HAL open science

Markov Modulated Bi-variate Gaussian Processes for Mobility Modeling and Location Prediction

Paulo Salvador, António Nogueira

To cite this version:

Paulo Salvador, António Nogueira. Markov Modulated Bi-variate Gaussian Processes for Mobility Modeling and Location Prediction. 10th IFIP Networking Conference (NETWORKING), May 2011, Valencia, Spain. pp.227-240, 10.1007/978-3-642-20757-0_18 . hal-01583408

HAL Id: hal-01583408

https://inria.hal.science/hal-01583408v1

Submitted on 7 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Markov Modulated Bi-variate Gaussian Processes for Mobility Modeling and Location Prediction

Paulo Salvador, António Nogueira

Department of Electronics, Telecommunications and Informatics
Institute of Telecommunications
University of Aveiro
E-mail:{salvador,nogueira}@ua.pt

Abstract. A general-purpose and useful mobility model must be able to describe complex movement dynamics, correlate movement dynamics with the nodes geographic position and be sufficiently generic to map the characteristics of the movement dynamics to general geographic regions. Moreover, it should also be possible to infer the mobility model parameters from empirical data and predict the location of any node based on known positions. Having the ability to model movement and predict future positioning of mobile nodes in complex environments can be very important to several operational and management tasks. Existing mobility modeling approaches are based on simple and limited models, specifically designed for particular application scenarios or requiring the complete knowledge of the mobility environment. These features make them unusable in complex scenarios with no (or partial) knowledge of the environment. This paper presents a discrete time Markov Modulated Bi-variate Gaussian Process that is able to characterize the position and mobility of any mobile node, assuming that the position within a generic sub-region can be described by a bi-variate Gaussian distribution and the transition between sub-regions can be described by an underlying (homogeneous) Markov chain. With this approach, it is possible to describe the mobile node movement within and between a set of geographic regions determined by the model itself and, due to the Markovian nature of the model, it is also possible to capture complex dynamics and calculate the future probabilistic position of a mobile node. The proposed approach can be applied to scenarios where the possible pathways are unknown or too complex to consider in a real model that must make a prediction in a very short time. The results obtained by applying the proposed model to real and publicly available data demonstrate the accuracy and utility of this approach: the model was able to efficiently describe the movement patterns of mobile nodes and predict their future position. Besides, the model has also revealed higher performances when compared to other modeling approaches.

Keywords: mobility modeling, location prediction, Markov model, modulated bi-variate Gaussian process.

1 Introduction

Being able to describe the movement and predict future positioning of mobile nodes in complex environments can be very important to different operational and management tasks: evaluation of different system design alternatives and implementation costs, simulation and study of new routing protocols for Mobile and Delay Tolerant Networks, knowledge of the nodes distribution for performance evaluation and message delivery optimization, prediction of the nodes preferred locations for connectivity issues, etc. Other research topics, like mobile computing, can also greatly benefit from the ability to track and predict the location of mobile devices: an accurate location predictor can significantly improve the performance or reliability of many applications in pervasive computing.

There are two different strands of the mobile modeling problem: the ability to describe or track the movement of mobile nodes, which is mainly used for simulation purposes, and the ability to predict the mobile nodes position, which can be useful to several management tasks. An efficient mobility model must be able to describe complex movement dynamics, correlate the movement dynamics with the nodes position and be sufficiently generic to map movement's dynamics characteristics to general geographic regions where the road/path topology is unknown or imprecise. The parameters of the mobility model should be easily inferred from real mobility data and the model should provide an almost real-time prediction of the nodes position.

This paper proposes a discrete time Markov Modulated Bi-variate Gaussian Process to characterize the position and mobility of any mobile node, assuming that the position can be described by a bi-variate Gaussian distribution and the transition between regions can be described by an underlying (homogeneous) Markov chain. This model is able to describe the mobile node movement within and between a set of geographic regions determined by the model itself. Due to the Markov nature of the model, it is possible to determine the future probabilistic position of a mobile node.

The results obtained by applying this model to real mobility data prove that this approach can accurately predict the future position of mobile nodes presenting, at the same time, higher performance than other prediction models that were selected for comparison.

The paper is organized as follows: section 2 presents some of the most relevant works related to the proposed model; section 3 presents the mathematical background on the multidimensional discrete time Markov Modulated Bi-variate Gaussian Process (dMMBVGP) and the description of the proposed model; section 4 presents the algorithmic details of the inference procedure; section 5 presents the results obtained in a proof of concept and a discussion of those results; finally, section 6 presents the main conclusions.

2 Related work

Multiple proposals for modeling movement of mobile nodes have already been published [5]. The proposed models can be mapped into four categories: (i) purely

random; (ii) with temporal dependency, when the movement of a mobile node is likely to be affected by its movement history; (iii) with spatial dependency between subjects, when mobile nodes tend to travel in a correlated manner, and (iv) with geographical dependency/restriction, when the movement of nodes is bounded by streets, freeways or obstacles.

