A BAYESIAN FRAMEWORK FOR INTENT PREDICTION IN OBJECT TRACKING

Bashar I. Ahmad, Patrick M. Langdon and Simon J. Godsill

Engineering Department, University of Cambridge, Trumpington Street, Cambridge, UK, CB2 1PZ

ABSTRACT

In this paper, we introduce a generic Bayesian framework for inferring the intent of a tracked object, as early as possible, based on the available partial sensory observations. It treats the prediction problem, i.e. not estimating the object state such as position, within an object tracking formulation. This leads to a low-complexity implementation of the inference routine with minimal training requirements. The proposed approach utilises suitable stochastic, namely linear Gaussian, models to capture long term dependencies in the object trajectory as dictated by intent. Numerical examples are shown to demonstrate the efficacy of this framework.

Index Terms— intent, tracking, destination prediction.

1. INTRODUCTION

1.1. Motivation and Scope

Knowing the destination, arrival time and future trajectory of a tracked object (e.g. pointing apparatus, vessel, pedestrian, jet, etc.) offers vital information on intent, enabling predictive functionalities, timely decision making and automation. Besides robotics [1, 2], it has numerous applications such as in
a) Human computer interaction (HCI): inferring the intended on-display item, whilst driving, reduces interaction effort [3].
b) Surveillance: predicting destination and trajectory of an object (e.g. vessel) can unveil anomalies or conflict [4, 5, 6].
c) intelligent vehicles (IV): delivering safer and personalised driving experience, e.g. by predicting motion of other road users and driver state [7, 8] or if driver is returning to car [9].

In this paper, we present a Bayesian framework that treats the tackled intent prediction task, i.e. *not* estimating the hidden state x_t such as the object's position and higher order kinematics, within a tracking formulation. The adopted approach capitalises on the premise that the trajectory followed by an object has long term underlying dependencies driven by intent. This is a shift away from the traditional viewpoint of scenes where objects are assumed to move in unpremeditated manners. It belongs to a higher system level compared with traditional *sensor-level* tracking algorithms for estimating x_t.

A key aspect of the introduced Bayesian framework is employing suitable stochastic models that capture the influence of intent on the object motion and devising inference algorithms to reveal it. This can be effectively achieved via the low-complexity and efficient *bridging distributions* formulation with Gaussian linear models. It permits the inference of the object destination, arrival time and future states, which are considered here to encapsulate intent.

1.2. Related Work and Proposed Framework

Several studies in the object tracking area consider the task of incorporating predictive, often known, information on the object's destination to improve the accuracy of the state estimates [10, 11], i.e. destination-aware tracking. This is in addition to plethora of well-established techniques for estimating the state x_t from noisy sensory data [12, 13]. Using mean reverting processes, such as Ornstein-Uhlenbeck (OU), for modelling a vessel motion in "global" maritime surveillance, e.g. [4, 5, 6], is gaining popularity as they yield more accurate x_t estimates. They are intrinsically driven by a defined mean, e.g. velocity or direction of travel in known high traffic routes. In this paper, the main objective is inferring the *unknown* intent of a tracked object, rather than only state estimations.

Various data driven prediction-classification methods rely on a dynamical model and/or *pattern-of-life* learnt from previously recorded data, e.g. [1, 2, 7, 8]. Whilst such techniques typically require substantial parameters training from extensive data sets (not always available) and can have high computational cost, a state-space modelling approach is adopted here. It uses known dynamical and measurements models, with a few unknown parameters, as is common in object tracking [12, 13]. Subsequently, an efficient inference approach, which requires minimal training, is proposed.

Meta-trackers for recognising intent are developed in [14, 15]. They use a discretised state-space and employ reciprocal processes or other models from natural language processing such as context-free grammars. Whilst the approach presented in this paper leads to notably less complex inference routines compared with [14, 15], it utilises continuous state space models and can treat asynchronous measurements.

Bridging-based inference was introduced in [16, 17], mainly for HCI [3], as an improvement to directly applying OU-type processes in [18]. It assumes that an object is heading to one of N possible endpoints. In this paper, we present a unified treatment and compact overview of this Bayesian (tracking-based) intent prediction approach with and without bridging. For example, the null hypothesis $\mathcal{H}_0$, i.e. the object is not returning to any of the destinations, is addressed unlike in [16, 17, 18]. Examples from two areas, maritime surveillance and IV, are shown to illustrate its efficacy.

