Team GME Meeting

Manon Mottier

CentraleSupélec - Avantix

April 27th, 2022

Context

Industrial thesis (CIFRE): Artificial Intelligence for Passive RADAR

- Phd student, 3rd year
- The Signals and System Laboratory of CentraleSupélec
  - Thesis director: Frédéric PASCAL
  - Co-supervisor: Gilles CHARDON
- AVANTIX (ATOS Group)

The company

AVANTIX

- French company specialized in Electronic Warfare
- Designs critical high-tech systems to provide homeland protection and military forces
- Provides strategic information to military platforms

The subject

- Use and development of machine learning tools for passive RADAR intelligence
- Improvement of the **RADAR Recognition process** thanks to Optimal Transport
- Passive RADAR(s):
  - Emit no electromagnetic waves
  - Only receivers
  - Detect, track and identify RADAR(s)

RADAR Recognition process

1. **RADAR pulses deinterlacing:**
   - ☛ Separation and grouping of mixed data from a RADAR signal containing data from an unknown number of RADARs
   - ? Methods: Pulse repetition intervals (Mardia - Milojevic), Clustering (Kang, Zhang, Gue and Tian), Deep Learning (Li, Xueqiong and Liu, Zhangmeng and Huang, Zhitao)...
   - ✘ Limits: Supervised methods, only efficient on simple RADAR, need a lot of data...
2. **RADAR Identification**:
   - ☛ Feature extraction and comparison with a reference database
   - ? Methods: Time-frequency analysis (Kawalec and Owczarek), Deep learning (Liu, Zhang-Meng and Philip, S Yu)...
   - ✘ Limits: Supervised methods, few classes, model retrained...

Our contribution: Use of Optimal Transport Theory

1. **RADAR pulses deinterlacing**:
   - **Clustering**: Cluster each observation of this signal under the constraint of not merging different radars in one cluster
   - **Cluster fusion**: Hierarchical agglomerative clustering combined with Optimal Transport Distances
2. **RADAR Identification**:
   - **Set of pulses and emitter class modelisation**: Representation of sets of pulses and emitter classes as probability distributions
   - **Selection**: Selecting the emitter class closest to the set of pulses as the true class with Optimal Transport Distances

Data description

- Work from simulated labeled data
- Each dataset is composed of RADAR pulses, described by only 4 features:
    - Frequency (F)
    - Pulse Width (PW)
    - Level (L)
    - Time of Arrival (TOA)

RADAR Signal example

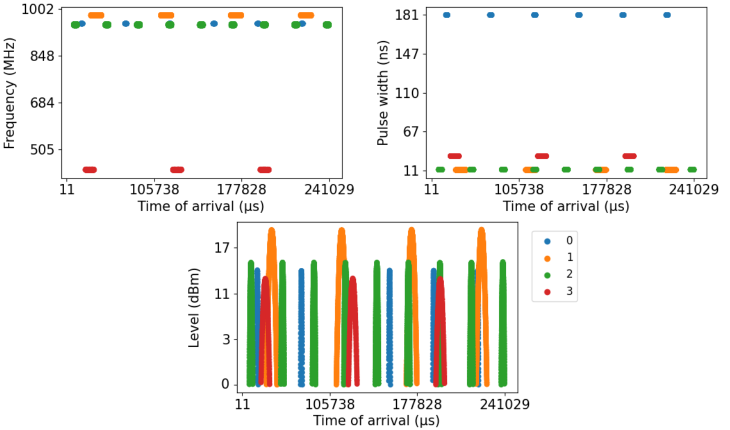


Figure: Example of a **simple** signal containing 4 RADARs.

