Density Based Clustering Data Association Procedure for Real-Time HFSWRs Tracking at OTH Distances

NIKOLA STOJKOVIC1,2, (Member, IEEE), DEJAN NIKOLIC2, (Member, IEEE), AND SNEZANA PUZOVIĆ2, (Member, IEEE)

1School of Electrical Engineering, University of Belgrade, 11120 Belgrade, Serbia
2Vlatacom Institute of High Technologies, 11070 Belgrade, Serbia

Corresponding author: Nikola Stojkovic (nikola.stojkovic@vlatacom.com)

This work was supported by the Vlatacom Institute of High Technologies under Project #89.1.

ABSTRACT In order to efficiently cover maritime areas at over the horizon (OTH) distances and thus increase marine safety in a nation’s exclusive economic zone (EEZ), a network of maritime sensors built around High Frequency Surface Wave Radars (HFSWR) can be an excellent choice. The critical parameter for success of the deployed sensor network is a real time tracking of all detected vessels. During the tracking process, data association (DA) is the first step and it defines the complexity and thus the speed of the whole tracking process. This paper presents a density based clustering DA procedure where the cluster complexity determines the applied DA procedure within the cluster itself. It is demonstrated that the great majority of clusters (over 98% of all clusters in the worst case) may be processed in a timely manner with an optimal DA procedure, or more precisely, a Joint Probability Data Association (JPDA). However, a small number of unusually large clusters (less than 2% of all clusters) requires the application of a sub-optimal DA procedure, more accurately, the Roecker’s suboptimal JPDA algorithm, in order to maintain real time performance of the whole tracking process. Moreover, unlike standard JPDA procedure which tends to be inapplicable for real time tracking in heavily cluttered environment, the density based clustering DA procedure presented here provides real time performances in the very same environment. The whole analysis is done on real HFSWR data obtained from two HFSWR, located in the Gulf of Guinea. The data set used for the experiments includes data obtained during a month and a half of constant HFSWR operation.

INDEX TERMS Clustering algorithms, data association, HF surface wave radar, JPDA, marine safety, OTH radar, radar remote sensing, radar tracking.

I. INTRODUCTION

According to the United Nations convention, the Exclusive Economic Zone (EEZ) [1] is a zone of a specific width that is stretched from a country’s territorial sea in the direction of open sea, up to 200 nautical miles (370 km), in which countries have exclusive rights to exploit the biological and mineral resources of the sea. With this grant comes a great responsibility to efficiently control this zone, which represents technological, financial and organizational challenges.

One of the solutions for complete EEZ monitoring may be a network of HFSWRs [2]. The benefits of this approach are described in [3], [4], but there are many open questions which are critical to the successful implementation of the integrated maritime surveillance system built around the HFSWR network. Such network has been installed in Gulf of Guinea, Africa. Primary surveillance role is given to two installed HFSWRs, whose locations, designated with sensor identifications 0 and 1, are given on Fig. 1. Deployed HFSWRs have nominal range of 80 nautical miles (around 150 km) during night-time and sea state of level 3 for smaller vessels. For larger ships and during the daylight, the range can extend up to 125 nm (approx. 230 km). Since stable target tracking at the over the horizon distances is the primary goal of the whole sensor network one of the most important questions is which Multi Target Tracking (MTT) procedure should
be implemented. Since all MTT procedures are dependent on the chosen the data association (DA) algorithm, the DA complexity and its suitability for real-time applications in a given HFSWR network will be the subject of this paper.

There are many joint MTT procedures such as MHT [5], JPDA [6] or Joint Integrated Probabilistic Data Association JIPDA [7]. Are any of them suitable for real–time application on HFSWR is the question which depends on the number of tracks and detections in every association step. Since during each association step there can be hundreds of tracks and individual measurements, it might show that the aforementioned joint procedures are unsuitable for practical use, because execution time can grow beyond integration periods used in HFSWR data acquisition. On the other hand, there are procedures with linear complexity (linear multi-target approach), such as LM IPDA [8] or LM Integrated Track Splitting (ITS) [9] whose execution time is growing much slower. Also, another approach is using heuristics and approximations of the aforementioned joint procedures, like integer programming [10], DFS [11], or Roecker-Phillips suboptimal JPDA [12] and its modifications such as the one described in [13] and others. Another aspect that affects the DA procedure choice is the tactical situation on the surveillance area. Since the experiment's results.

The rest of the paper is organized as follows: in Section II an overview of HFSWR data processing is given, with the description of the used HFSWR tracking procedure. Section III presents experimental results that demonstrate the effectiveness of the proposed solution as well as of an adopted DA procedure and the paper is concluded in Section IV.

