

Passive RADAR : Deinterleaving and Clustering unknown RADAR pulses

Manon Mottier
*Université Paris-Saclay, CNRS,
CentraleSupélec*
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
manon.mottier@centralesupelec.fr

Gilles Chardon
*Université Paris-Saclay, CNRS,
CentraleSupélec*
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
gilles.chardon@centralesupelec.fr

Frédéric Pascal
*Université Paris-Saclay, CNRS,
CentraleSupélec*
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
frederic.pascal@centralesupelec.fr

Abstract—In this paper, a two-step methodology is developed to deinterlace RADAR signals. We mainly worked from a data simulator to obtain a large diversity of signals representing typical RADARs. First, a clustering algorithm is used to separate the pulses in the F-PW (Frequency-Pulse Width) plane, then a phase of cluster agglomeration is performed using hierarchical agglomerative clustering combined with optimal transport distances. Results on labeled simulated data are given.

Index Terms—Passive RADAR, clustering, optimal transport

I. INTRODUCTION

The technological evolution of the last few years has allowed many innovations in the electronic field and more particularly in the field of electronic warfare. Electronic equipment is becoming increasingly complex and sophisticated, leading to a transformation of the RADAR signal processing chain (improvement of receivers and emitters, spreading of emissions over the spectrum, waveform modification, autonomous weapon system etc). All these changes represent a continuous challenge in order to propose new and more accurate techniques to classify and identify emitters in a signal.

Artificial intelligence now enables machines to learn from data and adapt very quickly their strategy. Identification algorithms are increasingly fast and efficient without the need for human interaction. They are able to adapt and evolve in the face of new situations. Over the next few years, the modernization of identification techniques will represent a major challenge for electronic warfare: they will give a considerable advantage to those actors capable of mastering them, enabling them to gain power and gain an advantage over the enemy. The amount of information to be processed increases exponentially with the input of a growing number of different sensors and by the improvement of their interception sensitivity. Only a complex decision system can highlight the essential information to build a strategic picture of the situation.

This is especially true for passive RADAR, although the ability to detect and intercept signals is greater, it is also more difficult to deinterleave the signals and identify the different transmitters present. The intercepted signals are increasingly large and contain a large amount of information to be processed; it is necessary to continuously develop new

algorithms capable of separating the pulses of a signal, to classify them correctly in order to identify the presence of emitters in them. RADARs now have agility characteristics on their primary parameters that make identification more complex.

Early conventional deinterleaving work mainly used DOA (Direction-Of-Arrival), F (central Frequency) and TOA (Time-Of-Arrival) for deinterleaving. The DOA and F were used as a first step to filter data and then they used the TOA difference histogram to identify pattern of pulse interval [1]. Many authors have contributed by proposing improvements to existing algorithms. For example, by using the cumulative difference histogram [2] or the sequential difference histogram [3]. Most of these methods are based on the search for PRI (Pulse Repetition Interval) in the signals and the use of several variants of TOA histograms.

Very quickly, new methods were used to better integrate the missing pulses and be more robust to noise. Many recent research papers use deep learning to separate the pulses of a signal and identify the transmitters present (see e.g., [4]). All these new methods require the setting of a very large number of parameters. They often do not take into account the nature of the data and are not necessarily easily replicated. In the RADAR environment, it is not easy to obtain large datasets and deep learning does not easily process signals with few pulses.

More recent work has begun to use a mixture of supervised and unsupervised methods for signal identification [5]. Supervised methods are applied to labeled data and then a hybrid classification model built from several algorithms is used to improve classification accuracy and have more robustness.

Data are truncated or partially observed making pattern identification difficult. GMMs (Gaussian Mixture Models) are increasingly used to classify and cluster signals because they allow latent variables to be introduced; Latent variables allow to manage this type of data and to take into account missing data [6].

Comparisons have shown that deep learning models are not necessarily better than more conventional and simpler models such as GMMs [7]. They have shown, for

example, that when the number of transmitters present in the signal was below a certain threshold, GMMs performed better.

