The Electromagnetic Signal Track Correlation Algorithm based on Attribute Clustering and Space-time Constraints

Bo Wu*, Jianan Wang, Zhaojun Wang, Jiecheng Yu, Yue Gao, Peiyun Li
Space Star Technology CO., LTD, Beijing, China
*Corresponding author's e-mail: wub@spacestar.com.cn

Abstract. In view of the problem that start-end correlation of ship electromagnetic signal, the electromagnetic signal track correlation algorithm, based on attribute clustering and space-time constraints, is proposed. The algorithm consists of K-means dynamic clustering stage, space-time constraints of the track association stage. By extracting the relevant properties of electromagnetic signals, through the clustering stage, obtaining an indefinite number of clusters, and through the space-time correlation stage, obtaining a number of tracks per cluster. The main contributions of this paper are: 1. the K-means dynamic clustering method is proposed; 2. the track correlation method of space-time constraints is proposed. In this paper, the results of the comparison experiment show that the accuracy of this algorithm is better than that of the comparison algorithm, and the practical feasibility is verified.

1. Introduction
As we all know, the problem that continuous monitoring of marine ship target, can be transformed into the problem, which ship target track start-end correlation, while the target correlation method is mostly based on the target position, and the accuracy of the target position is an important objective factor that affecting the accuracy of the target correlation result.

There are two solutions to the problem of correlating the start and end points of the track currently. One is by collecting electromagnetic signals from the ship’s radiation source, the other is by collecting remote sensing images of ship targets. However, in the real environment, the number of electromagnetic signals received is large, the position accuracy is low, and there are a large number of interference signals. At the same time, the costs of remote sensing images monitoring are high, not applicable to long-term tracking, which results in remote sensing images data sparse, long time span, non-real-time and other issues.

Therefore, how to solve these problems, finding the starting and ending tracks of credible ship targets, is the focus of this article. In this article, only using electromagnetic signals data and extracting the effective characteristics of electromagnetic signal, through a priori knowledge, the electromagnetic signal track correlation algorithm based on attribute clustering and space-time constraints is proposed. We call our algorithm is ESTCA-ACSC. The main innovations in this article are as follows:

1) The dynamic clustering method of K-means is proposed;
2) The track correlation method of space-time constraint is proposed.

2. ESTCA-ACSC
As we all know, for electromagnetic signals, the electromagnetic signal not only contains the location information of the radiation source, but also contains the specific properties of the radiation source, such as...
as carrier frequency, repetition frequency, pulse width and other attribute information. It makes that possible to cluster with the electromagnetic signal attribute information, clustering into different clusters, and then in each cluster, according to the specified space-time constraints, we can find the electromagnetic signal track. This improves the accuracy of the electromagnetic signal track correlation. Based on this, an electromagnetic signal correlation algorithm based on attribute clustering and space-time constraints is proposed.

2.1. K-means dynamic clustering method
In this paper, the property value of electromagnetic signal is standardized to obtain the standard value of property, and then we using the standard value of property for K-means dynamic clustering, the main steps are as follows:

1) We obtain the property values of the electromagnetic signals, including the maximum, minimum and average of carrier frequency, pulse width, repetition frequency, and we carry out Z-score standardization on the property values of the electromagnetic signals. The z-score as follows:

\[ Z \text{-score} = \frac{Y - \text{mean}(Y)}{\text{SD}(Y)} \]  
(Eq.1)

Where \( Y \) represents the initial field value, \( \text{mean}(Y) \) represents the mean of the initial field value, \( \text{SD}(Y) \) represents the standard deviation of the initial field value, and \( Z \)-score represents the standardized value;

2) We adopt to the method that K-means dynamic clustering with standardized values, using the angled cosine similarity measurement method in dynamic clustering, the similarity measurement threshold is \( \alpha \), the Eq as follow:

\[ s = \frac{\sum_{k=0}^{n} x_{ik} \times x_{jk}}{\sqrt{\sum_{k=0}^{n} (x_{ik})^2} \times \sqrt{\sum_{k=0}^{n} (x_{jk})^2}} \]  
(Eq.2)

where \( x_{ik} \times x_{jk} \) represents the \( k \) part of the signal \( i \), \( x_{jk} \times x_{jk} \) represents the \( k \) part of the signal \( j \), and \( s \) is similarity;

