Deinterleaving RADAR pulses

Manon MOTTIER

Laboratoire des signaux et systèmes
CentraleSupélec, Université Paris-Saclay

October 28th, 2021

Content

Content

Context

Industrial thesis (CIFRE)

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

Context

AVANTIX (ATOS Group)

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

Context

The subject

- Use and development of machine learning tools for passive RADAR intelligence to identify RADARs present in the environment.
- Improvement of the RADAR deinterlacing chain:
  1. Separation and grouping of pulses
  2. Identification of the RADAR present or updating of the database

Context

Work organisation

- Full time in the laboratory (Avantix is based in Aix-en-Provence)
- Weekly meeting with the company via teams
- Half-yearly meeting with the company, the supervisors and the DGA

Content

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)

Data description

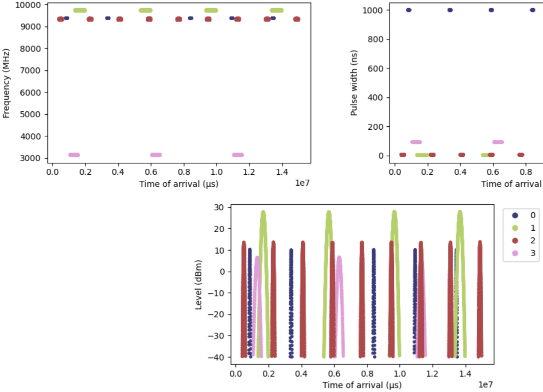


Figure: Example of a **simple** signal containing 4 RADARs.

Methodology

Implementing a two-step strategy:

1. **Deinterlacing pulses:**
   - **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**: Extraction of group characteristics and comparisons to a reference database to identify the transmitter

Clustering with HDBSCAN

Methodology:

- Application of clustering algorithm on two features: Frequency and Pulse Width
- Why HDBSCAN ?
  - Unsupervised algorithm that allows us to search for clusters of varying densities
  - Few parameters to optimize
- Over-estimation of the number of clusters
  - Allows to receive RADARs that do not transmit continuously
  - Each cluster must not contain the observations of different RADARs

Output of HDBSCAN

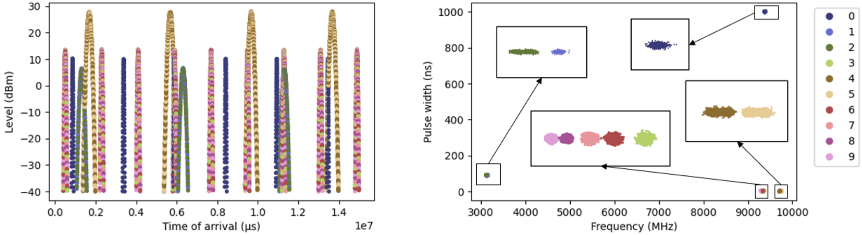


Figure: HDBSCAN Output. Colors represent the different clusters.

Cluster fusion

Methodology:

1. Dimensionality reduction: each cluster is represented by a histogram of its TOA and LEVEL
2. Compute distances between cluster histograms using optimal transport
3. Aggregation of the 2 closest clusters i.e., the 2 clusters with the lowest transport costs
4. Computation of new costs between old clusters and the new ones
5. Repeat steps 2 to 4 until a single cluster is obtained

Histograms

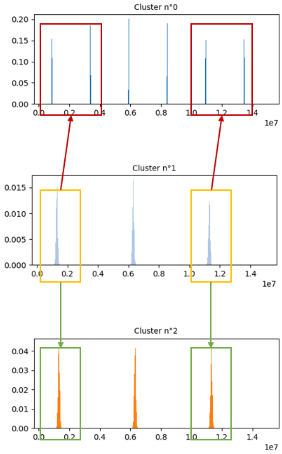


Figure: Histograms of the previous clusters.

Optimal Transport

- Transporting distribution to another has a cost:

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

$C(\mathrm{P}) =$ 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:

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

Dendrogram

- Decision model based on several unsupervised metrics to cut the dendrogram:
  - Silhouette score
  - Calinski-Harabasz score
  - Davies-Bouldin score

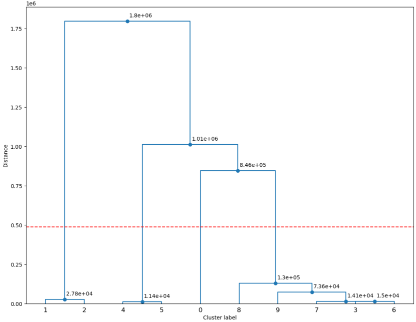


Figure: Cluster aggregations with Optimal Transport Distances.

