Bayesian Track-to-Graph Association for Maritime Traffic Monitoring

Raffaele Grasso, Leonardo M. Millefiori, Paolo Braca
NATO STO Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy
{raffaele.grasso, leonardo.millefiori, paolo.braca}@cmre.nato.int

***Abstract*—We present a hypothesis test to associate ship track measurements to an edge of a given graph that statistically models common traffic routes in a given area of interest. The association algorithm is based on the hypothesis that ship velocities are modeled by mean-reverting stochastic processes. Prior knowledge about the traffic is provided by the graph in form of probability density functions of the mean-reverting kinematic parameters for each node and edge of the graph, that are exploited in the formalization of the association algorithm. Tests on real Automatic Identification System (AIS) data show a qualitatively good association performance. Future developments of this work include the development of specific quantitative metrics to assess the association performance.**

***Index Terms*—Maritime surveillance, knowledge based tracking and prediction, statistical track association, graphs, mean-reverting stochastic processes.**

I. Introduction

In recent years Automatic Identification System (AIS) data has become heavily used for maritime surveillance and situational awareness, thanks also to a global and steadily increasing compliance with international regulations. As a consequence, networks of AIS receivers are growing and producing very large volumes of data day by day, which open up to interesting research opportunities. One notable example is the development of models and methods to extract and identify meaningful representative patterns, such as typical (or normal) ship traffic routes in a given area of interest [1], [2].

A network graph traffic model has been introduced recently [3], [4], in which way-point areas, i.e. regions where ships are more likely to change their direction or speed, are represented by graph nodes, and edges are representative of typical ship transitions from one way-point area to another. In a model like this, a crucial step before tracking and prediction is the association of ship track measurements to an edge or a node of the graph. The association is typically performed by means of distance-based methods.

In this paper we propose a statistical Bayesian approach that takes into account possible sources of uncertainty inherently related to the problem at hand. The approach is based on the mean-reverting model of ship trajectories presented in [3], [5], [6] and the use of priors on mean-reverting kinematic parameters learned during the construction of the traffic graph. A historical set of AIS data is used for estimation of the graph parameters, i.e. nodes, connectivity matrix and kinematic priors associated to each node and edge. The following section formalizes the problem as a maximum a posteriori statistical test assuming the graph is given. The learning of the graph is out of the scope of this paper, and interested readers can refer to [3], [4] for learning methods that can be used to build vessel traffic graph models.

A. Problem statement

Supposing that in a given area the normal traffic routes are modeled by a directed graph having M nodes, the association problem consists in deciding the edge $\boldsymbol{e}_{ij} = [i, j]^\intercal$, where $i, j = 1, \ldots, M$, along a ship is navigating, given the ship trajectory $\boldsymbol{s}_{0:N-1} = [\boldsymbol{s}_0^\intercal, \ldots, \boldsymbol{s}_{N-1}^\intercal]^\intercal$, where $\boldsymbol{s}_n = [\boldsymbol{x}^\intercal(t_n), \dot{\boldsymbol{x}}^\intercal(t_n)]^\intercal$ is the kinematic state at time t_n, $n = 0, \ldots, N-1$, and $\boldsymbol{x}(t_n)$ and $\dot{\boldsymbol{x}}(t_n)$ are the vessel position and instantaneous velocity, respectively, in a Cartesian coordinate reference system. The association criteria is the maximum a posteriori (MAP) hypothesis test as follows:

$$\hat{\boldsymbol{e}}_{ij} = \arg \max_{i,j=1,\ldots,M} \left[P\left(\boldsymbol{e}_{ij} | \boldsymbol{s}_{0:N-1}\right)\right], \tag{1}$$

where $P\left(\boldsymbol{e}_{ij} | \boldsymbol{s}_{0:N-1}\right)$ is the probability that the ship is navigating along the edge $\boldsymbol{e}_{ij}$, between an origin node i and a destination node j, given the measurements. A test for only the initial node can be derived by marginalizing (1) as follows:

$$\begin{aligned} \hat{i} &= \arg \max_{i=1,\ldots,M} \left[P\left(i | \boldsymbol{s}_{0:N-1}\right)\right] = \\ &= \arg \max_{i=1,\ldots,M} \left[\sum_{j=1}^{M} P\left(\boldsymbol{e}_{ij} | \boldsymbol{s}_{0:N-1}\right)\right]. \end{aligned} \tag{2}$$