II. HFSWR DATA PROCESSING OVERVIEW

The HFSWR sensor Constant False Alarm Rate (CFAR) algorithm provides measurement (detection) sets at a rate of approximately 33 seconds. Every detection (D) consists of the following parameters:

\[
D = [\rho, \theta_A, \dot{\rho}, \sigma_\rho, \sigma_{\theta_A}, \sigma_{\dot{\rho}}, \text{SNR}]
\]  

(1)

where \(\rho, \theta_A\) and \(\dot{\rho}\) present range, bearing (negative azimuth) and range rate (radial speed) of detection, respectively. These quantities represent sensor measurements. \(\sigma_\rho, \sigma_{\theta_A}\) and \(\sigma_{\dot{\rho}}\) are standard deviations of the respective quantities, while SNR is the signal to noise level in detection’s origin cell. It is assumed that there is no statistical correlation between any of the measured quantities and the measurement noise covariance will be represented as

\[
R_D = \text{diag}(\sigma_\rho, \sigma_{\theta_A}, \sigma_{\dot{\rho}})
\]  

(2)

Please note that the noise covariance matrix is not a stationary matrix, since standard deviations vary from measurement to measurement. Other important sensor parameters are measurement resolutions. For the installed HFSWR, these parameters are given in TABLE 1, with extreme value ranges of

| Quantity | Resolution | Extreme value range of \(\sigma\) | Number of used resolution cells | Used value range |
|----------|-----------|---------------------------------|-------------------------------|-----------------|
| Range    | 1.5 [km]  | 0.5-2                           | 200                           | 0 - 300 [km]    |
| Bearing  | 1 [°]     | 0.5-1.5                         | 121                           | \((\theta_A,\theta_A+60°)°, (\theta_A,\theta_A+60°)°\) | |
| Range rate\(\rho\) | 0.32 [m/s] | 0.5-2                           | 128                           | -20.5 - 20.5 [m/s] |

a) After conversion of Doppler frequency with HFSWR wavelength value

b) \(\theta_A\) is bearing of main radiation axis of H/FSWR, perpendicular to antenna array.
techniques have been developed earlier in [20] and are used in disturbances, even meteo-tsunami [19], etc. RFI suppression by strong gales on the open sea, strong atmospheric dis-

rence (RFI), such as those described in [18], clutter caused phenomena worth mentioning are radio-frequency interfer-

hours for HFSWR with a fixed working bandwidth. Other that they can even create complete blind zones for several hours for HFSWR with a fixed working bandwidth. Other phenomena worth mentioning are radio-frequency inter-

ference (RFI), such as those described in [18], clutter caused by strong gales on the open sea, strong atmospheric disturbances, even meteo-tsunami [19], etc. RFI suppression techniques have been developed earlier in [20] and are used in this system. Output of clutter density estimator block is associated with the measurement set and used in the data association procedure. Spatial Clutter Density Estimator (SCDE) with autoregressive averaging on L = 10 iterations is used (Eq. 36 in [21]). The first problem that needs to be solved is proper localization of interferences. For that purpose, the 3-D space \((\rho, \theta_A, \rho')\), limited with extreme observation values from TABLE 1, has been divided into cells of equal volume, similar to Eq. 5 and 6 in [21]. Extreme values for standard deviations of range, bearing and range rate from TABLE 1 have been adopted, to allow equispaced division, leading to cell dimensions expressed in a number of resolution cells and rounded to higher integer number: \(\Delta N_{\rho} = 7, \Delta N_{\theta A} = 6\) and \(\Delta N_{\rho} = 7\). Output from this block, in the form of a map \([P_D, P_f, \lambda](\rho, \theta_A, \rho')\) is delivered to DA block. Here \(P_D, P_f\) and \(\lambda\) represent probability of detection, probability of false alarm and estimated clutter density, respectively. Further details about SCDE operation can be found in [21].

### A. CLUTTER DENSITY ESTIMATOR

Clutter in the OTHR working environment presents a significant threat to the correctness of tracking procedures, since it can lead to the confirmation of non-existing tracks. Sometimes, these false tracks would appear on the user’s screen, which could cause, in some scenarios, unnecessary alarm with a higher level of engagement (e.g. monitoring zones of special interest, like oil platform areas). The nature of false alarms mostly lies in the inability of CFAR processing block and its associated logic to properly handle various clutter. For example, large areas in RDA cube (which is actually a CFAR input [16]) may be polluted by returns from E and F ionospheric layers. These returns can have such extreme values that they can even create complete blind zones for several hours for HFSWR with a fixed working bandwidth. Other phenomena worth mentioning are radio-frequency interference (RFI), such as those described in [18], clutter caused by strong gales on the open sea, strong atmospheric disturbances, even meteo-tsunami [19], etc. RFI suppression techniques have been developed earlier in [20] and are used in this system. Output of clutter density estimator block is associated with the measurement set and used in the data association procedure. Spatial Clutter Density Estimator (SCDE) with autoregressive averaging on \(L = 10\) iterations is used (Eq. 36 in [21]). The first problem that needs to be solved is proper localization of interferences. For that purpose, the 3-D space \((\rho, \theta_A, \rho')\), limited with extreme observation values from TABLE 1, has been divided into cells of equal volume, similar to Eq. 5 and 6 in [21]. Extreme values for standard deviations of range, bearing and range rate from TABLE 1 have been adopted, to allow equispaced division, leading to cell dimensions expressed in a number of resolution cells and rounded to higher integer number: \(\Delta N_{\rho} = 7, \Delta N_{\theta A} = 6\) and \(\Delta N_{\rho} = 7\). Output from this block, in the form of a map \([P_D, P_f, \lambda](\rho, \theta_A, \rho')\) is delivered to DA block. Here \(P_D, P_f\) and \(\lambda\) represent probability of detection, probability of false alarm and estimated clutter density, respectively. Further details about SCDE operation can be found in [21].

