Multi-Source and Multi-Target Tracking for Maritime Surveillance

Baptiste Morisse
Lead scientist
Aegir
baptiste.morisse@aegir.fr

Edouard Villain
Research engineer
Aegir
edouard.villain@aegir.fr

Joao Chueire
Research engineer
Aegir
joao.chueire@aegir.fr

Matthieu Vanicat
Chief scientific officer
Aegir
matthieu.vanicat@aegir.fr

***Abstract*—We present a multi-source, multi-target tracking system for maritime surveillance. This system integrates heterogeneous data sources—radar, satellite RF, AIS, EO/IR, sonar, and EW data—differing in nature, update rate, and accuracy. It is designed to process both out-of-sequence measurements or infrequent data, and data with varying levels of identification, from non-existing identifiers to strongly attributed sources. The system operates in real time, even in dense environments with several hundred concurrent tracks. It combines hypothesis management techniques, spatio-temporal data structures, and filters tailored to maritime dynamics. The system's performance is evaluated on synthetic data, and preliminary experiments on real AIS and satellite RF data have shown promising results.**

***Index Terms*—Target tracking; Sensor fusion; Bayesian methods; Maritime navigation; Multisensor systems; Kalman filters;**

I. Introduction

Multi-target tracking consists in performing two main tasks, given a set of partial and noisy measurements from one or more heterogeneous sensors:

- Grouping measurements generated by a same source, often referred to as a target or mobile. Such groups of measures generate a set of tracks, with ideally one track per source. One issue is the presence of sensor false positives (for instance, clutter). A track is a collection of spatio-temporally correlated measurements, carrying potentially an identifier defined at its creation as well as identification or classification information about the source. This step is referred to as *data association*.
- Estimating for each track the position, velocity, and associated uncertainties over time, based on the partial and noisy measurements that compose it. This step is referred to as *data assimilation*.

These two steps are performed jointly and enable the construction of the *Common Operational Picture (COP)*, i.e. the trajectories, current positions and velocities, together with identification and classification information for all mobiles within the sensors environment.

Multi-target tracking has been an active research area for several decades, evolving in response to industrial challenges and the steady increase in available computational power. The earliest industrial needs emerged in the 1960s around airborne radar systems, focusing on the detection and tracking of aircraft and missiles. The first multi-target tracking systems were born from the mathematical formalization of the Kalman filter [1], [2] and probabilistic data association in cluttered environments [3], as well as the advent of microprocessors.

These techniques matured with the introduction of methods such as Probabilistic Data Association (PDA) [4], Joint PDA (JPDA) [5], [6], and Multiple Hypothesis Tracking (MHT) [7]–[9] for robust data association and clutter handling, along with the Extended Kalman Filter (EKF) [10], [11] to handle nonlinear models. These approaches found applications in defense systems, in particular in combat systems and command-and-control (C2) architectures. Over time, their robustness and tracking accuracy were improved through techniques such as the Ensemble Kalman Filter (EnKF) [12], [13], Interacting Multiple Models (IMM) [14]–[16], and Particle Filters and Sequential Monte Carlo Methods (SMC) [17]–[19], enabling the tracking of maneuvering or evasive targets and the use of incomplete or imprecise data.

More recently, a new probabilistic paradigm based on Random Finite Sets (RFS) has emerged for describing the multi-target tracking problem [20]. This makes it possible to describe the tasks of data association and data assimilation in a unified manner within a rigorous and general Bayesian framework, to control the approximations and assumptions made in the modeling, and to state and prove optimality results. It led to algorithms such as PHD [21], [22], CPHD [23] and Multi-Bernouilli [24]–[26] filters, particularly suited to high-noise environments. In parallel, tracking algorithms coupled with raw signal processing [27]–[29] —often leveraging recent advances in machine learning (e.g., computer vision, point cloud analysis [30]–[33])—have been developed to address video, lidar and radar tracking challenges in domains such as Advanced Driver Assistance Systems (ADAS), autonomous driving, and autonomous navigation or decision-making for drones.

This paper addresses the problem of multi-source, multi-target fusion for maritime surveillance. Relevant applications include, but are not limited to: infrastructure monitoring (e.g. ports, offshore platforms), Exclusive Economic Zone (EEZ) surveillance by coast guards, generation of a common operational picture (COP) from onboard ships or drones, or more generally, within a network of distributed sensors and effectors. The system ingests heterogeneous data sources, such as AIS, radar, satellite RF, EO/IR, sonar, and EW data.

The primary challenges lie in handling (1) intermittent, irregular data with highly variable transmission delays, and (2) a large number of simultaneous targets (hundreds to thousands). Data intermittency—such as a vessel becoming undetectable for hours after disabling its AIS—requires the ability to maintain track continuity using large-scale behavioral models. Traditional velocity diffusion models (e.g. random acceleration) perform adequately with high-frequency updates (e.g. every second) but struggle to produce realistic presence probability estimates over longer time horizons (tens of minutes to hours) without measurements.

Highly variable transmission delays—such as satellite data becoming available several hours after acquisition—require the ability to ingest lukewarm data, often arriving out of sequence with respect to their acquisition time, with delays of up to several hours. Lastly, source heterogeneity requires the assimilation of measurements with diverse characteristics and varying levels of kinematic and identification content.