In purely random mobility models, mobile nodes move randomly and freely without restrictions, that is, the destination, speed and direction are all chosen randomly and independently of other nodes. The Random Waypoint model [4] and two of its variants, named Random Walk model [5] and Random Direction model [18], are the most important examples of this type of models. Obviously, these models fail to represent some mobility characteristics that are likely to exist in mobile ad hoc networks: the temporal and spatial dependency of velocity and the geographic restrictions of movement. Since, in general, the velocities of a single node at different time slots are "correlated", other models were proposed to capture this temporal dependency behavior: the Gauss-Markov mobility model [11], where the velocity of a mobile node is assumed to be correlated over time and modeled as a Gauss-Markov stochastic process and the Smooth Random Mobility Model [2], where the speed and direction of the node movement is incrementally and smoothly changed in order to generate more realistic movement behaviors. In [7], the authors proposed a time-variant community mobility model that realistically captures spatial and temporal correlations. Synthetic traces that match the characteristics of different mobility traces, including wireless LAN traces, vehicular mobility traces and human encounter traces, were generated, demonstrating the flexibility of the proposed model. Several Markovian approaches have also been proposed to address the mobility problem: reference [8] proposed the Markovian Waypoint model, where mobile nodes move along a straight line segment from one waypoint to another at a certain speed. In this approach, waypoints are obtained from a discrete time Markov process having a subset of $\mathbb{R}^2$ as the state space. Other mobility models, such as [2], are able to produce more smoothly turning trajectories.

In most real life applications, a node's movement is subject to the environment. Therefore, nodes may move in a pseudo-random way on predefined pathways in the environment field. Mobility models with geographical restrictions have been proposed to describe this kind of environments: in the Pathway Mobility Model [22] the node movement is restricted to pathways in a map, while in the Obstacle Mobility Model [9] the effects of obstacles on node mobility and radio propagation are taken into account. All these classes of models are unable to constrain/define the movement dynamics as a function of the mobile node positioning.

There has also been a significant interest in mobility models specially designed for partitioned networks, where end-to-end connectivity rarely or never exists. In [3], authors proposed the Area Graph Based Mobility Model that considers the inhomogeneity of spatial node distribution by using a directed graph with areas of different densities; in [6] similar approaches, called Weighted Waypoint Mobility and Clustered Mobility Models, respectively, were proposed

to capture node preferences in choosing possible destinations; the Community Based Mobility Model, proposed in [21], was designed to recreate the characteristics of mobility in sparse networks. All these models exhibit a skewed distribution of location visits, which could result in forming islands of connectivity at popular locations and being unable to stabilize. The Heterogeneous Random Walk model, proposed in [17], is a very parsimonious model where clusters emerge because of low node speed in popular areas. Musolesi et al. [13] presented a mobility model that incorporates sociological relationships between mobile nodes, resulting in a quite complex model. Recently, some studies have been devoted to model the realistic social dimension of human mobility. A human mobility model based on heterogeneous centrality and overlapping community structure in social networks was proposed in [23], being able to match characteristics observed from real human motion traces, especially heterogeneous human mobility popularity. This type of models requires pre-established division of the network geographic environment and are not able to adapt themselves to unknown generic scenarios.

Several approaches have been proposed to predict the future location of mobile entities. In [12], authors proposed a method recording a sequence of cell identifiers to predict future location of users in wireless ATM networks, while [1] use logical information for location prediction. Other approaches use road topology information to predict user location [19]. In [10], a semi-Markov model is used to describe the state transitions between stationary behavior and movement in a WLAN cell, where all model parameters can be obtained from movement monitoring. Most of the existing methods are restricted to a specific purpose, leading to straightforward implementation but limited effectiveness in generic network management tasks. Designing a robust flexible mobility and location model that can be used in generic environments is the motivation for this paper.

3 Model and Positioning Prediction

3.1 Markov Modulated Bi-variate Gaussian Processes Model

The proposed discrete time Markov Modulated Bi-variate Gaussian Process (dMMBVGP) model characterizes position and mobility of a subject based on the following assumptions: (i) within a restricted generic geographic region the position may be described by a bi-variate Gaussian distribution, (ii) the geographical space can be divided in N distinct regions (non-determined *a priori*) and (iii) the transition between regions can be described by an underlying (homogeneous) Markov chain, where each state maps the movement characteristics of a specific region. The dMMBVGP can then be described as a random process (Z) with a bi-variate Gaussian distribution, which determines a position in a two-dimensional environment, whose parameters are a function of the state (J) of the modulator Markov chain (Z, J) with N states. More precisely, the (homogeneous) Markov chain $(Z, J) = \{(Z_k, J_k),\ k = 0, 1, \ldots\}$ with state space $I\!R^2 \times U$, with $U = \{1, 2, \ldots, N\}$, is a dMMBVGP if and only if for $k = 0, 1, \ldots,$

$$P(Z_{k+1} = \mathbf{z},\ J_{k+1} = n | J_k = m) = p_{mn} \Gamma_n(\mathbf{z}) \tag{1}$$

where p_{mn} represents the probability of a transition from state m to state n of the underlying Markov chain in time interval $[k, k+1]$, and $\Gamma_n(\mathbf{z}) = \frac{1}{2\pi|\Sigma_n|^{1/2}} e^{-\frac{1}{2}(\mathbf{z}-\mathbf{\Lambda}_n)^T \Sigma_n^{-1}(\mathbf{z}-\mathbf{\Lambda}_n)}$ is the bi-variate Gaussian distribution in region n centered in Λ_n and having covariance matrix Σ_n.