In this work, $\boldsymbol{s}_{0:N-1}$ does not contain any way-point, i.e. the ship, in $[t_0 \ t_{N-1}]$, is navigating along an edge, without maneuvering and the kinematic state can be modeled by a stationary process. The normal trajectory of ships is assumed to be piece-wise linear, following a sequence of way-points. Each node of the graph is modeled by a distribution of way-point positions, $\boldsymbol{\omega}_i$, with probability density function (pdf) $p\left(\boldsymbol{\omega}_i\right)$. A generic edge of the graph, $\boldsymbol{e}_{ij}$, represents a linear segment of the ship trajectory between an origin node i and a destination node j (see Fig. 1). Each edge is modeled by the travel time, τ_{ij}, with pdf $p\left(\tau_{ij}\right)$ and by a vector of kinematic parameters, $\boldsymbol{\phi}_{ij}$, with pdf $p\left(\boldsymbol{\phi}_{ij}\right)$. A parameter vector for the edge $\boldsymbol{e}_{ij}$ can be defined as $\boldsymbol{\theta}_{ij} = \left[\boldsymbol{\omega}_i^\intercal, \boldsymbol{\omega}_j^\intercal, \boldsymbol{\phi}_{ij}^\intercal, \tau_{ij}\right]^\intercal$. In order to simplify the implementation of the association test, as a working hypothesis, the edge parameters are considered independent so that the pdf of $\boldsymbol{\theta}_{ij}$ can be written as $p\left(\boldsymbol{\theta}_{ij}\right) = p\left(\boldsymbol{\omega}_i\right) p\left(\boldsymbol{\omega}_j\right) p\left(\boldsymbol{\phi}_{ij}\right) p\left(\tau_{ij}\right)$. It is worthwhile to

ISBN 978-90-827970-1-5 © EURASIP 2018

note that the theoretical formulation of the test, reported in section II, is not affected by this strong hypothesis. The test implementation, see section II-A, that exploits such hypothesis is provided as an example. The same considerations apply to the Gaussian assumption of the edge parameter pdfs. However, such simplifications do not compromise the proposed association procedure tested on real-data.

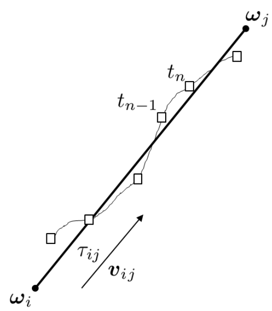

Fig. 1. Definition of a traffic graph edge.

A ship's kinematic state, $\boldsymbol{s}_n$, at time t_n, along an edge of the graph is modeled by a mean-reverting process as described in [5] and [6]. Under this hypothesis, the distribution of $\boldsymbol{s}_n$ is Gaussian, with mean and covariance given by the first- and second-order moments of the solution of a stochastic differential equation (SDE) with $\boldsymbol{s}_{n-1}$ as initial condition [5], [7]. In particular, the first-order moment, $\boldsymbol{\mu}_{\boldsymbol{s}_n}$, is given by:

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{s}_n} = \boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{\gamma}_{ij}, \boldsymbol{v}_{ij}, \Delta t_n) = \\ \boldsymbol{\Phi}\left(\Delta t_n, \boldsymbol{\gamma}_{ij}\right) \boldsymbol{s}_{n-1} + \boldsymbol{\Psi}\left(\Delta t_n, \boldsymbol{\gamma}_{ij}\right) \boldsymbol{v}_{ij}, \end{aligned} \tag{3}$$

where $\boldsymbol{\Phi}\left(\Delta t_n, \boldsymbol{\gamma}_{ij}\right)$ is the state transition matrix, $\boldsymbol{\Psi}\left(\Delta t_n, \boldsymbol{\gamma}_{ij}\right)$ is the control input matrix, $\Delta t_n = t_n - t_{n-1}$, $\boldsymbol{v}_{ij}$ is the long-run mean velocity, i.e. the desired velocity of a ship, and $\boldsymbol{\gamma}_{ij}$ is the vector of mean-reverting parameters [5]. The second-order moment, $\boldsymbol{\Sigma}_{\boldsymbol{s}_n} = \boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\tilde{\boldsymbol{C}}_{ij}, \boldsymbol{\gamma}_{ij}, \Delta t_n)$, defined in [5], is a function of Δt_n, $\boldsymbol{\gamma}_{ij}$ and the noise process covariance $\tilde{\boldsymbol{C}}_{ij}$. The expressions of all the involved matrices are not reported here for the sake of brevity. For details, the reader can refer to [5]. It is also worthwhile to mention that the Gaussian distribution of the state $\boldsymbol{s}_n$ given $\boldsymbol{s}_{n-1}$ is a direct consequence of assuming the dynamic of the ship as a mean-reverting process, a hypothesis which has been extensively validated using real data in [5].