We propose a multi-source, multi-target tracking system that addresses these challenges through several original components:

- A behavioral model for targets, capable of estimating realistic presence probabilities over a time period spanning several hours;
- A spatio-temporal KD-tree data structure enabling efficient retrieval of relevant tracks and time points for assimilating measurements, including out-of-sequence observations;
- Measurement likelihood models that account not only for kinematic parameters but also for various levels of identification features (e.g. RF signatures, MMSI identifiers).

The remainder of the paper is organized as follows. Section II gives an overview of the sensors and data that can be involved in maritime surveillance applications, describes real-world scenarios in which the system has been tested and details the synthetic data generation process used for development and quantitative evaluation of the system. Section III introduces notation sand definitions, and gives a detailed presentation of the algorithmic components. Section IV focuses on results and performance metrics of the tracking system. Finally, section V concludes with a discussion and outlines directions for future works.

II. Sensors and Datasets

A. Sensors and data for maritime surveillance

1) Sensors: A wide range of heterogeneous sensors can be deployed for maritime surveillance applications. Each provides complementary information, differing in coverage, precision, update rate, and robustness to adverse conditions.

- **Global Navigation Satellite Systems (GNSS).** GNSS receivers (such as GPS, Galileo, GLONASS or BeiDou) provide precise latitude, longitude, speed, and heading directly from the vessel instruments themselves. Their advantages include high accuracy (a few meters) and global coverage. However, GNSS is a cooperative signal. It requires the vessel to report its position, and it is vulnerable to spoofing, jamming, or intentional deactivation.
- **Radar.** They provide all-weather, day-and-night detection of vessels, with typical ranges from a few nautical miles (X-band navigation radars) to hundreds of nautical miles (long-range coastal or over-the-horizon radars). Their main characteristics include high update rates (seconds), good range accuracy, and medium angular resolution. However, radars are sensitive to sea clutter and multi-path, and may generate false alarms, particularly near coastlines or in heavy sea states.
- **Automatic Identification System (AIS).** AIS is a co-operative system where vessels broadcast their identity, position, speed, and other navigational information. Its main advantages are the richness of identification data and the high accuracy of reported positions. Limitations include voluntary deactivation, spoofing, delayed transmissions via satellite relay, and incomplete coverage in dense traffic or remote areas.
- **Satellite RF Sensors.** Passive RF payloads on satellites can detect and geolocate maritime transmissions such as AIS, VHF, or radar emissions. They offer wide-area coverage (regional to global) but with low temporal resolution (hours between revisits) Some limitations include high transmission delays and relatively coarse geolocation accuracy compared to terrestrial sensors.
- **Electro-Optical (EO) Cameras.** EO systems provide high-resolution imagery in the visible spectrum, enabling fine-grained classification (ship type, behavior) and situational awareness. They are limited by weather and daylight conditions, and typically offer a narrower field of view compared to radar.
- **Infrared (IR) Sensors.** IR sensors detect thermal emissions, allowing target detection and recognition at night or in reduced-visibility conditions. Their range is generally shorter than radar, and performance is strongly affected by atmospheric conditions (humidity, temperature gradients).
- **Electronic Warfare (EW) Antennas.** EW sensors detect and classify radar or communication signals emitted by vessels. They provide valuable identification cues and operate passively, without revealing the surveillance system. However, their performance depends on the emission behaviors of the target and is then limited in silent or emission-controlled environments.
- **Active Sonars.** Active sonars transmit acoustic pulses and analyze the returned echoes to detect underwater or surface targets. They provide range and bearing measurements with relatively high accuracy and are particularly effective in submarine detection. Main limitations are their limited coverage compared to passive arrays, susceptibility to environmental conditions (e.g. thermoclines), and the fact that emissions can reveal the presence of the surveillance platform.
- **Passive Sonars.** Deployed from fixed buoys, coastal arrays, gliders or naval platforms, passive sonars de-

tect acoustic signatures of vessels and submarines. They provide bearing-only measurements with potential long detection ranges in favorable propagation conditions. Limitations include strong dependence on the underwater environment and difficulty in resolving multiple targets.
- **Human Intelligence (HUMINT).** Reports from coastal patrols, aerial assets, or civilian observers can complement technical sensors. They provide flexible and context-rich information but are irregular in time and highly heterogeneous in reliability.

The core of our tracking system is designed to be sensor-agnostic and be able to operate with data from all of the sensors listed above. To tackle the issue of the variety and heterogeneity of the raw data, and the difficulty therefore to feed them directly to the tracker, we develop and make use of so-called *tactical data*. The idea is to provide a unified blueprint for information required by the tracker, and easily extracted from raw data. Integrating a new sensor therefore only requires specifying its measurement model, i.e. the tactical data it produces and the associated uncertainty. These measurement models are interchangeable components within the tracking system, used both for computing measurement likelihoods and for the filter update step (see Section III).

2) Tactical data: Tactical data are a unified blueprint for processed data, containing target detections and contain (possibly partial) information on position and velocity, as well as identification and classification attributes of the target. An example of tactial data from a raw data would be a raw image (e.g. radar, sonar, or video) that would require preprocessing to extract relevant detections. Following is a short description of such a blueprint:

- Position data: latitude, longitude, range, bearing
- Velocity data: heading, speed, radial velocity
- Classification data: electromagnetic or appearance signature, size, type, track ID, public ID

The capabilities of each sensor are summarized in figure 1.