Final Grouping

- Perfect grouping
- Statistics:
  - Homogeneity = 1
  - Completeness = 1
  - Unclassified rate = 2%

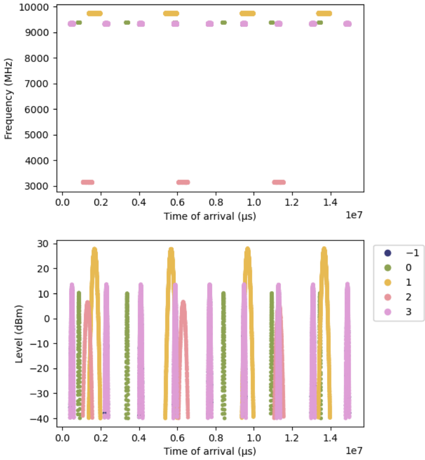


Figure: Final grouping

RADAR Identification

Methodology:

1. Construction of a measure that describes each RADAR belonging to a reference database based on its frequencies:

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

$N =$ the number of frequencies on which the RADAR transmits
$\alpha =$ the weight of the frequency
$\delta =$ dirac mass (with $\sum \alpha_n = 1$)

RADAR Identification

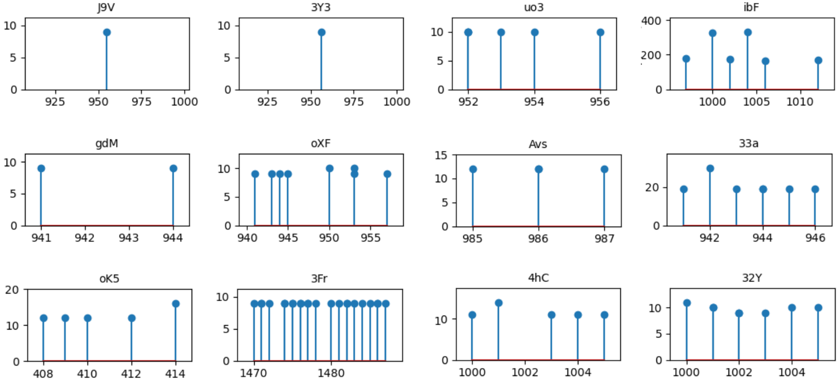


Figure: Referential Database representation. Example of the frequency representation of the RADAR present in the identification database. The Radar ID and value range have been changed.

RADAR Identification

2. Construction of a measure that describes each cluster from the frequency:

$$\gamma = \frac{1}{N} \sum_{m=1}^{N} \delta_{f_m} \tag{4}$$

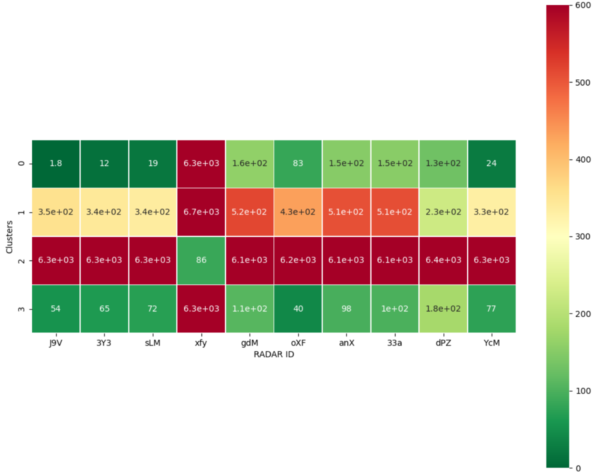
$N =$ the number of pulses in the cluster

3. Identification of the transmitter:

$$j^\star = \operatorname{argmin} d\left(\mu_j, \gamma\right) \tag{5}$$

Classification Output

- **Extract** of the cost matrix between the groups and the some waveforms.
- Red: High transport cost
- Green: Low transport cost



Classification Decision

Clusters ID	Rank 1		Rank 2		Rank 3	
	ID	Cost	ID	Cost	ID	Cost
0	eF2	1.62	J9V	1.83	sxh	6.02
1	BGF	1.58	d2B	1.54	smn	28.70
2	viY	1.43	xfy	85.75	oK5	98.75
3	D9V	2.14	36y	13.05	BF8	13.06

Table: Classification decision.