As an extension of these works, this article presents a new method to deinterleave a signal with HDBSCAN algorithm [8] and hierarchical agglomerative clustering with optimal distances [9]. It describes an unsupervised methodology to separate the RADAR pulses present in a signal and classify them in order to determine the number of transmitters present.

The paper is organized as follows. The methodology for the construction and acquisition of the database is developed in a first section. This section describes the process of data creation and acquisition. The classified nature of the data required us to work and validate with simulated data and then to apply our methods on real data. Then, the second section explains more precisely how the deinterleaving strategy works. More precisely, it first describes how the HDBSCAN clustering algorithm works and how we used it. Then, it describes how we combined optimal transport with hierarchical agglomerative clustering to group clusters. We combined supervised and unsupervised learning to build our method. The algorithms are developed in an unsupervised learning and applied on real data and validation was performed using the simulated labeled data. Finally, the last section contains present empirical results that validate and evaluate the proposed model in a real application context.

II. DATA DESCRIPTION

This section presents the datasets used to test our deinterleaving methods and the experimental procedure to obtain representative unclassified data. Given the difficulty in obtaining real-unclassified from legal instances, labeled and large volume data sets, we used simulated data. These simulated data come from a simulator built and certified by experts in the field. It allowed us to acquire a significant volume of signals, a necessary condition to obtain empirical results.

This simulator allowed us to have access to a large variety of signals from existing RADARs. They were simulated from a RADAR database containing the characteristics of 60 typical RADAR waveforms. This diversity results, in particular, in the acquisition of labeled mono or multi-sensor signals, signals with frequency and/or time modulation, measurement errors, outlier, missing data or the possibility to define the noise level of the signal.

Subsequently, all the methods were tested on real and labeled data sets in order to validate the methods and the results obtained from the simulated data. Our database contains a lot of signals each containing between 1 and 6 transmitters. The signals vary in size; some containing a very large number of pulses while others are much shorter. Receivers are assumed to be static and had a detection threshold around -60 dB, but we do not have any other

information on the characteristics of the transmitters and receivers. The simulated data are anonymized, but labeled. Therefore it is possible to compute robust classification statistics. Noise has been added to make the data more realistic and the size of the signals is heterogeneous. Receivers are omnidirectional, making the DOA unusable.

Pulses acquired by the passive RADAR system are segmented and described by a limited set of features. In this work, the N acquired pulses are described using the following four features (t_n, w_n, f_n, g_n) , for $1 \leq n \leq N$:

- t_n : time of arrival of the pulse (TOA, unit: seconds),
- w_n : pulse width (PW, seconds),
- f_n : central frequency (F, Hz),
- g_n : power (G, dB).

A fifth feature, difference of time of arrival (DTOA), can be defined by $\delta_n = t_n - t_{n-1}$.

Fig. 1 shows an example of simulated data. DTOA, PW, F, and G are plotted in the TOA plane. This signal represents the pulses from 4 different transmitters (each color represents a RADAR). Transmitter 3 emits only once during the observed period. The lobes corresponding to different RADAR can be mixed like transmitters 1 and 2. We can see an interlaced signal representing 4 transmitters.

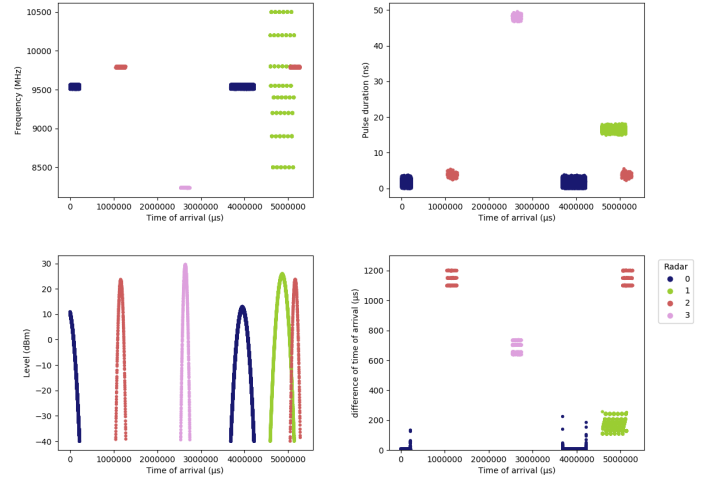


Fig. 1. Representation of a signal with 4 emitters. Each emitters is represented by a different color.