	GNSS	RADAR	EW	ACTIVE SONAR	PASSIVE SONAR	EO/IR	AIS	SAT RF	HUMINT
Lat/Lon	✓						✓	✓	(✓)
Bearing		✓	✓	✓	✓	✓			(✓)
Range		✓		✓		✓			(✓)
Heading	✓						✓		(✓)
Speed	✓						✓		(✓)
Radial velocity		✓		✓					
Signature			✓			✓		✓	
Size		(✓)			(✓)	(✓)	(✓)		(✓)
Type						(✓)	✓		(✓)
Track Id	✓	(✓)							
Public Id							✓		

Fig. 1. Capabilities provided by heterogeneous sensors for maritime surveillance. Checkmarks indicate typical availability; parenthesized checkmarks denote indirect/conditional availability.

B. *Real world experimental cases*

The system was tested on real data in two contexts:

- In real time, during the Dronathlon challenge organized by the French Navy. The setup included a USV equipped with a navigation radar, an AIS receiver, and an electro-optical suite, as well as an AUV equipped with a camera. The tracking system fused AIS streams, radar detections, video detections, and navigation information from the drones (MAVLINK). The operational area was limited in size (a few kilometers in radius) and restricted to challenge participants only, resulting in a relatively small number of entities to detect.
- Offline, on historical AIS data and satellite RF detections in the Gulf of Guinea. These data covered an area of several hundred kilometers in radius with several hundred vessels navigating simultaneously. Three satellite passes (one every 24 hours) were available.

These real-data experiments provide only a qualitative evaluation of the system, since omniscient ground-truth information on the operational situation was not available. It is therefore essential to quantitatively assess system performance on synthetic data, for which the ground truth is known. This is the focus of the following subsection.

C. *Synthetic data*

The generation of synthetic data is essential for the development of a tracking system. It offers several advantages:

- Scenario control: number, type, and trajectories of targets, as well as the type and performance of available sensors.
- Omniscient ground-truth knowledge: enabling rigorous evaluation of tracking system results.
- Mass generation of scenarios: allowing the study of parameter impacts such as noise level, false-positive rate, transmission delays, etc.

To generate synthetic data, we used a proprietary simulator, whose main characteristics are described below. The simulator relies on the Godot 3D game engine, which provides numerous utilities for estimating line-of-sight (ray casting), defining physical behavior laws, and handling 3D terrain models.

1) Simulated sensors:

- **AIS.** AIS messages include position information expressed in latitude and longitude (with Gaussian measurement noise of standard deviation 20 m) together with speed and heading, affected by Gaussian noise with standard deviations of 1 knot and 5°, respectively. Each message also carries the MMSI identifier, which is a persistent and unambiguous identifier of the source. Vessels broadcast AIS messages on average every 5 seconds, with a random transmission delay between 0 and 5 seconds. When a vessel is stationary, the emission rate decreases to one message every 5 minutes. Vessels may probabilistically disable AIS transmissions for random durations of up to 6 hours.
- **Radar.** Radar detections consist of distance (with multiplicative Gaussian noise of 0.5% standard deviation) and

azimuth (with Gaussian noise of 1.5° standard deviation) for all targets within the radar detection range. The radar performs one full revolution every 10 seconds. False negatives (missed detections) occur with a given probability, while false positives (spurious echoes unrelated to any real target) are generated uniformly across the coverage area, also with a given probability per revolution. The radar may attach an ephemeral identifier (tracklet) to detections, modeling the presence of an internal tracking system.

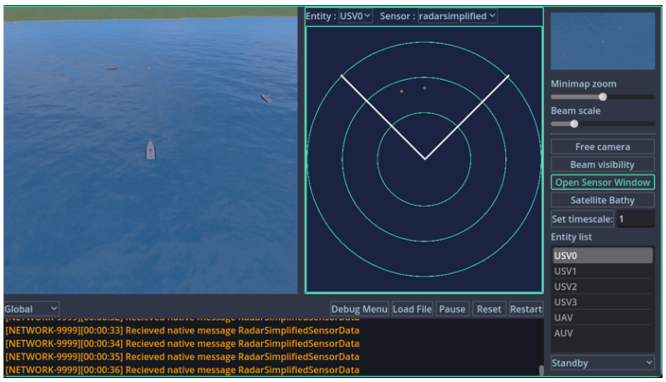

Fig. 2. Simulator user interface, with graphical view of the theater and graphical outputs of a radar sensor

2) Entities and behavior laws: Three categories of entities, each with specific behavior models, are included:

- **Civilian.** Pleasure craft operate in shallow waters (bathymetry between 10 m and 100 m), following random trajectories at speeds between 2 and 25 knots.
- **Fishing vessels.** Fishing boats depart from their home ports, transit to fishing grounds at depths greater than 200 m, and then perform random movements within a restricted one-nautical-mile radius area at speeds between 2 and 6 knots for several hours (fishing operations) before returning to port.
- **Merchant ships.** Merchant vessels perform transits between ports within the theater (or exit points) selected randomly, cruising at a constant speed of 20 knots, except when approaching ports, where they reduce speed.