- Correct identification
- Problem with mono frequency RADAR
- Difficulty to identify sub-modes
- Frequency not sufficient

Content

Unlabeled data

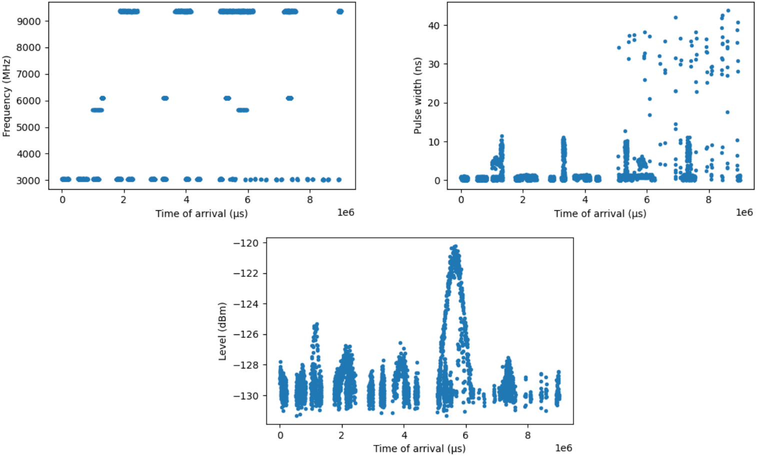


Figure: Unlabeled data with 5 RADARs.

Clustering with HDBSCAN

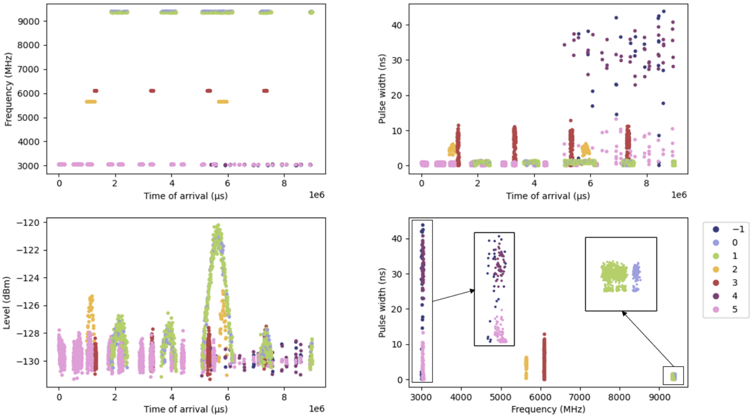


Figure: HDBSCAN Output. Colors represent the different clusters (including outliers' class).

Histograms

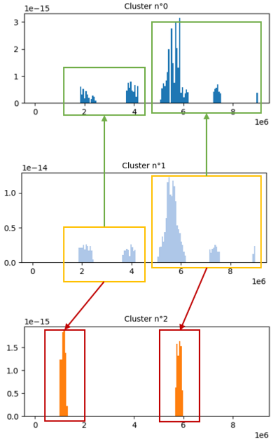


Figure: Histograms of the previous clusters.

Final Grouping

- Pulses of the 2 transmitters around 3GHz are mixed
- Hdbscan can not differentiate them at step 1
- Statistics:
  - Homogeneity = 0,94
  - Completeness = 1
  - Unclassified rate = 0,9%

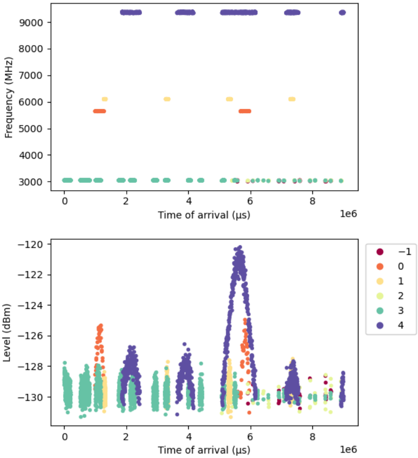


Figure: Final grouping.

Identification

Clusters	Rank 1		Rank 2		Rank 3	
	Id	Cost	Id	Cost	Id	Cost
0	32P	2.35	yfN	22.60	YuK	29.49
1	5fd	2.51	38M	30.84	32P	455.27
2	XAQ	7.23	xfy	10.45	oK5	14.04
3	xfy	2.72	XAQ	16.28	oK5	20.11
4	uo3	1.96	3Rs	1.99	Ykn	6.50

Table: Classification decision.

- Correct identification
- Problem with mono frequency RADAR
- Difficulty to identify sub-modes
- Frequency not sufficient

Content

Conclusions

- Works on moderately complex cases
- Identification phase gives encouraging results
- Communication:
  - Publication of a Conference paper: "Deinterleaving and Clustering unknown RADAR pulses" - Manon Mottier; Gilles Chardon; Frédéric Pascal
  - International Conference: IEEE RadarConf 2021, Atlanta, GA, USA
  - International Seminar: Atos Scientific Community

Next objectives

- Improve the proposed approach:
  - Deinterleaving Radar pulses:
    - Define a metric to identify identical RADARs in Frequency and Pulse Width
    - Improve model decision to cut the dendrogram
  - Identification:
    - Add Pulse width to improve identification
    - Define a distance between Groups
    - Detection of new transmitter
- Publication of an article in a specialized journal

Thank you for your attention

For more information please contact:

Manon MOTTIER

manon.mottier@atos.net