Fig. 2 shows a pulse spread diagram in the F-PW plane with several clusters, each corresponding to a given F-PW pair of each RADAR. We note that a RADAR can be represented by multiple clusters, and that a particular cluster can be shared by multiple RADARs. In this figure, one can see the distinct appearance at several places in the F-PW plane of RADAR 1. The transmitters 0, 2 and 3 will only appear in one place in the signal and are perfectly identifiable in the F-PW plane.

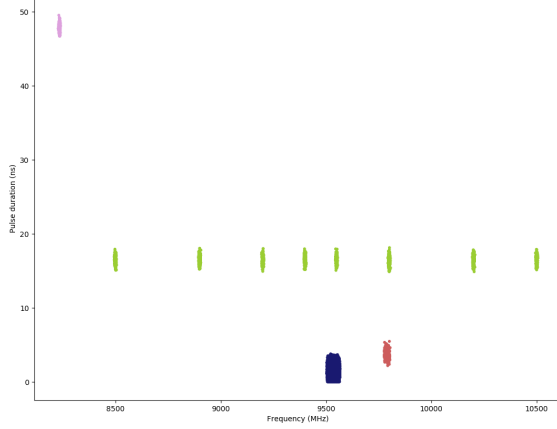


Fig. 2. Zoom on the F-PW plane.

III. CLUSTERING ALGORITHM

We will now assume that an F-PW cluster is associated to a unique RADAR system. Our goal will be to group F-PW clusters generated by the same RADAR. The clustering algorithm consists of three steps:

- pre-processing: data normalization (i.e. non-dimensionalization) in the F-PW plane;
- clustering in the F-PW plane;
- grouping of the F-PW clusters using temporal characteristics.

A. Pre-processing

The physical dimension of the frequency (Hz) and pulse width data (s) are inconsistent. Non-dimensionalization is necessary in order to compute distances between features.

Two normalization methods are tested:

- Quantile normalization: use of inter-quantile intervals to scale the data.
- GMM pre-clustering: estimation of the intra-cluster variances with respect to the two features, based on GMM clustering with a overestimated number of clusters.

We finally chose to use quantile normalization since it preserves the data distribution.

B. Clustering in the F-PW plane

In this section, we present the first step of the clustering process introduced in this paper, based on the HDBSCAN algorithm [8]. HDBSCAN is a hierarchical version of the so-called DBSCAN algorithm [10]. DBSCAN is a clustering algorithm, allowing for the presence of outliers. The main idea is to regroup points that live in a same dense area. More precisely, the algorithm defines an ϵ -neighbourhood for each vector $\mathbf{x}_i$ to be clustered, as follows $\mathcal{N}_\epsilon(\mathbf{x}_i) = \{\mathbf{x}_k \in \mathcal{X} | d(\mathbf{x}_k, \mathbf{x}_i) \leq \epsilon\}$, for a given distance $d(\cdot, \cdot)$. Then, depending on the size of $\mathcal{N}_\epsilon(\mathbf{x}_i)$, namely $\#\mathcal{N}_\epsilon(\mathbf{x}_i) \geq \text{MinPts}$, $\mathbf{x}_i$ is either

clustered or labelled as an outlier. The DBSCAN algorithm is recalled in Algorithm 1.

Algorithm 1 DBSCAN algorithm

Input: $\mathcal{X}$ (set of the vectors to be clustered), ϵ , MinPts

Output: Z (labels of vectors of $\mathcal{X}$)

For all $\mathbf{x}_i$ of $\mathcal{X}$

- Verify that $\mathbf{x}_i$ has not been visited by the algorithm, else $\mathbf{x}_i$ is marked as “visited”
 - Identify $\mathcal{N}_\epsilon(\mathbf{x}_i)$, an ϵ -neighbourhood of $\mathbf{x}_i$.
 - **If** $\#\mathcal{N}_\epsilon(\mathbf{x}_i) \leq \text{MinPts}$ label $\mathbf{x}_i$ as an isolated point
Else create a class C containing $\mathbf{x}_i$ and apply *class_extension*($C, \mathbf{x}_i, \epsilon, \text{MinPts}$) (Algorithm 2)
-

Algorithm 2 “Class_extension”.