3) Scenario and test dataset: Using the simulator, we generated one dataset to evaluate the tracking system. The results are presented in Section IV. The test scenario represents the case of some coastal surveillance with a 10 km range radar (without an internal tracker) and AIS data. The theater covers a 70 by 60 nautical-mile area, includes six ports (serving as departure and arrival points for fishing and merchant vessels) At any given tim, around 200 active entities are in navigation. The dataset consists of 4 hours of simulation. In addition to the tactical sensor data, ground-truth positions, velocities, and identifiers of all targets are recorded to enable quantitative evaluation of tracking performance.

III. Methods

A. *Definitions and Overall Operation*

The main objective of the system is to track entities, that is to make available in real time to the user their position, velocity, and potential metadata (typically, identifiers), and the amount of certainty regarding those intel (quality of data used, noise estimation, precision of the underlying algorithm...). The available information from those entities comes from sensor measurements, which are by nature heterogeneous, potentially delayed, noisy, and so on. The system feed on those measurements and fuse them in a smart way to recover only the needed information.

Our system relies on Targets and Tracks to do so: a *Target* mimics a real entity, via its metadata; and a *Track* represents the kinematic state of a Target. We will see in the following that one Target may have multiple Tracks associated to it at one given time, using *hypotheses*. The design of our system is then twofold:

- **Data Association** – assigns a measurement to a Track, possibly creating a new one.
- **Track Estimation** – estimates a Target's position and velocity, with some confidence scores regarding the estimation.

These two points are highly connected, as Data Association influences the next Track estimations, and track estimation is used to compare Tracks and data.

Data Association is based on a *Multi-Hypothesis Tracking* (MHT) approach. One hypothesis represents a possible interpretation of the observed data, assuming specific associations between measurements and existing Tracks. Multiple hypotheses must be maintained, especially when there is ambiguity about the assignment of a measurement to a Track. The idea is to delay the choice of the "correct" association until future measurements help resolve the ambiguity.

Track Estimation relies on Bayesian inference techniques to estimate the Target state (position, velocity, and associated uncertainties) from partial and noisy measurements. These techniques use a stochastic motion model (i.e. a prior on the Target's dynamics and behavior) along with models for each and every one existing sensors. The *Extended Kalman Filter* (EKF) performs this estimation in the case of Gaussian uncertainties on both motion and measurement noise.

The system maintains a set of Targets, Tracks, and hypotheses. Intuitively, a hypothesis is a proposed description of reality, containing a set of Targets, each with a defined state. A Track is a mathematical object that characterizes a Target's state and ensures its temporal continuity.

Formally:

- Let $\{M_1, M_2, \ldots, M_I\}$ denote the set of Targets.
- Let $\{T_1, T_2, \ldots, T_J\}$ denote the set of Tracks.
- Let $\{H_1, H_2, \ldots, H_K\}$ denote the set of current hypotheses.

Targets form a partition of the Tracks (i.e. a Track belongs to exactly one Target):

$$M_i = \{T_{i1}, T_{i2}, \dots\}, \quad \{T_1, T_2, \dots, T_J\} = \bigcup_i M_i.$$

The Tracks $\{T_{i1}, T_{i2}, \dots\}$ represent the Target M_i in the different description of the reality carried by the hypotheses. A hypothesis is a set of $(\text{Target}, \text{Track})$ pairs:

$$H_k = \{(M_{in}, T_{jn})\}$$

where $T_{jn} \in M_{in}$ and all M_{in} are distinct. A hypothesis defines the set of Targets present and the Tracks that represent them. Each hypothesis has a score S_k, interpreted as the log-likelihood of the hypothesis.

The system ingests data *frame-by-frame*, where a *Frame* is a set of measurements $\{z_1, z_2, \dots, z_m\}$ from a single sensor, assumed to originate from different Targets. A Frame is a way to make unity in a world of chaos: where the data comes from several, different, heterogeneous, delayed, noisy sensors, a Frame represents a batch from one sensor and as homogeneous as possible (with respect to time and some potential hyperparameters). This allows for fine-grained, sensor specific and time-wise analysis of a batch of Data, improving greatly the fusion and uses of the data.

Data fusion proceeds through:

1) Preselection of relevant Tracks
2) Estimation of Track positions at the measurement time
3) Computation of association likelihoods
4) Determination of the best associations
5) Update of hypotheses and Tracks

B. Detailed processing

1) Gating and Preselection of Relevant Tracks: Incoming measurements allow the system to preselect relevant Tracks, that is the most likely to be associated with the measurements. Tracks that cannot explain any measurement with sufficient certainty are discarded. This greatly reduces the computational complexity of the following steps.

This selection is done by computing distances between measurements and Tracks, keeping only those within a gating threshold derived from the sensor and motion models. To perform this efficiently, a KD-tree data structure is used to retrieve the nearest Tracks in logarithmic complexity.

2) Motion Model and Position Estimation: Each Track estimates the Target's position at any time using a motion model and Bayesian filtering (EKF).