In the mean-reverting motion model, the instantaneous velocity of the ship, when perturbed, is forced toward $\boldsymbol{v}_{ij}$ in a certain amount of time that depends on a characteristic time constant. In this work, $\boldsymbol{v}_{ij}$ depends on the graph edge, $\boldsymbol{e}_{ij}$, while the remaining parameters, $\boldsymbol{\gamma}_{ij}$ and $\tilde{\boldsymbol{C}}_{ij}$ are assumed constant in a given region of interest (ROI) and are estimated by following the procedure in [5]. In the next section, since $\boldsymbol{\gamma}_{ij}$ and $\tilde{\boldsymbol{C}}_{ij}$ are constant, the first- and second-order moments of $\boldsymbol{s}_n$ are written as $\boldsymbol{\mu}_{\boldsymbol{s}_n} = \boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{v}_{ij}, \Delta t_n)$ and $\boldsymbol{\Sigma}_{\boldsymbol{s}_n} = \boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\Delta t_n)$, while the mean-reverting kinematic parameter vector is set to $\boldsymbol{\phi}_{ij} = \boldsymbol{v}_{ij}$.

The graph parameters, i.e. the node number, the connectivity matrix and $p\left(\boldsymbol{\theta}_{ij}\right)$, are estimated by processing a set of AIS data collected over the ROI for a long period of time. This paper is focused on the derivation of the statistical association test (1) given the graph and the target track data. The reader can find examples on how to build the traffic graph in [3] and [4], where the hypothesis of mean-reverting ship motion is exploited in a kinematic parameter change detector to detect ship way-points and estimate the graph parameters.

B. Paper organization

The paper is organized as follows. Section II details the derivation of the MAP test and, specifically, the posterior $P\left(\boldsymbol{e}_{ij}|\boldsymbol{s}_{0:N-1}\right)$. Then, two implementations of the test are described, one based on Monte Carlo simulation and the other on the unscented transform (UT). Section III shows results obtained by processing a set of real AIS data collected in the Eastern Atlantic. Section IV is devoted to conclusion and future work.

II. Track-to-graph association

The hypothesis test is iteratively performed by using track measurements in a sliding window of N samples, collected at time instants in the set $T_n \equiv \{t_{n-N}, t_{n-N+1}, \ldots, t_n\}$. The associated edge $\hat{\boldsymbol{e}}_{ij}$ becomes formally a function of time, i.e. $\hat{\boldsymbol{e}}_{ij} = \hat{\boldsymbol{e}}_{ij}\left(t_n\right)$, with time dependency hereafter dropped for the sake of clarity.

The posterior $P(\boldsymbol{e}_{ij}|\boldsymbol{s}_{0:N-1})$ can be factored as

$$P(\boldsymbol{e}_{ij}|\boldsymbol{s}_{0:N-1}) = \frac{P(\boldsymbol{e}_{ij})\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij})}{p(\boldsymbol{s}_{0:N-1})}, \tag{4}$$

where $\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij}) = p(\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij})$ is the likelihood function of the measurements given $\boldsymbol{e}_{ij}$, $P(\boldsymbol{e}_{ij})$ is the prior of the edge vector, which is estimated during the analysis of the historical data set of ship trajectories to learn the graph parameters ($P(\boldsymbol{e}_{ij}) = 0$ means that i and j are not connected), and $p(\boldsymbol{s}_{0:N-1})$ is the prior of the measurements. The likelihood depends on the edge parameters, $\boldsymbol{\theta}_{ij}$, which can be considered as nuisance parameters. Considering the vessel state dynamic as a Markov process, the likelihood can be written as follows:

$$\begin{aligned} \mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij}) = p(\boldsymbol{s}_0|\boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij}) \\ \times \prod_{n=1}^{N-1} p(\boldsymbol{s}_n|\boldsymbol{s}_{n-1}, \boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij}), \end{aligned} \tag{5}$$

