

Abnormal Ship Behavior Detection after the Closure of AIS Based on Radar Data

Tao Zhang (✉ zhangtaozt01@163.com)

Beijing University of Posts and Telecommunications <https://orcid.org/0000-0001-5697-681X>

Chuangchang Liu

Beijing University of Posts and Telecommunications

Bodong Wen

Beijing 101 Middle School

Research Article

Keywords: Abnormal ship behavior detection, radar trajectory, abnormal speed, abnormal direction, the closure of AIS

Posted Date: June 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-551597/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abnormal Ship Behavior Detection after the Closure of AIS Based on Radar Data

Tao Zhang^{1*}, Chuanchang Liu¹, Bodong Wen²

¹State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, 100876, China

²Beijing 101 Middle School, Beijing, 100091, China

*Corresponding Author: zhangtaozt01@163.com

Abstract: In marine transportation, most ships are equipped with AIS devices. The AIS data sent by these devices can help maritime authorities to manage ships in relevant sea areas. However, AIS is a self-reporting system, when a ship is engaged in illegal activities, the AIS device may be turned off. Therefore, after the AIS is closed, if the ship's behavior during a certain period of time is different from the ship's behavior before the closure of AIS, the different behavior is likely to represent that the ship is conducting illegal activities. This behavior is considered abnormal and needs to be detected in time. Based on radar trajectory data, the detection of abnormal ship behavior is studied from two aspects: speed and direction. In order to improve the intelligent level of abnormal ship behavior detection, the abnormal speed behavior detection algorithm combined with rules and clustering (ASBD-RC) and the abnormal direction behavior detection algorithm combined with partition and the earth mover's distance (ADBBD-PE) are proposed. The ASBD-RC algorithm can reduce the influence of noise and sea clutter on abnormal speed behavior detection. The ADBBD-PE algorithm can effectively partition and identify trajectory segments with abnormal direction. In the experiment, based on the real and simulated radar trajectories, the abnormal behaviors of ships under different scenarios are generated. The experimental results show that in most scenarios, the ASBD-RC algorithm and the ADBBD-PE algorithm can effectively detect abnormal ship behavior. And compared with other algorithms, the proposed two algorithms have better and more stable detection results.

Keywords: Abnormal ship behavior detection; radar trajectory; abnormal speed; abnormal direction; the closure of AIS

1. Introduction

Sea freight is very important to the world economy. It carries more than 90% of global trade (Wang et al. 2020). At present, most ships are equipped with Automatic Identification System (AIS) devices. AIS devices can continuously provide dynamic and static information of ships. Dynamic information mainly includes the ship's position, speed, direction at a certain moment, and the static information mainly includes the ship's Maritime Mobile Service Identify (MMSI) number, type, length and width. After AIS data is processed, AIS trajectories can be generated. This information can help maritime authorities to better manage the ships in relevant sea areas.

When ships break down or engage in illegal activities, their behaviors are often different from the normal pattern, and these abnormal behaviors will threaten the navigation safety of ships. Therefore, it is an important task to detect abnormal ship behavior based on trajectory

data. Moreover, it is feasible to detect the abnormal behavior of moving targets through trajectory data (Piciarelli et al. 2008; Brun et al. 2014; Laxhammar and Falkman 2014). At this stage, there are some researches on detecting abnormal ship behavior based on AIS trajectory data. (Lei 2016) proposed a framework named MT-MAD for maritime trajectory modeling and anomaly detection. In this paper, anomalous behavior of ship was identified by three outlying features, including spatial feature, sequential feature, and behavioral feature. (Castaldo et al. 2016) presented a framework based on Bayesian networks for the surveillance of public transportation sites, and this framework was tested with trajectory data of ships. In the experiment, the authors studied the identification of abnormal and possibly dangerous situations in a canal. (Tu et al. 2018) thought that ship anomaly can be classified into three types: position anomaly, speed anomaly, and time anomaly. (Arguedas et al. 2018) proposed a maritime traffic knowledge discovery and representation system, and the generated maritime traffic network can be used to detect abnormal behaviors of vessels. (Yan et al. 2020) proposed a maritime route extraction method based on AIS data, and the maritime traffic routes extracted can be used for abnormal ship behavior detection. (Wang et al. 2020) proposed a parameter-light rare behavior detection method and presented experiments on simulated data sets and real AIS data sets. Experiments showed that the method can identify rare behaviors such as off-course vessel tracks. (Mascaro 2014) explored the use of Bayesian networks for analyzing vessel behavior and detecting anomalies, and the effectiveness of the method is verified by the anomalous data generated. (Fu et al. 2017) proposed a novel abnormal trajectory detection method, this method can detect abnormal vessel trajectories via feature learning algorithm. (Kazemi 2013) demonstrated that using open data can improve the efficiency of the maritime surveillance systems, and the abnormal behavior of ships was defined by experts. (Rong et al. 2020) established the normalcy models of the maritime traffic based on historical AIS data, and according the normalcy models, ship trajectory anomaly detection can be conducted. (Huang et al. 2020) proposed a semantic analysis methodology to identify pattern knowledge from ship trajectories, which facilitates the maritime administrators to identify abnormal behavior. (Sidibé and Shu 2017; Riveiro et al. 2018) summarized the researches on abnormal ship behavior detection.