The dMMBVGP may be also regarded as a Markov random walk [14] where the two-dimensional positions increments (Δ) in each instant k have a distribution given by the difference of two bi-variate Gaussian distributions whose parameters are a function of the state of the modulator Markov chain

$$\begin{aligned} P(Z_{k+1} =& \mathbf{z} + \Delta,\, J_{k+1} = n | Z_k = \mathbf{z},\, J_k = m) = \\ &= p_{mn} \frac{1}{2\pi\left|\Sigma_n + \Sigma_m\right|^{1/2}} e^{-\frac{1}{2}(\Delta - \mathbf{\Lambda}_n + \mathbf{\Lambda}_m)^T (\Sigma_n + \Sigma_m)^{-1} (\Delta - \mathbf{\Lambda}_n + \mathbf{\Lambda}_m)} \end{aligned}$$

Whenever (1) holds, we say that (Z, J) is a dMMBVGP with a set of modulating states with size N and parameter matrices $\mathbf{P}$, $\mathbf{S}$, and $\mathbf{L}$. The matrix $\mathbf{P}$ is the transition probability matrix of the modulating Markov chain J,

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1N} \\ p_{21} & p_{22} & \dots & p_{2N} \\ \dots & \dots & \dots & \dots \\ p_{N1} & p_{N2} & \dots & p_{NN} \end{bmatrix} \tag{2}$$

matrix $\mathbf{L}$ defines the centers of each region,

$$\mathbf{L} = \left[\mathbf{\Lambda}_1\ \mathbf{\Lambda}_2\ \dots\ \mathbf{\Lambda}_N\right], \mathbf{\Lambda}_n = \begin{bmatrix} \lambda_n^{(x)} \\ \lambda_n^{(y)} \end{bmatrix} \tag{3}$$

and matrix $\mathbf{S}$ contains the covariance (sub-)matrices of each region:

$$\mathbf{S} = \left[\Sigma_1\ \Sigma_2\ \dots\ \Sigma_N\right], \mathbf{\Sigma}_n = \begin{bmatrix} (\sigma_n^{(x)})^2 & \rho_n \sigma_n^{(x)} \sigma_n^{(y)} \\ \rho_n \sigma_n^{(x)} \sigma_n^{(y)} & (\sigma_n^{(y)})^2 \end{bmatrix}. \tag{4}$$

Moreover, we denote by $\mathbf{\Pi} = [\pi_1, \pi_2, \dots, \pi_N]$ the stationary distribution of the underlying Markov chain.

Specifying $\mathbf{z} = \begin{bmatrix} x & y \end{bmatrix}^T$, where x and y represent longitude and latitude, respectively, and defining the general (not weighted) bi-variate Gaussian distribution within a region n by

$$\Gamma_n(x, y) = \frac{1}{2\pi\sigma_n^{(x)}\sigma_n^{(y)}\sqrt{1-\rho_n^2}} e^{Q_n} \tag{5}$$

with $Q_n = -\frac{1}{2(1-\rho_n^2)}\left(\frac{(x-\lambda_n^{(x)})^2}{(\sigma_n^{(x)})^2} + \frac{(y-\lambda_n^{(y)})^2}{(\sigma_n^{(y)})^2} - \frac{2\rho_n(x-\lambda_n^{(x)})(y-\lambda_n^{(x)})}{\sigma_n^{(x)}\sigma_n^{(y)}}\right)$, the stationary probability density map of position (x, y) can be determined by a weighted sum of N bi-variate Gaussian distributions:

$$f^N(x, y) = \sum_{n=1}^{N} \pi_n \Gamma_n(x, y) \tag{6}$$

obtaining a position probability density map.

3.2 Positioning prediction

Defining $\mathbf{q}_k = \{q_{k,1}, q_{k,2}, \ldots, q_{k,N}\}, k = 0, 1, \ldots, K$, where $q_{k,n}$ is the probability that, at time interval k (within a total of K observations), the mobile node is in region n (i.e. the probability of the underlying Markov chain of the mobility model is in state n at time instant k) and based on equation (6) we can define the predicted position probability map at time instant k as:

$$\bar{f}_k(x, y) = \sum_{n=1}^{N} q_{k,n} \Gamma_n(x, y) \tag{7}$$

with

$$\mathbf{q}_k = \mathbf{q}_0 \mathbf{P}^k \tag{8}$$

where $\mathbf{q}_0$ represents the probability vector of the last known position (x_0, y_0) belonging to the different regions, and can be inferred as

$$q_{0,n} = \frac{\pi_n \Gamma_n(x_0, y_0)}{f(x_0, y_0)} \tag{9}$$

4 dMMBVGP Inference Procedure

Hereafter, we will refer to the empirical position probability density map as $f^e(x, y)$ and $\epsilon_{n-1}(x, y) = f^e(x, y) - f^n(x, y)$ will represent the difference function between the empirical and fitted position probability maps with n bi-variate Gaussian distributions (f^n).

