DATA DRIVEN VESSEL TRAJECTORY FORECASTING USING STOCHASTIC GENERATIVE MODELS

Murat Üney, Leonardo M. Millefiori, Paolo Braca

NATO STO Centre for Maritime Research and Experimentation
Viale San Bartolomeo 400, 19126 La Spezia SP, Italy
{ murat.uney, leonardo.millefiori, paolo.braca } @ cmre.nato.int

ABSTRACT

In this work, we propose a data driven trajectory forecasting algorithm that utilizes both recorded historical and streaming trajectory observations. The algorithm performs Bayesian inference on a directed graph the walks on which represent stochastic change point models of trajectory classes. Parameter distributions of these models are learnt from recorded trajectories. Forecasting is then made by calculating the class – or, walk– probabilities and corresponding predictive distributions for a given stream of location and velocity observations. This approach is tailored for the maritime domain and automatic identification system (AIS) data exploitation through the use of an Ornstein-Uhlenbeck process driven stochastic process model that captures vessel motion characteristics. We demonstrate the efficacy of this approach on a real data set.

Index Terms— Change point models, Ornstein-Uhlenbeck processes, predictive models, Maritime traffic analysis

1. INTRODUCTION

Long term forecasting of vessel routes is a highly desired capability for safety analysis and planning in maritime situational awareness [1–3]. This work aims to perform trajectory forecasting given trajectory samples in the form of online received Automatic Identification System (AIS) messages from a certain vessel based on a trajectory corpus comprised of recorded AIS streams. These messages report the location and velocity of vessels in an asynchronous and intermittent fashion. Recorded AIS data facilitates learning of patterns of trajectories and their quantitive summaries offline. These patterns are then used for forecasting future position of a vessel by assessing how well the live stream fits into the learnt patterns and accordingly combining predictions output by each pattern.

We propose a hierarchical generative model that comprises a stochastic process (SP) with temporal change points and associated parameter distributions modelling classes of continuous state trajectories. Such change point parameters can be viewed as an encoding of a piece-wise continuous function that maps time to the parameter values to which the underlying SP model is conditioned, thereby capturing piecewise stationary phenomenon. We capture the variability in the trajectories from the same class by using random variable change points and parameters.

The change point parameters for a given trajectory class constitute a directed Markov chain. Bayesian trajectory forecasting hence involves computing the prediction density over the state for selected future time instants over these chains (or, walks over the graph obtained as the join of these chains [4]) and computing the class likelihoods (or, the walk likelihoods) given the observations received so far. We provide explicit formulae and Monte Carlo (MC) computational procedures that specify a data driven trajectory forecasting algorithm.

Previous work on traffic and trajectory analysis has considered Gaussian process models and their sparse variants for modelling continuous trajectories in surveillance, robotic perception, and, anomaly detection (see, e.g., [5–7]). In [8], GP regression is used to learn a distribution over continuous state trajectories using intermittent observations. The cubic computational complexity of GP regression with the number of data samples and the non-stationary nature of maritime trajectories, however, complicate their use.

Parametric models that follow from kinematics, on the other hand, have lead to efficient and interpretable trajectory representations. For example, [9] proposes a family of stochastic models that are conditioned on a future position intended as the destination. We use an Ornstein-Uhlenbeck (OU) process model for capturing vessel velocity characteristics stemming from restricted maneuverability in the maritime domain. These models have proved useful in the maritime domain for streaming-only AIS based prediction [10, 11] and traffic visualisation [12]. This work accommodates OU driven kinematic models in a hierarchical random changepoint model for enabling data driven forecasting.

The article is organised as follows: Sec. 2 provides the mathematical problem definition. The proposed hierarchical trajectory model and approximate Bayesian computations are

This work is supported by the NATO Allied Command Transformation (ACT) via the project Data Knowledge Operational Effectiveness (DKOE) at the NATO Science and Technology Organization (STO) CMRE.

detailed in Sec. 3 and 4, respectively. We discuss learning of relevant distributions in Sec. 5. Then, we demonstrate the proposed approach with real data in Sec. 6 and conclude.

