Key Radar Signal Sorting and Recognition Method Based on Clustering

Combined with PRI Transform Algorithm

Kai Kang¹, Yixao Zhang¹, Wenpu Guo¹, Luogeng Tian²

¹High-Tech Institute of Xi'an, Xi’an 710025, China
²Information and Communication College, National University of Defense Technology, Xi’an 710106 China

Abstract: In this paper, we investigate the problem of key radar signal sorting and recognition in electronic intelligence (ELINT). Our major contribution is the development of a combined approach based on clustering and PRI transform algorithm, to solve the problem that the traditional methods based on Pulse Description Word (PDW) were not exclusively targeted at tiny particular signals and were less time-efficient. We achieve this in three steps: Firstly, PDW presorting is carried out by the DBSCAN clustering algorithm, and then PRI estimates of each cluster are obtained by the PRI transform algorithm. Finally, by judging the matching between various PRI estimates and key targets, it is determined whether the current signal contains key target signals or not. Simulation results show that the proposed method should improve the time efficiency of key signal recognition and deal with the complex signal environment with noise interference and overlapping signals.

Keywords: radar signal recognition; PDW; DBSCAN clustering; PRI transform

1. Introduction

Battlefield information superiority is the focus of both sides in information warfare. In the electromagnetic environment of the modern battlefield, the signals are increasingly dense and complex, so the electronic reconnaissance system faces great difficulties in sorting and identifying the signals. The sorting and recognition of complex signals is an important function of Electronic support measures (ESM). It is a process of comparing the measured signal parameters with those accumulated in advance to confirm the original properties of the radiation source. The ultimate goal is to locate and analyze the observed and intercepted radar signals and to provide battlefield situation information for commanders to assist in decision-making.

The existing radar countermeasure and reconnaissance equipment mainly carry out radiation source sorting and identification based on Pulse Description Word (PDW) of detected signals. PDW refers to the Time of Arrival (TOA), Direction of Arrival (DOA), Radio Frequency (RF), Pulse Width (PW), Pulse Amplitude (Pulse Amplitude, PA), etc. extracted by the radar reconnaissance receiver from a single radar Pulse. Wherein, Pulse Repetition Interval (PRI) can be calculated according to the TOA parameter.

Early studies mostly adopted PRI-based single-parameter signal sorting algorithms, and the classical algorithms include extended association algorithm, histogram algorithm, and PRI transform algorithm[1]. Among them, the PRI transform algorithm has a good sorting effect on complex modulation signals. However, because the type and number of intercepted signals cannot be determined directly in the high-density non-cooperative radar signal environment, the single-parameter signal sorting algorithm cannot directly complete the non-cooperative radar signal sorting[2–4].

With the widespread use of machine learning methods, artificial intelligence techniques, and their applications in different areas[17], researchers
found that the unsupervised clustering analysis in data mining is very suitable for solving the problem of radar signal sorting without prior information, and proposed the radar signal sorting algorithm based on the clustering algorithm\(^5\). Typical representatives of clustering methods are distance-based clustering algorithms, hierarchical clustering algorithms, density-based clustering algorithms\(^6\), etc. Such as k-means clustering algorithm\(^7\,8\), BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies)\(^9\), etc. However its clustering number must be determined in advance, and the noise data has a great impact on it. The algorithm DBSCAN (Density-Based Spatial Clustering of Applications with Noise) was proposed by Ester M.\(^10\), which is a typical density-based spatial clustering algorithm\(^11\), and can find clusters of arbitrary shapes by clustering according to the tightness between data points.

At present, radar signal sorting and recognition generally follow the normal process of ‘signal sorting-parameter extraction-data comparison’. However, in practical applications, more attention is often paid to whether the current pulse contains known key targets. At this time, it is not necessary to identify all radiation source signals and requires high efficiency in identifying key targets\(^12\). Sorting and identifying by the conventional process will cause a waste of computing time and data resources, and seriously affect the timeliness of finding key targets.

Based on the above problems, this paper proposes a key radar signal sorting and identification method using DBSCAN clustering algorithm combined with PRI transform. Firstly, PDW presorting is carried out based on the DBSCAN clustering algorithm. Then, various PRI estimates are obtained by the PRI transform algorithm. Finally, the matching between various PRI estimates and key targets is judged to determine whether the current signal contains key target signals. This method can be used to process PDW data of pulse signals in real-time, which is highly targeted and improves the recognition time of key signals.