The inference procedure can be divided in two steps: (i) inference of matrices $\mathbf{L}$ and $\mathbf{S}$ and vector $\mathbf{\Pi}$ by adjusting $f^e(x, y)$ by a weighted sum of bi-variate Gaussian distributions and (ii) inference of matrix $\mathbf{P}$ by analyzing the transitions between the underlying Markov chain inferred in step (i).

The first step of the inference procedure relies on the minimization process defined by

$$\min |f^n(x, y) - f^e(x, y)|$$

where $f^n(x, y)$, as defined in (6), is a weighted sum of n bi-variate Gaussian distributions and n is determined as part of the inference process and does not need to be defined *a priori*.

The iterative fitting procedure starts with $n = 1$, finds the global maximum of $\epsilon_0(x, y) = f^e(x, y)$ and defines that point as the center of the first mobility region. Each mobility region (n) will have associated a bi-variate Gaussian distribution with parameters $\lambda_n^{(x)}, \lambda_n^{(y)}, \sigma_n^{(x)}, \sigma_n^{(y)}, \rho_n$ and weight π_n. The n-th mobility region center (i.e. the bi-variate Gaussian distribution mean) is determined by:

$$(\lambda_n^{(x)}, \lambda_n^{(y)}) = \epsilon_{n-1}^{-1} \left(\max(\epsilon_{n-1}(x, y))\right) \quad , n \geq 1. \tag{10}$$

The inference of the remaining parameters of the bi-variate Gaussian distributions are estimated by analyzing the bi-dimensional local decay behavior of

$f^e(x,y)$ in the vicinity of the local maximum points along an orthogonal plane. Therefore, based on the bi-variate Gaussian distribution function (5), it is possible to determine the ratio between the partial derivative of $\pi_n \Gamma_n(x,y)$ with respect to the variable x in the plane defined by $y = \lambda_n^{(y)}$ and $\pi_n \Gamma_n(x, \lambda_n^{(y)})$ as

$$\frac{\frac{\pi_n \partial \Gamma_n(x,\lambda_n^{(y)})}{\partial x}}{\pi_n \Gamma_n(x,\lambda_n^{(y)})} = -\frac{(x-\lambda_n^{(x)})}{(1-\rho_n^2)(\sigma_n^{(x)})^2} \Leftrightarrow (\sigma_n^{(x)})^2 = \frac{\gamma_n^{(x)}}{1-\rho_n^2} \tag{11}$$

and the ratio between the partial derivative of $\pi_n \Gamma_n(x,y)$ with respect to the variable y in the plane defined by $x = \lambda_n^{(x)}$ and $\pi_n \Gamma_n(\lambda_n^{(x)}, y)$ as

$$\frac{\frac{\pi_n \partial \Gamma_n(\lambda_n^{(x)},y)}{\partial y}}{\pi_n \Gamma_n(\lambda_n^{(x)},y)} = -\frac{(y-\lambda_n^{(y)})}{(1-\rho_n^2)(\sigma_n^{(y)})^2} \Leftrightarrow (\sigma_n^{(y)})^2 = \frac{\gamma_n^{(y)}}{1-\rho_n^2} \tag{12}$$

where

$$\gamma_n^{(x)} = -\frac{(x-\lambda_n^{(x)})\Gamma_n(x,\lambda_n^{(y)})}{\frac{\partial \Gamma_n(x,\lambda_n^{(y)})}{\partial x}} \qquad \text{and} \qquad \gamma_n^{(y)} = -\frac{(y-\lambda_n^{(y)})\Gamma_n(\lambda_n^{(x)},y)}{\frac{\partial \Gamma_n(\lambda_n^{(x)},y)}{\partial y}}. \tag{13}$$

Based on definitions (5) and (13) it is possible to redefine the n-th weighted bi-variate Gaussian distribution function

$$\Gamma_n(x,y) = \Gamma_n(\lambda_n^{(x)}, \lambda_n^{(y)}) e^{Q'_n} \tag{14}$$

with $\Gamma_n(\lambda_n^{(x)}, \lambda_n^{(y)}) = \pi_n \frac{\sqrt{1-\rho_n^2}}{2\pi\sqrt{\gamma_n^{(x)}\gamma_n^{(y)}}}$ and
$Q'_n = -\frac{1}{2}\left(\frac{(x-\lambda_n^{(x)})^2}{\gamma_n^{(x)}} + \frac{(y-\lambda_n^{(y)})^2}{\gamma_n^{(y)}} - \frac{2\rho_n(x-\lambda_n^{(x)})(y-\lambda_n^{(x)})}{\sqrt{\gamma_n^{(x)}\gamma_n^{(y)}}}\right)$.