2. PROBLEM STATEMENT AND SYSTEM MODEL

Let $\mathbb{D} = \{\mathcal{D}_i : i = 1, 2, ..., N\}$ be the set of N nominal endpoints (e.g. harbours where a vessel can dock or selectable on-display icons or locations of interest a pedestrian can walk to) of a tracked object (e.g. vessel, pointing apparatus or pedestrian). Each destination can be an extended region. For simplicity, a Gaussian distribution $\mathcal{D}_i \backsim \mathcal{N}(a_i^d, \Sigma_i^d)$ is assumed to model an endpoint where the mean a_i^d and covariance Σ_i^d represent the centre and orientation-extent of $\mathcal{D}_i$, respectively.

We aim to *sequentially* calculate the probability of each of $\mathcal{D}_i \in \mathbb{D}$ being the intended destination $\mathcal{D}_I$. This is stated as determining the probability of $N+1$ hypotheses at t_k:

$$\mathcal{P}(t_k) = \{p(\mathcal{H}_I = \mathcal{H}_i \mid y_{1:k}) : i = 0, 1, ..., N\},$$

where $\mathcal{H}_i$ is the hypothesis that $\mathcal{D}_i$ is the intended destination, $\mathcal{H}_I$ is for the *unknown* $\mathcal{D}_I$ and $\mathcal{H}_0$ is the null hypothesis such that $\mathcal{D}_I \notin \mathbb{D}$. Based on $\mathcal{P}(t_k)$ a decision on $\mathcal{D}_I$ can be made, e.g. using a MAP, threshold or other decision criterion. Available (noisy) sensory observations at time instant t_k are $y_{1:k} = \{y_1, y_2, ..., y_k\}$ for times $\{t_1, t_2, ..., t_k\}$. Each y_k is assumed to be derived from a true, underlying, object state x_{t_k} at t_k. The introduced formulation is generally agnostic to the employed sensing technology and it can treat asynchronous imprecise measurements via continuous-time models. In this paper, we also seek inferring the posterior $p(T|y_{1:k})$ of the arrival time T at $\mathcal{D}_I$, if any, and the posteriors $p(x_{t^*}|y_{1:k})$ of the future states of the tracked object such that $t^* > t_k$.

Approximate motion models that enable inferring intent, i.e. not necessarily the exact modelling of the object motion, can suffice. Thus, Gaussian Linear Time Invariant (LTI) models are adopted since they lead to computationally efficient predictors, compared with non-linear non-Gaussian models [13, 19]. The system state $x_{t_k} \in \mathbb{R}^s$ at t_k can be written as

$$x_{t_k} = F(h)x_{t_{k-1}} + M(h) + \varepsilon_k, \tag{1}$$

with $\varepsilon_k \sim \mathcal{N}(0, Q(h))$ is a Gaussian dynamical noise. Matrices F and Q as well as vector M, which define the state transition, are functions of the time step $h = t_k - t_{k-1}$. Similarly, noisy observations $y_k \in \mathbb{R}^l$ are modelled by

$$y_k = Gx_{t_k} + \nu_k, \tag{2}$$

where G is a matrix mapping from the hidden state to the observed measurement and noise component is $\nu_k \sim \mathcal{N}(0, V_k)$.

The class in (1) encompasses many models widely used in tracking applications, e.g. the (near) Constant Velocity (CV) and others that describe higher order kinematics as well as the Linear Destination Reverting (LDR) models discussed below. It results from integrating the motion model Stochastic Differential Equation (SDE) over the time interval $\tau_k = [t_{k-1}, t_k]$.

3. INTENT INFERENCE

Within a Bayesian framework, we have

$$\begin{aligned} p(\mathcal{H}_I = \mathcal{H}_i|y_{1:k}) &\propto p(\mathcal{H}_I = \mathcal{H}_i)p(y_{1:k} \mid \mathcal{H}_I = \mathcal{H}_i) \qquad (3)\\ &= p(\mathcal{H}_I = \mathcal{H}_i)p(y_{1:k-1} \mid \mathcal{H}_I = \mathcal{H}_i)p(y_k \mid y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i). \end{aligned}$$