Content

Content

Step 1: Clustering with HDBSCAN

Methodology:

- Application of HDBSCAN on two features: Frequency and Pulse Width
- Why HDBSCAN ?
  - $\Rrightarrow$ Unsupervised algorithm
  - $\Rrightarrow$ Search clusters of varying densities
  - $\Rrightarrow$ Few parameters to optimize
- Over-estimation of the number of clusters
  - $\Rrightarrow$ Some RADARs don't transmit continuously
  - $\Rrightarrow$ Each cluster must not contain the observations of different RADARs

Output of HDBSCAN

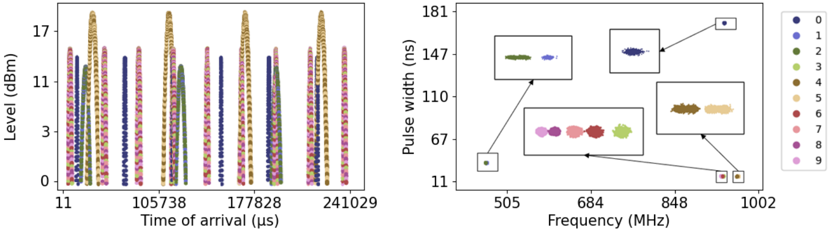


Figure: HDBSCAN output. The colors represent the different clusters.

Step 2: Cluster fusion with Hierarchical Agglomerative Clustering combined with Optimal Transport Distances

Methodology:

1. Cluster representation with histogram of their TOA and LEVEL
2. Compute distances between cluster histograms using **optimal transport**
3. Aggregation of the two closest clusters with the smallest transport cost
4. Updated distances
5. Repeat until a single cluster is obtained

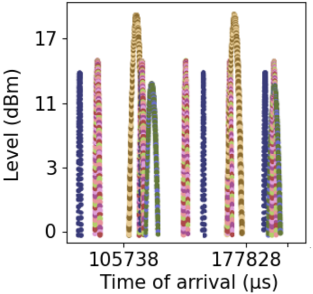


Figure: Lobes of the HDSBCAN Clusters.

Histograms of the clusters

- Cluster 0 and 1 **don't belong** to the same transmitter
  - ✘ **Different** distribution
  - ✘ **High** transport cost
- Cluster 1 and 2 **belong** to the same transmitter
  - ✔ **Similar** distribution
  - ✔ **Low** transport cost

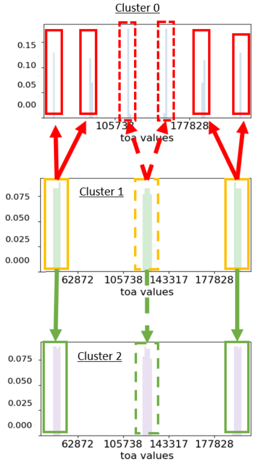


Figure: Histograms of clusters 0, 1 and 2 made from TOA and LEVEL.

Optimal Transport explanation

- Transporting distribution to another has a cost:

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

$C(\mathbf{P})$ = Total cost

C_{nm} = cost transporting unit mass from x_n to y_m (transport cost)

P_{nm} =quantity of mass taken from x_n to y_m (transport plan)

- The optimal transport plan is a simple minimization problem:

$$\mathbf{P}^{\star} = \underset{\mathbf{P} \in \mathbf{R}_{+}^{N \times M}}{\text{argmin}} \langle \mathbf{C}, \mathbf{P} \rangle \text{ such that } \mathbf{P}\mathbf{1}_M = \mathbf{a}, \mathbf{P}^T \mathbf{1}_N = \mathbf{b}^T. \tag{2}$$

Hierarchical Approach: Dendrogram

- The decision model is mainly based on the use of 3 metrics applied to the entire dendrogram:
  - Silhouette score
  - Calinski-Harabasz score
  - Davies-Bouldin score

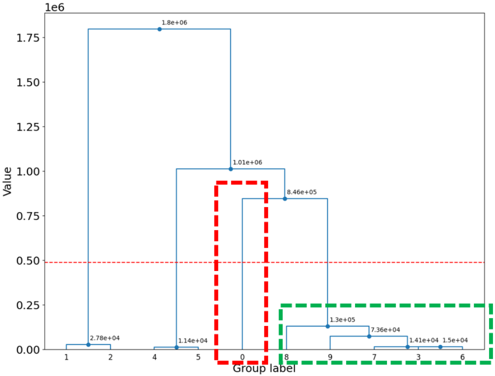


Figure: Cluster aggregations with Optimal Transport Distances.