Input: Class C to be extended, vectors $\mathbf{x}_i \in C$, MinPts, ϵ .

Output: Z labels of vectors in $\mathcal{N}_\epsilon(\mathbf{x}_i)$

For all $\mathbf{x}_j$ in $\mathcal{N}_\epsilon(\mathbf{x}_i)$

- Verify that $\mathbf{x}_j$ has not been visited by the algorithm, else $\mathbf{x}_j$ is marked as “visited”
 - Identify $\mathcal{N}_\epsilon(\mathbf{x}_j)$, an ϵ -neighbourhood of $\mathbf{x}_j$.
 - **If** $\#\mathcal{N}_\epsilon(\mathbf{x}_j) \geq \text{MinPts}$
 $\mathcal{N}_\epsilon(\mathbf{x}_i) = \mathcal{N}_\epsilon(\mathbf{x}_i) \cup \mathcal{N}_\epsilon(\mathbf{x}_j)$
 - **If** $\mathbf{x}_j$ is not already classified, add $\mathbf{x}_j$ to C .
-

Of course, for DBSCAN the two key parameters MinPts and ϵ have to properly be estimated to ensure good clustering performance. To handle this problem, HDBSCAN has been recently introduced. This algorithm relies on a hierarchical approach allowing to omit the crucial ϵ parameter, by providing dendrograms for all DBSCAN clustering solutions. Then, the “best” value of ϵ is chosen thanks to an optimization over the trees (details on the HDBSCAN algorithm can be found at <https://hdbscan.readthedocs.io/>).

The HDBSCAN algorithm has been set up to overestimate the number of clusters found for possibly capturing RADARs with few pulses or emitting very little over time. Underestimating the number of clusters could lead to grouping the pulses from several different RADARs.

C. Hierarchical clustering with optimal transport distance

RADAR systems that use several frequencies and/or pulse widths will not be described by a unique cluster in the F-PW plane. A final hierarchical clustering is used to agglomerate clusters belonging to a RADAR system [9].

This step is based on the assumption that clusters from a RADAR will be active at the same time periods, i.e. the time intervals where the RADAR is pointed towards the receiver. This is formalized using optimal transport distances, defined between probability measures.

In the case of two discrete probability measures $\nu = \sum_{n=1}^N a_n \delta_{x_n}$ and $\mu = \sum_{m=1}^M b_m \delta_{y_m}$ defined on a set X ,

with $\mathbf{a} = (a_1, \dots, a_N) \in \mathbf{R}_+^N$, $\mathbf{b} = (b_1, \dots, b_M) \in \mathbf{R}_+^M$, and $x_n, y_m \in X$, a transport plan from ν to μ is described by a $N \times M$ matrix $\mathbf{P}$, with coefficients P_{nm} describing the quantity of mass taken from x_n to y_m . Consistency with the weights $\mathbf{a}$ and $\mathbf{b}$ of the measures μ and ν is enforced by the linear conditions

$$\mathbf{P}\mathbf{1}_M = \mathbf{a}, \mathbf{P}^T\mathbf{1}_N = \mathbf{b}^T. \quad (1)$$

where $\mathbf{1}_M$ and $\mathbf{1}_N$ are vectors of ones with appropriate dimensions, and $\cdot^T$ denotes transposition.

The total cost $C(\mathbf{P})$ of a transport plan is assumed to be linear, i.e.

$$C(\mathbf{P}) = \sum_{n=1}^N \sum_{m=1}^M C_{nm} P_{nm} = \langle \mathbf{C}, \mathbf{P} \rangle \quad (2)$$

The coefficients $C_{nm} = c(x_n, y_m)$ of the $N \times M$ cost matrix $\mathbf{C}$ model the cost of transporting a unit of mass from x_n to y_m , for a given cost function $c(\cdot, \cdot)$.