The prior $p(\mathcal{H}_I = \mathcal{H}_i)$ on the i^{th} hypothesis is independent of observations $y_{1:k}$ and is considered to be known, e.g. from relevant contextual information or previously learnt pattern-of-life. The likelihood $L_{k-1}^{\mathcal{H}_i} = p(y_{1:k-1} \mid \mathcal{H}_I = \mathcal{H}_i)$ pertains to the previous time instant t_{k-1} and is available at t_k. Estimating the Prediction Error Decomposition (PED) defined by: $\ell_k^{\mathcal{H}_i} = p(y_k \mid y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i)$ suffices to *sequentially* estimate the likelihood $L_k^{\mathcal{H}_i} = p(y_k \mid \mathcal{H}_I = \mathcal{H}_i) = \ell_k^{\mathcal{H}_i} \times L_{k-1}^{\mathcal{H}_i}$ and consequently the hypotheses probabilities in (3) at t_k.

The sought PED is given by: $p\left(y_k|y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i\right) = \int p\left(y_k|x_{t_k}, \mathcal{H}_I = \mathcal{H}_i\right) p\left(x_{t_k}|y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i\right) dx_{t_k}$. Since the system model is Gaussian LTI, it can be calculated using

$$\ell_k^{\mathcal{H}_i} = \mathcal{N}(y_k; G\hat{x}_{k|k-1}, G\Sigma_{k|k-1}^{xx}G' + V_k), \tag{4}$$

such that $\hat{x}_{k|k-1}$ and $\Sigma_{k|k-1}^{xx}$ are the the mean and the covariance of the distribution of the predictive state $p(x_{t_k}|y_{1:k-1}) = \int p(x_{t_k}|x_{t_{k-1}})p(x_{t_{k-1}}|y_{1:k-1})dx_{t_{k-1}}$ [13]. Thereby, we can *conveniently* utilise the Kalman filtering equations to calculate the PEDs in (4) for all considered $N+1$ hypotheses. Whilst $\hat{x}_{k|k-1}$ and $\Sigma_{k|k-1}^{xx}$ are the output of the predict step of the Kalman filter, the mean and covariance of the state (i.e. outcome of the correct step with the Kalman gain) are used at the next time step t_{k+1}. The calculated hypotheses probabilities are normalised $\hat{p}(\mathcal{H}_I = \mathcal{H}_i|y_{1:k}) = p(\mathcal{H}_I = \mathcal{H}_i|y_{1:k})/\vartheta_k$ and $\vartheta_k = \sum_{i=0}^{N} p(\mathcal{H}_I = \mathcal{H}_i|y_{1:k})$ to ensure they add to 1.

However, this simple prediction procedure fundamentally relies on a formulation that enables capturing the influence of $\mathcal{D}_I$ on the object behaviour. In other words, how to condition on hypotheses $\{\mathcal{H}_i\}_{i=1}^N$ for the N nominal destinations whilst using (1) and (2). Next, we first describe motion models that naturally incorporate endpoints, then bridging distributions.

3.1. Destination Prediction with LDR Models

The state evolution of a linear destination reverting model is intrinsically driven by an endpoint $\mathcal{D}_i$ as per the SDE

$$dx_t = \Lambda\left(\mu_i - x_t\right)dt + \sigma dw_t, \tag{5}$$

such that μ_i is the end-state set by $\mathcal{D}_i$, matrix Λ (a design parameter) stipulates the nature-strength of the reversion of x_t towards μ_i and w_t is a Wiener process. They are based on an OU process and x_t as well as μ_i can include position, velocity and higher order kinematics. For instance, whilst the state of the mean reverting diffusion model consists only of position (e.g. in 2D and $s = 2$) with $\mu_i = a_i^d$ equal to the location of $\mathcal{D}_i$, the state x_t and μ_i of the equilibrium reverting velocity model includes velocity [17]. Integrating (5) over interval τ_k produces (1) with $M(h)$ a function of the endpoint, i.e. μ_i.

Hence, N LDR models can be written, one per nominal endpoint, with μ_i for the corresponding $\mathcal{D}_i$. This introduces the conditioning on $\{\mathcal{H}_i\}_{i=1}^N$ and N Kalman filters can be utilised to calculate the PED for each endpoint as in (4). A none LDR-type model is used for $\mathcal{H}_0$. However, LDR models can be sensitive to fine parameter tuning [17] and other popular motion models, e.g. CV with $M(h) = 0$ in (1), implicitly ascertain that the object movements are unpremeditated.