## 2. PDW presorting based on DBSCAN

DBSCAN algorithm has three input parameters, including data set to be clustered \(D\), neighborhood radius \(\varepsilon\), and density threshold \(\text{minPts}\)\(^13\). DBSCAN classifies data points by expanding and aggregating adjacent areas with similar density efficiently handles irregular and abnormal data. But its performance highly depends on the manual selection of two input parameters \(\varepsilon\) and \(\text{minPts}\).

Set \(x \in D\),
\[
N_{\varepsilon}(x) = \{ y \in D : \text{dist}(y, x) < \varepsilon \} \quad (1)
\]
\(N_{\varepsilon}(x)\) is the \(\varepsilon\)-neighborhood of \(x\), and the density is defined as:
\[
\rho(x) = |N_{\varepsilon}(x)| \quad (2)
\]

In density-based clustering algorithms, a cluster is considered as a dense region of objects and an object of one cluster must be surrounded by objects of the same cluster. The main idea of DBSCAN is to partition density-connected objects into the same cluster\(^14\). Some basic concepts of the algorithm DBSCAN are listed below.

- \(Eps\)-neighbourhood: an \(d\)-dimensional hypersphere region with a given object \(p\) as its core and \(Eps\) as its radius.
- Core point: If the density of point \(x_i\), \(\rho(x) > \text{minPts}\), the point is called the Core point. Not a Core point, but a point within the epsilon and loss region of a Core point is called a boundary point. Otherwise, it is called a non-Core point or noise.

Directly density-reachable: Set \(D_{Core}\) as the set of all Core points. If \(x \in D_{Core}\), it is called that \(y\) is directly density-reachable for \(x\). Suppose \(P_{1}, P_{2}, \ldots, P_{m} \in D\), where \(m > 2\), if \(P_{m+1}\) meets the requirement directly density-reachable from \(P_{m}\).
then it is called that $P_m$ is density-reachable from $P_1$. Let $x, y, z \in D$, if $y$ and $z$ are both density-reachable to $x$, then it is called and $z$ are density-connected.

DBSCAN checks the $Eps$-neighbourhood of each data point in $D$ to search the centroids (core points) of clusters, iteratively aggregates objects that are directly reachable from these core objects, and finally output the collection of clusters when no new points are added to any cluster.

In general, the parameters of RF, PW, and DOA of a single radar have a limited range of variation, and the characteristics of data clustering are relatively obvious. Therefore, RF, PW, and DOA parameters in PDW are selected in this paper for sorting. However, RF, PW, and DOA parameters are not in the same dimension. To ensure the scientific nature of the algorithm, the normalized transformation of parameters is required (that is, all radar characteristic parameter data are converted into data between [0, 1]). The range normalization formula is as follows:

$$
\text{rf}_i = \frac{RF_i - \min RF}{\max RF - \min RF}, \quad \text{pw}_i = \frac{PW_i - \min PW}{\max PW - \min PW}, \quad \text{doa}_i = \frac{DOA_i - \min DOA}{\max DOA - \min DOA}
$$

(3)

$i = 1, 2, \ldots n, n$ is the PDW sample number of radar signals used for clustering.

$pdw_i=(\text{rf}_i, \text{pw}_i, \text{doa}_i)$ is the normalized PDW data sample.

The distance formula from data point $pdw_i$ to point $pdw_j$ is:

$$
dist(pdw_i, pdw_j) = \sqrt{\sum_{t=1}^{n} (pdw_i(t) - pdw_j(t))^2}
$$

(4)

It is not economical to use all the detection parameters for clustering sorting when the key detection target parameters are known and the detection target is clear. However, the variation range of carrier frequency RF and pulse width PW of the same radiation source signal must be limited. Therefore, screening conditions can be set according to the known key target parameters, and only the detected signal parameters satisfying and can be used for clustering sorting. Where $RF_0$ and $PW_0$ is the setting reference parameter, $\Delta RF$ and $\Delta PW$ is the parameter screening range.

3. PRI estimation and signal recognition

After the PDW presorting is completed for a certain key target signal, it is necessary to further determine whether the PRI parameter characteristics contained in the selected data points are consistent (or partially consistent) with the PRI parameter characteristics of the targeted signal.

Among the various parameters of radar signals, PRI has the largest variation type and range, but the PRI of the same radar radiation source signal has a strict time correlation, so PRI is the most important parameter in signal identification.

Common PRI sorting algorithms are mainly divided into the histogram method\cite{15} and the PRI transformation method. The histogram method includes the cumulative difference histogram (CDIF) method and sequence difference histogram method (SDIF). Among them, the SDIF is an improvement of the CDIF, which has a fast computing speed, but cannot suppress the second and higher harmonics, and PRI misjudgment is easy to occur when the pulse loss rate is high.