3) We do that randomly select a record to be the initial cluster point;

4) We cycle each record to calculate the similarity of the current record to all cluster points that calculate result is \( S = \{s_1, s_2, ..., s_s\} \); 

5) If \( s \geq \alpha (s \in S) \), then \( s \in S' \) ( \( S' \) is a collection of similarity for all eligible conditions), and the record is grouped as a cluster represented by the maximum value of \( S' \); Cycle through all records to get \( N \) clusters;

6) We do that updating the cluster points with the average value of points in each cluster at the end of the loop;

7) We calculate the similarity between the cluster points of all clusters, if the similarity between the cluster points of any two clusters \( D1 \) and \( D2 \) is higher than the similarity threshold, then we judge the similarity between all records in cluster \( D1 \) and cluster points of cluster \( D2 \), and whether the similarity between all records in cluster \( D2 \) and the cluster points of cluster \( D1 \) meets the similarity measurement criteria, if the similarity measurement conditions are met, the two clusters are merged into the same cluster, otherwise they are not merged;

8) We determine whether the results after each merge cluster are the same as before the merge, and if so, end, otherwise repeat step 4;

9) Finally, the result of the clustering result is \( N \) different clusters, namely: \( A = \{X_1, X_2, ..., X_n\} \)

\( A = \{X_1, X_2, ..., X_n\} \).

2.2. Trace association based on space-time constraints method
According to the clustering results, we sort signal by radar type and time obtaining the sequenced signal
collection, according to the position information of the signal, calculate the velocity between the signals, angle characteristics. Under the condition of meeting the speed constraint, the track points with the highest speed and angle similarity are associated into a track. The main steps are as follows:

1) We use the results based on the clustering, where \( A = \{x_1, x_2, \ldots, x_n\} \) and \( X_i = \{x_i, x_{i+1}, \ldots, x_{n}\} \), \( x_i \) represents a collection of properties for the signal \( i, i \in [1, 2, \ldots, n] \). The \( X_i \) is that the bubble sort algorithm \( f(X) \) in chronological order, to obtain a new \( X_i' \);

2) In each \( X_i' \), we look for \( m \) signals in one cycle as the set of starting points of the track, \( M = \{x_1, x_2, \ldots, x_m\} \) and the other \( n - mn - m \) signals in the collection are the other point set \( N = \{x_{m+1}, x_{m+2}, \ldots, x_n\} \);

3) According to the position attribute of the signal, we calculate the velocity characteristics and angle characteristics between the signals, and the velocity equation is as follows:

\[
v = \frac{2 \text{arcsin} \left( \frac{\text{lat}_i - \text{lat}_j}{2} \right) + \cos(\text{lat}_i) \times \cos(\text{lat}_j) \times \sin \left( \frac{\ln g_i - \ln g_j}{2} \right)}{\left| t_i - t_j \right|} \quad (\text{Eq.3})
\]

where \( v \) represents the average velocity of the movement of the signal \( j \) to the signal \( i \), \( \text{lat}_i, \text{lat}_j \) indicates the latitude of the signal \( i, j \), \text{lng}_i, \text{lng}_j \) represents the longitude of the signal \( i, j \), \( R \) represents the radius of the earth, \( t_i, t_j \) represents the time of the signal \( i, j \).

The angle equation is as follows:

\[
\varphi = \frac{\text{arcsin} \left( \sin \left( \frac{90 - \text{lat}_i}{180} \pi \right) \sin \left( \frac{\ln g_i - \ln g_j}{180} \pi \right) \right)}{\pi} \times 180 \quad (\text{Eq.4})
\]

\[
\theta = \begin{cases} 
0 & (\text{lat}_i \geq \text{lat}_j, \ln g_i = \ln g_j) \\
90 & (\text{lat}_i = \text{lat}_j, \ln g_i > \ln g_j) \\
180 & (\text{lat}_i < \text{lat}_j, \ln g_i = \ln g_j) \\
270 & (\text{lat}_i = \text{lat}_j, \ln g_i < \ln g_j) \\
\varphi & (\text{lat}_i > \text{lat}_j, \ln g_i > \ln g_j) \\
180 - \varphi & (\text{lat}_i < \text{lat}_j, \ln g_i \neq \ln g_j) \\
360 + \varphi & (\text{lat}_i > \text{lat}_j, \ln g_i < \ln g_j)
\end{cases} \quad (\text{Eq.5})
\]