In most researches, the main idea is to discover the normal behavior pattern of ships through historical trajectory data, and then the behavior that deviates from the normal pattern is detected as abnormal behavior. The data used in most researches are AIS data of ships, but AIS is a self-reporting system. When a ship is engaged in illegal activities, the crew may turn off the AIS device (Mazzarella et al. 2017). At this time, maritime authorities can no longer obtain the ship's AIS data and need to use other sensors to monitor the ship's navigation status. Among them, radar is often used as an alternative source for surveillance near the ports (Habtemariam et al. 2012). The radar can obtain the position and radial velocity of ships. After the radar data is processed, radar trajectories can also be generated.

Since there is no uniform definition of abnormal ship behavior, the abnormal ship behavior studied in the paper should be defined. As mentioned above, it is possible for a ship to close its AIS device when conducting illegal activities. Therefore, after the AIS is closed, if the ship's behavior during a certain period of time is different from the behavior before the closure of AIS, we have reason to believe that the different behavior during this period may represent that the ship is engaged in illegal activities. This behavior is considered abnormal

and should be detected in time. Since the AIS trajectory can no longer be obtained after the AIS is closed, this paper mainly studies the detection of abnormal ship behavior based on radar trajectory data. Unless otherwise specified, the trajectory mentioned in the following content is the radar trajectory. In this paper, our research object is each radar trajectory of the ship. Research on each radar trajectory can avoid the use of a large amount of historical trajectory data to obtain the normal behavior pattern of ships, which can improve the efficiency of abnormal ship behavior detection.

According to (Lei 2016), the behavior of a ship can be represented by a behavioral vector, which mainly includes speed and direction. Therefore, in this paper, abnormal ship behavior is divided into abnormal speed behavior and abnormal direction behavior. It is inefficient and inaccurate to detect abnormal ship behavior only by manual observation. Therefore, in order to improve the intelligent level of abnormal ship behavior detection, this paper designs ASBD-RC and ADBD-PE algorithms, and proves the effectiveness of the two algorithms through experiments. The main contributions of this paper are as follows.

(1) Due to the influence of noise and sea clutter, the positions of ships detected by radar are not all accurate, and the computed speed of trajectory points based on the positions will also fluctuate greatly. At this time, the trajectory points with abnormal speed detected only by setting rules are not accurate. In order to solve this problem, combining rule-based algorithm and clustering algorithm, the abnormal speed behavior detection algorithm is proposed, abbreviated as ASBD-RC. First, a rule-based algorithm is used to detect trajectory points with abnormal speed. Then the time interval between trajectory points is regarded as the distance between them, and the DBSCAN clustering algorithm is used to cluster the detected trajectory points. In this way, on the one hand, the detected points which are relatively concentrated in time are more likely to be connected into abnormal trajectory segments. On the other hand, the trajectory points which are relatively scattered in time are more likely to be identified as noise. Therefore, the ASBD-RC algorithm can alleviate the impact of speed fluctuation on abnormal speed behavior detection.

(2) When a ship changes its direction while sailing, the trajectory will be curved. In order to take advantage of this feature, when detecting abnormal direction behavior, the radar trajectory can be partitioned first, so that the trajectory can be divided into different segments according to different directions. Then, whether the direction of each trajectory segment is significantly different from the direction of the normal trajectory segment is judged. Since there are many points in each trajectory segment, and each point has a direction, the direction of each trajectory segment is represented by a set containing multiple direction values. At this time, when computing the difference between the directions of two trajectory segments, a simple statistical comparison is not accurate. Therefore, combining with trajectory partition and the Earth Mover's distance, the abnormal direction behavior detection algorithm is proposed, abbreviated as ADBD-PE. In the ADBD-PE algorithm, the Earth Mover's distance is used to measure the difference between the directions of two trajectory segments. The ADBD-PE algorithm can detect the abnormal direction behavior of ships more accurately.

(3) In real scenes, there are very few ship trajectory segments with abnormal behavior, which brings difficulties to the verification of the algorithms. In order to solve this problem, based on real and simulated radar trajectory data, we generate datasets with abnormal ship

behavior under different combinations of conditions. These datasets can help us verify the effectiveness of the algorithms more comprehensively.