The optimal transport plan $\mathbf{P}^*$ is the solution of the following linear optimization problem:

$$\mathbf{P}^* = \underset{\mathbf{P} \in \mathbf{R}_+^{N \times M}}{\operatorname{argmin}} \langle \mathbf{C}, \mathbf{P} \rangle \text{ such that } \mathbf{P}\mathbf{1}_M = \mathbf{a}, \mathbf{P}^T\mathbf{1}_N = \mathbf{b}^T. \quad (3)$$

Optimal transport distance between two clusters is computed as follows: a probability measure

$$\mu_j = \frac{1}{N_j} \sum_{k \in C_j} \delta_{t_k} \quad (4)$$

is associated to a cluster C_j , and the unit transport cost between two times of arrival τ_1 and τ_2 is taken as

$$c(\tau_1, \tau_2) = |\tau_1 - \tau_2|. \quad (5)$$

In most cases, transport of mass between two clusters from the same RADAR will occur in the short time interval of a RADAR pass, while mass of two clusters of two different RADAR will be transported over larger time differences. This method has the advantage of not making assumptions about the regularity of the pulse repetition period.

Clusters are fused in a hierarchical way, by iteratively agglomerate the clusters with smallest optimal transport distance. After fusion the distance between the fused clusters and the other clusters is updated. The process is halted when all the clusters are merged and there is only one cluster left. From these results the dendrogram is built and allows to visualize the groupings made. Several metrics are used to create a decision model to determine where to cut the dendrogram and specify the final grouping.

The decision model was mainly built from 3 metrics:

- Silhouette score [11]: Measurement using the average intra-cluster distance and that of the nearest cluster;
- Davies-Bouldin score [12]: Measures the similarity between clusters;

- Calinski-Harabasz Score [13]: Measures the proportion of inter-group variance to intra-group variance (also called variance ratio criterion).

The exact constraints (1) are often too strict when matching clusters. Indeed, it is frequent that all lobes from a particular RADAR pass do not contain the same proportions of pulses from each sub-cluster. In this case, mass has to be transported from a lobe to another, raising the cost of the optimal transport.

The optimal transport problem can be relaxed by using approximate fit between the actual weights, and the weights used to compute the transport plan. This relaxation yields the following optimization problem

$$\mathbf{P}^* = \underset{\mathbf{P} \in \mathbf{R}_+^{N \times M}}{\operatorname{argmin}} \langle \mathbf{C}, \mathbf{P} \rangle + \lambda_1 D(\mathbf{P}\mathbf{1}_M, \mathbf{a}) + \lambda_2 D(\mathbf{P}^T\mathbf{1}_N, \mathbf{b}) \quad (6)$$

where, as an example, D can be the Kullback-Leibler distance between probability distributions. The parameters λ_1 and λ_2 control the fidelity of the transport plan to the actual weights.

Some signals are very large which makes the optimal transport very long and very expensive in memory. Cluster pulses were not used directly in the optimal transport but histograms were built from the pulses to reduce computation time and make it numerically more stable. The size of the bins has been determined so as not to lose information on the distribution of impulses by clusters.

IV. RESULTS

A. Results

In order to select the most efficient clustering algorithm for deinterleaving a signal, we performed comparative studies between several unsupervised clustering algorithms in particular between DBSCAN, OPTICS [14] and HDBSCAN. We applied each algorithm on several labeled simulated data and compared the clustering quality. In the majority of cases the HDBSCAN algorithm obtained the best results.

Our methodology does not work when the signal groups together RADAR pulses with similar characteristics. For instance, when 2 RADARs are identical in frequency and pulse duration, HDBSCAN is not able to separate the pulses from the transmitters and considers it as a single RADAR. Some types of signals are complicated to analyze, such as signals that pick up pulses from fast-scanning RADARs: their frequencies may hide those of other transmitters.