Among them, the direction angle from the signal \( j \) to the signal \( i \) is represented by \( \theta \), the latitude of the signal \( i \) is represented by \( \text{lat}_i, \text{lat}_j \), the latitude of the signal \( j \) is represented by \( \text{lat}_j, \text{lat}_j \), the longitude of the signal \( i \) is represented by \( \ln g_i, \ln g_j \), the longitude of the \( \ln g_j, \ln g_j \) is the longitude of the signal \( j \).

And the speed constraints are as follows:

\[
v_{\text{min}} \leq v \leq v_{\text{max}} \quad (\text{Eq.6})
\]
4) Each signal where our loop point set M and each signal where point set N, respectively, seek speed and angle, under the condition of meeting the speed constraints, then we judge the speed and angle of the most similar track as a confirmation track, wherein the signal is the starting point of the track, after the end of the cycle, obtain the N tracks that meet the conditions.

3. Experiments and analysis
In order to verify the performance of ESTCA-ACSC, two groups of comparative experiments are set up in this section. The first group of experiments compares ESTCA-ACSC with the track starting algorithm based on logical rules and the track starting algorithm based on Hough transformation on the simulated data set, and the second group of experiments uses a real electromagnetic signal data set to verify the time validity of ESTCA-ACSC. In two groups of experiments, the parameters of the algorithm are set to: angle cosine threshold $\theta_s \geq 0.5$, the velocity range is $v_{\text{min}} = 0, v_{\text{max}} = 60$.

3.1. Simulated simulation experiments

3.1.1. Simulation experimental conditions
The simulation scenario is as follows: in the sea area $100 \times 100 km^2$, it is assumed that there is an aircraft carrier, 2 cruisers, 2 frigates, 2 destroyers, 2 supply ships, different ships are equipped with different types of radar, the ship target in the rectangular coordinate system of the initial position as follows:

$\{(15,15);(6,47);(35,31);(35,18);(35,40);(64,64);(35,78);(64,32);(3,30)\}$

The target does a uniform-speed linear motion during the observation time, the observation time is 120 minutes, the target motion speed is between 15~35 km/h, assuming that the electromagnetic signals emitted by the targets can be detected all. The parameters captured by the ship's radar at different times are shown in the table as follows:

| Ship Type          | RF Mean (MHZ) | RF Max (MHZ) | RF Min (MHZ) | PRI Mean (μs) | PRI Max (μs) | PRI Min (μs) | PW Mean (ns) | PW Max (ns) | PW Min (ns) |
|--------------------|---------------|--------------|--------------|---------------|--------------|--------------|--------------|-------------|-------------|
| aircraft carrier   | 8750          | 8800         | 8700         | 666           | 668          | 664          | 100          | 120         | 80          |
| cruisers           | 8850          | 8900         | 8800         | 1724          | 1728         | 1720         | 500          | 502         | 498         |
| frigates           | 9500          | 9700         | 9300         | 1612          | 1620         | 1604         | 500          | 508         | 492         |
| destroyers         | 7250          | 7500         | 7000         | 769           | 770          | 768          | 250          | 260         | 240         |
| supply ships       | 3750          | 4000         | 3500         | 625           | 628          | 622          | 150          | 155         | 145         |

For the simple reasons, the system noise and the observed noise of this experiment are set to the average value as 0, the co-variance is $R_w$ and the Goss white noise is $Q_w$; assuming that the probability of clutter at different times assign the Poisson distribution, in the observation area assign uniformly distribution, and the signal parameters of clutter are set as follows:

| Table 2. clutter radiation source signal parameters |
|-----------------------------------------------|
| parameters         | mse     | mean       |
| RF                | 11460MHZ | 20.15GHZ   |
| PRI               | 2777.78μs | 1041.67μs |
| PW                | 570ns   | 720ns      |
3.1.2. Analysis of simulation results
In order to fully verify the performance of this algorithm, we compare the algorithm with the representative track association based on logical rules and the track association algorithm based on Hough transformation. For the target situation distribution map in Figure 1(a), ESTCA-ACSC as shown in Figure 1(b), track association results based on logical rules as shown in Figure 1(c), and track association results based on Hough transformation as shown in Figure 1(d), ESTCA-ACSC successfully associated 9 target tracks, after comparison, the track association of logical rules has a target loss, Hough transformation association results are false.