2. Problem Statement

2.1 Definition of Speed and Direction

The behavior of a ship in navigation can be represented by speed and direction. In this paper, the speed and direction of the ship at a certain time are computed based on the positions of adjacent radar trajectory points. *Point 2* in Figure 1 is taken as an example, and the time of *Point 2* is *time 2*. Considering that when computing the speed and direction of the ship at *time 2*, the position of the ship at the next time may not be known. Therefore, the previous point of *Point 2* (*Point 1*) is used, and the time of *Point 1* is *time 1*. The speed of the ship at *time 2* is $d12/(time\ 2-time\ 1)$, and $d12$ is the distance between *Point 1* and *Point 2*. The angle rotated clockwise from true north to the vector from *Point 1* to *Point 2* is regarded as the direction of the ship at *time 2*, as shown by angle *A* in Figure 1.

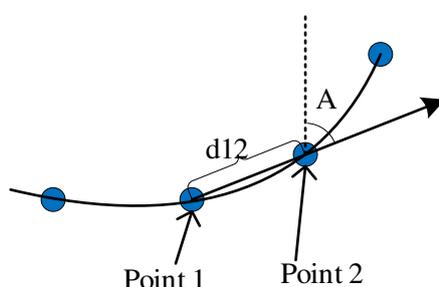


Fig. 1 Definition of speed and direction

2.2 Abnormal Speed Behavior

Figure 2(a) shows an example of abnormal speed behavior. In the initial stage of the ship's navigation, the ship's AIS device is in normal operation. At this time, both AIS receiver and radar can detect the position of the ship. The AIS trajectory (as shown by the red line in Figure 2(a)) and the radar trajectory (as shown by the blue line in Figure 2(a)) are formed. When the ship travels to *A*, the AIS device is turned off, at this time, the AIS receiver can no longer obtain the position of the ship, and the AIS trajectory is interrupted. The radar can still detect the position of the ship, and the radar trajectory continues to be generated. When the ship sails to *B*, the speed of the ship is changed (deceleration). This behavior is obviously different from the behavior before the closure of AIS. This means that the ship may be engaged in illegal activities, such as illegal fishing and illegal surveillance.

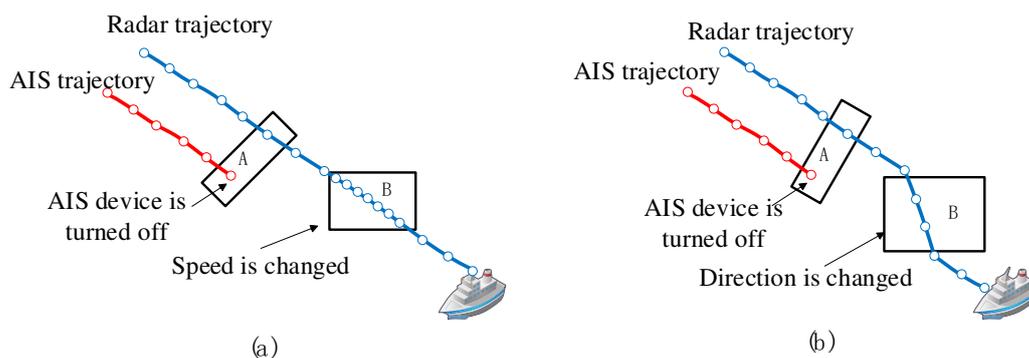


Fig. 2 Abnormal ship behavior: (a) Abnormal speed behavior. (b) Abnormal direction behavior

2.3 Abnormal Direction Behavior

Figure 2(b) shows an example of abnormal direction behavior. Similar to the previous example of abnormal speed behavior, when the ship reaches A , the AIS device is turned off. When the ship sailed to B , the ship's sailing direction is changed. This behavior may represent that the ship is involved in illegal activities, such as illegal entry into the prohibited area and deviation from the standard sea lane.

3 Trajectory Preprocessing

The i (th) radar trajectory is represented as TRR_i , $TRR_i = \{PR_{i(0)}, PR_{i(1)}, \dots, PR_{i(j)}, \dots, PR_{i(n-1)}\}$. $PR_{i(j)}$ is the $(j+1)$ th radar trajectory point in TRR_i . $PR_{i(j)} = \{trId, lon, lat, time\}$. Among them, $trId$ is the unique identification of TRR_i , lon , lat and $time$ are longitude, latitude and timestamp of $PR_{i(j)}$ respectively.