2. PROBLEM DEFINITION

We consider AIS messages reporting position and velocity of vessels labelled with a time tag and a unique vessel identification number. Let us denote all messages retrieved from vessel with ID l by $\mathbf{d}^{(l)}$. Each message $d \in \mathbf{d}^{(l)}$ is a pair $d = (x, t)$ where t is a time tag and x captures a state observation $x \triangleq [s, \dot{s}]$ such that s is the location and $\dot{s}$ is the velocity. These messages are collected with irregular time intervals. Let us denote all such messages by $\mathcal{D} \triangleq \{\mathbf{d}^{(l)}\}$.

Now, let us consider an online AIS stream from some vessel and denote the most recent messages by $\mathbf{d} \triangleq (\bar{\mathbf{t}}, \bar{\mathbf{x}})$ and a future state at t_f by x_f. Our goal is to specify a (forecast) distribution for the (associated) random variable $\boldsymbol{X}_f$ that is conditioned on the recorded data $\mathcal{D}$ and the streaming data $\mathbf{d}$ and generates $\bar{\mathbf{x}}$. Let us consider the probability density function (PDF) of this distribution:

$$
\begin{aligned}
p(x_f|\mathbf{d}, \mathcal{D}; t_f) &= \sum_{y \in \mathcal{Y}} p(x_f|\bar{d}, \mathcal{D}, y; t_f) p(y|\mathcal{D}), \\
&= \sum_{y \in \mathcal{Y}} \frac{p(x_f, \bar{\mathbf{x}}|\mathcal{D}_y, y; \bar{\mathbf{t}}, t_f)}{p(\bar{\mathbf{x}}|\mathcal{D}_y, y; \bar{\mathbf{t}}, t_f)} p(y|\mathcal{D}), \quad (1)
\end{aligned}
$$

where y is a class variable taking values from a finite set of (trajectory) classes $\mathcal{Y}$, and, $\mathcal{D}_y \subset \mathcal{D}$ is the set of trajectories of class y. Here, the product of the PDFs inside marginalisation in the first line is often referred to as a generative model. In the second line, both the numerator and the denominator are evaluations of a class conditional density $p(\mathbf{x}|., y)$ at $[x_f, \bar{\mathbf{x}}]$ and $\mathbf{x} = \bar{\mathbf{x}}$, respectively. This PDF is similar to the "posterior predictive distribution" in a regression context [13].

The partitioning of the data $\mathcal{D}$ into classes can be performed in an unsupervised/automated fashion using one of the many methods in the literature (see, e.g., [14] for a review and [15] for a variational method). We give details of our approach later in Sec. 5.

In the rest of this article, we will avoid explicit conditioning over the observation times unless it is inevitable, for the sake of simplicity in notation.

Now, let us focus on the class conditional density in (1) and assert the assumption that the trajectory observations admit a parametric representation. It follows that

$$
\begin{aligned}
p(\mathbf{x}|\mathcal{D}_y, y) &= \int p(\mathbf{x}|\boldsymbol{\psi}, \mathcal{D}_y, y) p(\boldsymbol{\psi}|\mathcal{D}_y, y) \mathrm{d}\boldsymbol{\psi} \\
&= \int p(\mathbf{x}|\boldsymbol{\psi}) p(\boldsymbol{\psi}|\mathcal{D}_y, y) \mathrm{d}\boldsymbol{\psi} \quad (2)
\end{aligned}
$$

where $\boldsymbol{\psi}$ capture parameters that specify the distribution on the state observations. The second line above follows from the assumption that the state observations are fully explained by $\boldsymbol{\psi}$ and hence are conditionally independent from the other variables.

Our problem is to realise (1) by substituting from (2) for data driven trajectory forecasting. In the next section, we detail the PDFs inside the integral in (2). Then, in Sec. 4, we introduce a MC approximation for its evaluation.

3. THE CLASS CONDITIONAL DENSITY

3.1. Route generating model $p(\mathbf{x}|\psi)$

A route is a spatial trajectory $s(t)$ that a vessel follows $t_S \leqslant t \leqslant t_E$ where t_S and t_E indicate the starting and end times, respectively. This trajectory concatenated with the velocity vector of the vessel $\dot{s}(t)$ constitute a state trajectory $x(t) \triangleq [s(t), \dot{s}(t)]$ for $t_S \leqslant t \leqslant t_E$. We treat $x(t)$ as a realisation of a random variable associated with a stochastic process, i.e., x is a realisation of a stochastic process with the boundary conditions given by $x(t_S), x(t_E)$, respectively.