![Simulation results](a) target situation distribution map  ![Simulation results](b) ESTCA-ACSC  ![Simulation results](c) track association results based on logical rules  ![Simulation results](d) track association results based on Hough transformation

Figure 1. Simulation results

It can be seen from the experimental results that ESTCA-ACSC has a good track correlation effect compared with the logic rule method and the Hough transformation method.

3.2. Real data experiments
In order to verify the ability of the algorithm to solve practical problems and the applicability of the real environment, the actual data provided by X is used for experiments. The data set consists of real radar attribute information, longitude and latitude information, time and other information, and the experimental results of the data are as follows:
The experimental results of the real data are shown in Figure 2. Figure 2 (a) is the target situation distribution of the real data, Figure 2(b) is the result graph of ESTCA-ACSC, Figure 2(c) is the track association result graph based on logical rules, and Figure 2 (d) is the result map of the track association based on the Hough transformation.

From the experimental results, it can be seen that in the real scene, the electromagnetic signal is mixed with a large number of false alarm signals, and it is difficult to determine its track by the correlation of position information alone. In this case, ESTCA-ACSC proposed in this paper can successfully associate 7 target tracks, and the correct rate is higher than that of logical rule-based correlation method, while the Hough transformation method only associates 5 successful tracks. It can be seen that in the real environment of a large number of clutter waves, the algorithm can still get better correlation results. This experiment verifies the validity of the algorithm in this paper.

4. Related Work
At present, the vast majority of correlation algorithms are based on location information research, high-precision location information is the reliability and effectiveness of the correlation algorithm guarantee. Therefore, in the electromagnetic signal correlation algorithm, the primary problem to be solved is the extraction of high-precision position information. At present, the main track association algorithm is divided into two categories[1], one is to detect the track according to the order, the representative algorithm is: the track association based on logical rules[2-4], heuristic rule algorithm[5,6], etc., and the other is the starting track algorithm using batch processing technology, such as: the method based on Hough transformation[7-11].
4.1. The Track Association Based On Logical Rules

The track association algorithm based on logical rules is realized according to the motion characteristics and measurement error of the target. We assume it that $Z_m(i(t))$ is the jst state weight of the mth dot $Z_m(i(t))$ of the $i(t)$th moment, then the jst state weight $D_m(t(2))$ of $D_m(t(2))$ as follow:

$$D_m(t(2)) = \max \left( 0, Z_m(t(2)) - Z_m(t(1)) - V_{\text{max}}(t(2) - t(1)) \right)$$

$$+ \max \left( 0, Z_m(t(2)) - Z_m(t(1)) - V_{\text{max}}(t(2) - t(1)) \right)$$

(Eq.7)

We assume it that $Z_m(t(i))$ the co-variance of $Z_m(t(i))$ measurement error is $R_m(t(i))$ $R_m(t(i))$, then we get follow:

$$d_{mn}(t(2)) = \left[ D_m(t(2)) \right] \left[ R_m(t(1)) + R_m(t(2)) \right]^{-1} \left[ D_m(t(2)) \right]$$

(Eq.8)

We assume it that $d_{mn}(t(2)) = R_m(t(2))$ is the degree to which dot $Z_m(t(1))$ is associated with $Z_m(t(2))$. Since that, the general steps of the logical rule algorithm are as follows:

1) Using the point set information obtained from the first 2 cycles and the correlation correct rate constraints, we infer the possible starting track, and estimate the location information of the next scan, according to the time interval between points and points;

2) After completing the position estimate of the last step, the third scan is carried out, and the associated door is established with the prediction error, if the new point set is in the associated door, the track is disassembled into multiple tracks, and if the new point is not in the associated door, the track is terminated;

3) If more than 3 (including 3) or more points in time set are successfully associated, the location of the next interval is predicted by the proposed merge of the trail using the quadratic polynomial;

4) Constantly scan and repeat step 3.