### B. MOVEMENT MODEL AND ESTIMATOR OPERATIONS

Before being used by the track management, the input measurement set is converted to a more suitable form within the measurement conversion block. Measurements are converted from polar to Cartesian coordinates in a process described in [22]. The coordinate system used for tracking procedures of all sensors is oriented towards true north direction (TN) and it is shown in Fig. 3. It is centered at the position of the first antenna in a receiving array. Following example described in [23], new measurement value shall be defined as a product of range and range rate quantities \(\eta = \rho \dot{\rho}\), thus defining measurement conversion process of the measurement value and its noise covariance as:

\[
\begin{bmatrix}
\rho & \theta_A & \dot{\rho}
\end{bmatrix} \rightarrow \begin{bmatrix} x & y & \eta \end{bmatrix}
\]

\[
R_D = diag(\sigma_\rho, \sigma_\theta, \sigma_\dot{\rho}) \rightarrow R_C = \begin{bmatrix}
R_{a11} & R_{a12} & \sigma_{x\eta} \\
R_{a12} & R_{a22} & \sigma_{y\eta} \\
\sigma_{x\eta} & \sigma_{y\eta} & \sigma_\eta^2
\end{bmatrix}
\]

Coefficients in converted noise covariance \(R_C\) can be found in [23], taking into account the 2-D movement scenario.
Cartesian coordinates of the converted coordinate system are denoted as x and y.

Due to the characteristics of target movement and practical aspects, in which HFSWR is used to follow relatively larger vessels whose movement parameters slightly change between measurement cycles, near-constant speed target model have been adopted to model target movement. Details of this modelling can be found in [24].

C. DATA ASSOCIATION PROCEDURE

With measurement conversion done and clutter levels estimated, DA procedures may be performed. The simple gating density-based clustering method has been adopted as method of choice. Firstly, gating around each measurement needs to be defined. Gating is based on the squared Mahalanobis distance and performed via an ellipsoid gate [5]. The description of the used gating process can be found in [24]. Since calculation of the Mahalanobis distance is costly and it is possible to obtain hundreds of measurements within a single scan, a suitable gating strategy must be adopted. In this case, a simple spacial selector is deployed to preselect measurements for chosen track’s predicted range and bearing values. A (min, max) search against every measurement position is performed to obtain a preselected list for the gating test. Execution time of the gating procedure is thus greatly reduced and have no significant impact on overall performance. However, entire procedure can be further enhanced by deploying kd-Tree [25] (binary search methods), to drastically reduce the number of (min, max) comparisons.

Let every track maintain a list of indexes of its associated measurements and every measurement maintains a list of associated track indexes. The clustering process is relatively straight-forward and starts from the list of all available tracks, see Fig. 4. If the current track was already part of another cluster, it is skipped. On the first non-assigned track, a new cluster is formed, the track is considered used and its index and associated measurement indexes are added in the list of cluster members, marking it as used at the same time. The process is continued by recursively iterating through the measurement’s lists of unused associated tracks and beyond, until all available tracks and measurements have been used and assigned to a cluster members list. This algorithm is depicted in Fig. 5. Note that clusters with one track and no assigned measurements are considered trivial in terms of data association and will be ignored in future analyses. Unassociated measurements do not participate in the DA procedure. They are delivered to track management stage as candidates for new tracks. Better inter-group isolation yields less impact on data association errors. Cluster isolation depends on the statistical distances of elements from other near-by clusters. It also depends on the volume of gating area, determined by the probability of valid observation satisfying gating criteria, $P_G$ [5]. Squared Mahalanobis distance $d_m^2$ is statistically described as chi-square distribution with 3 degrees of freedom $\chi^2_3$. Choosing a fairly large gating threshold $\gamma = 61$ yields gating criteria $d_m^2 < 61$ and the gate rejection probability for valid measurement $\bar{P}_G = 1 - P_G = P(\chi^2_3 > \gamma) \approx 3 \cdot 10^{-13}$.

In other words, it may be considered that there are practically no links between the clusters. If squared Mahalanobis distance between two arbitrary measurements is calculated using extreme values for $\sigma$ from TABLE 1. in different scenarios and after the measurement conversion, the size of the gating area volume is revealed. In case of two measurements that differ only by range, it can be shown that the allowed range value differences span from 8 km to 32 km for extreme precision cases. Similarly, bearing differences span from 5° to 16°, while radial velocity spans from 1.8 m/s to 8 m/s. This clearly indicates very large volumes of gate area and that measurements outside track’s gates will have poor association likelihood values. Association and clustering errors should be minimal, thus further increasing the cluster isolation. For this reason, mentioned gating threshold is adopted and used in experiment that will follow.

Let the number of tracks in the cluster be denoted as $N_{tc}$, number of measurements $N_{mc}$ and $N_{ac}$ number of measurement-to-track associations. The pair $(N_{tc}, N_{mc})$ will be named size of the cluster. It is assumed that clusters form nearly independent regions and that applying JPDA as cluster DA procedure on them would produce optimal weighting values inside the cluster. In that case, the usual constraints in the formation of joint events from the gating validation matrix are also applied: each measurement is assigned to at most one
If JPDA is applied on a cluster of reasonably small size, then its computation time, observing the worst case scenario, might be acceptable. The worst case scenario assumes all 1’s in gating validation matrix for the cluster. Since complexity is exponentially dependent on the number of cluster tracks or measurements, assume that the maximal number of tracks allowed in cluster is $N_{tc_{\text{max}}}$ and the maximal number of measurement associations is $N_{mc_{\text{max}}}$. Every cluster which fulfills these conditions will have JPDA applied immediately, leading to the optimal values of association weights. Suitable criteria for choosing these two quantities is a number of possible joint assignments, denoted as $C_{\text{max}}$, for calculating association probabilities between tracks and measurements, since it almost proportionally affects execution time, needed for solving a single cluster. Please note that number of joint events $C$ in cluster is the same if $N_{tc}$ and $N_{mc}$ switch values and that neither $N_{tc}$ nor $N_{mc}$ are critical parameters by itself.