3.1 Trajectory Smooth

Due to the influence of noise and clutter, the data detected by radar will have measurement errors and false alarms. This will cause the "sawtooth" phenomenon in the generated radar trajectory. Therefore, the moving average (MA) algorithm is used to smooth the radar trajectory. The formulas are shown below.

$$PR_{i(j)}^s.lon = \frac{1}{2k+1} * (PR_{i(j+k)}.lon + PR_{i(j+k-1)}.lon + \dots + PR_{i(j-k)}.lon) \quad (1)$$

$$PR_{i(j)}^s.lat = \frac{1}{2k+1} * (PR_{i(j+k)}.lat + PR_{i(j+k-1)}.lat + \dots + PR_{i(j-k)}.lat) \quad (2)$$

$$PR_{i(j)}^s.time = PR_{i(j)}.time \quad (3)$$

$PR_{i(j)}^s$ is the smoothed trajectory point of $PR_{i(j)}$. $k=1,2,\dots$. The i th radar trajectory after smoothing is represented as TRR_i^s , $TRR_i^s = \{PR_{i(0)}^s, PR_{i(1)}^s, \dots, PR_{i(j)}^s, \dots, PR_{i(n-1)}^s\}$.

3.2 Setting of Equal Time Interval

In the generated radar trajectory, the time interval between adjacent points is different. If the time interval between trajectory points is too small, according to the speed computation formula, measurement errors and false alarms will have a great impact on the accuracy of speed. Therefore, mean filtering and interpolation are used to set the time interval between adjacent radar trajectory points to be equal.

Assuming that the starting time of the radar trajectory is $T0$ and the set time interval is t , the timestamps of the points in the radar trajectory after setting are $T0, T0+t, T0+2*t, \dots, T0+l*t, \dots, T0+(m-1)*t$. Among them, $T0+l*t$ is the timestamp of the $(l+1)$ th point in the new trajectory, and m is the number of points in the new trajectory. The i th new trajectory is represented by TR_i , $TR_i = \{P_{i(0)}, P_{i(1)}, \dots, P_{i(l)}, \dots, P_{i(m-1)}\}$. The algorithm for computing the longitude and latitude of $P_{i(l)}$ in the new trajectory is shown in Algorithm 1. First, the points in the trajectory TRR_i^s whose timestamps are between $(T0+l*t-t/2)$ and $(T0+l*t+t/2)$ are searched and stored in *radarList* (line 1 in Algorithm 1). If *radarList* is not empty, the mean filtering is used, that is, the mean longitude and mean latitude of all radar

points in *radarList* are used as the longitude and latitude of $P_{i(l)}$, respectively (lines 2-4 in Algorithm 1). If *radarList* is empty, the interpolation is used to compute the longitude and latitude of $P_{i(l)}$. The point whose timestamp is before $T0+l*t$ and closest to $T0+l*t$ in TRR_i^s is recorded as PR_b , and the point whose timestamp is after $T0+l*t$ and closest to $T0+l*t$ in TRR_i^s is recorded as PR_a (lines 6-7 in Algorithm 1). The longitude (latitude) of $P_{i(l)}$ is the weighted sum of the longitude (latitude) of PR_b and the longitude (latitude) of PR_a (lines 8-9 in Algorithm 1). From the formulas of $a1$ and $a2$ (lines 10-11 in Algorithm 1), it can be known that when computing the longitude and latitude of $P_{i(l)}$, the weight of the point whose timestamp is close to $(T0+l*t)$ is larger.

Algorithm 1 Algorithm for computing the longitude and latitude of $P_{i(l)}$.

Require: The $(i)th$ radar trajectory after smoothing, TRR_i^s ; Timestamp of trajectory point $P_{i(l)}$, $T0+l*t$; The set time interval, t .

Output: Longitude and latitude of trajectory point $P_{i(l)}$: $P_{i(l)}.lon$, $P_{i(l)}.lat$.

```

1: radarList ← searchRadarPoints( $TRR_i^s, T0+l*t-t/2, T0+l*t+t/2$ )
2: if radarList is not null then
3:    $P_{i(l)}.lon$  ← computeMeanLon(radarList)
4:    $P_{i(l)}.lat$  ← computeMeanLat(radarList)
5: else
6:    $PR\_b$  ← searchBeforePoint( $TRR_i^s, T0+l*t$ )
7:    $PR\_a$  ← searchAfterPoint( $TRR_i^s, T0+l*t$ )
8:    $P_{i(l)}.lon$  ←  $a1 * PR\_b.lon + a2 * PR\_a.lon$ 
9:    $P_{i(l)}.lat$  ←  $a1 * PR\_b.lat + a2 * PR\_a.lat$ 
10:   $a1$  ←  $(PR\_a.time - (T0+l*t)) / (PR\_a.time - PR\_b.time)$ 
11:   $a2$  ←  $((T0+l*t) - PR\_b.time) / (PR\_a.time - PR\_b.time)$ 
12: end if