Therefore, this paper chooses the PRI transformation method to estimate the PRI value contained in the radar pulse flow, and judge whether it contains pulse sequences consistent with PRI parameters of key targets.

3.1 Basic principles of PRI transform algorithm

The principle of the PRI transform algorithm is: set the arrival time of each radar pulse signal $ast_n=0, 1, 2, \ldots N-1, n$ is the number of sampling
impulses, and the sampling impulse can be expressed as the sum of unit impulse function, namely
\[ g(t) = \sum_{n=0}^{N} \delta(t - t_n) \]  

Phase factor \( \exp(2\pi j\tau / \tau) \) is introduced into the integral transformation formula of \( g(t) \), and PRI transform formula of a pulse is defined as
\[ D(\tau) = \int_{-\infty}^{\infty} g(t)g(t+\tau)\exp(2\pi j\tau / \tau)dt \]  

In the above equation, \( \tau > 0 \), \( D(\tau) \) is the PRI transform spectrum.

Substitute equation (5) into the equation, and get
\[ D(\tau) = \sum_{n=1}^{N} \sum_{m=0}^{\infty} \delta(t - t_n + t_m)\exp(2\pi j(t_n - t_m)/(t_n - t_m)) \]  

The analysis of the equations shows that there will be a peak value of the PRI transform spectrum at the potential PRI parameters.

To facilitate analysis and calculation, PRI transformation needs to be discretized. \( \{\tau_{min} \leq \tau_{max}\} \) is set as the PRI change range, which is divided into \( K \) intercells. The width of each intercell is \( \tau_b = (\tau_{max} - \tau_{min}) / K \) , and the center of the \( K \)th intercell is set as:
\[ \tau_k = (k - \frac{1}{2})\tau_b + \tau_{min}, \quad k = 1,2,\cdots,K \]  

After determining the value range of PRI and the center value of each PRI small box, \( D_k \) can be expressed as:
\[ D_k = \int_{-\infty}^{\infty} D(\tau)d\tau = \sum_{t_1 < \tau_k, t_1 < t_2} \int_{t_1}^{t_1+\frac{1}{2}\tau_b} D(\tau)d\tau \]  

When the \( b \to 0 \). The \( n \) PRI values whose \( |D_k| \) exceeds the detection threshold are taken as the estimated PRI values and expressed as, where \( n \) is the estimated number of PRI values.

3.2 Improvement of PRI transform algorithm

The study [16] proposed an improved PRI variation algorithm. Firstly, the phase error caused by the PRI jitter was reduced by dynamically changing the starting point of time. The second is to make PRI small box range overlap, to avoid the corresponding error caused by PRI jitter small box. Space is too limited to elaborate.

3.3 Target identification based on PRI estimates

The advantage of the PRI transform algorithm lies in its ability to accurately judge the PRI characteristic parameter value, while the disadvantage is that it is not suitable for real-time processing of large amounts of data due to its complex calculation and a large amount of computation.

DBSCAN-based presorting in the previous paper has completed the directional screening of PDW data. Based on the TOA value corresponding to the screened PDW data points, more accurate PRI values can be obtained, which eliminates a lot of interference and greatly simplifies the PRI data volume, making the included PRI features more prominent and significantly reducing the calculation amount. The problem of low efficiency of the PRI transformation method is solved.

4. The process of sorting and identification

Combining the DBSCAN based clustering signal presorting and PRI-based key signal recognition, the key radar signal sorting and recognition method is formed. The specific process is as follows:

Suppose there are \( N \) types of key target signals, and they are numbered according to respectively. In the following steps, the key target of class \( j \) is taken as an example to illustrate that all the key signals of class \( N \) are processed in parallel according to the same steps.

**Step 1:** according to the RF and PW parameter characteristics of the key targets in category \( j \) set the reference parameters \( \Delta RF \) and \( \Delta PW \), and the parameter screening range, and Set DBSCAN cluster neighborhood radius \( \text{Eps} \) and density threshold \( \text{minPts} \).

**Step 2:** input pulse parameter PDW, and put the pulse parameter satisfying \( |RF - RF_k| \leq \Delta RF \) and \( |PW - PW_k| \leq \Delta PW \) into cluster \( C_1 \).
Step 3: DBSCAN clustering was performed on the data in cluster $C_1$. Suppose the clustering results divided $C_1$ into $k$ classes and put the pulse parameters corresponding to each core point and boundary point into the corresponding category cluster $C_{2k}$.