If the cluster has such boundaries that the number of possible joint assignments is larger than $C_{\text{max}}$, more robust approximation must be deployed. Good choices for this suboptimal procedure are algorithms presented in [10], [11], [12] and in more recent work [27]. In general, good candidates are algorithms with tendency to linear complexity, since size of clusters and number of internal associations have no bounds. For that reason, a robust, simple and fast Roecker’s suboptimal JPDA method [12] is used in these cases, with the bias term $B$ from Eq. 29 in [12] calculated using Helmick’s proposal from [28].

The stated procedure should provide optimal track-measurement assignments within the established cluster boundary limitations. These limitations are highly dependable on the adopted algorithm implementations, used programming language, compilers and hardware. While implementation of Roecker’s suboptimal JPDA is relatively straightforward and clear, the JPDA implementation has a key role in entire procedure, since it determines the boundaries for the real time implementation. Conventional implementations assume forming of gating validation matrix from the list of track and measurement associations for particular cluster. The matrix is used for extraction of feasible joint events or hypothesis, under the rules stated earlier. Hypothesis extraction is easily implemented by recursive selection of exclusive 1’s in matrix columns, (exception is the first column, which represents clutter assignment), until all measurements are assigned to only one track (and vice versa) or declared clutter. The process is branched on every step with multiple combinations and represents traversing a node tree problem with additional limitations. This process also has its iterative implementation, which might provide a slightly faster execution in some cases. There are implementations that avoid usage of feasible event matrices, e.g. Projection Based JPDA [29], that is reported to provide faster execution times than conventional methods. However, the conventional recursive implementation has been adopted and limitations it presents for real time execution will be a subject of investigation.

D. TRACK MANAGEMENT RULES

Adopted track management rules are based on previous work and practices summarized and described in [5]. Tracks are passing 4 life states: tentative, preliminary, confirmed and deleting state, controlled by confirming and deleting M/N rules and quality counter. Further details can be found in [24].

III. EXPERIMENTAL RESULTS

The purpose of the data association analyses is to investigate characteristics of HFSWR tracking scenario and to support the hypothesis that majority of associations are formed in small clusters, which could be easily solved by an optimal DA procedure. This assumption is derived from tactical supposition that HFSWRs are used for open sea surveillance at over the horizon distances, where traffic is generally dispersed over a great area thus making the dense concentration of vessels a seldom case. Additionally, this means that HFSWR poor resolution and inability to resolve dense target concentrations will have a minor effect to the quality of tracking. Another important aspect is the execution time of the whole tracking procedure, which determines its practical usability. For this investigation the continuous set of detection records from both HFSWRs are used. Each data set is consisted of at least 100,000 individual detection sets per HFSWR.
A. FIELD EXPERIMENTS

Here, some interesting cases will be examined. Firstly, coverage areas of both HFSWRs with all targets which passed over the area during the observation period of 24h will be presented in Fig. 6 and discussed below.

The total HFSWRs coverage area is slightly larger than 100,000 km$^2$ covering the western part of the Gulf of Guinea, namely Bight of Benin. Note that the coverage areas of single HFSWRs are overlapping significantly in order to ensure constant surveillance over critical areas. During a 24h period there were over 750 and 600 confirmed tracks created by HFSWR 0 and 1 respectively. Tracks created by HFSWR 1 are marked with a green colour, while tracks created by HFSWR 0 are marked with a red colour in Fig. 6. Taking into account that a single track may consist of thousands of independent HFSWR detections it is clear that the DA procedure needs to process a fairly large amount of data. Add that to the previous number of detections which originated from false alarms and the already demanding situation becomes even more demanding. Also, in Fig. 6 it can be noted that there are tracks marked with a white color. Those tracks represent tracks reported by AIS [30], both land and satellite. Please note, that in this region of the world there are no regulating bodies which enforce proper usage of AIS devices on-board vessels. This leads to the situations where AIS data can significantly differ from the HFSWR data for the same vessel, simply because the onboard AIS device is turned off. Additionally, satellite AIS data can have significant latency, which, again, leads to the significant discrepancy between HFSWR and AIS data. This paper will not delve deeply into this matter, since it was already considered in [31], but it needed to be mentioned in order to fully describe Fig. 6.

In the end there were 833 vessels detected during the observation period (24h), which may lead to the conclusion that there is one vessel at every 100 km2. These by itself will completely diminish the needs for good sensor resolution and ease task for DA procedures, since simple single-track PDA procedure may be used.

However, a bit more detailed examination of Fig. 6 will show that vessels are generally concentrated in the areas close to the major harbors, oil platforms or at the usual maritime routes. Since HFSWRs are designed for deep sea monitoring, the region right in front of the harbors are not covered by the HFSWR network, since surveillance of these areas is a job for micro-wave radars. However, approaching routes to the harbors are covered by the HFSWR network, while oil platforms and maritime routes may be in the deep sea, so they are covered as well.