Final Grouping

- Perfect grouping
- Statistics to evaluate the regroupment:
  - Homogeneity = 1
  - Completeness = 1
  - Unclassified rate = 2.1%

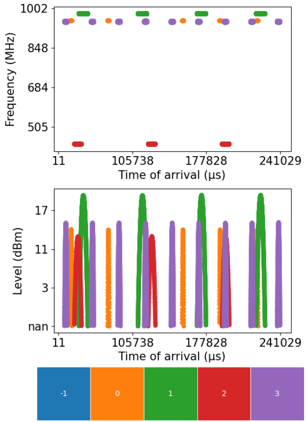


Figure: Final grouping. The colors represent the different sets of pulses with an outlier class (-1).

Content

Methodology

- **Input:** Deinterlaced Sets of pulses obtained by the first step of the RADAR Recognition process
- **Hypothesis:** 1 set of pulses = 1 RADAR
- **Modelisation:** Emitter classes and sets of pulses are represented as probability distributions
- **Identification:** Selection of the emitter class with the lowest transport cost between its probability distribution and the distribution of the set of pulses as the true class

Emitter class model

- Database with more than **60 classes**
- Construction of a measure that describes each RADAR belonging to a reference database based on its frequencies and pulse width:

$$\mu_j = \sum_{n=1}^{N} \alpha_n \delta_{f_n, pw_n} \tag{3}$$

$N =$ the number of frequencies and pulse width on which the RADAR transmits
$\alpha =$ the weight of the frequency and pulse width
$j =$ the index of the transmitter in the database
$\delta =$ dirac mass (with $\sum \alpha_n = 1$)

Emitter Database analysis

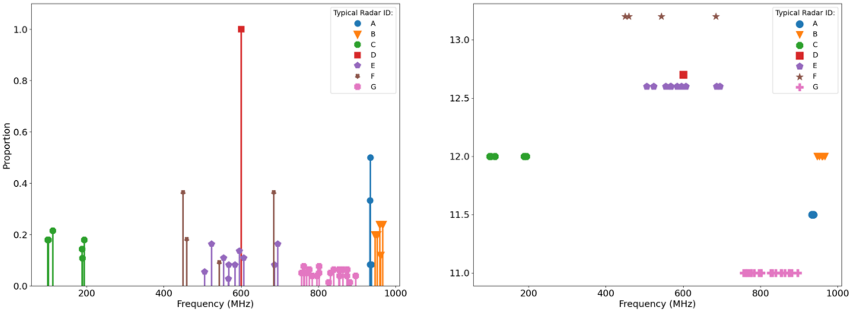


Figure: Representation of **simulated** RADARs classes. The graph represents the RADARs in the frequency plane and the F-PW plane. Each color represents a simulated emitter class.

Optimal transport plan between emitters classes

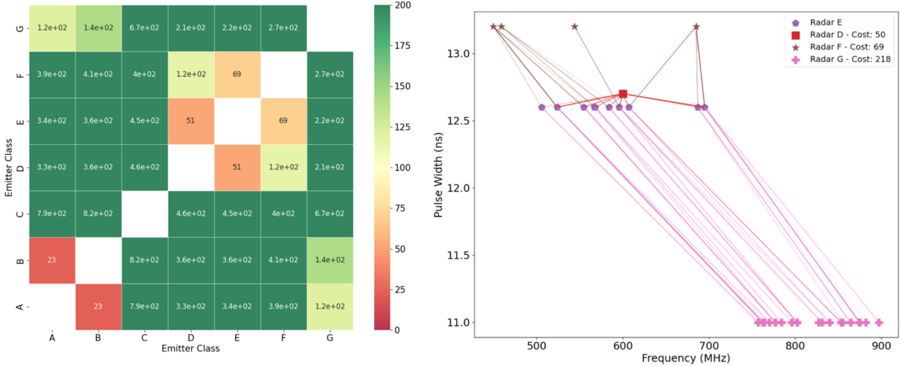


Figure: Cost matrix between Emitters Classes (left plot) - Transport plan between purple emitter toward red, brown and pink emitters (right plot).