Each realisation x involves M change-points modelling changes of path as the journeys are typically divided into legs. These change times are denoted by $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_M]$ and satisfy $t_S < \tau_1 < \ldots < \tau_M < t_E$. The state trajectory is observed at N points $\mathbf{x} = [x(t_1), \ldots, x(t_N)]$ for $t_S < t_1 < \ldots < t_N < t_E$ and at the boundary points $\partial\boldsymbol{x} \triangleq (x(t_S), x(t_E))$. Let us define the time boundaries as $\partial\boldsymbol{t} \triangleq [t_S, t_E]$ and denote the data collection times by $\boldsymbol{t} = [t_1, \ldots, t_N]$. Then, the observation likelihood factorises in accordance with a change-point model [16]:

$$
p(\mathbf{x}|\boldsymbol{\tau}, \boldsymbol{\theta}; \boldsymbol{t}, \partial\mathbf{x}, \partial\boldsymbol{t}) = \prod_{i=1}^{M+1} p(\mathbf{x}_i|\theta_i; \boldsymbol{t}_i, \partial\mathbf{x}_i, \partial\boldsymbol{\tau}_i), \quad (3)
$$

where $\mathbf{x}_i \triangleq \{x(t_k)|\tau_{i-1} < t_k < \tau_i\}$ are the observations associated with the ith leg excluding the boundary points $\partial\mathbf{x}_i \triangleq (x_{\tau_{i-1}}, x_{\tau_i})$ suggesting that $x(\tau_{i-1}) = x_{\tau_{i-1}}$ and $x(\tau_i) = x_{\tau_i}$, repsectively. Here, $\tau_0 = t_S$ and $\tau_{M+1} = t_E$. The parameter vector θ_i specify the likelihood induced by the selected stochastic process family leading to the journey parameters $\boldsymbol{\theta} \triangleq [\theta_1, \ldots, \theta_{M+1}]$. The likelihood is also a function of the observation times $\boldsymbol{t}_i \triangleq \{t_k|\tau_{i-1} < t_k < \tau_i\}$, and, the time boundaries $\partial\boldsymbol{\tau}_i \triangleq [\tau_{i-1}, \tau_i]$. Note that $\boldsymbol{\theta}$, $\partial\boldsymbol{\tau}_i$s and $\partial\mathbf{x}_i$s constitute ψ for an arbitrary parametric model. We explicitly specify this likelihood for an Ornstein-Uhlenbeck process driven kinematics model in the next subsection.

3.2. Ornstein-Uhlenbeck driven stochastic kinematics

Let us consider the data likelihood in (3) for a state trajectory generated by an Ornstein-Uhlenbeck process $\dot{s}$. An OU process is equivalently specified by the state dynamics given by

$$
\ddot{s}(t) = \boldsymbol{\Gamma}\left(\boldsymbol{v} - \dot{s}(t)\right) + \mathbf{H}\dot{\boldsymbol{n}}(t) \quad (4)
$$

for some $t \in [\tau_{i-1}, \tau_i]$ where $\boldsymbol{n}$ is a sample generated by a 2-D Wiener process. Here, $\boldsymbol{v} = [v_x, v_y]^T$ is the expectation of $\dot{s}$ over time and also acts as a control input thereby leading to a mean-reverting behaviour. The characteristics of this behaviour is determined by $\mathbf{\Gamma}$ and the *process* noise covariance $\mathbf{H}\mathbf{H}^T$. We assume that the eigen-vectors of $\mathbf{\Gamma}$ are the unit vector along $\boldsymbol{v}$ and its $\pi/2$ rotated version $\boldsymbol{v}^{\perp}$ and the noise terms are independent along these directions [17], i.e.,

$$\mathbf{E} = [\boldsymbol{v}, \boldsymbol{v}^{\perp}], \mathbf{\Gamma} = \mathbf{E}\left[\begin{array}{cc}\gamma_1 & 0 \\ 0 & \gamma_2\end{array}\right]\mathbf{E}^T, \mathbf{H} = \mathbf{E}\left[\begin{array}{cc}\sigma_1 & 0 \\ 0 & \sigma_2\end{array}\right]. \quad (5)$$

The state dynamics which relate the location and velocity components for the case is given by

$$\dot{\boldsymbol{x}}(t) = \left[\begin{array}{cc}\mathbf{0}, & \mathbf{I} \\ \mathbf{0}, & -\mathbf{\Gamma}\end{array}\right]\boldsymbol{x}(t) + \left[\begin{array}{c}\mathbf{0} \\ \mathbf{\Gamma}\end{array}\right]\boldsymbol{v} + \left[\begin{array}{c}\mathbf{0} \\ \mathbf{H}\end{array}\right]\dot{\boldsymbol{n}}(t), \quad (6)$$

where $\mathbf{I}$ and $\mathbf{0}$ are the 2×2 identity and null matrices, respectively.