¿From equation (14), it is possible to define the n-th bi-variate Gaussian distribution correlation:

$$\rho_n = \frac{\sqrt{\gamma_n^{(x)}\gamma_n^{(y)}}}{(x-\lambda_n^{(x)})(y-\lambda_n^{(y)})} r_n(x,y) \tag{15}$$

where
$r_n(x,y) = ln\left(\frac{\Gamma_n(x,y)}{\Gamma_n(\lambda_n^{(x)},\lambda_n^{(y)})}\right) + \frac{(x-\lambda_n^{(x)})^2}{2\gamma_n^{(x)}} + \frac{(y-\lambda_n^{(y)})^2}{2\gamma_n^{(y)}}$, and the n-th bi-variate Gaussian distribution weight (i.e. the stationary probability of the underlying Markov chain n-th state) is given by:

$$\pi_n = \frac{2\pi\Gamma_n(\lambda_n^{(x)}, \lambda_n^{(y)})\sqrt{\gamma_n^{(x)}\gamma_n^{(y)}}}{\sqrt{1-\rho_n^2}} \tag{16}$$

It is assumed that within a square region with width $2\Delta_n^{(x)}$, height $2\Delta_n^{(y)}$ and center at the n-th local maximum of $f^e(x,y)$ (equation (10)), the position is

modeled exclusively by the n-th weighted bi-variate Gaussian distribution, i.e., $\Gamma_n(x,y) \approx \epsilon_{n-1}(x,y), x \in [\lambda_n^{(x)} - \Delta^{(x)}, \lambda_n^{(x)} + \Delta^{(x)}] \wedge y \in [\lambda_n^{(y)} - \Delta^{(y)}, \lambda_n^{(y)} + \Delta^{(y)}]$. Therefore, it is possible to find estimators for parameters $\gamma_n^{(x)}$ and $\gamma_n^{(y)}$ based on the average value of equations (13) calculated in the vicinities of the $f^e(x,y)$ local maximums:

$$\bar{\gamma}_n^{(x)} = -\frac{1}{2\Delta^{(x)}} \int_{\lambda_n^{(x)} - \Delta^{(x)}}^{\lambda_n^{(x)} + \Delta^{(x)}} \frac{(x - \lambda_n^{(x)})\epsilon_{n-1}(x, \lambda_n^{(y)})}{\frac{\partial \epsilon_{n-1}(x, \lambda_n^{(y)})}{\partial x}} \, dx \tag{17}$$

and

$$\bar{\gamma}_n^{(y)} = -\frac{1}{2\Delta^{(y)}} \int_{\lambda_n^{(y)} - \Delta^{(y)}}^{\lambda_n^{(y)} + \Delta^{(y)}} \frac{(y - \lambda_n^{(y)})\epsilon_{n-1}(\lambda_n^{(x)}, y)}{\frac{\partial \epsilon_{n-1}(\lambda_n^{(x)}, y)}{\partial y}} \, dy. \tag{18}$$

The estimator for the n-th bi-variate Gaussian distribution correlation is defined as the average value of (15) calculated over a square region in the vicinity of the $f^e(x,y)$ local maximums:

$$\bar{\rho}_n = \frac{\sqrt{\bar{\gamma}_n^{(x)} \bar{\gamma}_n^{(y)}}}{4\Delta^{(x)}\Delta^{(y)}} \int_{\lambda_n^{(x)} - \Delta^{(x)}}^{\lambda_n^{(x)} + \Delta^{(x)}} \int_{\lambda_n^{(y)} - \Delta^{(y)}}^{\lambda_n^{(y)} + \Delta^{(y)}} \frac{\bar{r}_n(x,y)}{(x - \lambda_n^{(x)})(y - \lambda_n^{(y)})} \, dydx \tag{19}$$

where $\bar{r}_n(x,y) = ln\left(\frac{\epsilon_{n-1}(x,y)}{\epsilon_{n-1}(\lambda_n^{(x)}, \lambda_n^{(y)})}\right) + \frac{(x - \lambda_n^{(x)})^2}{2\gamma_n^{(x)}} + \frac{(y - \lambda_n^{(y)})^2}{2\gamma_n^{(y)}}$.

The weight of the n-th bi-variate Gaussian distribution can be calculated using the estimators obtained with equations (17), (18) and (19):

$$\bar{\pi}_n = \frac{2\pi \epsilon_{n-1}(\lambda_n^{(x)}, \lambda_n^{(y)}) \sqrt{\bar{\gamma}_n^{(x)} \bar{\gamma}_n^{(y)}}}{\sqrt{1 - \bar{\rho}_n^2}} \tag{20}$$

The n-th bi-variate Gaussian distribution variances can be obtained based on equations (11), (12), (17), (18) and (19):

$$(\bar{\sigma}_n^{(x)})^2 = \frac{\bar{\gamma}_n^{(x)}}{(1 - \bar{\rho}_n^2)} \quad \text{and} \quad (\bar{\sigma}_n^{(y)})^2 = \frac{\bar{\gamma}_n^{(y)}}{(1 - \bar{\rho}_n^2)} \tag{21}$$

The final iteration step is to re-calculate the difference function and update $n = n + 1$:

$$\epsilon_n(x,y) = \epsilon_{n-1}(x,y) - \bar{\pi}_n \Gamma_n(x,y). \tag{22}$$

The iterative process finishes after a predefined number of iterations or when $\min |f^n(x,y) - f^e(x,y)|$ reaches a predefined fitting error; if that goal is not achieved, the value n is increased by one and the iterative process continues recomputing equations (10) to (22). At the end of the iteration process it is possible to determine the model matrices $\mathbf{L}$ and $\mathbf{S}$.