Step 4: Arrange the TOA corresponding to $C_{2k}$ data points from small to large, identify PRI characteristic parameters according to the PRI transformation method, and discuss in three cases:

1. The identified PRI feature parameter values are partially the same as the PRI parameters of key target signals, but they do not completely include the PRI parameters of the target. At this time, it can be considered that the $j$-class target is suspected to be discovered, PDW parameters within $C_{2k}$ are saved, and the next batch of PDW data is processed, focusing on the $j$-class target.

2. The identified PRI feature parameter values fully include PRI parameters of key target signals. At this time, it can be considered that the $j$th target is successfully discovered, PDW parameters within $C_{2k}$ are saved, and timely reported;

3. The identified PRI feature parameter values are completely different from the PRI parameters of key target signals, and it is considered that the $j$th class target has not yet appeared, so the next batch of PDW data will be processed.

Step 5: According to the identification results, save PDW data corresponding to the $j$th signal, and complete the accurate sorting of detected PDW parameters.

5. Simulation analysis

5.1 Simulation setting

According to the actual radar signals received, pulse flow containing multiple radar signals is simulated. Pulse simulation data is shown in Table 1, and its PDW 3d schematic diagram is shown in Fig. 1. The starting time is 100us in sequential order, and pulse loss caused by pulse overlap is not considered. Among them, RF, PRI, PW, and DOA add 10%gaussian noise respectively. 500 interference pulses are added to the PDW data stream, and the added noise parameters are uniform random Numbers within the value range of mixed radiation source signals (using rand function to generate evenly distributed pseudo-random Numbers). The 3d schematic diagram of simulation PDW data obtained after adding interference is shown in Fig. 2.

The pulse signals generated by the simulation include complex signals such as frequency agility, frequency jump, frequency fluctuation, frequency jitter, etc. PDW data of each radar are seriously overlapped and contain similar signals with the same parameters and different directions.

Key signal parameters are shown in Table 2. There are five radars in total. Radar1 has the largest parameter range and overlaps with Radar2 and Radar4. Radar2 signals appear from two different directions in the simulation pulse signal, but Radar5 signals do not appear.

<table>
<thead>
<tr>
<th>Serial number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>interference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOA/ (°)</td>
<td>36–39</td>
<td>33–35</td>
<td>40–43</td>
<td>43–44</td>
<td>42–45</td>
<td>43–45</td>
<td></td>
</tr>
<tr>
<td>RF/ (MHz)</td>
<td>6 450–6 650 agility</td>
<td>6 600 fixing</td>
<td>6 200–6 500 agility</td>
<td>6 400,6 500, 6 450,6 300 hopping</td>
<td>6 600–6 800</td>
<td>6 600 fixing</td>
<td></td>
</tr>
<tr>
<td>PW/ (µs)</td>
<td>5.10–11.40</td>
<td>9.65</td>
<td>14.5</td>
<td>15.0–20.0</td>
<td>10.5–15.5</td>
<td>9.65</td>
<td></td>
</tr>
<tr>
<td>PRI/ (µs)</td>
<td>800<em>5/830</em>5/ 400//440 fixing</td>
<td>670 fixing</td>
<td>1 500 fixing</td>
<td>1 200//1 250</td>
<td>400//440 fixing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>900<em>5/780</em>5/ jitter</td>
<td>1 500 fixing</td>
<td>1 200//1 250</td>
<td>400//440 fixing</td>
<td>jitter</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Radar pulse simulation parameters
Table 2: Key target signal parameters

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Radar 1</th>
<th>Radar 2</th>
<th>Radar 3</th>
<th>Radar 4</th>
<th>Radar 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF/ (MHz)</td>
<td>6 450–6</td>
<td>6 600 fixing</td>
<td>6 200–6</td>
<td>6 600–6</td>
<td>6 300 fixing</td>
</tr>
<tr>
<td></td>
<td>650 agility</td>
<td></td>
<td>500 agility</td>
<td>800 agility</td>
<td></td>
</tr>
<tr>
<td>PW/ (μs)</td>
<td>5.10–11.40</td>
<td>9.65</td>
<td>14.5</td>
<td>10.5–15.5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>800<em>5/830</em>5/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRI/ (μs)</td>
<td>900<em>5/780</em>5/</td>
<td>400/440</td>
<td>670 fixing</td>
<td>1 850/1 900</td>
<td>1 200/1 250</td>
</tr>
<tr>
<td></td>
<td>820<em>5/930</em>5/</td>
<td>jitter</td>
<td>jitter</td>
<td>jitter</td>
<td>jitter</td>
</tr>
<tr>
<td></td>
<td>irregular</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1 Diagram of the original PDW data set