Since HFSWR resolution is not particularly good, it comes to the DA procedure to distinguish between close spaced
targets. In the Fig. 7 a–c, a situation when two vessels are passing by each other is presented. Furthermore, this is all happening at the end of HFSWR 1 coverage area, which makes for very imprecise detection, creating an unstable track and thus putting the DA procedure to a more demanding test. As it can be seen, despite unfavorable circumstances presented DA procedure is performing admirably, assigning detections to the right tracks and thus maintaining them flawlessly. Taking this into account it is clear that simple single-track PDA procedures may cause significant errors, hence more complex JPDA algorithm is used.

B. STATISTICAL ANALYSIS

To perform statistical analyses an additional block is added to the DA procedure. That block collects statistical data regarding cluster information and the time of execution during every individual data scan. The data of interest is:

- \(N_{tc}\) – number of tracks in cluster,
- \(N_{mc}\) – number of measurements received during last update in a cluster,
- \(N_{ac}\) – number of measurement-to-track associations in cluster
- \(C\) – number of joint assignments and
- \(t_{exec}\) – time needed to execute all needed operations in cluster.

The experiment is performed on an Intel Quad Core i5-7200 @2.5 GHz CPU machine, assigning each HFSWR tracking process to one CPU core. The program is written in C# programming language without the use of parallelism features and Single Instruction Multiple Data concepts for internal algorithm calculations. In order to perform all data processing in a real-time (i.e. during one update period – 33 s), the whole tracking procedure (DA included) needs to be finished in no more than 15% of the total update period (i.e. 5 s), to allow other processing stages (i.e. sensor fusion) to finish its work in a timely manner.

Taking into account that the average number of clusters that need to be processed is around 50, the average time for a single cluster processing should not exceed 100 ms.

On our experimental PC, using an optimal data assignment procedure for a given time and following the recommendations from previous sections, the total number of joint assignments \(C_{\text{max}}\) for an arbitrary cluster size in the worst case scenario must be less than \(2^{16}\) (represented with the red dashed line in Fig. 8).

Fig. 8 shows the relationship between the number of joint assignments \(C\) for clusters of various cluster sizes \((N_{tc}, N_{mc})\) and the given threshold \(C_{\text{max}}\). TABLE 2 shows the cluster sizes on which JPDA algorithm may be directly applied. Note that any combination of \((N_{tc}, N_{mc})\) and vice-versa from TABLE 2. can be addressed as the boundary size for the cluster.

It may seem that the margin for a cluster containing 6 tracks is a bit too strict (7), since that cluster can have 37,633 joint assignments. However, if one of the boundary values is increased only by one, the number of joint assignments rises to 130,922 for a 7 x 7 cluster or 93,286 for 8 x 6 cluster, which is well beyond the boundary value (65,536).

The collected results for both HFSWRs are firstly grouped by the \(C_{\text{max}}\) threshold criteria. Then, those clusters that fulfill these criteria are classified by the number of tracks in clusters. Fig. 9 depicts this distribution, while the last record (> \(C_{\text{max}}\)) shows the percentage of clusters where the number of joint associations is greater than the boundary value, forcing the usage of a suboptimal DA method in order to keep the total tracking time under the boundary value (5 s). It can be seen that one track - multiple measurement clusters dominate in calculations, with 70.33 % for HFSWR 0 and 74.46 % for HFSWR 1 of all clusters. Clusters that fulfill conditions for direct application of JPDA make 98.04 % and 98.98 % of all clusters. There is a small percentage of clusters that contain an unusually large number of objects and it represents in most cases the result of ionospheric effects. Furthermore, sporadic regional clutter outbursts are quite frequent in the HFSWR environment, adding to the formation of large clusters and making additional difficulties for tracking procedures. These
effects are the usual reason for the formation of large clusters with a very high number of track-measurement associations that will be solved in a sub-optimal way. On the other hand, large clusters can also be created in areas close to large harbors or oil platforms, since there will be a large concentration of vessels in relatively small area. In order to quantify this influence, an additional experiment is performed, taking into account the DA operations with confirmed tracks. For every confirmed track during its life, its associated cluster size is recorded in every step. In total, there were 25,991 and 32,826 confirmed tracks per HFSWR in a sample set. Note that the term “confirmed track” may either be a false or true target, as defined in [21]. Fig. 10 depicts the cluster size distribution for confirmed tracks, through which all confirmed tracks were updated. On first glance it can be noted that for both HFSWRs the total number of clusters with a non-optimal DA is slightly increased, but this analyses simply counts updates through clusters, which indicates that there are multiple confirmed tracks that were updated through the same clusters with an applied non-optimal DA procedure. In some cases, this situation could indicate the appearance of false targets in some areas. However, total number of optimal
assignments to confirmed tracks is very high: confirmed tracks are updated in an optimal manner within its clusters in 97.37\% and 94.68\% of cases, respectively.

Now, let’s consider the time needed to execute DA procedures. If even one cluster whose number of joint assignments exceeds the boundary value, the total processing time needed for optimal association will exceed boundary value (5s). For example, a cluster with 8 tracks and 9 measurements in the worst case has over 4.5 million of joint assignments and requires nearly 13s to be processed with the JPDA procedure. It is clear that this cluster alone can’t be processed in a timely manner. Not to mention that there are other clusters as well, whose number of joint associations can be under or above the boundary value during the same update period.