We introduce an original motion model adapted to maritime dynamics over long time horizons: A Target has a probability $\lambda \cdot dt$ of changing heading and $\mu \cdot dt$ of changing speed during a time interval dt. As $dt \to 0$, the number of heading and speed changes follows Poisson distributions with parameters λ and μ respectively. When a heading change occurs, the increment is drawn uniformly from $[-\pi/2, \pi/2]$. When a speed change occurs, the new speed is drawn uniformly from $[v_{\min}, v_{\max}]$.

The first and second moments of the state vector $(x(t), y(t), v(t), \theta(t))$—where $x(t)$ and $y(t)$ are Cartesian coordinates, $v(t)$ the speed, and $\theta(t)$ the heading—can be computed analytically, yielding a Gaussian approximation of the stochastic dynamics, compatible with an EKF.

3) Likelihood Computation: Measurement-to-track associations are based on an association cost combining kinematic proximity and identification information.

Kinematic proximity is computed from the predicted Track state at the measurement time, with the sensor model providing the log-likelihood.

Identification information is used to refine this score. Identifiers may be:

- Strong/unambiguous: MMSI, radar track ID, combat system track ID, ...
- Partial: RF signature, visual features, ...
- Absent: basic radar, ...

Identifiers may be persistent or change over time. The diversity of cases justifies the need for sensor-specific exploitation strategies and the careful design of Frames.

4) Association: For each hypothesis H_k, an association cost matrix C_{ij} is built, where i indexes measurements and j indexes Tracks in H_k.

From this matrix, the best global association functions f are computed, with $f(i)$ giving the Track index for measurement i, ensuring $f(i) \neq f(i')$ for $i \neq i'$. The best associations minimize:

$$L_f = \sum_i C_{i,f(i)}.$$

They are computed using the Jonker-Volgenant (JV) algorithm and Murty's algorithm.

5) Update: The best associations for each of the K hypotheses are combined to form the new K best hypotheses. Specifically, we determine the K pairs (k, f) minimizing $S_k + L_f$, where k is the current hypothesis index and f is a global association for that hypothesis.

These pairs define the new hypotheses and update the Tracks. For (k, f), let $H_k = \{(M_{in}, T_{jn})\}$ be the Target-Track content of hypothesis k. All Tracks T_{jn} are updated using the EKF update step with the measurement assigned to j_n by f.

C. Cleaning up and Outputs

As we said at the beginning of this Section, the goal of the Tracker is to output the states of the tracked entities. The core algorithms of our Tracker produce many Tracks and Targets, often more than the "real" number of entities. To maintain those numbers in check and not overwhelm the end user, our Tracker has several way to clean up Tracks:

- Confirmation: a Track is said to be confirmed if it contains some signed data, or enough unsigned data ; an unconfirmed Track is discarded.
- Activation: a Track not receiving data for more than some threshold defined in advance is put to sleep. This avoids Track with too much uncertainty on their state, which could attract too many sound data from other entities.

IV. Results

Simulated data and the tracker described in Sections II and III were used to obtain the following results. A Tracker such as the one designed and studied here may be analysed and studied via various methods. We focus here on two of them: classification tools and trajectories evaluation.

In Subsection IV-A we study our Tracker as a classifier regarding the Data Association process. One advantage using simulated data over real data is the knowledge of the ground truth, via the GPS and medatadata of each entity. As a Track is associated with only one Target, which is defined by a unique set of metadata (in the present situation: AIS metadata), we can pass on the "truth" from the the data to the Tracks themselves. **A Track is said to be the entity j if it contains AIS data from the entity j.** We can then say if a radar data has been correctly associated to the right Track or not.

In Subsection IV-B we study our Tracker through Tracks as whole: does one Track reproduces with high precision the path of the underlying entity? To simplify here the analysis, we compare only Tracks associated with signed (AIS) data, meaning, we already know which entity to compare to.

We partition all Tracks produced by the Tracker as follow, using in particular the confirmed/unconfirmed distinction explained in the previous Section:

- **Signed Track:** is confirmed and contains signed (AIS) data.
- **Ghost Track:** is confirmed but lacks identifier.
- **Unconfirmed Track:** is unconfirmed.

In a real situation, a Ghost Track could be for instance a jet-ski, a drone, or any object without identification (as a malicous or not intent). In our dataset, all entites are signed thanks to AIS metadata, hence no Ghost Tracks should exist if our Tracker were perfect.

A Track is considered pure relatively to one Sensor if all data from the Sensor associated to the Track comes from only one entity.

A. Data Association analysis

We focus here on the quality of the radar data associations. The AIS data association is automatically perfect, as we assume here a high confidence on the AIS metadata (i.e. no spoofing or other techniques).

Radar data are twofold: echoes from actual entities, and false positive like echoes from birds, coastlines or waves. In our dataset, there are 26131 echoes from entities, and 11 false positive. As explained in the previous Section, false positive of the radar should directly be associated with an unconfirmed Track. Finally, a radar data from entity j should be associated to the Track j ; several type of errors may occur: association to another signed Track, or a Ghost Track, or an Unconfirmed Track. Table I sums up the discussion.

Table II presents the distribution of signed, unconfirmed, and ghost tracks, along with other statistics.