B. Experimentation

In this section, we present an application case using a signal from our data simulator. We apply the different steps of the proposed methodology. A signal representing the pulses from 4 different emitters has been selected. The signal is not represented in its entirety in the figures. Plots show a zoom of this signal. Note that a relatively clean and low noise signal was deliberately chosen.

Fig. 3 represents the data of this signal. The size of the signal is more 153000 observations and have 4 main parameters. The DTOA has been calculated as the TOA t_n and t_{n-1} . DTOAs cannot be used at this stage since they are not necessarily from the same RADAR. The G-TOA plane shows an overlapping of the lobes.

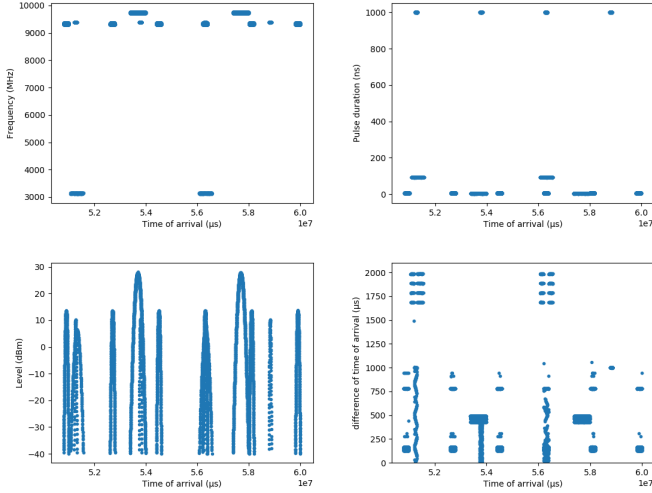


Fig. 3. The figure represents a signal without any indication of the number of transmitters present. The signal is represented using its primary parameters (F, PW, G, DTOA and TOA). The signal has not been represented entirely, a zoom has been applied to facilitate its visualization. The DTOA has been calculated on the whole signal.

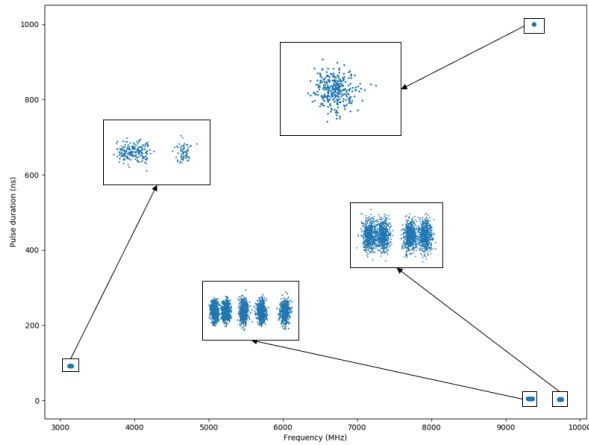


Fig. 4. Zoom on the F-PW plane of the previous signal.

Fig. 4 shows a zoom in the F-PW plane. The impulses are grouped together and form several distinct blocks. An emitter can be represented by several clusters as the pulse packets at the bottom of the figure show it.

The HDBSCAN algorithm was applied in the F-PW plan. Fig. 5 shows us the results of clustering. HDBSCAN identifies 10 clusters from this signal. Clusters have heterogeneous sizes, as some have less than 1000 observations while others have more than 32000 observations. The clustering results are consistent and show that one transmitter is represented by several clusters. We can see on the plot in the G-TOA plane