However, as it can be seen from Fig. 9 and Fig. 10 that the great majority of clusters (over 98\% in the worst case) may be processed in a timely manner with an optimal DA procedure, while a small number (less than 2\%) requires the application of a sub–optimal DA procedure. The sub-optimal procedure used for processing of 2\% of clusters which can’t be processed in real – time with the optimal DA was Roecker’s algorithm. Roecker’s algorithm itself poses no execution problem for large clusters. On the given experimental sample, it was noticed that the average execution time for the
whole tracking procedures is less than 300 ms, with outliers not exceeding 1.2 s, which is well below the given real-time boundary of 15% of update cycle.

At the end, it should be mentioned that utilizing CPUs with a higher base clock, while still performing single core data processing, will bring little to no benefit to the overall processing time. For example, the Intel Core i7-7740X has nearly twofold better single core performances than the tested CPU. However, utilizing it, even if it speeds up the processing time twofold, will still not bring real-time performances for all clusters (just look at the example above), but would significantly increase the price of a system. On the other hand, the presented clustering-based approach possess a high degree of parallelism, in terms of solving particular sets of clusters at the same time on different CPU threads. By utilizing a proper execution plan (division of jobs), higher number of available CPU cores and implementation in more optimal programming languages, such as C++ with Single Instruction Multiple Data support, the presented DA procedure may speed up even further.

Experiment results showed that maritime environment MTT problem is ideal for deploying existing clustering solutions, due to the high dispersion of measurements, due to the high dispersion of measurements. Particular clusters could be solved in majority of cases with known optimal joint procedures, like JPDA, whose direct usage is otherwise impractical for current stage of computational equipment in the presence of large number of targets and high clutter rates.

IV. CONCLUSION
This paper presents the sub-optimal tracking procedure for maritime surveillance at OTH distances based around
HFSWRs. It uses a density-based clustering approach to overcome real-time execution problems with optimal JPDA data association in a heavily cluttered environment, rich with real targets. Tactical aspects of maritime surveillance at OTH distances, like dispersed target appearance, help mutual isolation of clusters thus allowing usage of optimal cluster-based DA procedures on majority of clusters in order to move entire DA procedure close to the optimal one. Results from experiments showed that it is possible to achieve the majority of optimal track-to-measurement assignments in a data association process with partial use of JPDA on a majority of clusters (over 98 %) and to easily achieve real-time performance with the help of faster sub-optimal Roecker’s algorithm in less than 2 % of all cluster cases. Moreover, unlike the standard JPDA procedure which tends to be inapplicable for real time tracking in a heavily cluttered environment, the density based clustering DA procedure presented here provides real time performances in the very same environment. For the future research, multithreaded implementation in C++ with SIMD support is planned in order to move the margin for real-time optimal DA even closer to 100 % and further reduce latency caused by tracking procedure execution. Additionally, during the future research, implementation of simple single-track PDA procedure will be considered in areas where track density is low in order to further reduce complexity and increase processing speed.

APPENDIX I

In this appendix, more details about selected gating threshold is given. Comparison between selected gating threshold and the one more commonly used in the literature ($\gamma = 9.85$) is provided. The use of large gating threshold may seem at first glance as unreasonably high, however, examples given in Section II indicate largely expanded gate volumes. It should be noted that these values are calculated using extreme error conditions, which should not be considered as the usual situation. Other authors, when considering similar problems, often use the gating threshold values that provide gate validation probabilities in the range 0.99-0.999 [32]. In order to confirm the previous experiment results, a comparison of usual and used gating thresholds seems necessary since gate volume expansion in our case might, quite justifiably, lead to the conclusion that the distribution of the cluster sizes is more in favor of complex cluster structures. An additional experiment is made with the adopted gating threshold $\gamma = 9.85$, which yields a gating validation probability of $P_G = 0.9916$. The results are compared for both HFSWRs with the previous case of the expanded gate threshold $\gamma = 61$ and presented on Fig. 11 and Fig. 12. These comparisons show that in both gate threshold cases the cluster size distribution experiences minor changes: number of clusters with 1 associated track raises for 1% or 2%, while the number of more complex clusters slightly drops. As it can be seen, the cluster structure is very similar to the case with the much larger threshold used in Section II. According to the results, it seems that higher values for the gating threshold should be used since better isolation of the clusters is achieved in this way, while the percentage of the complex clusters is almost identical as in the case of gating threshold of $\gamma = 9.85$.

APPENDIX II

In this Appendix, the detection accuracy of HFSWR is observed. Accuracy of the tracking procedure can be estimated by comparing it with AIS data, used as a ground truth. Fusion and integration algorithms, presented in [31] and [33], are deployed to complete this task. Every fused track contains the underlying HFSWR track information. Once AIS integration is achieved, instead of AIS-fused track data pairs, AIS data with HFSWR track data is paired. AIS position parameters are taken as accurate. Its velocity data, reported position and known position of HFSWRs are used to calculate range, bearing and radial velocity components of the AIS target, to be directly compared with HFSWR track data output.