	Signed j	Signed k	Ghost	Unconfirmed
radar from j	True	False	False	False
false positive	False	False	False	True

TABLE I
Truthness table

	Signed	Ghost	Unconfirmed
Count of Tracks	601	57	18
Count of pure Tracks	532	49	16
# Radar	25607	511	24
# correct Radar	25397	0	11

TABLE II
Details of tracks

The Table I allows to compute for instance the (global) accuracy of our Tracker regarding radar data association:

$$acc = \frac{25397 + 0 + 11}{25607 + 511 + 24} = 97.2\%$$

Though the accuracy is a very classic metric for classifiers, in our case a more detailed analysis is necessary to understand how the Tracker fails.

Signed Tracks comprise a total of 601 tracks, of which 88.5% of are pure, and contained most of the radar data (98% total). The accuracy is as high as 99%, meaning radar data associated to Signed Tracks are almost always associated to the right Signed Track. Some entities generated multiple tracks, typically for entities remaining in ports for some time. This is a byproduct of the deactivation of Tracks as explained in the previous Section (see Figures 3 and 4 for an example of an entity tracked by two tracks).

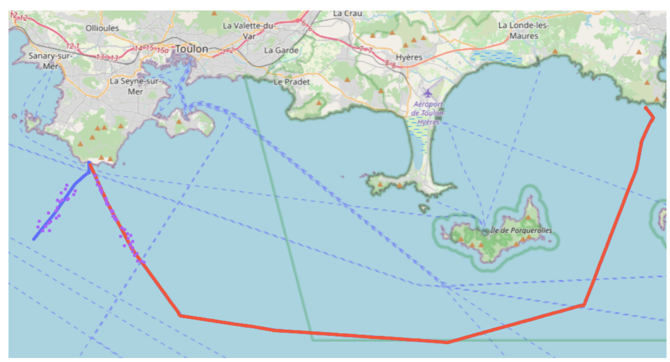

Fig. 3. First part of an entities tracked by two tracks (with red : tracker output; purple : radar input and blue : GPS ground truth)

Unconfirmed Tracks account for 18 tracks in total and less than 0.1% of the radar data. Eleven of these correspond to false-positive radar data points, resulting in 11 pure, unconfirmed Tracks, meaning all unsound radar data have been correctly discarded as unconfirmed Tracks.

Ghost Tracks amount to 57 in total, comprising 511 radar data points (1.9% of all radar data). Among these, 49 were classified as pure, representing 424 radar data points, while the remaining eight accounted for 87 points. Of particular note, 498 out of these 511 radar data points originated from the

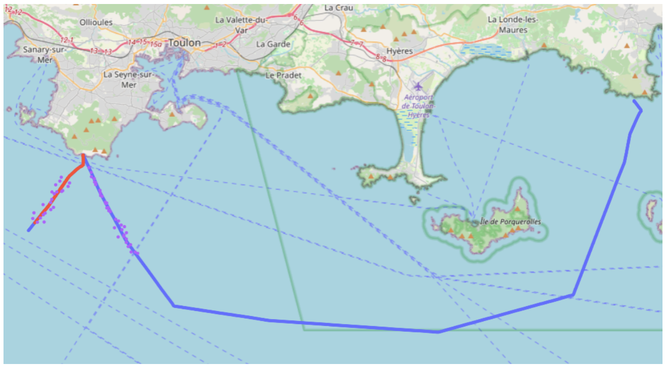

Fig. 4. Second part of an entities tracked by two tracks (with red : tracker output; purple : radar input and blue : GPS ground truth)

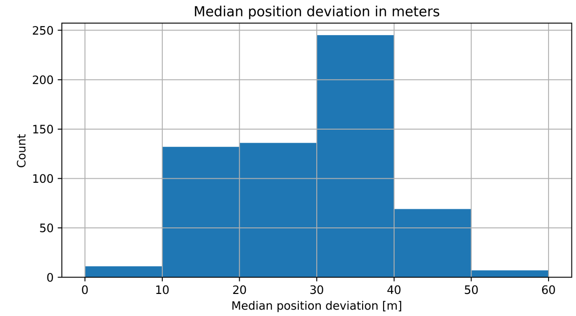

Fig. 5. Median position deviation histogram between tracker prediction and ground truth

three entities that had deactivated their AIS while in motion, thereby simulating malicious behaviour; however, they were still detected by the radar sensor during this AIS interruption. Indeed these ghost tracks contain unsigned data and are used to raise an alert for the end user.

Finally, the tracker was run for 45 minutes in order to assimilate four hours worth of data (see II-C for input data details).

B. Trajectories analysis

As a complement to the previous analysis focused on Data Association, this Subsection focuses on the study of trajectories, comparing Signed Tracks to the path of the corresponding entities.

To this end we use Trajectopy, a python module dedicated to trajectories comparison [34]. A spatio-temporal matching procedure is apply to select the GPS points of the entity's path that best correspond to the associated Signed Track trajectory. The Absolute Trajectory Error (ATE) is then computed, providing metrics that describe the similarity between the predicted trajectory and the ground truth. The three most relevant ATE metrics are used : the minimum, median, and maximum position deviation between the tracker output and the ground truth, in meters. Table III presents the minimum, median and maximum values of Signed Tracks for the three ATE metrics.