The final step of the inference procedure is to infer matrix $\mathbf{P}$, i.e. the transition probabilities between the regions/states defined in the first step. The task is achieved by probabilistically mapping each known position $(x_k, y_k), k = 0, 1, \ldots, K$ to one region and then average the probabilistic transitions between states.

The position in time interval k (x_k, y_k) will be probabilistically mapped to one of the N regions according to a probability vector $\mathbf{q}_k = \{q_{k,1}, q_{k,2}, \ldots, q_{k,N}\}, k = 0, 1, \ldots, K$ defined by:

$$q_{k,n} = \frac{\pi_n \Gamma_n(x_k, y_k)}{f(x_k, y_k)} \tag{23}$$

Since the underlying Markov chain at instant k is in state m with probability $P(J_k = m)$, it is possible to define a partial estimator for p_{mn} at all k time instants as (taking equations 1 and 2 into account):

$$\begin{aligned} \bar{p}_{mn,k} &= P(J_k = m)p_{mn} = P(J_k = m)P(J_{k+1} = n|J_k = m) = \\ &= P(J_k = m)P(J_{k+1} = n) = q_{k,m}q_{k+1,n} \end{aligned} \tag{24}$$

Therefore, the overall estimator for p_{mn} can then be inferred by averaging over k the partial estimators defined in (24),

$$\bar{p}_{mn} = \frac{1}{K-1} \sum_{k=0}^{K-1} \bar{p}_{mn,k}, \quad m, n = 1, \ldots, N \tag{25}$$

Rewriting (25) in matrix form, it is possible to obtain matrix $\mathbf{P}$:

$$\mathbf{P} = \frac{1}{K-1} \sum_{k=0}^{K-1} \mathbf{q}_k^T \times \mathbf{q}_{k+1} \tag{26}$$

5 Results

As proof of concept, we applied the proposed model to mobility data of taxicabs in San Francisco, USA. Data was provided by Exploratorium, the San Francisco museum of science, art and human perception, through the cabspotting project (http://cabspotting.org). A pre-processed data set is freely available [16] and contains GPS coordinates, sampled at approximately 1 per minute, of 500 taxis collected over 30 days of May 2008 in the San Francisco Bay Area. Fig. 1 depicts a sample trajectory of one taxi in the San Francisco Bay area. In this paper, in order to infer and test the proposed model we have only considered the movements within a rectangular region in the San Francisco downtown, which is depicted in Fig. 1 within the white rectangle, with longitude limits between -122.447° and -122.388° and latitude limits between 37.77° and 37.81°. The data subset was randomly divided in two sub-sets (of 250 cars each), one to infer the model parameters and the other to test the model.

Fig. 1. San Francisco map with one sample taxi movement (base map source: GoogleMaps).

Applying the proposed mobility model to the test data sub-set, a model with 64 states (sub-regions) was obtained. The model inference time interval was less than 30 minutes with an octave/Matlab implementation in a generic PC. Fig. 2 shows the position probability distribution map for a taxi in San Francisco downtown and Fig. 3 shows the stationary position probability distribution map of the dMMBVGP inferred model. It is possible to observe that the dMMBVGP model was able to incorporate and model the complexity of the cars movement/location in San Francisco Downtown. With a more detailed analysis, it is possible to observe that the inference algorithm centered the sub-regions in the more visited/crowded San Francisco's landmarks.

In order to evaluate the location prediction performance of the model, we used the test sub-set to predict the node future position based on a known position, using different time-lags (1 to 20 timeslots, approximately 1 to 20 minutes). The performance metric was the (average) distance (in meters) between the predicted and real future node position. After inferring the model parameters, position prediction using the dMMBVGP, equation (7), is performed in hundreds of microseconds (average value $\sim 921\mu$s), using also an octave/Matlab implementation in a generic PC. Note that the prediction time interval depends on the number of future time slots for which the prediction is made.

As a complementary evaluation of the proposed model performance, we also evaluate other four mobility and location prediction models in the same scenario:

- Complete random model, where node location follows a uniform distribution over the complete geographic region under study.
- Grid random model, where node location is probabilistically determined in two phases; firstly, the location sub-region is determined according to the empirical location distribution map inferred over a uniform grid and, secondly,

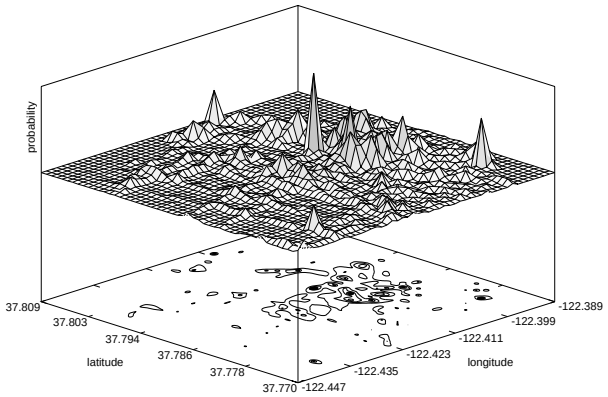

Fig. 2. Real position probability map, San Francisco downtown only.