3.2. Inference with Bridging Distributions

For $\mathcal{H}_i$, $i \neq 0$, the path of the tracked object, albeit random, must end at the intended destination at arrival time T. Accordingly, a Markov bridge to each nominal endpoint is built, facilitating intent prediction with any Gaussian linear motion model, including CV. This can be modelled by an *artificial* prior distribution for x_T equal to that of $\mathcal{D}_i$, i.e. $p(x_T \mid \mathcal{H}_I = \mathcal{H}_i) = \mathcal{N}\left(x_T; a_i^d, \Sigma_i^d\right)$, or pseudo-observation $\tilde{y}_T^i$ with $p(\tilde{y}_T^i = a_i^d | x_T) = \mathcal{N}(a_i^d; x_T, \Sigma_i^d)$; see [17] for more details. For simplicity, the former construct is adopted below,

$$p(y_k \mid y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i, T) = p(y_{1:k} \mid y_{1:k-1}, x_T, T),$$

i.e. conditioning on prior x_T for $\mathcal{D}_i$ entails conditioning on arrival time. Thus, we estimate *arrival-time-conditioned* PED $\ell_{k,n}^{\mathcal{H}_i} = p(y_k \mid y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i, T = T_n)$ and likelihood $L_{k,n}^{\mathcal{H}_i} = p(y_{1:k} \mid \mathcal{H}_I = \mathcal{H}_i, T = T_n) = \ell_{k,n}^{\mathcal{H}_i} \times L_{k-1,n}^{\mathcal{H}_i}$ for T_n.

However, the arrival time is unknown in practice and a prior distribution on T is assumed, e.g. from contextual data. For instance, arrival might be expected uniformly within $\mathcal{T} = [t_a, t_b]$ and $p(T \mid \mathcal{H}_i) = \mathcal{U}(t_a, t_b)$, e.g. when object is within a certain range from $\mathcal{D}_i$. We can then marginalise arrival time

$$L_k^{\mathcal{H}^i} = p(y_{1:k} \mid \mathcal{H}_I = \mathcal{H}_i) = \int_{T \in \mathcal{T}} p(y_{1:k} \mid \mathcal{H}_I = \mathcal{H}_i, T) \times p(T \mid \mathcal{H}_I = \mathcal{H}_i) dT, \quad (6)$$

to obtain the likelihood in (3). A numerical approximation is applied here to solve the integral in (6) since the arrival time is a one-dimensional quantity [20], e.g. Simpson's rule denoted by quad(.). This approximation requires q evaluations of the arrival-time-conditioned PEDs, and thereby likelihoods for various arrival times $T_n \in \mathbb{T} = \{T_1, T_2, ..., T_q\}$ drawn from $p(T \mid \mathcal{H}_I = \mathcal{H}_i)$, e.g. for q quadrature points.

For hypothesis $\mathcal{H}_i$ and T_n, one approach to introduce bridging is by augmenting the system state x_{t_k} with the artificial prior $x_T \backsim \mathcal{N}(a_d^i, \Sigma_d^i)$ forming an extended state $z_{t_k}^{i,n} = [x_{t_k}' \; x_T']'$ and $z_{t_k}^{i,n} \in \mathbb{R}^{2s}$. This can be shown to lead to the extended linear Gaussian state model

$$z_{t_k}^{i,n} = R_k^{i,n} z_{t_{k-1}}^{i,n} + \tilde{m}_k^{i,n} + \gamma_k^{i,n}, \quad (7)$$

[17] where $\gamma_k^{i,n} \sim \mathcal{N}\left(0, U_k^{i,n}\right)$, $P_T = [0_{s \times s} \quad I_s]$,

$$R_k^{i,n} = \begin{bmatrix} H_k^{i,n} \\ P_T \end{bmatrix}, \quad \tilde{m}_k^{i,n} = \begin{bmatrix} m_k^{i,n} \\ 0_{s \times 1} \end{bmatrix}, \quad U_k^{i,n} = \begin{bmatrix} C_k^{i,n} & 0_{s \times s} \\ 0_{s \times s} & 0_{s \times s} \end{bmatrix},$$

$H_k^{i,n} = \left[C_k^{i,n} Q^{-1}(h) F(h) \quad C_k^{i,n} F'(\tilde{h}) Q^{-1}(\tilde{h})\right]$, $\tilde{h} = T_n - t_k$, $\tilde{m}_k^{i,n} = C_k^{i,n}(Q^{-1}(h)M(h) - F'(\tilde{h})Q^{-1}(\tilde{h})M(\tilde{h}))$ such that $C_k^{i,n} = \left(Q^{-1}(h) + F'(\tilde{h})Q^{-1}(\tilde{h})F(\tilde{h})\right)^{-1}$; $Q(.)$ is from (1), $0_{s \times l}$ is an $s \times l$ zeros matrix and I_s is an $s \times s$ identity matrix. The observation model can be expressed by

$$y_k = \tilde{G} z_{t_k} + \nu_k, \quad (8)$$

with $\tilde{G} = [G \quad 0_{l \times s}]$ where G and ν_k are from (2).