Not all AIS-track point data pairs are suitable for accuracy evaluations. The first problem is the timing of the data. The maximal time difference between AIS and track data pairs is limited to 5 seconds, to avoid the extrapolation of the AIS data in time for longer periods. It is considered safe to assume that such time differences shouldn’t cause any
significant errors due to the low maneuverability and speed of the vessels. The second problem are integrations with lower probabilities. When AIS-track integration probability is dropping towards the integration threshold, which is set to 51%, as described in [31], it becomes more likely that the AIS and HFSWR tracks are to be disintegrated, having significantly different moving parameters and a constantly increasing mutual distance. To avoid these possible miss-integration situations, only the AIS-track pairs that have a reported integration probability larger than 55% are used for evaluation.

Over a 15000 AIS - HFSWR track pairs are used for HFSWR track accuracy analyses. Parameter error distributions are presented in Fig. 13-16. In total, 16424 AIS-track pairs have been recorded for HFSWR 0 and 19100 for HFSWR 1. There were 90 different AIS MMSI integrated, followed by 134 and 150 different confirmed tracks from HFSWR0 and HFSWR 1, respectively. Error parameters are summarized in TABLE 3. It can be noted that range and bearing errors are well inside the resolution cells of the sensor. Distance errors, on average, are less than 1 km, which indicates precision improvement as a consequence of the applied tracking procedure.

ACKNOWLEDGMENT
The authors would like to thank Mr. Stefan Mrkic for English language editing and anonymous reviewers for the constructive comments.

REFERENCES
[1] Law of the Sea, Part V—Exclusive Economic Zone, United Nations, New York, NY, USA, Aug. 2011.
[2] G. Fabrizio, High Frequency Over-the-Horizon Radar: Fundamental Principles, Signal Processing, and Practical Applications. New York, NY, USA: McGraw-Hill, 2013.
Multitarget-Multisensor Tracking: Applications and Advances

[1] L. Sevgi, A. Ponsford, and H. C. Chan, “An integrated maritime surveillance system based on high-frequency surface-wave radars. 1. Theoretical background and numerical simulations,” IEEE Antennas Propag. Mag., vol. 43, no. 4, pp. 28–43, Aug. 2001.

[2] A. Ponsford, L. Sevgi, and H. C. Chan, “An integrated maritime surveillance system based on high-frequency surface-wave radars. 2. Operational status and system performance,” IEEE Antennas Propag. Mag., vol. 43, no. 5, pp. 52–63, Oct. 2001.

[3] S. Blackman, and R. Popoli, Design and Analysis of Modern Tracking Systems. Norwood, MA, USA: Artech House, 1999.

[4] Y. Bar-Shalom, and T. E. Fortmann, Tracking and Data Association. Orlando, FL, USA: Academic, 1988.

[5] D. Musicki and R. Evans, “Joint integrated probabilistic data association: JIPDA,” in Proc. FUSION Conf., Annapolis, MD, USA, vol. 2, Jul. 2002, pp. 1120–1125.

[6] D. Musicki and B. La Scala, “Multi-target tracking in clutter without measurement assignment,” IEEE Trans. Aerosp. Electron. Syst., vol. 44, no. 3, pp. 877–896, Jul. 2008.

[7] D. Musicki and R. J. Evans, “Multiscan multitarget tracking in clutter with integrated track splitting filter,” IEEE Trans. Aerosp. Electron. Syst., vol. 45, no. 4, pp. 1432–1447, Oct. 2009.

[8] S. H. Rezatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, and I. Reid, “Joint probabilistic data association revisited,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 3047–3055.

[9] B. Zhou and N. K. Bose, “Multitarget tracking in clutter: Fast algorithms for data association,” IEEE Trans. Aerosp. Electron. Syst., vol. 29, no. 2, pp. 352–363, Apr. 1993.

[10] J. A. Roecker and G. L. Phillis, “Suboptimal joint probabilistic data association,” IEEE Trans. Aerosp. Electron. Syst., vol. 29, no. 2, pp. 510–517, Apr. 1993.

[11] E. Aoki and K. Kienitz, “Suboptimal JPDA for tracking in the presence of clutter and missed detections,” in Proc. 12th IEEE Int. Conf. Inf. Fusion, Seattle, WA, USA, Jul. 2009, pp. 818–825.

[12] Charu C. Aggarwal and Chandan K. Reddy, Data Clustering: Algorithms and Applications (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series). Boca Raton, FL, USA:, Chapman & Hall, 2014.

[13] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density based algorithm for discovering clusters in large spatial databases with noise,” in Proc. AAAI, 1996, pp. 226–231.

[14] N. Stojiljkovic, D. Nikolic, P. Petrovic, N. Tocic, I. Gluvacevic, N. Stojiljkovic, and N. Lekic, “An implementation of DBF and CFAR models in OTHR signal processing,” in Proc. IEEE 15th Int. Conf. Logol. Signal Process. Appl. (CSPA), Penang, Malaysian, Mar. 2019, pp. 7–11.

[15] P. Petrovic, N. Gricic, N. Stojikovic, D. Nikolic, and N. Lekic, “Implementation of algorithm for excision of point targets from distributed radar detections,” in Proc. 6th iETRAN, Srebrno Jezero, Serbia, Jun. 2019, pp. 1000–1004.

[16] N. M. Tosic, A. Samcovic, D. Nikolic, D. Drajic, and N. Lekic, “An algorithm for detection of electromagnetic interference in high frequency radar range-Doppler images caused by LEDs,” IEEE Access, vol. 7, pp. 84413–84419, Jun. 2019.

