}_{dyn}$ model on average by 31% and achive optimal results at ATPE($\mathcal{M}_{bat}$) = 45 and RTPE($\mathcal{M}_{bat}$) = 0.37, which are superior to the results of the rule-based baseline predictor at ATPE($\mathcal{M}_{rule}$) = 82 and RTPE($\mathcal{M}_{rule}$) = 0.57.

4.6.2 Evaluation of Spatial Prediction Performance

Table 1 shows the spatial prediction performance of the different models and the rule-based baseline predictor in terms of OSAP, TSPC, and FSPC measures. Observations that are similar the observations in Section 4.6.1 can be made about the spatial prediction performance of the models w.r.t. one another. In particular, only the ensemble models employing the dynamic weighting scheme can successfully utilize the specializations of their submodels and deliver an overall spatial prediction accuracy that is compa-

Table 1: Spatial prediction performance of different prediction models.

Model	OSPA	TSPC	FSPC
$\mathcal{M}_{tod}$	0.42	0.41	0.21
$\mathcal{M}_{dow}$	0.25	0.52	0.19
$\mathcal{M}_{ww}$	0.42	0.42	0.17
$\mathcal{M}_{sta}$	0.36	0.39	0.16
$\mathcal{M}_{dyn}$	0.59	0.763	0.13
$\mathcal{M}_{bat}$	0.67	0.79	0.12
$\mathcal{M}_{rule}$	0.62	0.68	0.19

rable or superior to the performance of the rule-based baseline predictor, i.e., OSPA($\mathcal{M}_{dyn}$) = 0.59 < OSPA($\mathcal{M}_{rule}$) = 0.62 < OSPA($\mathcal{M}_{bat}$) = 0.67. The TSPC and FSPC measures show that predictions of an ensemble model employing the dynamic weighting scheme are also more preferable in a cost-sensitive setting.

4.6.3 Effects of Preprocessing Parameters

Due to space limitations the results of the experiments investigating the effects of the parameters of the preprocessing steps of the proposed method are summarized as follows.

The grid cell size, $glen \in [25, 100]$ meters does not significantly change the number, location and size of the extracted pmd-regions and hence has negligible effects on the overall prediction performance. The minimum relative prevalence parameter, $min\_rp \in [0.7, 0.9]$, does not significantly effect the locations and size of the extracted pmd-regions. Merely, an additional 1-2 new pmd-regions are extracted compared to the average number of pmd-regions of 7.8 pmd-regions per user at $min\_rp = 0.7$. For $min\_rp > 0.9$ the number of additional pmd-regions increases drastically. However, the newly extracted pmd-regions likely represent transition points or traffic jams during trips as they collocate with transit stops and intersections and have short staytimes; such pmd-regions are regarded as noise in the given context. The min_t_{st} and max_t_{int} thresholds used in the pmd-region arrival and departure validation, respectively, have a profound effect on the prediction performance. The threshold min_t_{st} can be effectively used to filter out visits to any pmd-region that has been extracted for high values of min_rp thereby improving the prediction accuracy and certainty. The max_t_{int} threshold can be used to make sure that a single visit to a pmd-region is not split into several consecutive visits due to noise in GPS measurements, thereby providing more accurate temporal prediction results. To ensure optimal prediction results, the value max_t_{int} is determined as the smallest value for which during the conversion from the object's grid-based trajectory to it's region-based trajectory no consecutive visits to the same pmd-region are generated.

5. CONCLUSION AND FUTURE WORK

The paper proposed the use of a dynamically weighted ensemble of ICTM models to simply but effectively capture the temporal-, periodic- and sequential regularities in movements of an object to predict *when* and *where* the object will move *next*. The empirical evaluations show that proposed method has a prediction performance that is superior to the state-of-the-art rule-based predictor. In particular, since the effects of the "batch learning" advantage are expected fade as the length of the region-based trajectory of an object increases over time, the prediction performance of the dynamically weighted ensemble of ICTM models is expected to be able to predict (1) the departure time on average to be within 45 minutes of the actual departure time and (2) the next region correctly in 67% of the cases.