4.2. The Method Based On Hough Transformation

The track association algorithm based on Hough transformation uses batch processing technology to accumulate a certain set of points over time, transforming from space domain to parameter domain, so as to determine the degree of correlation between points and points, and achieve the goal of track starting.

The extreme coordinates are Hough transformed so that the measured $Z_i = [x_i, y_i]$ is converted to $p_i = x_i \cos \theta_i + y_i \sin \theta_i$. The Hough transformation transforms Descartes coordinate system point information $(x_i, x_i, y_i)$ into spatial parameter coordinates $(\rho_i, \theta_i)$, and as shown in Figure 3, the line in the data space can be defined by the distance from the coordinate origin to the line $\rho_0 \rho_0$ and the angle $\theta_0 \theta_0$ of the $\rho_0 \rho_0$ and x-axis.

![Figure 3. Hough transform map](image-url)
Assuming that point set \( \{Z(t(1)), Z(t(2)), \ldots, Z(t(N))\} \) corresponds to the same target, traces in the point set sequence is connected together by an approximate straight line, which indicates that all traces will intersect at a point when they are transformed into parameter domain space, which represents the traces of the corresponding point set sequence in its space domain.

5. Conclusions
In order to further improve the accuracy of the electromagnetic signal track correlation algorithm, this paper proposes the electromagnetic signal track correlation algorithm based on attribute clustering and space-time constraints.

The method effectively utilizes the space-time constraint between the characteristics of the electromagnetic signals with carrier frequency, repetition frequency, pulse width and the point set. Finally obtains the optimal trace in the same cluster. Through the experiments about simulation data and real data, the algorithm of this paper is compared with the track association based on logical rules and the track association algorithm based on Hough transformation.

The performance advantage of this algorithm over the other two algorithms is verified.

References
[1] Poore A B. Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking[J]. (1994)Computational Optimization and Applications, 3(1): 27-57.
[2] Bar-Shalom Y, Fortman T E. Tracking and Data Association[J]. (1990)The Journal of the Acoustical Society of America, 87(2).
[3] Hu Z, Leung H, Blanchette M. Statistical performance analysis of track initiation techniques[J]. (1997)IEEE Transactions on Signal Processing, 45(2): 445-456.
[4] Yong P, Fu-Yu Y, Yan M. Analysis for Common Track Initiation Techniques[J]. (2006)Command Control & Simulation, 28(1): 98-101.
[5] Van Keuk G. Sequential track extraction[J]. (1998)IEEE Trans.aerosp. & Electron.syst, 34(4): 1135-1148.
[6] Tantan L, Ming L. Heuristic Algorithm-Based Initiation Method of Probability Hypothesis Density Filter for Target Tracking[J]. (2018)JOURNAL OF SHANGHAI JIAO TONG UNIVERSITY, 52(001): 63-69.
[7] Carlson, B. D, Evans, et al. Search radar detection and track with the Hough transform. III. Detection performance with binary integration[J]. Aerospace and Electronic Systems, (1994)IEEE Transactions on,.
[8] Carlson B D, Evans E D, Wilson S L. Search radar detection and track with the Hough transform. II. detection statistics[J]. (1994)IEEE Transactions on Aerospace and Electronic Systems, 30(1): 109-115.
[9] Chen J, Leung H. A modified probabilistic data association filter in a real clutter environment[J]. (1996)IEEE Transactions on Aerospace & Electronic Systems, 32(1): 300-313.
[10] Yun-Bo K, Xin-Xi F, Chuan-Guo L, et al. Track Initiation Algorithm for Passive Sensor System[J]. (2011)Opto-Electronic Engineering, 08: 64-70.
[11] Jian-Feng T, Xiu T. Passive sonar track initiation research based on hough transform[J]. (2008 )Ship Electronic Engineering, 028(001): 141-142.