Given the linear Gaussian nature of the system described by (7) and (8), the arrival-time-conditioned PED for $\mathcal{H}_i$ and T_n can be sequentially calculated utilising Kalman filtering equations with the extended state $z_{t_k}^{i,n}$ in lieu of x_{t_k} in (4), i.e. $\ell_{k,n}^{\mathcal{H}_i} = \mathcal{N}(y_k; \tilde{G}\hat{z}_{k|k-1}^{i,n}, \tilde{G}\Sigma_{k|k-1}^{i,n}\tilde{G}' + V_k)$ such that $\hat{z}_{k|k-1}^{i,n}$ and $\Sigma_{k|k-1}^{i,n}$ are the mean and covariance of the predictive state distribution $p(z_{t_k}^{i,n} | y_{1:k-1}, \mathcal{H}_I = \mathcal{H}_i, T_i)$, respectively.

In summary, to obtain the likelihood of hypothesis $\mathcal{H}_i$ at t_k given the available sensory observations $y_{1:k}$:

1) Run Kalman filtering calculations q times, one per quadrature point in $T_n \in \mathbb{T}$ to estimate arrival-time-conditioned PED $\ell_{k,n}^{\mathcal{H}_i}$ and likelihood $L_{k,n}^{\mathcal{H}_i} = \ell_{k,n}^{\mathcal{H}_i} \times L_{k-1,n}^{\mathcal{H}_i}$.

2) Apply a numerical approximation to estimate the likelihood in (6), e.g. $\hat{L}_k^{\mathcal{H}^i} = \text{quad}(L_{k,1}^{\mathcal{H}_i}, L_{k,2}^{\mathcal{H}_i}, ..., L_{k,q}^{\mathcal{H}_i})$.

The above two steps are repeated for each of $\mathcal{D}_i \in \mathbb{D}$. It is noted that for none LDR motion models, e.g. CV, bridging can be implemented more efficiently as explored in [21].

3.3. Time of Arrival and Future Trajectory Estimation

First, we have: $p(T \mid y_{1:k}) \propto \sum_{i=1}^{N} p(y_{1:k} \mid T, \mathcal{H}_I = \mathcal{H}_i) \times p(T \mid \mathcal{H}_I = \mathcal{H}_i) p(\mathcal{H}_I = \mathcal{H}_i)$ by integrating over the discrete set $\mathbb{D}$ of endpoints. The conditioned likelihood $L_{k,n}^{\mathcal{H}_i}$, for all q quadrature points $T_n \in \mathbb{T}$ are readily available from the endpoint prediction calculations. For hypothesis $\mathcal{H}_i$, $i \neq 0$, we have the approximate posterior: $p(T \mid \mathcal{H}_I = \mathcal{H}_i, y_{1:k}) \approx \frac{1}{\kappa_{i,k}} \sum_{n=1}^{q} L_{k,n}^{\mathcal{H}_i} p(T_n \mid \mathcal{H}_I = \mathcal{H}_i) \delta_{\{T_n\}}$ such that the normalisation factor is $\kappa_{i,k} = \sum_{n=1}^{q} L_{k,n}^{\mathcal{H}_i} p(T_n \mid \mathcal{H}_I = \mathcal{H}_i)$ and $\delta_{\{T_n\}}$ is a Dirac delta located at T_n. Assuming that the arrival times are the same for all $\mathcal{D}_i \in \mathbb{D}$, we obtain

$$p(T \mid y_{1:k}) \approx \frac{1}{\eta_k} \sum_{n=1}^{q} \psi_{k,n} \delta_{\{T_n\}}, \quad (9)$$

where $\psi_{k,n} = \sum_{i=1}^{N} L_{k,n}^{\mathcal{H}_i} p(T_n \mid \mathcal{H}_I = \mathcal{H}_i) p(\mathcal{H}_I = \mathcal{H}_i)$ and $\eta_k = \sum_{n=1}^{q} \sum_{i=1}^{N} L_{k,n}^{\mathcal{H}_i} p(T_n \mid \mathcal{H}_I = \mathcal{H}_i) p(\mathcal{H}_I = \mathcal{H}_i)$. Posterior can be analogously estimated for varying T priors.