Future work is considered along to main directions. First, other statistical based, dynamical weighting schemes will be investigated. Second, because the movements of an individual are largely influenced by the movements of other individuals via social relations, future research will investigate how the ICTM model can be adapted to perform predictions for a group of socially related individuals.

6. REFERENCES

[1] A. Asahara, A. Sato, K. Maruyama, and K Seto. Pedestrian-movement Prediction based on Mixed Markov-chain Model. In *Proc. of ACM-GIS*, pp. 25-33, 2011.

[2] D. Ashbrook and T. Starner. Using GPS to Learn Significant Locations and Predict Movement across Multiple Users. *Personal and Ubiquitous Computing* 7(5):275-286, 2003.

[3] A. Bhattacharya and S. K. Das. LeZi-Update: An Information-Theoretic Approach to Track Mobile Users in PCS Networks. In *Proc. of MobiCom*, pp. 1-12, 1999.

[4] F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi. Trajectory Pattern Mining. In *Proc. of KDD*, pp. 330-339, 2007.

[5] G. Gidófalvi, M. Kaul, C. Borgelt, and T. Bach Pedersen. Frequent Route Based Continuous Moving Object Location- and Density Prediction on Road Networks. In *Proc. of ACM-GIS*, pp. 381-384, 2011.

[6] M. Greenwood. The Natural Duration of Cancer. *Reports on Public Health and Medical Subjects* 33:1-26, 1926.

[7] Y. Ishikawa, Y. Tsukamoto, and H. Kitagawa. Extracting Mobility Statistics from Indexed Spatio-Temporal Dataset. In *Proc. of STDBM*, pp. 9-16. 2004.

[8] H. Jeung, Q. Liu, H. T. Shen, and X. Zhou. A Hybrid Prediction Model for Moving Objects. In *Proc. of ICDE*, pp. 70-79, 2008.

[9] Kaplan, E. L. and Meier, P.. Nonparametric Estimation from Incomplete Observations. *J. Amer. Statist. Assn.* 53:457-481, 1958.

[10] X. Liu. The Principle of Scaling of Geographic Space and its Application in Urban Studies. *Ph.D. Thesis at the KTH*, ISBN: 978-91-7501-277-3, 2012.

[11] A. Monreale, F. Pinelli, R. Trasarti and F. Giannotti. WhereNext: a Location Predictor on Trajectory Pattern Mining. In *Proc. of KDD*, pp. 637-645, 2009.

[12] S. M. Ross, Introduction to Probability Models. Academic Press, Sixth Edition, ISBN-10: 0125984707, 1997.

[13] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of Predictability in Human Mobility. *Science* 327:1018-1021, 2010.

[14] F. Verhein and S. Chawla. Mining Spatio-Temporal Association Rules, Sources, Sinks, Stationary Regions and Thoroughfares in Object Mobility Databases. In *Proc. of DASFAA*, pp. 187-201, 2006.

[15] Y. Zheng, L. Zhang, X. Xie, and W.Y. Ma. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In *Proc. of WWW*, pp. 791-801, 2009.

[16] G. Yavas, D. Katsaros, O. Ulusoy, and Y. Manolopoulos. A Data Mining Approach for Location Prediction in Mobile Environments. *DKE* 54(2):121-146, 2005.

[17] Y. Ye, Y. Zheng, Y. Chen, J. Feng and X. Xie. Mining Individual Life Pattern Based on Location History. In *Proc. of MDM*, pp. 1-10, 2009.

[18] J. J.-C. Ying, W.-C. Lee, T.-C. Weng and V. S. Tseng. Semantic Trajectory Mining for Location Prediction. In *Proc. of ACM-GIS*, pp. 34-43, 2011.