where $p(\boldsymbol{s}_n|\boldsymbol{s}_{n-1}, \boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij})$ is the vessel state transition density and $p(\boldsymbol{s}_0|\boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij})$ is the prior of $\boldsymbol{s}_0$, both conditioned on $\boldsymbol{e}_{ij}$ and function of the edge parameters $\boldsymbol{\theta}_{ij}$. As stated in section I-A, for a mean-reverting process the transition and prior densities are Gaussian with mean and covariance given by the first and the second order moment of the solution of a linear SDE with initial conditions given by $\boldsymbol{s}_{n-1}$. The moments are calculated given the parameters $\boldsymbol{\theta}_{ij}$. Then, the transition densities can be written as:

$$\begin{aligned} p(\boldsymbol{s}_n|\boldsymbol{s}_{n-1}, \boldsymbol{e}_{ij}; \boldsymbol{\theta}_{ij}) = p(\boldsymbol{s}_n|\boldsymbol{s}_{n-1}, \boldsymbol{e}_{ij}; \boldsymbol{v}_{ij}) \\ = \mathcal{N}_{\boldsymbol{s}_n}[\boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{v}_{ij}, \Delta t_n), \boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\Delta t_n)], \end{aligned} \tag{6}$$

where $\boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{v}_{ij}, \Delta t_n)$ is the first order moment of $\boldsymbol{s}_n$

which is a function of the initial state $\boldsymbol{s}_{n-1}$, the mean velocity $\boldsymbol{v}_{ij}$ and the time interval $\Delta t_n = t_n - t_{n-1}$, while $\boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\Delta t_n)$ is the second order moment which depends on Δt_n. The notation $\mathcal{N}_{\boldsymbol{x}}(\boldsymbol{\mu}_{\boldsymbol{x}}; \boldsymbol{\Sigma}_{\boldsymbol{x}})$ indicates the multivariate normal pdf of the $N_{\boldsymbol{x}}$-dimensional random variable $\boldsymbol{x} \in \Re^{N_{\boldsymbol{x}}}$ having mean vector and covariance matrix given by $\boldsymbol{\mu}_{\boldsymbol{x}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{x}}$, respectively.

The prior density of $\boldsymbol{s}_0$ is also Gaussian:

$$p(\boldsymbol{s}_0|\boldsymbol{e}_{ij};\boldsymbol{\theta}_{ij}) = \mathcal{N}_{\boldsymbol{s}_0}[\boldsymbol{\mu}_{\boldsymbol{s}_0}(\boldsymbol{\omega}_i, \boldsymbol{v}_{ij}, \tau_{i0}), \boldsymbol{\Sigma}_{\boldsymbol{s}_0}(\tau_{i0})], \quad (7)$$

where τ_{i0} is the time it takes the vessel to reach $\boldsymbol{s}_0$ from the way-point position, $\boldsymbol{\omega}_i$, associated to the source node i (see Fig. 2). As an approximation, τ_{i0} is considered proportional to the edge transit time τ_{ij}, $\tau_{i0} = \alpha\tau_{ij}$, where the proportionality factor $\alpha = \min[1, \|\boldsymbol{x}(t_0) - \boldsymbol{\omega}_i\| \,/\, \|\boldsymbol{\omega}_j - \boldsymbol{\omega}_i\|]$ is the fraction of the path traveled from the starting way-point area to $\boldsymbol{x}(t_0)$ with respect to the whole length of the edge.

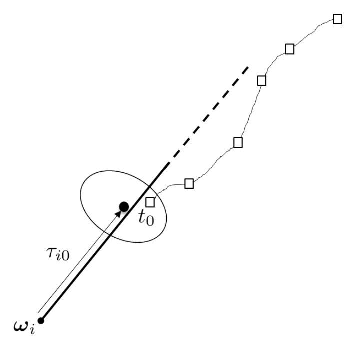

Fig. 2. Prior distribution of the vessel state at t_0.

Given (6) and (7), the likelihood (5) can be expressed as follows:

$$\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij}) = p(\boldsymbol{s}_0|\boldsymbol{e}_{ij};\boldsymbol{\theta}_{ij})\Lambda_{\boldsymbol{s}_{1:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{v}_{ij}) \quad (8)$$

where

$$\begin{aligned}\Lambda_{\boldsymbol{s}_{1:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{v}_{ij}) &= \prod_{n=1}^{N-1} p(\boldsymbol{s}_n|\boldsymbol{s}_{n-1}, \boldsymbol{e}_{ij}; \boldsymbol{v}_{ij}) = \\ &= \prod_{n=1}^{N-1} \mathcal{N}_{\boldsymbol{s}_n}[\boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{v}_{ij}, \Delta t_n), \boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\Delta t_n)].\end{aligned} \quad (9)$$