ATE Metrics	Minimum	median	Maximum
Minimum position deviation [m]	0.05	1.0	48.8
Median position deviation [m]	8.9	30.6	65.3
Maximum position deviation [m]	25.1	116.8	3811.9

TABLE III
ABSOLUTE TRAJECTORY ERROR SCORES ACHIEVED ON SIGNED TRACKS (IN METERS)

Figure 5 complements Table III by detailing the distribution of the median position deviation, in metres, for signed tracks. It is of particular importance that most of the signed tracks exhibit a median position deviation of less than 30 metres.

Based on the aforementioned ATE metrics and scores, a few tracks were selected. Figure 6 presents one of the best signed tracks in this study, with 100% accuracy on radar data. The ATE metrics for this track show minimum and maximum position deviations of 0.21 and 67.7 meters, respectively, while the median position deviation is less than 11.2 meters.

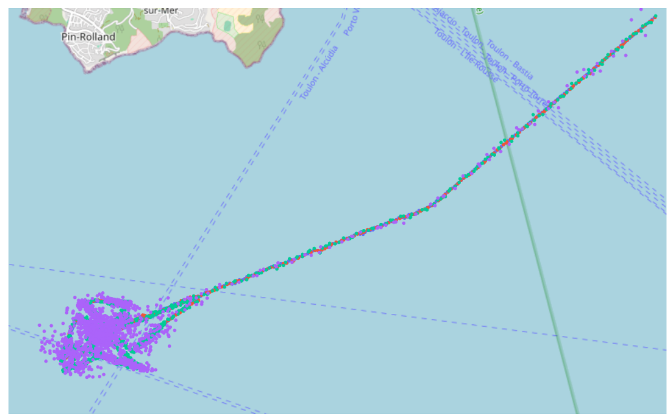

Fig. 6. Track with 100 % accuracies on AIS and radar data (with red : tracker output; green : AIS input; purple : radar input and blue : GPS ground truth)

Figure 7 presents one of the worst signed track in this study, with a data point 3,811 meters from the ground truth. It is worth noting, however, that this track accounts for only 3 errors in radar data, alongside 232 AIS and 99 radar good associations.

V. DISCUSSIONS

This work paves the way towards a generic system for a Common Operational Picture (COP) in the maritime domain. It is designed to integrate any sensor delivering tactical data (that is, kinematic information, optionally enriched with identification cues) and to scale up to complex, high-density scenarios involving hundreds or thousands of tracks. Detailed performance results are provided on synthetic data in a coastal surveillance scenario involving both radar and AIS streams. The system has also been tested on real data in two contexts:

- in real time, during the *Dronathlon* challenge organized by the French Navy, by fusing AIS streams with video detections and navigation information from friendly drones (MAVLINK);

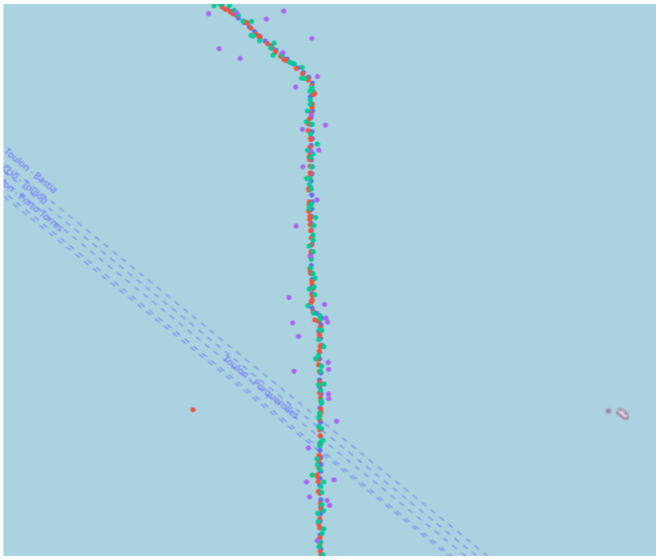

Fig. 7. Track with the highest maximum position deviation (3811 m) (with red : tracker output; green : AIS input; purple : radar input and blue : GPS ground truth)

- offline, on historical AIS data and satellite RF detections in the Gulf of Guinea.

However, several directions remain open for further development.

1) On the theoretical side: There are several elements which may enhance robustness, performance, and applicability of the current system:

- Incorporating ensemble Kalman filters could improve track initialization robustness, particularly in bearing-only scenarios (e.g. passive sonar or EW antennas).
- Leveraging Interacting Multiple Models (IMM) would allow for better tracking of highly maneuverable targets (e.g. USVs, jet-skis), especially in infrastructure protection or onboard COP applications - situations where both spatial and temporal scales are smaller.
- Developing more realistic sensor models and more robust track initiation strategies would help mitigate the impact of spatio-temporally consistent false detections (e.g. false echoes generated by ship wakes);
- Accounting for extended targets that generate multiple detections on a sensor's frame.

2) On the implementation side: Several performance bottlenecks could benefit from parallelization, including filter prediction and update steps, as well as the computation of optimal data association hypotheses. The latter is more delicate and could be addressed by decomposing the bipartite measurement-to-track association graph into connected components for parallel processing, or by parallelizing the partitions evaluated in Murty's algorithm.