For a given arrival time $T = T_n$ and destination $\mathcal{D}_i$, the future state at times $t^* > t_k$ can be attained directly from the extended model in (7). It is a Gaussian distribution with mean $\hat{z}_{t^*}^{i,n} = R_{t^*}^{i,n} \hat{z}_k^{i,n} + \tilde{m}_{t^*}^{i,n}$ and covariance $\Sigma_{t^*}^{i,n} = R_{t^*}^{i,n} \Sigma_k^{i,n} (R_{t^*}^{i,n})' + U_{t^*}^{i,n}$ such that $h = t^* - t_k$. By integrating over the discrete sets of all endpoints in $\mathbb{D}$ and arrival times in $\mathbb{T}$, it is given by the Gaussian mixture

$$p(x_{t^*} \mid y_{1:k}) \approx \sum_{i=1}^{N} \sum_{n=1}^{q} u_{i,n}^k \mathcal{N}\left(x_{t^*}; \Gamma \hat{z}_{t^*}^{i,n}, \Gamma \Sigma_{t^*}^{i,n} \Gamma'\right) \quad (10)$$

where $u_{i,n}^k = L_{k,n}^{\mathcal{H}_i} p(\mathcal{H}_I = \mathcal{H}_i) p(T_n \mid \mathcal{H}_I = \mathcal{H}_i) / \varpi_k$ and $\varpi_k = \sum_{i=1}^{N} \sum_{n=1}^{q} L_{k,n}^{\mathcal{H}_i} p(T_n \mid \mathcal{H}_I = \mathcal{H}_i) p(\mathcal{H}_I = \mathcal{H}_i)$; pruning matrix $\Gamma = [I_s \quad 0_{s \times s}]$. It is noted that the uncertainty associated with the state prediction typically grows arbitrarily large as t^* increases with classical trackers. As shown below, the reliability of the future state estimates can be notably improved by exploiting the endpoint prediction results.

4. NUMERICAL EXAMPLES

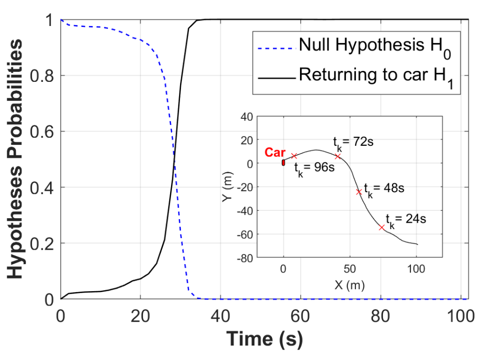
The following two examples are shown here:

a) ***Intelligent vehicle*** **(Figs. 1 and 2)**: predict if and when a driver is returning to car ($N = 1$) in a car park from his positions in 2-D provided by a smartphone (GPS) location service every 2s. Full track (converted to meters and centered/rotated at the car) is displayed. Predicted future states from the predict step of a Kalman filter (i.e. no bridging) are depicted.

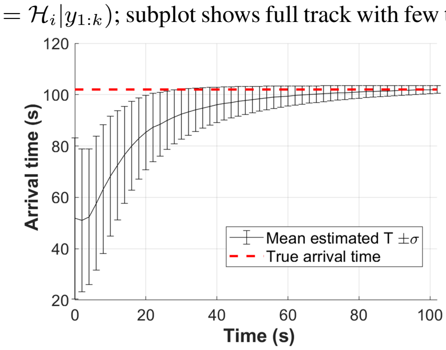
b) ***Maritime surveillance*** **(Fig. 3)**: predict a vessel destination out of five possible harbours in a bay, $N = 5$, from partial noisy observations of its 2-D locations, e.g. AIS-based. Nine synthetically generated trajectories are considered; all start from a rendezvousing area off the coast of the bay. Portions of the track when the true endpoint is correctly inferred are highlighted, i.e. a hypothesis fulfills the MAP criterion (most probable) for at least three successive time steps.