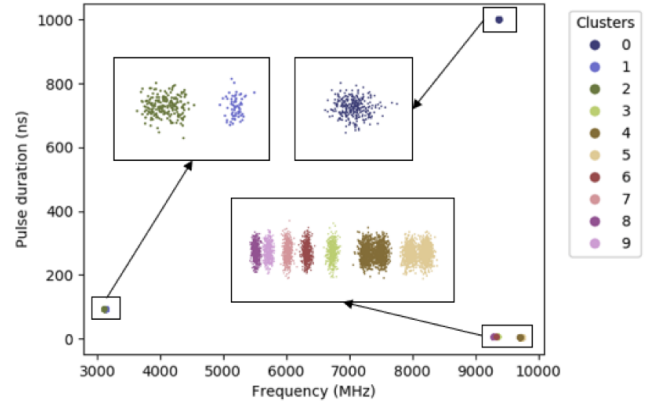
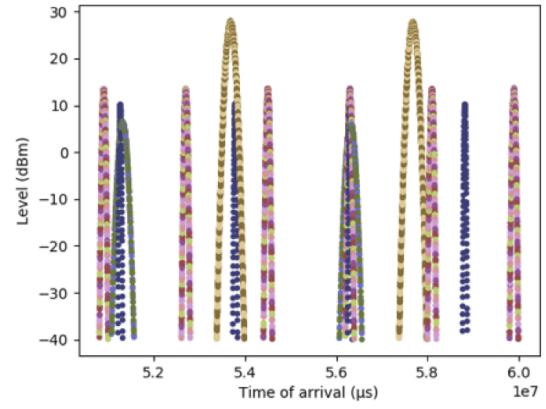


Fig. 5. Results of HDBSCAN clustering in the F-PW plan. A zoom has been applied in the F-PW plane. Colors represent the 10 different clusters returned by the algorithm. Clusters are distributed along the lobes in the G-TOA plane.

that several lobes overlap around 5.6 μ s.

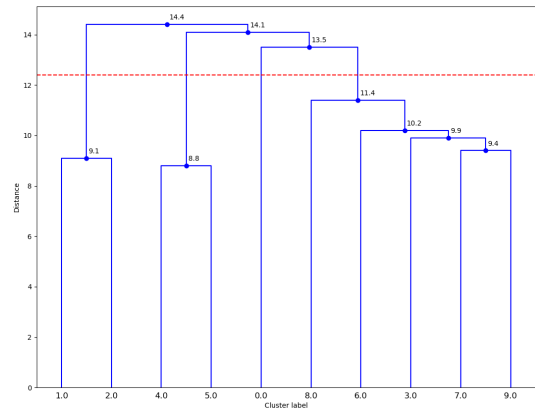


Fig. 6. Hierarchical agglomerative clustering results in logarithmic form with Optimal transport measures. We constructed decision rules from the optimal transport values using several metrics to determine where to cut the dendrogram. The red line indicates where to cut the dendrogram. Here the metrics tell us that our signal contains pulses from 4 different transmitters.

From the HDBSCAN clusters, a hierarchical ascending clustering was built using the optimal transport distances as shown in Fig. 6. At each step, the optimal transport distance between each pairs of clusters is computed. The pair of clusters with the smallest distance is merged into a single cluster. This step is repeated until the clusters are fully aggregated. The dendrogram presents in a simplified way the aggregations of clusters at each step. The results are displayed in a logarithm scale. The dendrogram was cut using decision rules built from several metrics and from knowledge of the RADAR environment. The decisional system indicates that the dendrogram must be cut so that 4 clusters remain. The red line is used to identify clusters.