Let us introduce $\theta_i = [\boldsymbol{v}_i, \gamma_{1,i}, \gamma_{2,i}, \sigma_{1,i}, \sigma_{2,i}]$ and consider the data likelihood in (3). The state dynamics (6) with boundary conditions $\partial\boldsymbol{x}_i = [x_{\tau_{i-1}}, x_{\tau_i}]$ at $\partial\boldsymbol{\tau}_i = [\tau_{i-1}, \tau_i]$ induce a likelihood that is first-order Markov conditioned on θ_i and the boundary conditions [10]:

$$\begin{aligned} p(\mathbf{x}_i|\theta_i; \boldsymbol{t}_i, \partial\mathbf{x}_i, \partial\boldsymbol{\tau}_i) &= p(\mathbf{x}_{i,1}|\partial\mathbf{x}_{i,1}, \theta_i; \boldsymbol{t}_{i,1}, \partial\boldsymbol{\tau}_{i,1}) \\ &\times \prod_{k=1}^{D_i} p(\mathbf{x}_{i,k}|\mathbf{x}_{i,k-1}, \theta_i; \boldsymbol{t}_{i,k}, \boldsymbol{t}_{i,k-1}) \\ &\times p(\partial\mathbf{x}_{i,2}|\mathbf{x}_{i,D_i}, \theta_i; \partial\boldsymbol{\tau}_{i,2}, \boldsymbol{t}_{i,D_i}) \quad (7) \end{aligned}$$

where D_i is the number of entries in $\mathbf{x}_i$ (and, $\boldsymbol{t}_i$) and $\mathbf{x}_{i,k}$ denotes the kth entry. Here, all the factors are of the same form with the first and last terms taking into account the boundary conditions and they are Gaussians. Specifically, for some state values x_2, x_1 and observation times t_2, t_1,

$$\begin{aligned} &p(x_2|x_1, \theta; t_2, t_1) = \\ &\quad \mathcal{N}(\boldsymbol{x}_2; \mathbf{\Phi}(t_2 - t_1, \mathbf{\Gamma})x_1 + \mathbf{\Psi}(t_2 - t_1, \mathbf{\Gamma})\boldsymbol{v}, \mathbf{\Sigma}_n), \quad (8) \end{aligned}$$

where $\mathbf{\Phi}$, $\mathbf{\Psi}$, and, $\mathbf{\Sigma}_n$ are given by [10]:

$$\begin{aligned} \mathbf{\Phi}(\Delta t, \mathbf{\Gamma}) &= \tilde{\mathbf{E}}\left[\begin{array}{cc}\mathbf{I}, & \big(\mathbf{I} - \exp(-\mathbf{\Gamma}\Delta t)\big)\mathbf{\Gamma}^{-1} \\ \mathbf{0}, & \exp(-\mathbf{\Gamma}\Delta t)\end{array}\right]\tilde{\mathbf{E}}^T, \\ \mathbf{\Psi}(\Delta t, \mathbf{\Gamma}) &= \tilde{\mathbf{E}}\left[\begin{array}{c}\Delta t\mathbf{I} - \big(\mathbf{I} - \exp(-\mathbf{\Gamma}\Delta t)\big)\mathbf{\Gamma}^{-1} \\ \mathbf{I} - \exp(-\mathbf{\Gamma}\Delta t)\end{array}\right]\tilde{\mathbf{E}}^T \end{aligned}$$

and, $\tilde{\mathbf{E}} = \mathbf{I}\otimes\mathbf{E}$ with $\otimes$ denoting the Kronecker product. The covariance of concern is given by

$$\mathbf{\Sigma}_n = \tilde{\mathbf{E}}\Big(\mathbf{\Sigma}_1 \circ \mathbf{\Sigma}_2(\Delta t)\Big)\tilde{\mathbf{E}}^T\Big|_{\Delta t = t_2 - t_1} \quad (9)$$

where $\circ$ denotes the Hadamard product of the two matrices and $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2(\Delta t)$ are defined in [10, Eq.s (41),(52)–(55)].