```

4 Algorithm

As mentioned above, the $(i)th$ radar trajectory can be represented by TR_i , $TR_i = \{P_{i(0)}, P_{i(1)}, \dots, P_{i(l)}, \dots, P_{i(m-1)}\}$. Assuming that the corresponding radar trajectory point when the AIS is closed is $P_{i(l)}$, where $0 < l < (m-1)$. Then the radar trajectory segment $\{P_{i(0)}, P_{i(1)}, \dots, P_{i(l-1)}\}$ is the radar trajectory before AIS is closed (*rdTrBeforeAC*), and the radar trajectory segment $\{P_{i(l)}, P_{i(l+1)}, \dots, P_{i(m-1)}\}$ is the radar trajectory after AIS is closed (*rdTrAfterAC*). The behavior of the ship in *rdTrBeforeAC* is considered normal, and we need to detect whether the ship has abnormal behavior in *rdTrAfterAC*. Abnormal ship behavior is divided into abnormal speed behavior and abnormal direction behavior. The overall framework of the abnormal ship behavior detection algorithm is shown in Figure 3.

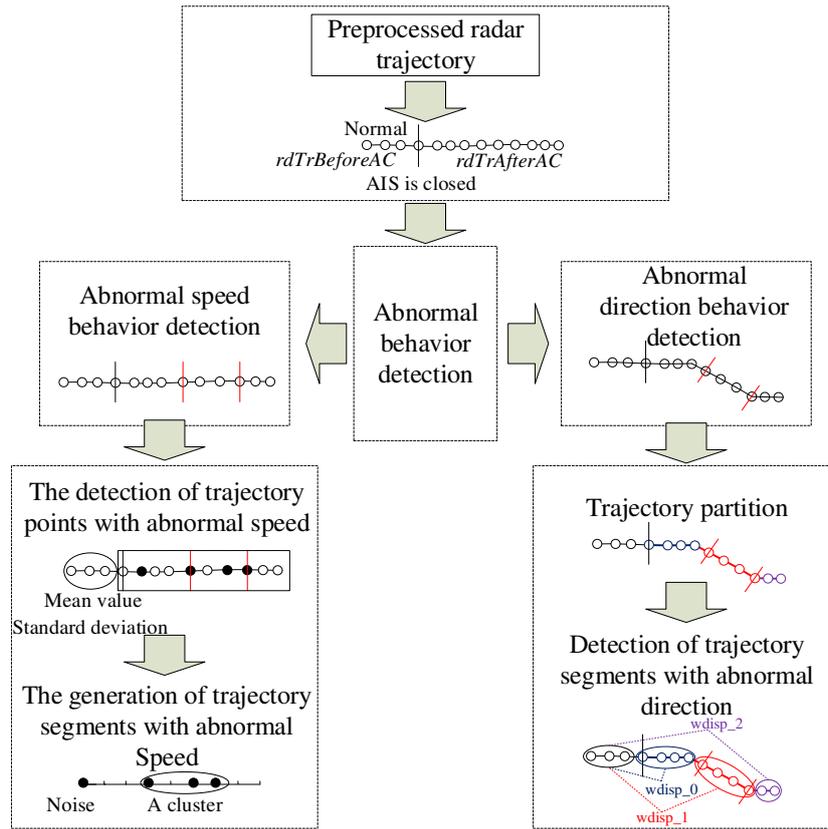


Fig. 3 The overall framework of the abnormal ship behavior detection algorithm

4.1 Abnormal Speed Behavior Detection Algorithm

According to the definition of abnormal ship behavior, when detecting abnormal speed behavior, first, the speed in $rdTrBeforeAC$ is considered normal, and we need to detect whether there are trajectory segments in $rdTrAfterAC$ whose speeds are significantly different from the normal speed. An intuitive way is to detect based on rules. For example, the normal speed range can be set according to the distribution of normal speed. Trajectory points whose speed values are outside the normal range are recognized as abnormal. However, as mentioned above, due to the influence of noise and sea clutter, the computed speed fluctuates greatly, and the trajectory points with abnormal speed detected only by the rules are inaccurate and discrete. Therefore, combining rule-based algorithm and clustering algorithm, the abnormal speed behavior detection algorithm is proposed, abbreviated as ASBD-RC. The ASBD-RC algorithm is divided into two steps, as shown in Figure 3. The first step is the detection of trajectory points with abnormal speed. This step is mainly achieved by setting rules. The second step is the generation of trajectory segments with abnormal speed. This step is mainly realized by clustering the detected trajectory points based on the DBSCAN algorithm. Algorithm 2 is the pseudo code of the ASBD-RC algorithm.

Algorithm 2 The ASBD-RC algorithm.

Require: Radar trajectory before AIS is closed, $rdTrBeforeAC$; Radar trajectory after AIS is closed, $rdTrAfterAC$; Parameter when detecting abnormal speed trajectory points, $ratio$; Parameters of the DBSCAN algorithm, Eps , $MinPts$.