The edge nuisance parameters can be now averaged out as follows:

$$\begin{aligned}\bar{\mathcal{L}}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}} &= E_{\boldsymbol{\theta}_{ij}}[\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij})] = \\ &= E_{\boldsymbol{\theta}_{ij}}[p(\boldsymbol{s}_0|\boldsymbol{e}_{ij};\boldsymbol{\theta}_{ij})\Lambda_{\boldsymbol{s}_{1:N-1}|i}(\boldsymbol{v}_{ij})] = \\ &= \int_{\boldsymbol{\theta}_k} \mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij})p(\boldsymbol{\theta}_{ij}),\end{aligned} \quad (10)$$

where $p(\boldsymbol{\theta}_{ij})$ is the prior density of the edge parameters as defined in section I-A.

The posterior can be redefined by substituting the mean likelihood (10) in (4) so that:

$$\begin{aligned}&P(\boldsymbol{e}_{ij}|\boldsymbol{s}_{0:N-1}) \propto P(\boldsymbol{e}_{ij})\bar{\mathcal{L}}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}} = \\ &= P(\boldsymbol{e}_{ij})E_{\boldsymbol{\theta}_{ij}}[p(\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij};\boldsymbol{\theta}_{ij})] = \\ &= P(\boldsymbol{e}_{ij})E_{\boldsymbol{\theta}_{ij}}\{\mathcal{N}_{\boldsymbol{s}_0}[\boldsymbol{\mu}_{\boldsymbol{s}_0}(\boldsymbol{\omega}_i, \boldsymbol{v}_{ij}, \tau_{i0}), \boldsymbol{\Sigma}_{\boldsymbol{s}_0}(\tau_{i0})] \\ &\times \prod_{n=1}^{N-1} \mathcal{N}_{\boldsymbol{s}_n}[\boldsymbol{\mu}_{\boldsymbol{s}_n}(\boldsymbol{s}_{n-1}, \boldsymbol{v}_{ij}, \Delta t_n), \boldsymbol{\Sigma}_{\boldsymbol{s}_n}(\Delta t_n)]\}.\end{aligned} \quad (11)$$

The MAP association test is finally obtained by substituting (11) into (1) or (2).

A. Test implementation

The hypothesis tests is first implemented by a Monte Carlo Method (MCM) as described in section II-A1. An approximated method will be considered such as the one based on the Unscented Transform (UT) which will be detailed later in section II-A2.

1) MCM implementation: The calculation of the posterior in (11) using an MCM requires sampling from the distributions of the traffic model parameters, $p(\theta_{ij})$, defined in section I-A. As in [4], the way-point positions, $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$, and the mean velocity, $\boldsymbol{v}_{ij}$, are sampled from the Gaussian mixture models (GMM) learned during the construction of the traffic graph. The travel time, τ_{ij}, is sampled from an Erlang distribution, whose parameters are estimated during the graph training too. If the edge travel time samples used in the graph learning step are not sufficient to estimate the parameters of the Erlang pdf, the travel time is sampled from the empirical distribution with replacement, as in a bootstrap procedure. Given the edge $\boldsymbol{e}_{ij}$, with $i, j = 1, \ldots, M$, a set of N_p particles, $\mathcal{P} \equiv \{\boldsymbol{\theta}_{ijl}\}_{l=1}^{N_p}$, is generated by randomly sampling from the distribution of the edge parameter vector $\boldsymbol{\theta}_{ij}$. The mean likelihood (10) is then approximated by averaging the likelihood samples calculated for each particles in $\mathcal{P}$. The mean likelihood approximation is then weighted by $P(\boldsymbol{e}_{ij})$ and used in the association test.