3.3. Trajectory class parameter distributions

Let us consider the class parameter PDF $p(\psi|\mathcal{D}_{(y)},\ y)$ for all the parameters introduced in the previous sections, i.e.,

$$\psi = [\theta_1, \partial\mathbf{x}_1, \partial\boldsymbol{\tau}_1 \ldots, \theta_{M+1}, \partial\mathbf{x}_{M+1}, \partial\boldsymbol{\tau}_{M+1}]. \quad (10)$$

The parameters admit an ordering with respect to time. The boundary states are generated by the previously introduced SP model. We introduce the modelling assumption that parameters associated with different segments of the change-point model are independent. Thus, the parameter PDF factorises as

$$p(\psi|y) = p(x_{\tau_0}, \tau_0|y)\prod_{i=1}^{M+1} p(\Delta\tau_i, \theta_i)p(x_{\tau_i}|\theta_i, x_{\tau_{i-1}}, \Delta\tau_i) \quad (11)$$

where $\Delta\tau_i = \tau_i - \tau_{i-1}$ are the length of time intervals between parameter switches, and, conditioning on data is not shown for simplicity in the notation.

Note that the boundary state random variables generating x_{τ_i}s admit a Markov chain with transition densities specified by the SP model parameters and the time elapse of the segments, i.e., $\Delta\tau_i$s. The location components of these states are *journey way-points*. Therefore, (11) admits a Markov graph in the form of a chain.

As a result, the class parameter density in (11), the OU driven SP model in (7)–(9) and the change point model in (3) uniquely specify the class conditional density in (2) which is instrumental to data driven trajectory forecasting in (1). The integral over the parameters, however, do not admit analytic expressions for this density. In the next section, we provide a MC approximation to this density which leads to a Gaussian mixture forecast density.

4. MONTE CARLO APPROXIMATION TO THE FORECAST DENSITY

Let us rewrite the class conditional forecast density by omitting conditioning on the data and the observation times for the sake of simplicity:

$$p(x_{t_f}|\bar{\mathbf{x}}, y) = \frac{1}{p(\bar{\mathbf{x}}|y)}\int p(x_{t_f}|\bar{\mathbf{x}}, \boldsymbol{\psi})p(\bar{\mathbf{x}}|\boldsymbol{\psi})p(\boldsymbol{\psi}|y)\mathrm{d}\boldsymbol{\psi}. \quad (12)$$

The MC approximation [18] to the above integration is obtained by first sampling from the class parameter distribution and obtaining the sample set

$$\psi^{(l)} \sim p(\psi|y) \quad i = 1, \ldots, L \quad (13)$$

and using the MC method summation formula [18, Chp. 3] which leads to the approximation given by

$$p(x_{t_f}|\bar{\mathbf{x}}, y) \approx \frac{1}{\sum_{l'=1}^{L}\omega^{(l')}}\sum_{l=1}^{L}\ \omega^{(l)}p(x_{t_f}|\bar{\mathbf{x}}, \boldsymbol{\psi}^{(l)}), \quad (14)$$

$$\omega^{(l)} \triangleq p(\bar{\mathbf{x}}|\boldsymbol{\psi}^{(l)}). \quad (15)$$

where the PDF inside the summation in (14) is nothing but (3),(7)–(9) evaluated for the argument x_{t_f} and the boundary conditions, change point times and parameters selected in accordance with $\bar{\mathbf{x}}, \boldsymbol{\psi}^{(l)}$. Similarly, the weights in (15) are the observation likelihood of trajectory observation $\bar{\mathbf{x}}$ for class y.

The parameter samples used are generated from the class parameter distribution in accordance with the factorisation in (11). First, we set $\tau_0 = 0$ and then (with a notation that uses PDFs in place of the associated random variables) proceed as follows:

$$
\begin{aligned}
x_{\tau_0}^{(l)} &\sim p(x_{\tau_0}|y), \\
&\text{for} \quad i = 1, \ldots, M+1 \\
(\Delta\tau_i^{(l)}, \theta_i^{(l)}) &\sim p(\Delta\tau_i, \theta_i|y), \\
x_{\tau_i}^{(l)} &\sim p(x_{\tau_i}|\Delta\tau_i^{(l)}, x_{\tau_{i-1}}^{(l)}, \theta_i^{(l)}) \qquad (16)
\end{aligned}
$$

As a result, the proposed computational procedure for trajectory forecasting starts with this sampling procedure and proceeds by evaluating (14) and (15), and substituting them in (1) in place of the class conditional PDF.