Output: Trajectory segments with abnormal speed, $abnormalSSegs$.

```

1:  $meanSpeedB \leftarrow getMeanSpeed(rdTrBeforeAC)$ 
2:  $stdSpeedB \leftarrow getStdSpeed(rdTrBeforeAC)$ 
3:  $meanSpeedA \leftarrow getMeanSpeed(rdTrAfterAC)$ 
4:  $abnormalPts \leftarrow []$ 
5: if  $meanSpeedA > meanSpeedB$  then
6:   for each point  $radarP$  in  $rdTrAfterAC$  do
7:     if  $radarP.speed > meanSpeedB + ratio * stdSpeedB$  then
8:       add  $radarP$  into  $abnormalPts$ 
9:     end if
10:  end for
11: else
12:  for each point  $radarP$  in  $rdTrAfterAC$  do
13:    if  $radarP.speed < meanSpeedB - ratio * stdSpeedB$  then
14:      add  $radarP$  into  $abnormalPts$ 
15:    end if
16:  end for
17: end if
18:  $clusters \leftarrow DBSCAN(abnormalPts, Eps, MinPts)$ 
19:  $abnormalSSegs \leftarrow []$ 
20: for each cluster  $CL$  in  $clusters$  do
21:   Get the minimum serial number in  $CL$ :  $startNum$ 
22:   Get the maximum serial number in  $CL$ :  $endNum$ 
23:   The points between  $startNum$  and  $endNum$  are added to  $abnormalSSegs$ 
24: end for

```

4.1.1 The Detection of Trajectory Points with Abnormal Speed

When detecting trajectory points with abnormal speed, a rule-based algorithm is used. First, the mean value and standard deviation of the speed of the points in $rdTrBeforeAC$ need to be computed, and they are named $meanSpeedB$ and $stdSpeedB$ respectively (line 1 and line 2 in Algorithm 2). Then, the mean value of the speed of the points in $rdTrAfterAC$ is computed, which is named $meanSpeedA$ (line 3 in Algorithm 2). If $meanSpeedA$ is greater than $meanSpeedB$, it means that the ship is more likely to accelerate after the AIS is closed, at this time, the points in $rdTrAfterAC$ whose speeds are greater than $(meanSpeedB + ratio * stdSpeedB)$ are regarded as trajectory points with abnormal speed; Otherwise, it means that the ship is more likely to decelerate, at this time, the points in $rdTrAfterAC$ whose speeds are less than $(meanSpeedB - ratio * stdSpeedB)$ are regarded as trajectory points with abnormal speed (lines 5-17 in Algorithm 2). Among them, $ratio$ is a parameter that needs to be set.

Figure 4 shows an example of a ship with abnormal speed behavior, and the ship's navigation direction is shown by the arrow in Figure 4. When the ship sails to the black line

segment, its AIS device is turned off; when it reaches the first red line segment, its speed starts to change; when it travels to the second red line segment, the change in speed is over.

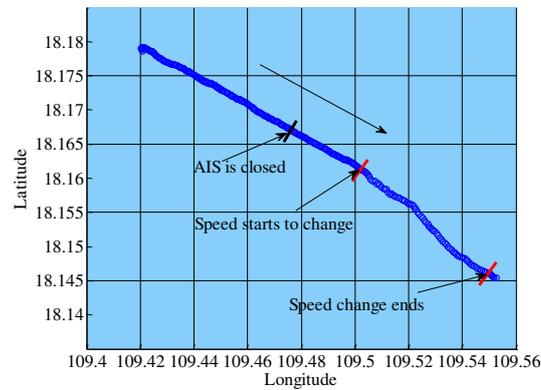


Fig. 4 An example of a ship with abnormal speed behavior

Figure 5 shows the detection process of trajectory points with abnormal speed in the radar trajectory in Figure 4. X axis in Figure 5 represents the serial number of the points in the radar trajectory, and the serial number starts from 0. Y axis represents the speed value, and its unit is m/s . The X value of the black vertical line in Figure 5 is the serial number of the corresponding radar trajectory point when AIS is closed. The radar trajectory points between the two red vertical lines are the points where the speed changes, that is, the trajectory segment composed of these points is abnormal and needs to be detected. The Y value of the red horizontal solid line is $meanSpeedB$, and the Y value of the red horizontal dashed line is $(meanSpeedB+ratio*stdSpeedB)$, at this time, $ratio=0.7$. Since $meanSpeedA$ is greater than $meanSpeedB$, the points in $rdTrAfterAC$ whose speeds are greater than $(meanSpeedB+ratio*stdSpeedB)$ are regarded as abnormal trajectory points, as shown by the red dots in Figure 5. In Figure 5, we use black dots to mark the serial numbers of the detected abnormal trajectory points, as shown in the black rectangle. It can be seen from Figure 5 that the detected points are not continuous in the sequence.