2) UT implementation: In this subsection, an alternative approach to the implementation of the track-to-graph association algorithm is proposed in order to improve the algorithm processing time performance. The implementation is based on the calculation of the mean likelihood by using the UT [8], which is an approximated method to calculate the first- and second-order moments of a random vector after a non-linear transformation function, given the first- and second-order moments of the random vector before. It builds on the idea that it is easier to approximate a smooth pdf than a generic non-linear function [8]. The mean likelihood (10) is evaluated by approximating the distribution of the graph parameter vector, $\boldsymbol{\theta}_{ij} \in \Re^{N_\theta}$, by a multivariate Gaussian:

$$\begin{aligned}\bar{\mathcal{L}}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}} &= E_{\boldsymbol{\theta}_{ij}}[\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij})] = \\ &= \int_{\boldsymbol{\theta}_{ij}} \mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij})p(\boldsymbol{\theta}_{ij}) \approx \\ &\approx \int_{\boldsymbol{\theta}_{ij}} \mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\theta}_{ij}) \cdot \mathcal{N}_{\boldsymbol{\theta}_{ij}}[\boldsymbol{\mu}_{\boldsymbol{\theta}_{ij}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_{ij}}]\end{aligned} \quad (12)$$

ISBN 978-90-827970-1-5 © EURASIP 2018

where $\boldsymbol{\mu}_{\boldsymbol{\theta}_{ij}} = [\boldsymbol{\mu}_{\boldsymbol{\omega}_i}^\top, \boldsymbol{\mu}_{\boldsymbol{\omega}_j}^\top, \boldsymbol{\mu}_{\boldsymbol{v}_{ij}}^\top, \mu_{\tau_{ij}}]^\top$ and

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}_{ij}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\omega}_i} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{2\times 2} & \boldsymbol{\Sigma}_{\boldsymbol{\omega}_j} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \boldsymbol{\Sigma}_{\boldsymbol{v}_{ij}} & \mathbf{0}_{2\times 1} \\ \mathbf{0}_{1\times 2} & \mathbf{0}_{1\times 2} & \mathbf{0}_{1\times 2} & \sigma_{\tau_{ij}}^2 \end{bmatrix}, \tag{13}$$

with the mean vectors, $\boldsymbol{\mu}$, and covariances, $\boldsymbol{\Sigma}$, that are estimated during the graph learning phase [4], using a historical data set of AIS trajectories. The integral (12) is then approximated by a weighted sum as follows:

$$\bar{\mathcal{L}}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}} \approx \frac{1}{N_{UT}} \cdot \sum_{l=0}^{2N_\theta} w_l \cdot \mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\chi}_l), \tag{14}$$

where $\boldsymbol{\chi}_l$ are the so called UT sigma points, with $\boldsymbol{\chi}_0 = \boldsymbol{\mu}_{\boldsymbol{\theta}_k}$ and

$$\begin{aligned} \boldsymbol{\chi}_l &= \boldsymbol{\mu}_{\boldsymbol{\theta}_{ij}} - \boldsymbol{L}_l, \quad l = 1, \ldots, N_\theta \\ \boldsymbol{\chi}_l &= \boldsymbol{\mu}_{\boldsymbol{\theta}_{ij}} + \boldsymbol{L}_l, \quad l = N_\theta + 1, \ldots, 2N_\theta, \end{aligned} \tag{15}$$

where $\boldsymbol{L}_l$ is the l-th row of the Cholesky factorization matrix $\boldsymbol{L}$ of $(N_\theta + \kappa) \cdot \boldsymbol{\Sigma}_{\boldsymbol{\theta}_{ij}}$, with $\boldsymbol{L}^\top \cdot \boldsymbol{L} = (N_\theta + \kappa) \cdot \boldsymbol{\Sigma}_{\boldsymbol{\theta}_{ij}}$, where κ is a scaling parameter.

The method relies on a deterministic sampling scheme of the input space which provides the sigma points in (15). These peculiar points for the input distribution are propagated through the non-linear likelihood function to calculate the output sigma points, $\mathcal{L}_{\boldsymbol{s}_{0:N-1}|\boldsymbol{e}_{ij}}(\boldsymbol{\chi}_l)$, which are used to evaluate the mean likelihood by (14). In this way the average can be evaluated deterministically by using a fixed number of samples that is of the order of the input vector dimension, resulting in a considerable reduction in the required processing time.

III. Results

The track-to-graph statistical association algorithm has been tested by using real AIS data sets collected in different areas and time frames. AIS tracks have been used to construct the traffic graph in the ROI and as ground-truth to test the association algorithm. In this section a real-world data set is considered (see Fig. 3), which has been collected in the Eastern Atlantic, specifically off the coast of Portugal (min/max lat=35.5/42.5 deg, min/max lon=-11/-6.5 deg) from 01-Jun-2016T10:07:03+00 to 31-Jul-2016T23:58:16+00. The traffic graph for this area has been estimated by using the procedure in [4]. The two association algorithm implementations, described in section II-A, have been tested showing similar results. The performance of the association test is qualitatively evaluated by checking the coherence between way-point positions and mean velocity distributions of the associated edge, and the present and future kinematic state of the ground truth ship track with respect to the data window under test. In general, applying this qualitative criteria, the association test performs well for the chosen data sets. The results reported make use of the UT association test implementation. As an example, Fig. 4 provides a spatial representation of the traffic graph with nodes located at the mean values of the way-point position clusters and connections between nodes i and j traced if $P(\boldsymbol{e}_{ij}) \neq 0$. Fig. 5 shows some examples