We used a CV model with bridging distributions and Simpson's quadrature scheme with $q = 41$ points from a conservative uniform prior $p(T|\mathcal{H}_I = \mathcal{H}_i)$, e.g. $\mathcal{U}(30s, 140s)$ in Fig. 1 as a pedestrian walks at an average speed of 1.5ms^{-1}. All hypotheses are equally probable, $p(\mathcal{H}_I = \mathcal{H}_i) = 1/(N+1)$.

These examples clearly demonstrate the effectiveness of the proposed framework to predict the object intent, e.g. early destination inference and accurate arrival time estimates. The state predictions are noticeably more certain compared with classical Kalman filtering, especially for $t^* \gg t_k$ (see row 2 in Fig. 2). It is noted that estimating T via (9) does not entail additional computations as likelihoods $L_{k,n}$ are calculated for determining $\mathcal{D}_I$, a small q typically suffices [17], and computations (for each T_n and $\mathcal{H}_i$) can be easily parallelised.



(a) $p(\mathcal{H}_I = \mathcal{H}_i|y_{1:k})$; subplot shows full track with few timestamps.



(b) Mean and standard deviation of estimated $p(T|y_{1:k})$.

Fig. 1: Inference results as user walks back to car from a building.

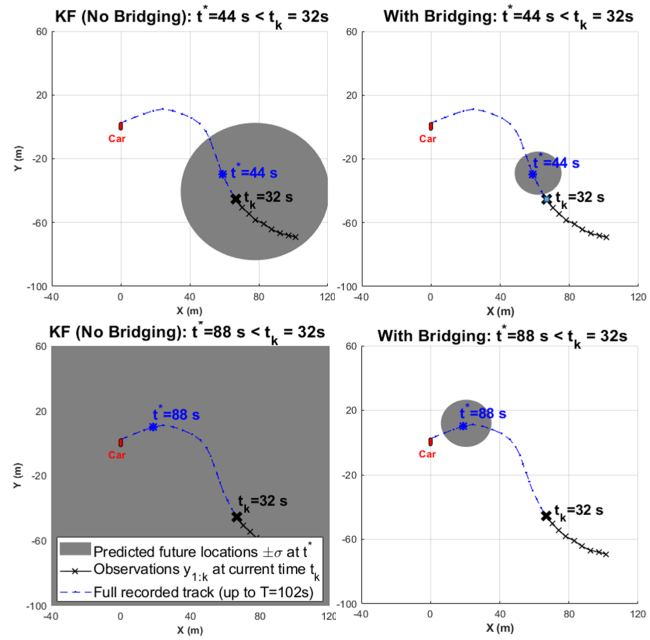


Fig. 2: Future state estimates at $t_k = 32s$ (bold cross is t_k). Shading shows the covariance of the posterior $p(x_{t^*}|y_{1:k})$, centered on its mean, at two future time instants $t^* = 44$s (top row) and $t^* = 88$s (bottom row) each marked by a blue asterisk. Left and right columns are for classical filtering and bridging distributions, respectively.

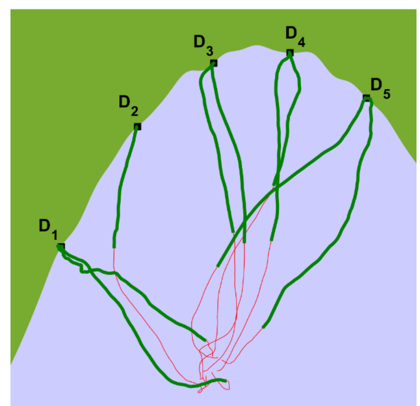


Fig. 3: Destination inference for the nine vessel tracks. Thick green lines shows the track portion when the true endpoint is inferred.

5. FINAL REMARKS

The introduced "simple" intent prediction framework leads to low-complexity, Kalman-filtering-type, inference algorithms with minimal parameter training requirements. It can be combined with other data-driven approaches, such as those that focus on learning patterns-of-life, within the flexible Bayesian formulation. The proposed framework has several extensions, e.g. using nonlinear and/or non-Gaussian models, interacting multiple models to represent varying motion behaviours, a group tracking formulation and others. This paper serves the purpose of motivating such future work.

6. REFERENCES

[1] Hao-Tien Chiang, Nathanael Rackley, and Lydia Tapia, "Stochastic ensemble simulation motion planning in stochastic dynamic environments," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2015.