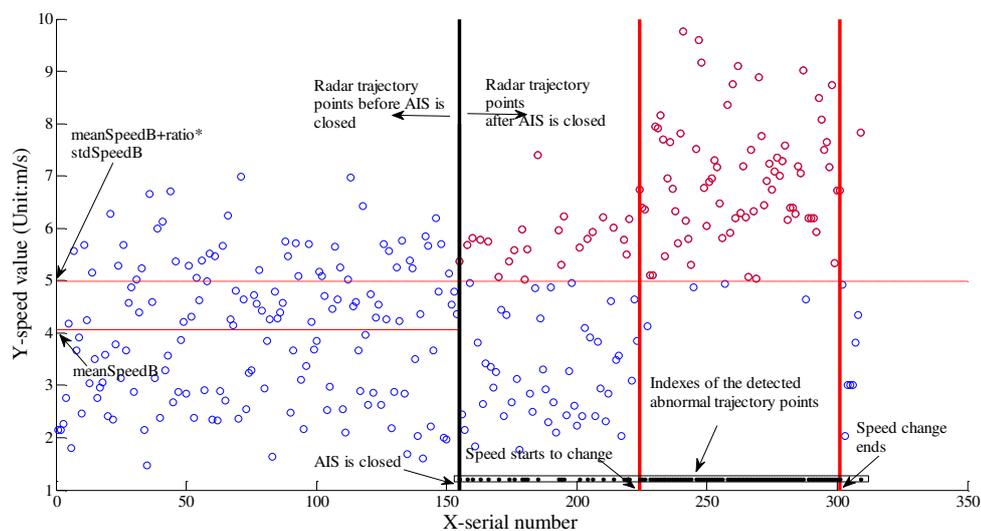


Fig. 5 The detection process of trajectory points with abnormal speed

4.1.2 The Generation of Trajectory Segments with Abnormal Speed

In actual navigation, after a ship changes its speed, it will often continue to sail for a period of time at the changed speed, and a trajectory segment with abnormal speed is generated. However, due to the influence of noise and sea clutter, the positions of ships detected by radar are not all accurate, and the computed speed of trajectory points based on the positions will also fluctuate greatly, as shown in Figure 5. If we judge whether the ship has abnormal speed behavior only by the set rules, the detected abnormal trajectory points will not perfectly match the real abnormal trajectory points. Abnormal trajectory points can also be detected in the normal trajectory segment, and the distribution of detected points is scattered. In the real abnormal trajectory segment, the distribution of detected points is more concentrated, as shown in Figure 5. Therefore, the abnormal trajectory points detected need to be further processed, on the one hand, abnormal trajectory points that are close in time should be connected into abnormal trajectory segments, and on the other hand, isolated trajectory points should be eliminated. At this time, we can take the time interval between trajectory points as the distance between them, and use DBSCAN algorithm (Ester et al. 1996) to cluster the detected trajectory points. The DBSCAN (Density Based Spatial Clustering of Applications with Noise) algorithm is a density-based clustering algorithm, which can divide high-density areas into clusters and detect noise. The DBSCAN algorithm can meet our above-mentioned needs.

Figure 6 is an example of clustering abnormal trajectory points using the DBSCAN algorithm. During the clustering process, we can take the time interval between trajectory points as the distance between them. Since the time interval between adjacent trajectory points is set to be equal, for the convenience of computation, we can regard the interval between the serial numbers of trajectory points as the distance between them. The numbers in Figure 6 represent the serial numbers of the points in the radar trajectory. The trajectory points are represented by P_0, P_1, \dots, P_8 , and the black dots represent the abnormal trajectory points detected. In this example, the scan radius (Eps) is set to 2, and the minimum number of points included ($MinPts$) is set to 3. Suppose the trajectory points are visited in order of serial number. The process of clustering is as follows.

(1) P_0 is visited, $|N_2(P_0)| < MinPts$, therefore, P_0 is set as noise. (2) P_3 is visited, $|N_2(P_3)| < MinPts$, therefore, P_3 is set as noise. (3) P_5 is visited, $N_2(P_5) = \{P_3, P_5, P_6\}$, $|N_2(P_5)| = MinPts$, therefore, P_3, P_5, P_6 are classified into *Cluster 0*. At this time, the cluster should be expanded. $N_2(P_6) = \{P_5, P_6, P_8\}$, $|N_2(P_6)| = MinPts$, therefore, P_8 is classified into *Cluster 0*. $|N_2(P_8)| < MinPts$, the cluster cannot be expanded anymore, and the clustering process ends.