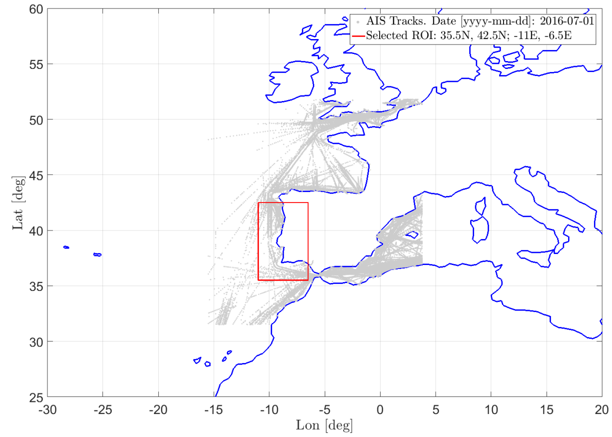

Fig. 3. Eastern Atlantic AIS track data set and selected ROI.

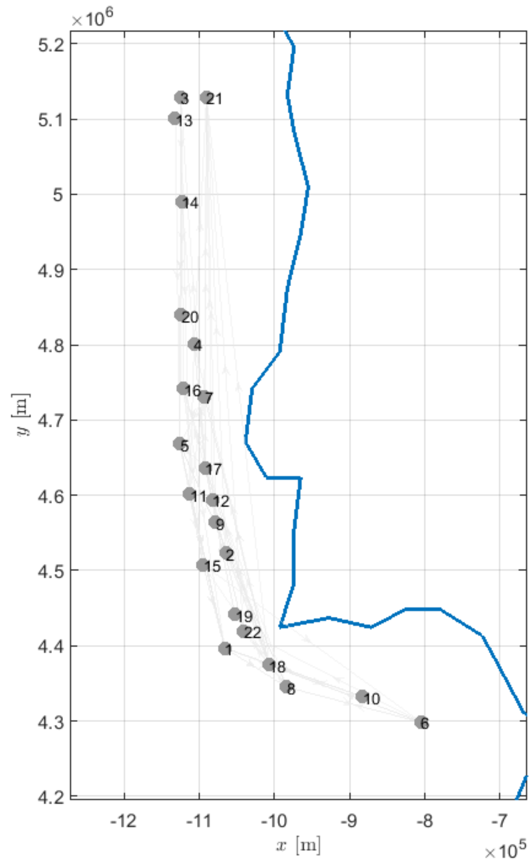

Fig. 4. Traffic graph model spatial representation. The nodes are the mean of each way-point position cluster.

of simulated trajectories using the mean-reverting model with parameters distributed according to the traffic graph priors, $p(\boldsymbol{\theta}_{ij})$. The graph-to-track association has been tested using a data window of 5 samples. Fig. 6 and Fig. 7 show the association results for two different real tracks. The associated initial node is represented by the way-point position cluster mean and covariance, $\boldsymbol{\mu}_{\boldsymbol{\omega}_i}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\omega}_i}$, respectively, with the average of $\boldsymbol{v}_{ij}$, that is $\boldsymbol{\mu}_{\boldsymbol{v}_{ij}}$, superimposed.

ISBN 978-90-827970-1-5 © EURASIP 2018

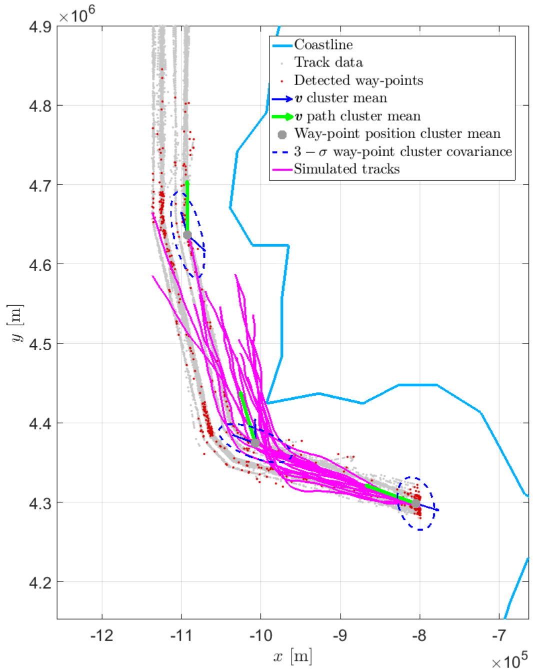

Fig. 5. Example of traffic state sequence and simulated vessel trajectories.