5. LEARNING CLASS PARAMETER DENSITIES

The learning procedure begins with the partitioning of the trajectory corpus $\mathcal{D}$, as aforementioned. For the purpose of position forecasting, we classify these trajectories with respect to their initial and final state. We use a fixed spatial grid $\{\mathcal{S}_s\}$ such that the cells $\mathcal{S}_s$ are mutually exclusive and their union cover the spatial region the corpus lies in. Therefore $\mathbf{d}^{(l)} \in \mathcal{D}$ is labelled as (b, d) if and only if $x(t_S) \in \mathcal{S}_b$ and $x(t_E) \in \mathcal{S}_d$. All trajectories with the same label constitute an associated $\mathcal{D}_y$. The class posterior $p(y|\mathcal{D})$ in (1) is found by using an equal-probability prior and the cardinality ratio of $\mathcal{D}_y$ and $\mathcal{D}$ as the likelihood.

Next, we perform approximate learning of the boundary condition distributions by using change point detection for each trajectory in the class labelled data set. Specifically, for each class y, we use the algorithm in [19] and obtain estimates of change point timings $\hat{\boldsymbol{\tau}}^{(l)}$ for each trajectory $\mathbf{d}^{(l)}$ in $\mathcal{D}_y$. Then we fit a Gaussian mixture model for these timings using expectation maximisation [13]. The resulting number of components yield the number of change points M for each class. We also have obtained segments of trajectories which we use to find maximum likelihood estimates of the parameters $\hat{\psi}^{(l)}$ in (10) [17]. We use $\hat{\boldsymbol{\tau}}^{(l)}, \hat{\psi}^{(l)}$ from all $\mathbf{d}^{(l)} \in \mathcal{D}_y$ and find kernel density estimates [20] approximating $p(\Delta\tau_i, \theta_i|y)$s in (16). The samples of boundary conditions are generated from the SP model given parameter samples from these distributions.

6. EXAMPLE

In this section, we demonstrate the proposed algorithm using real AIS data collected over a two months period between $01.04.2016 - 01.06.2016$ from a region at the Ionian Sea. The grid based trajectory clustering explained in Sec. 5 is used with a 15×15 grid over this region which results with 128 distinct classes depicted using a colour codes in Fig. 1(a). The learning procedure detailed in Sec. 5 results with the chains

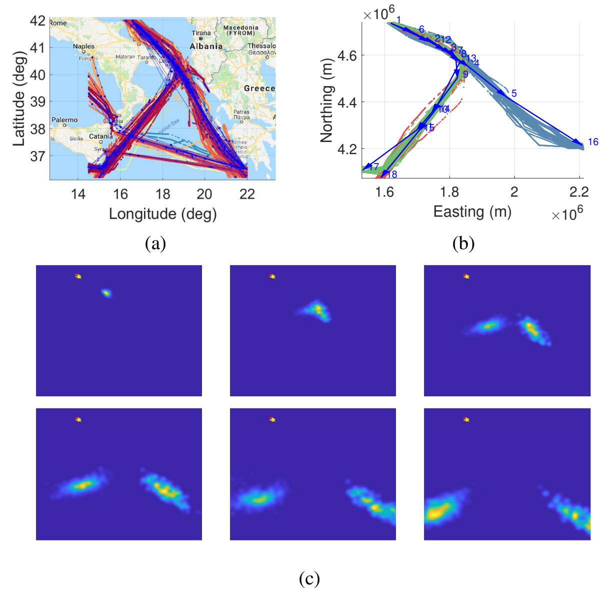

Fig. 1. Illustration of the data driven forecast: (a) AIS data and trajectory classes (colour coded). (b) Subgraph queried for trajectory forecasting (c). Querying trajectory (yellow crosses) and forecast densities obtained with 2e4 s steps.

whose join is the graph $\mathcal{G}$ seen in this figure. This graph has 265 nodes and 128 walks over 318 edges. The node locations here are found by taking the expectation of the marginal position distributions of the associated boundary conditions x_v induced by the parameter distributions learnt.