3) From an application perspective: Adding an intelligent analysis layer on top of the tracking outputs would enable several valuable functionalities, such as:

- Automatically raising alerts when a vessel disables its AIS, or when a detection cannot be correlated with any AIS message.
- Detecting data inconsistencies (e.g. spoofing attempts, sensor biases).
- Computing metrics for each sensor, such as effective coverage area, consistency index relative to other sensors, update frequency, and transmission latency.
- Computing metrics for each track, such as classification uncertainty, historical richness, regularity, and diversity of contributing sources.

4) In terms of testing and qualification: The system needs to be evaluated more extensively on both synthetic and real data. For synthetic data, it is necessary to:

- Implement within the simulator all sensors listed in the first paragraph of section II.
- Integrate "bearing-only" sensors into the scenario, such as EW antennas or passive sonars.
- Include satellite sensors with transmission delays of several hours, such as satellite RF and satellite imagery.
- Study the tracker's sensitivity to radar false-positive rates and to AIS outages.

It would be beneficial to expand the evaluation tools and performance metrics in order to analyze the tracker's behavior in greater detail and to facilitate its tuning across different application contexts. It is also important to experiment with and validate the system on real data, involving a wide variety of sensors and data sources:

- Radar sensors and EO/IR streams, which may produce large numbers of false positives, sometimes with spatio-temporally correlated distributions (e.g. ship wakes, echoes from obstacles or coastlines).
- EW and acoustic data.
- Fusion of tracks from a combat system or from a tactical data link network (Link 16 or Link 22).
- Integration of human intelligence reports.

Acknowledgments

The authors would like to thank the French Navy for the opportunity to participate in the Dronathlon, and Unseenlabs for providing satellite RF data.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[3] D. L. Snyder, "Filtering and detection for doubly stochastic poisson processes," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 91–102, 1972.

[4] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.

[5] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE Journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, 1983.

[6] Y. Bar-Shalom, *Multitarget-Multisensor Tracking: Applications and Advances. Volume I.* Artech House, 1990.

[7] D. B. Reid, “An algorithm for tracking multiple targets,” *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[8] S. S. Blackman, “Multiple hypothesis tracking for multiple target tracking,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.

[9] S. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[10] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[11] A. Gelb, *Applied Optimal Estimation*. MIT Press, 1974.

[12] G. Evensen, “Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics,” *Journal of Geophysical Research: Oceans*, vol. 99, no. C5, pp. 10 143–10 162, 1994.

[13] ——, “The ensemble kalman filter: theoretical formulation and practical implementation,” *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, 2003.

[14] H. A. P. Blom and Y. Bar-Shalom, “The interacting multiple model algorithm for systems with markovian switching coefficients,” *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[15] X. R. Li and Y. Bar-Shalom, “Multiple-model estimation with variable structure—part i: Model algorithms and part ii: Mode probability computation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 448–468, 1996.

[16] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley-Interscience, 2001.

[17] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, “Novel approach to nonlinear/non-gaussian bayesian state estimation,” *IEE Proceedings F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[18] A. Doucet, S. Godsill, and C. Andrieu, “On sequential monte carlo sampling methods for bayesian filtering,” *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[19] A. Doucet, N. de Freitas, and N. J. Gordon, *Sequential Monte Carlo Methods in Practice*. Springer, 2001.

[20] R. P. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[21] ——, “Multi-target bayes filtering via first-order multi-target moments,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[22] B.-N. Vo and W.-K. Ma, “The gaussian mixture probability hypothesis density filter,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[23] R. P. Mahler, “Phd filters of higher order in target number,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

[24] B.-N. Vo and W.-K. Ma, “The multiple target member filter: A new approach to multi-target tracking,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1226–1245, 2005.

[25] B.-N. Vo, B.-T. Vo, and A. Cantoni, “The cardinality balanced multi-target multi-bernoulli filter and its implementations,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 409–423, 2009.

[26] B.-N. Vo, B.-T. Vo, and D. Phung, “Labeled random finite sets and the bayes multi-target tracking filter,” *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6554–6567, 2014.

[27] W. Koch, “Bayesian approach to extended object and cluster tracking using random matrices,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 880–886, 1999.

[28] R. P. Mahler, “Statistics 101 for data fusion: Tracking and detection before tracking,” *Proceedings of the 12th International Conference on Information Fusion*, pp. 1409–1416, 2009.

[29] B.-N. Vo, B. Ristic, and B.-T. Vo, “A particle method for multi-target detection and tracking from image observations,” in *13th International Conference on Information Fusion*, 2010, pp. 1–8.

[30] N. Wojke, A. Bewley, and D. Paulus, “Simple online and realtime tracking with a deep association metric,” in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3645–3649.

[31] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, “Fairmot: On the fairness of detection and re-identification in multiple object tracking,” *International Journal of Computer Vision*, vol. 129, no. 11, pp. 3069–3087, 2021.

[32] X. Weng and K. Kitani, “A baseline for 3d multi-object tracking,” in *European Conference on Computer Vision (ECCV)*, 2020, pp. 20–35.

[33] T. Yin, X. Zhou, and P. Krähenbühl, “Center-based 3d object detection and tracking,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 11 784–11 793.

[34] G. Tombrink, A. Dreier, L. Klingbeil, and H. Kuhlmann, “Spatio-temporal trajectory alignment for trajectory evaluation,” *Journal of Applied Geodesy*, 2024. [Online]. Available: https://doi.org/10.1515/jag-2024-0040