[2] Kris Kitani, Brian Ziebart, James Bagnell, and Martial Hebert, "Activity forecasting," in *Proc. of European Conf. on Computer Vision*, 2012, pp. 201–214.

[3] Bashar I. Ahmad, James K. Murphy, Simon Godsill, Patrick Langdon, and Robert Hardy, "Intelligent interactive displays in vehicles with intent prediction: A Bayesian framework," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 82–94, 2017.

[4] Giuliana Pallotta, Steven Horn, Paolo Braca, and Karna Bryan, "Context-enhanced vessel prediction based on Ornstein-Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results," in *Proc. Int. Conf. on Information Fusion*, 2014, pp. 1–7.

[5] Enrica d'Afflisio, Paolo Braca, Leonardo M Millefiori, and Peter Willett, "Maritime anomaly detection based on mean-reverting stochastic processes applied to a real-world scenario," in *In Proc. of 21st International Conference on Information Fusion (FUSION)*, 2018, pp. 1171–1177.

[6] Murat Uney, Leonardo M Millefiori, and Paolo Braca, "Prediction of rendezvous in maritime situational awareness," in *in Proc. of 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 622–628.

[7] Benjamin Völz, Holger Mielenz, Igor Gilitschenski, Roland Siegwart, and Juan Nieto, "Inferring pedestrian motions at urban crosswalks," *IEEE Transactions on Intelligent Transportation Systems*, 2018.

[8] Takashi Bando, Kana Takenaka, Shogo Nagasaka, and Takafumi Taniguchi, "Unsupervised drive topic finding from driving behavioral data," in *Proc. of IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 177–182.

[9] Bashar I Ahmad, Patrick M Langdon, and Simon J Godsill, "A meta-tracking approach for predicting the driver or passenger intent," in *Proc. of 21st International Conference on Information Fusion*. IEEE, 2018, pp. 1–5.

[10] David A Castanon, Bernard C Levy, and Alan S Willsky, "Algorithms for the incorporation of predictive information in surveillance theory," *International Journal of Systems Science*, vol. 16, no. 3, pp. 367–382, 1985.

[11] Enzo Baccarelli and Roberto Cusani, "Recursive filtering and smoothing for reciprocal Gaussian processes with Dirichlet boundary conditions," *IEEE Trans. on Signal Processing*, vol. 46, no. 3, pp. 790–795, 1998.

[12] Yaakov Bar-Shalom, Peter Willett, and Xin Tian, *Tracking and Data Fusion: A Handbook of Algorithms*, YBS Publishing, 2011.

[13] Anton J Haug, *Bayesian Estimation and Tracking: A Practical Guide*, John Wiley & Sons, 2012.

[14] Mustafa Fanaswala and Vikram Krishnamurthy, "Spatiotemporal trajectory models for metalevel target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 1, pp. 16–31, 2015.

[15] Vikram Krishnamurthy and Sijia Gao, "Syntactic enhancement to VSIMM for roadmap based anomalous trajectory detection: A natural language processing approach," *IEEE Trans. on Signal Processing*, 2018.

[16] Bashar I. Ahmad, James K. Murphy, Patrick M. Langdon, Simon J. Godsill, and Robert Hardy, "Destination inference using bridging distributions," in *Proc. of the 40th IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP '15)*, 2015, pp. 5585–5589.

[17] Bashar I Ahmad, James K Murphy, Patrick M Langdon, and Simon J Godsill, "Bayesian intent prediction in object tracking using bridging distributions," *IEEE Trans. on Cybernetics*, vol. 48, no. 1, pp. 215–227, 2018.

[18] Bashar I Ahmad, James K Murphy, Patrick M Langdon, Simon J Godsill, Robert Hardy, and Lee Skrypchuk, "Intent inference for hand pointing gesture-based interactions in vehicles," *IEEE Trans. on Cybernetics*, vol. 46, no. 4, pp. 878–889, 2016.

[19] Olivier Cappé, Simon J Godsill, and Eric Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 899–924, 2007.

[20] Philip J Davis and Philip Rabinowitz, *Methods of numerical integration*, Courier Corporation, 2007.

[21] Tohid Ardeshiri, Bashar I Ahmad, Patrick M Langdon, and Simon J Godsill, "Efficient bridging-based destination inference in object tracking," in *Proc. of ICASSP*. IEEE, 2017, pp. 4406–4410.