The proposed forecasting algorithm is used with a subgraph $\mathcal{H}$ that captures 3 walks given in Fig. 1(b). The walks lengths are found as $M = 3, 4, 4$, from node 1 towards nodes $16, 17$ and 18. The posterior class probabilities $p(y|\mathcal{D})$ are approximately $0.6, 0.2$ and 0.2, respectively. We query this graph $\mathcal{H}$ for a forecast with 10 initial samples of a trajectory depicted with yellow crosses in Fig. 1(c). The proposed algorithm (13)–(16) is used for a future time t_f that is 2e4 s after the last observation time stamp (with $L = 500$ samples for MC computations). We repeat these computations with 2e4 s steps into the future yielding the forecasts in Fig. 1(c). The forecasts demonstrate that the "Y" junction in the data associated with $\mathcal{H}$(Fig. 1(b)) is learnt by our model and reflected in the forecasts of the proposed algorithm.

7. CONCLUSION

In this article, we proposed a novel SP change point model for data driven trajectory forecasting in the maritime domain. We introduced MC computational procedures and obtained Gaussian mixture forecast densities. We demonstrated the efficacy of our approach on a real data set. Future work will focus on learning class parameter distributions from data.

8. REFERENCES

[1] B. Ristic, B. La Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in AIS data: Anomaly detection and motion prediction," in *2008 11th International Conference on Information Fusion*, June 2008, pp. 1–7.

[2] Giuliana Pallotta, Michele Vespe, and Karna Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.

[3] D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, "A multi-task deep learning architecture for maritime surveillance using ais data streams," in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, Oct 2018, pp. 331–340.

[4] Reinhard Diestel, *Graph Theory (Graduate Texts in Mathematics)*, Springer, August 2005.

[5] D. Ellis, E. Sommerlade, and I. Reid, "Modelling pedestrian trajectories with Gaussian processes," in *Proceedings of the 9th International Workshop on Visual Surveillance*, 2009.

[6] Sean Anderson, Timothy D. Barfoot, Chi Hay Tong, and Simo Särkkä, "Batch nonlinear continuous-time trajectory estimation as exactly sparse Gaussian process regression," *Autonomous Robots*, vol. 39, no. 3, pp. 221–238, Oct 2015.

[7] C. Zor and J. Kittler, "Maritime anomaly detection in ferry tracks," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 2647–2651.

[8] Kihwan Kim, Dongryeol Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 1164–1171.

[9] B. I. Ahmad, J. K. Murphy, P. M. Langdon, and S. J. Godsill, "Bayesian intent prediction in object tracking using bridging distributions," *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 215–227, Jan 2018.

[10] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2313–2330, October 2016.

[11] M. Uney, L. M. Millefiori, and P. Braca, "Prediction of rendezvous in maritime situational awareness," in *2018 21st International Conference on Information Fusion (FUSION)*, July 2018, pp. 622–628.

[12] P. Coscia, P. Braca, L. M. Millefiori, F. A. N. Palmieri, and P. Willett, "Multiple Ornstein-Uhlenbeck processes for maritime traffic graph representation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. PP, no. 99, pp. 1–1, 2018.

[13] Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, The MIT Press, 2012.

[14] Yu Zheng, "Trajectory data mining: An overview," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 3, pp. 29:1–29:41, May 2015.

[15] Alp Kucukelbir, Dustin Tran, Rajesh Ranganath, Andrew Gelman, and David M. Blei, "Automatic differentiation variational inferece," *Journal of Machine Learning Research*, vol. 18, pp. 1–45, January 2017.

[16] Idris A. Eckley, Paul Fearnhead, and Rebecca Killick, *Analysis of changepoint models*, chapter 10, p. 205224, Cambridge University Press, 2011.

[17] M. Uney, L. M. Millefiori, and P. Braca, "Maximum likelihood estimation in a parametric stochastic trajectory model," in *Proceedings of the Sensor Signal Processing for Defence (SSPD) Conference 2019*, 2019, to appear.

[18] George Casella and Christian P. Robert, *Monte Carlo Statistical Methods*, Springer, second edition, 2005.

[19] L. M. Millefiori, P. Braca, and G. Arcieri, "Scalable distributed change detection and its application to maritime traffic," in *2017 IEEE International Conference on Big Data (Big Data)*, Dec 2017, pp. 1650–1657.

[20] B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, 1986.