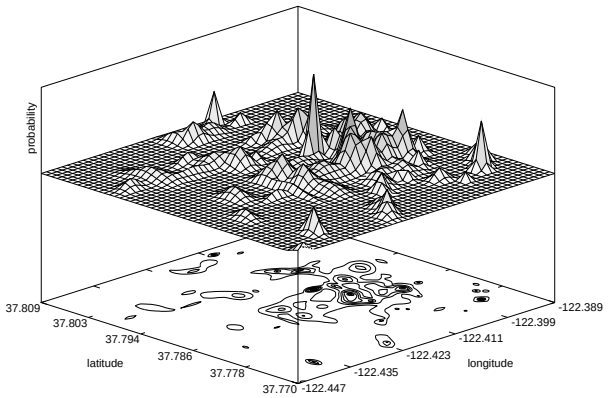

Fig. 3. Fitted dMMBVGP position probability map, San Francisco downtown only.

the exact position within the chosen sub-region is determined according to a uniform distribution.

- Random walk model [5] with reflection at the region borders [15], where the node direction of movement is uniformly distributed in $[0, 2\pi]$ and the speed has a Gaussian distribution with mean and standard deviation inferred from empirical data.
- Markov grid model, which is a sub-region based Markov family model [20] where location is determined by an underlying Markov chain that determines the transition between sub-regions. We choose to define the sub-regions as an uniform grid. Within a sub-region, node location has an uniform distribution.

The complete random model has no geographic dependency and has no information of the node movement history. The model location prediction does not depend on the current node position. This model is used as the worst case reference value. The grid random model prediction does not depend also on the

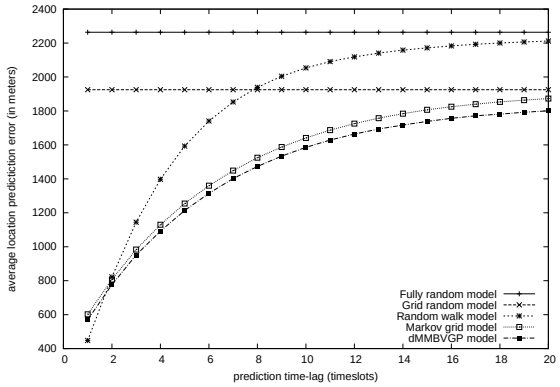

Fig. 4. Models relative location prediction performance.

current node position, however it incorporates some geographic dependency inferred from empirical historic data. A 8x8 grid (64 equal sub-regions) was used to infer the empirical location distribution map. The random walk model has no geographic dependency but the prediction process utilizes the current node position and extrapolates the future position based on a known constant speed and direction. The model speed was modeled by a Gaussian distribution with mean 302.1m/minute ($\approx$ 18.1Km/h or $\approx$ 11.2mph) and standard deviation of 346.7m/minute. For the inference of the Markov grid model, a 8x8 grid (64 equal sub-regions) was used in order to obtain a model with the same level of complexity as the dMMBVGP. Note that the Markov grid model differs from the dMMBGP model in the way regions are defined and position is modeled within each region; the time dependability (based on an underlying Markov chain) is modeled and inferred in a similar way in both models.

Figure 4 depicts the models prediction performance. It is possible to observe that the dMMBGP outperforms all the other models and it is possible to obtain relatively small prediction errors, in a region so large and complex as the one under study, even in long range predictions (20 timeslots $\sim$ 20 minutes). As expected, the worst performing models are the ones that do not receive any real-time input at the time of the prediction, and therefore the prediction error is independent of the time-lag. The random walk model is able to surpass the others in very short time-lag predictions ($\leq$ 2 minutes); however the model is not able to capture the complex dynamics of the taxis movements and the prediction error greatly increases when the prediction range is larger than 1 slot and rapidly approximates the error of a completely random location choice. The Markov grid model, due to its nature, is able to capture the complexity and dynamism of the movements; however, the definition of generic regions restricts the model performance. Therefore, the dMMBVGP model is able to model the complexity of the movements due to its Markovian nature, while its adaptable definition of sub-regions allows it to maintain the model complexity at low levels and achieve very good location prediction results.

6 Conclusions

Being able to model movement and predict future positioning of mobile nodes in complex environments can be very important to several operational and management tasks, like evaluation of different system design alternatives, simulation and study of new routing protocols, message delivery and nodes connectivity optimization or mobile computing. Since existing mobility modeling approaches are based on simple and limited models that were specifically designed for particular application scenarios or require the complete knowledge of the mobility environment, this paper presented a discrete time Markov Modulated Bi-variate Gaussian Process that is able to characterize the position and mobility of any mobile node by assuming that the position within a generic sub-region can be described by a bi-variate Gaussian distribution and the transition between sub-regions can be described by an underlying (homogeneous) Markov chain. This approach can be applied to scenarios where the possible pathways are unknown or too complex to consider in a real model that must make a prediction in a very short time. The results obtained by applying the proposed model to real and publicly available mobility data of San Francisco taxi-cabs demonstrate the accuracy and utility of this approach: the model was able to efficiently describe the movement patterns of the mobile nodes and predict their future position. Besides, the model has also revealed higher performances when compared to other modeling approaches.