Fig. 2 Diagram of the simulated PDW data set

5.2 Simulation of presorting based on DBSCAN

Set $\varepsilon = 1$, density threshold $\min Pts = 10$, parameter screening range $\Delta RF = 200$. The signals of the 5 key radars are sorted according to the process, and the sorting effect is shown in Fig. 3.
The analysis of clustering results shows that: Because Radar1 has a wide range of parameters, its RF and PW parameters completely cover Radar2 and partially cover Radar4, so it is divided into 3 categories in different DOA ranges for Radar1 clustering. Radar2 itself appeared twice, and RF and PW parameters overlapped partially with Radar1, so it was also divided into 3 categories for Radar2 clustering. Although Radar5 does not appear, its RF and PW parameter range is included in the Radar1 parameter range, so there are corresponding clustering results.

5.3 Simulations of PRI estimates

Based on the sorting results, PRI estimation is carried out by the PRI transform algorithm, and the obtained PRI spectrum is shown in Fig. 4.
PRI identification results of each key radar target are shown in Table 3.

Table 3 The result of identification of PRI

| Key radar target | PRI estimation | |Dk| peak | PRI of key radar | Conclusion |
|------------------|----------------|---------|-----------|------------------|------------|
| Radar1           | 423            | 108     | 800,83    | 0,900            | Suspected identify |
|                  | 772,852,944    | 124     | 780,82    | 0,930            |            |
|                  | 1170~1253      | 110     |           |                  |            |
| Radar2           | 418            | 155     | 404–4     | 40               | Successful identify |
|                  | 852,946        | 57      |           |                  |            |
| Radar3           | 670            | 261     | 670       |                  | Successful identify |
| Radar4           | 1196~1258      | 212     | 1200~1250 |                  | Successful identify |
|                  | 850            | 45      |           |                  |            |

Analysis of the identification results shows that the method can accurately identify PRI parameter value, and even in the case of multiple signal parameters and the existence of |Dk|, the main signal and mixed-signal can be distinguished by the size of |Dk|, so the identification accuracy is very high. However, it is also found that, for the PRI irregular signals such as Radar1, this method cannot guarantee to find all PRI values and identify their arrangement.

5.4 PDW re-sorting according to the estimated PRI

After the PRI identification of each key target is completed, PDW data measured can be accurately sorted based on the identification results, and PDW data from the same radar radiation source can be classified into one category. It is important to note that multiple radiation sources of the same type of radar can exist simultaneously in multiple directions. Therefore, when combining PDW data of the same type of radar, it should be prevented to combine PDW data that is manifested as two types (or more) in the pre-scoring selection into one.

The final 3d schematic diagram of PDW sorting results is shown in Fig. 5.

5.5 Comparative analysis of algorithm

To verify the improvement effect of the algorithm proposed in this paper, the sorting accuracy formula is defined as
The sorting accuracy of the proposed sorting recognition algorithm in this paper is compared with the signal sorting method based on pulse train samples and the classical k-means clustering sorting method to verify the improvement effect. Simulation parameters are set as in Table 1 but interference pulses are removed. The k-means clustering algorithm, the signal sorting method based on pulse train samples, and the improved algorithm in this paper were used 50 times respectively to calculate the statistical average of sorting accuracy.

To achieve the comparison effect, the number of known clusters is set as 5 in this paper, and the initial clustering center of the k-means algorithm is randomly generated.

Under the above conditions, the k-means algorithm, the signal sorting method based on pulse train samples, and the algorithm in this paper have their respective sorting accuracy as shown in Table 4.

It can be seen from Table 4 that, compared with the k-means algorithm and signal sorting method based on pulse train samples, the sorting accuracy of the method proposed in this paper is significantly improved.

6. Conclusion
This paper studied the separation and recognition of key target signals based on PDW and proposed a new separation and recognition method based on clustering and PRI transform algorithm. Simulation experiments showed that the algorithm had the following advantages: Firstly, it was insensitive to the influence of noise and parameter overlap and could be applied to the complex electromagnetic environment composed of non-cooperative radar signals. Secondly, it reduced the amount of computation, avoids sorting and operating of all data, and it significantly improved the efficiency of radar signal recognition. However, this method could not effectively judge PRI arrangements. If the radar parameter range and PRI center value were similar or the same as the key radar target, but the arrangement mode was different, the error identification result would be produced, and further research is needed to solve this problem.

Reference