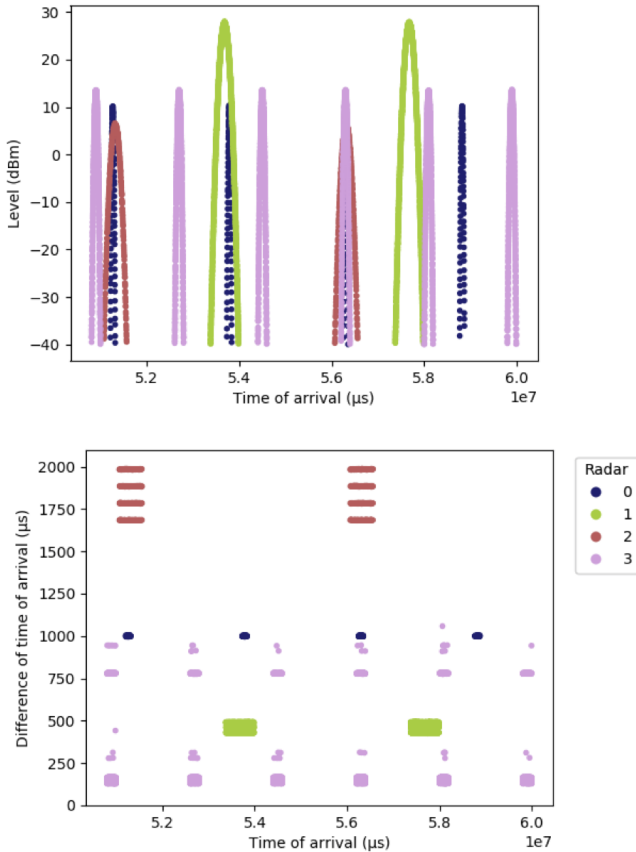


Fig. 7. Final grouping of clusters according to HAC results with POT measurements. We used the dendrogram to group clusters and obtained this representation. The colors represent the final clusters and the 4 emitters present. The DTOA was calculated from the pulses of each emitter.

Finally, Fig. 7 shows the result of cluster agglomeration. Here the clustering has worked perfectly and correctly identifies the 4 transmitters present in the signal. In particular, the lobes present around $5.6 \mu s$ were perfectly reconstructed and the optimal transport enabled the clusters distributed over these lobes to be grouped together correctly. The graph in the DTOA-TOA plane shows the DTOA reconstructed from the pulses of each RADARs.

V. CONCLUSION AND PERSPECTIVES

In this paper, we proposed a methodology to deinterleave RADAR signals from its primary parameters. It is able to identify the number of transmitters present in the signal by deinterleaving the pulses. This method gave very good results on simulated data sets and encouraging results on real data.

In some cases, RADARs can have the same characteristics and be confused in frequency and pulse duration. This is the case, for instance, in ports or airports where several similar models are used. Some clusters do not contain enough pulses to be properly used in optimal transport. The next objectives are to investigate the use of optimal transport to tackle these challenging problems.

VI. ACKNOWLEDGEMENTS

This work was supported by ATOS which provided the data and its RADAR expertise.

REFERENCES

- [1] D. Wilkinson and A. Watson, "Use of metric techniques in esm data processing," in *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 132, no. 4. IET, 1985, pp. 229–232.
- [2] H. Mardia, "New techniques for the deinterleaving of repetitive sequences," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 136, no. 4. IET, 1989, pp. 149–154.
- [3] D. Milojević and B. Popović, "Improved algorithm for the deinterleaving of radar pulses," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, no. 1. IET, 1992, pp. 98–104.
- [4] Z. Zhou, G. Huang, H. Chen, and J. Gao, "Automatic radar waveform recognition based on deep convolutional denoising auto-encoders," *Circuits, Systems, and Signal Processing*, vol. 37, no. 9, pp. 4034–4048, 2018.
- [5] X. Zhang, P. Luo, and X. Hu, "A hybrid method for classification and identification of emitter signals," in *2017 4th International Conference on Systems and Informatics (ICSAI)*. IEEE, 2017, pp. 1060–1065.
- [6] G. Revillon, A. Mohammad-Djafari, and C. Enderli, "Radar emitters classification and clustering with a scale mixture of normal distributions," *IET Radar, Sonar & Navigation*, vol. 13, no. 1, pp. 128–138, 2018.
- [7] X. Gong, H. Meng, and X. Wang, "A gmm-based algorithm for classification of radar emitters," in *2008 9th International Conference on Signal Processing*. IEEE, 2008, pp. 2434–2437.
- [8] R. J. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," in *Pacific-Asia conference on knowledge discovery and data mining*. Springer, 2013, pp. 160–172.
- [9] S. Chakraborty, D. Paul, and S. Das, "Hierarchical clustering with optimal transport," *Statistics & Probability Letters*, p. 108781, 2020.
- [10] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Kdd*, vol. 96, no. 34, 1996, pp. 226–231.
- [11] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.
- [12] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224–227, 1979.
- [13] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
- [14] G. Jajoo, Y. Kumar, S. K. Yadav, B. Adhikari, and A. Kumar, "Blind signal modulation recognition through clustering analysis of constellation signature," *Expert Systems with Applications*, vol. 90, pp. 13–22, 2017.