Set of pulses model and Identification

- Construction of a probability distribution from the deinterlacing set of pulses based on frequency and pulse width:

$$\nu = \frac{1}{M} \sum_{m=1}^{M} \delta_{f_m, pw_m} \tag{4}$$

with M the number of pulses in the set.

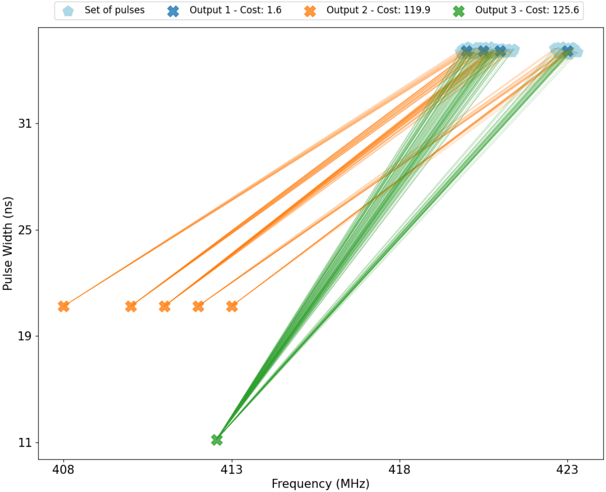
- Identification of the emitter class by assigning the closest RADAR class in the optimal transport distance sense to each set of pulses:

$$j^\star = \text{argmin}\, d\left(\mu_j, \nu\right). \tag{5}$$

Classification output for set of pulses 2

- Transport plan between the pulses and the three closest emitter classes identified by the algorithm.

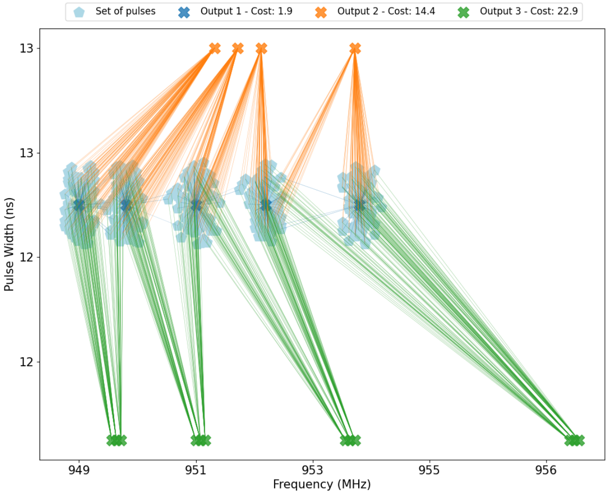
✔ Correct identification



Classification output for set of pulses 3

- Transport plan between the pulses and the three closest emitter classes identified by the algorithm.

✔ Correct identification



Content

Conclusion and Perspectives

- Works on moderately complicated cases with few data
- Methodology can handle a large number of classes to identify
- And after ?
  - Improvement of the deinterlacing methodology (better decision model, better pulse separation...)
  - Integration of a new feature to improve RADAR classification results
  - Detection of new RADAR transmitters

Publications

- **Current Publications**
  - Deinterleaving and Clustering unknown RADAR pulse *(IEEE Radar Conference (RadarConf 2021)*
  - RADAR Emitter Classification with Optimal Transport Distances *(EUSIPCO 2022)*
  - Désentrelacement et classification de signaux RADAR basés sur des distances de transport optimal *(GRETSI 2022)*
- **Next publications**
  - CAID *(Conference on Artificial Intelligence for Defense)*
  - AESS *(Article in a scientific journal)*

Thank you for your attention