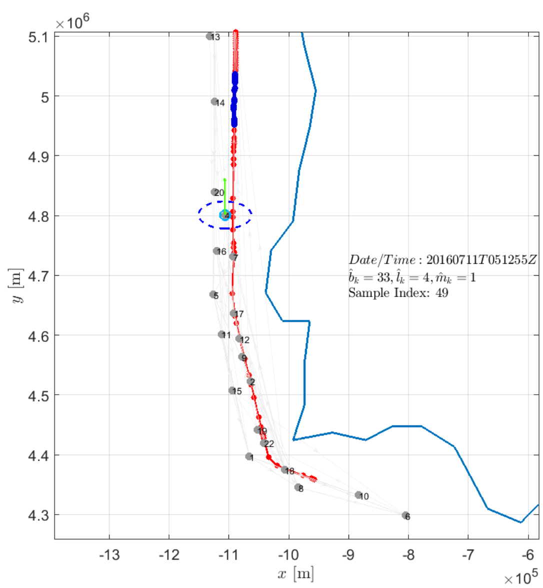

Fig. 6. Association test example 1.

IV. Conclusion and Future Work

The MAP rule has been proposed to associate a ship trajectory to an edge (or node) of a given graph that statistically models typical vessel traffic routes in a ROI. The approach is based on the assumption that ship velocities are modeled by mean-reverting processes. Prior knowledge of mean-reverting kinematic parameters, which are learned during the construction of the graph from historical data, are exploited to compute the MAP rule. The paper proposes two implementations of the algorithm, one based on MCM, the other one on the UT.

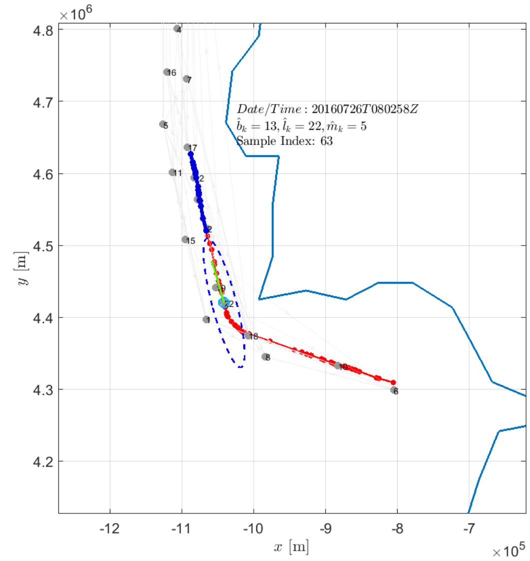

Fig. 7. Association test example 2.

Tests on real AIS data show a qualitatively good association performance. Future developments of this work include the definition and the evaluation of suitable performance metrics.

V. Acknowledgment

This work has been funded by DRDC Maritime Information Warfare Project on Next Generation Naval C2 Systems under the project *Patterns of Life Model Parametrization for Exploitation in Command and Control Systems*, and this support is acknowledged with thanks to Steven Horn. There is parallel support to the NATO ACT project *Data Knowledge Operational Effectiveness*.

References

[1] G. Pallotta, M. Vespe, and K. 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. [Online]. Available: http://www.mdpi.com/1099-4300/15/6/2218

[2] B. Ristic, B. L. 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.

[3] 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*, to be published.

[4] R. Grasso, L. M. Millefiori, P. Braca, and S. Horn, "A Bayesian model for maritime ship route learning based on mean-reverting stochastic processes," in progress.

[5] 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.

[6] L. M. Millefiori, P. Braca, and P. Willett, "Consistent estimation of randomly sampled Ornstein-Uhlenbeck process long-run mean for long-term target state prediction," *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1562–1566, Nov 2016.

[7] B. K. Øksendal, *Stochastic differential equations: an introduction with applications*. Springer, 2013.

[8] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar 2004.

ISBN 978-90-827970-1-5