[17] A. Dzvonkovskaya, D. Nikolic, V. Orlic, M. Peric, and N. Tocic, “Remote observation of a small meteotsunami in the bight of Benin using HF radar operating in lower HF band,” IEEE Access, vol. 7, pp. 88601–88608, Jul. 2019.

[18] K.-W. Gurgel, Y. Barbin, and T. Schlick, “Radio frequency interference suppression techniques in FMCW modulated HF radars,” in Proc. OCEANS Eur., Jun. 2007, pp. 1–4.

[19] T. Hanselmann, D. Musicki, and M. Panaliswami, “Adaptive target tracking in slowly changing clutter,” in Proc. 9th Int. Conf. Inf. Fusion, Florence, Italy, Jul. 2006, pp. 1–8.

[20] Y. Bar-Shalom and X. R. Li, Multitarget-Multisensor Tracking: Applications and Advances. Bradford, UK.: YBS Publishing, 1995.

[21] S. Zolo and B. Ristic. On the Choice of the Coordinate System and Tracking Filter for the Track-While-Scan Mode of an Airborne Pulse Doppler Radar. DSTO Research Laboratory, Salisbury, SA, Australia. Accessed: Oct. 20, 2019. [Online]. Available: https://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA375838

[22] N. Stojikovic, D. Nikolic, B. Dzolnic, N. Toci, V. Orlic, N. Lekic, and B. M. Todorovic, “An implementation of tracking algorithm for over-the-horizon surface wave radar,” in Proc. 24th Telecommun. Forum (TELFOR), Belgrade, Serbia, Nov. 2016, pp. 1–4.

[23] J. K. Uhlmann, “Introduction to the algorithms of data association in multiple-target tracking,” in Handbook of Multisensor Data Fusion. Boca Raton, FL, USA: CRC Press, 2017, pp. 89–106.

[24] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, “Sonar tracking of multiple targets using joint probabilistic data association,” IEEE J. Ocean. Eng., vol. 8, no. 3, pp. 173–184, Jul. 1983.

[25] J. Williams and R. Lau, “Approximate evaluation of marginal association probabilities with belief propagation,” IEEE Trans. Aerosp. Electron. Syst., vol. 50, no. 4, pp. 2942–2959, Oct. 2014.

[26] R. Helmick, “IMM estimator with nearest-neighbor joint probabilistic data association,” Multitarget-Multisensor Tracking: Applications and Advances vol. 3, Y. Bar-Shalom, and W. D. Blair, Eds. Norwood, MA, USA: Artech House, 2000, ch. 3.

[27] B. J. Van Wyk, M. A. Van Wyk, and G. Noel, “A projection-based joint probabilistic data association algorithm,” in Proc. IEEE AFRICON, Gaborone, Botswana, Sep. 2004, pp. 13–17.

[28] D. Nikolic, N. Stojikovic, and N. Lekic, “Maritime over the horizon sensor integration: High frequency Surface-Wave-Radar and automatic identification system data integration algorithm,” Sensors, vol. 18, no. 4, p. 1147, Apr. 2018, doi: 10.3390/s18041147.

[29] P. Braca, R. Grasso, M. Vespe, S. Maresca, and H. Horstmann, “Application of the JPDA-UKF to HFSW radars for maritime situational awareness,” in Proc. 15th Int. Conf. Inf. Fusion (FUSION), pp. 2585–2592, Jan. 2012.

[30] D. Nikolic, N. Stojikovic, Z. Popovic, N. Tocic, N. Lekic, Z. Stankovic, and N. Doncov, “Maritime over the horizon sensor integration: HFSWR data fusion algorithm,” Remote Sens., vol. 11, no. 7, p. 852, Apr. 2019, doi: 10.3390/rs11070852.

NIKOLA STOJKOVIC (Member, IEEE) was born in Belgrade, Serbia, in 1978. He received the Dipl.Ing. degree in electronics from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2005, where he is currently pursuing the Ph.D. degree in electrical engineering. He is currently a Senior Software Engineer at the Vlatacom Institute of High Technologies, Belgrade. His main interests are telecommunication protocols, object tracking, radar data processing, sensor fusion, algorithm optimization, software-defined radio, and FPGA design. He has authored or coauthored more than 15 papers in referenced journals, international conferences, and technical solutions.

DEJAN NIKOLIC (Member, IEEE) received the B.S. degree in electrical engineering from the University of Defense, Serbia, in 2005, and the M.S. degree in electrical engineering from the University of Belgrade, Serbia, in 2013. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Nis, Nis, Serbia. Since 2015, he has been a Researcher with Vlatacom Institute, Belgrade. His research interests include the development of high frequency surface wave radars with the emphasis on constant false alarm rate algorithms and performance modeling as well as maritime multisensor data fusion algorithms. He has authored or coauthored more than 20 papers in international journals and conferences as well as several technical solutions.

SNEZANA PUZOVIC (Member, IEEE) received the Dipl.Ing. degree in telecommunications from the University of Belgrade, Belgrade, Serbia in 2006, and the Ph.D. degree in telecommunications from the University of Siena, Siena, Italy, in 2011. She is currently with Vlatacom Institute, Belgrade. Her research interests include performance modeling for radar systems, resource management in broadband wireless and heterogeneous networks, quality of service, and cross-layer optimizations.