References

1. Akyildiz, I., Wang, W.: The predictive user mobility profile framework for wireless multimedia networks. IEEE/ACM Transactions 12(6), 1021–1035 (2004)
2. Bettstetter, C.: Mobility modeling in wireless networks: Categorization, smooth movement and border effects. ACM Mobile Computing and Communications Review 5(3), 55–67 (2001)
3. Bittner, S., Raffel, W.U., Scholz, M.: The area graph-based mobility model and its impact on data dissemination. In: Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications Workshops (2005)
4. Broch, J., Maltz, D.A., Johnson, D.B., Hu, Y.C., Jetcheva, J.: A performance comparison of multi-hop wireless ad hoc network routing protocols. In: Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom98) (Oct 1998)
5. Camp, T., Boleng, J., Davies, V.: A survey of mobility models for ad hoc network research. Wireless Communications and Mobile Computing - Special Issue in Mobile Ad Hoc Networking: Research, Trends and Applications 2(5), 483–502 (2002)
6. Hsu, W., Merchant, K., Shu, H., Hsu, C., Helmy, A.: Weighted waypoint mobility model and its impact on ad hoc networks. SIGMOBILE Mobile Computing Communication Review 9(1), 59–63 (2005)
7. Hsu, W.J., Spyropoulos, T., Psounis, K., Helmy, A.: Modeling spatial and temporal dependencies of user mobility in wireless mobile networks. Networking, IEEE/ACM Transactions on 17(5), 1564 –1577 (2009)

8. Hyytia, E., Lassila, P., Virtamo, J.: A markovian waypoint mobilitymodel with application to hotspot modeling. In: Proceedings of the IEEE ICC (2006)
9. Johansson, P., Larsson, T., Hedman, N., Mielczarek, B., Degermark, M.: Scenario-based performance analysis of routing protocols for mobile ad-hoc networks. In: Proc. of the International Conference on Mobile Computing and Networking (MobiCom'99) (1999)
10. Lee, J.K., Hou, J.C.: Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application. In: MobiHoc '06: Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing. pp. 85–96. ACM, New York, NY, USA (2006)
11. Liang, B., Haas, Z.J.: Predictive distance-based mobility management for PCS networks. In: Proceedings of IEEE Information Communications Conference (INFOCOM 1999) (Apr 1999)
12. Liu, T., Bahl, P., Chlamtac, I.: Mobility modeling, location tracking and trajectory prediction in wireless networks. IEEE Journal on Special Areas in Communications 16(6), 922–936 (1998)
13. Musolesi, M., Mascolo, C.: Designing mobility models based on social network theory. SIGMOBILE Mobile Computing Communication Review 11(3), 59–70 (2007)
14. Pacheco, A., Tang, L.C., Prabhu, N.U.: Markov-modulated processes & semiregenerative phenomena. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2009)
15. Papoulis, A.: Probability, random variables, and stochastic processes. McGraw-Hill (1984)
16. Piorkowski, M., Sarafijanovic-Djukic, N., Grossglauser, M.: CRAWDAD data set epfl/mobility (v. 2009-02-24). Downloaded from http://crawdad.cs.dartmouth.edu/epfl/mobility (February 2009)
17. Piorkowski, M., Sarafijanovoc-Djukic, N., Grossglauser, M.: A Parsimonious Model of Mobile Partitioned Networks with Clustering. In: The First International Conference on COMmunication Systems and NETworkS (COMSNETS) (January 2009), http://www.comsnets.org
18. Royer, E.M., Melliar-Smith, P.M., Moser, L.E.: An analysis of the optimum node density for ad hoc mobile networks. In: Proceedings of the IEEE International Conference on Communications (ICC) (Jun 2001)
19. Soh, W., Kim, H.: Dynamic bandwidth reservation in cellular networks using road topology based mobility predictions. In: Proceedings of the IEEE Infocom. vol. 4, pp. 2766–2777 (2004)
20. Song, L., Kotz, D., Jain, R., He, X.: Evaluating location predictors with extensive wi-fi mobility data. In: INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies. vol. 2, pp. 1414–1424 (7-11 2004)
21. Spyropoulos, T., Psounis, K., Raghavendra, C.S.: Performance analysis of mobility-assisted routing. In: Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing (2006)
22. Tian, J., Hahner, J., Becker, C., Stepanov, I., Rothermel, K.: Graph-based mobility model for mobile ad hoc network simulation. In: Proceedings of 35th Annual Simulation Symposium, in cooperation with the IEEE Computer Society and ACM (Apr 2002)
23. Yang, S., Yang, X., Zhang, C., Spyrou, E.: Using social network theory for modeling human mobility. Network, IEEE 24(5), 6 –13 (2010)