It can be seen from the clustering results that P_0 is a noise point, and P_3, P_5, P_6 and P_8 belong to the same cluster. After clustering the abnormal trajectory points using the DBSCAN algorithm, we obtain the minimum and maximum serial numbers of trajectory points in each cluster, and the trajectory segment between the two serial numbers is regarded as the abnormal trajectory segment (lines 20-24 in Algorithm 2). In Figure 6, the trajectory segment between P_3 and P_8 is regarded as the trajectory segment with abnormal speed.

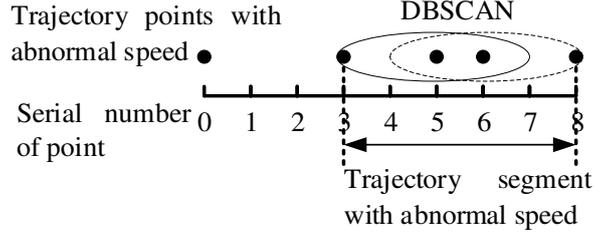


Fig. 6 Application of DBSCAN

4.2 Abnormal Direction Behavior Detection Algorithm

Similar to the detection of abnormal speed behavior, when detecting abnormal direction behavior, first, the direction in $rdTrBeforeAC$ is considered normal, and we need to detect whether there are trajectory segments in $rdTrAfterAC$ whose directions are significantly different from the normal direction. As mentioned above, the shape of the trajectory will change after the ship changes its sailing direction. In order to make full use of this feature and determine whether the direction of each trajectory segment is abnormal more accurately, combining with trajectory partition and the Earth Mover's distance, the abnormal direction behavior detection algorithm is proposed, abbreviated as ADBD-PE. The ADBD-PE algorithm is also divided into two steps, as shown in Figure 3. The first step is the trajectory partition based on the Douglas-Peucker algorithm. This step is mainly to divide the trajectory into different segments according to different directions. The second step is the detection of trajectory segments with abnormal direction. This step is mainly based on the Earth Mover's distance to determine whether the trajectory segment has abnormal direction behavior. Algorithm 3 is the pseudo code of the ADBD-PE algorithm.

Algorithm 3 The ADBD-PE algorithm.

Require: Radar trajectory before AIS is closed, $rdTrBeforeAC$; Radar trajectory after AIS is closed, $rdTrAfterAC$; Parameter when detecting trajectory segments with abnormal direction, $ratio$.

Output: Trajectory segments with abnormal direction, $abnormalDSegs$.

```

1:  $wdis \leftarrow computeEMD(rdTrBeforeAC, rdTrAfterAC)$ 
2:  $rdTrAfterACSegs \leftarrow partitionByDP(rdTrAfterAC)$ 
3:  $abnormalDSegs \leftarrow []$ 
4: for each segment  $rdTrAfterACSeg_i$  in  $rdTrAfterACSegs$  do
5:    $wdisp_i \leftarrow computeEMD(rdTrBeforeAC, rdTrAfterACSeg_i)$ 
6:   if  $wdisp_i > ratio * wdis$  then
7:     Add all points in  $rdTrAfterACSeg_i$  to  $abnormalDSegs$ 
8:   end if
9: end for

```

4.2.1 Trajectory Partition

Figure 7 shows an example of a ship with abnormal direction behavior, specifically, when the ship is driving to the black line segment, its AIS device is turned off; when it reaches the first red line segment, its direction starts to change; when it travels to the second

red line segment, its direction change ends. It can be seen from Figure 7 that when the ship changes direction while sailing, the trajectory will be curved. Therefore, when detecting whether the ship changes direction after the closure of AIS, the radar trajectory after the closure of AIS can be partitioned first, and the trajectory points with changes in direction can be partitioned from other trajectory points, which is convenient for further analysis. When partitioning the trajectory, the Douglas-Peucker algorithm (Douglas et al. 1973) is used. The specific partitioning process is as follows.

(1) As mentioned above, the radar trajectory after the closure of AIS is named *rdTrAfterAC*, and a straight line *AB* is connected between the starting point *A* and the ending point *B* of *rdTrAfterAC*. (2) The distance from each point in *rdTrAfterAC* to the straight line *AB* is computed, where the maximum distance is represented by *maxD* and the corresponding point is represented by *C*. (3) If $maxD \leq threshold$, then point *A* and point *B* are the partition points of *rdTrAfterAC*, and the partitioning process is complete. (4) If $maxD > threshold$, point *C* is regarded as a partition point, and the trajectory is divided into two segments, namely the trajectory segment from *A* to *C* and the trajectory segment from *C* to *B*. Then the processing steps (1)-(4) are performed on the two trajectory segments. (5) Until all the trajectory segments are processed, the algorithm ends.

After the Douglas-Peucker algorithm is used to partition the radar trajectory in Figure 7, the partition points can be obtained, as shown by the red dots. It can be seen from Figure 7 that the trajectory segment between *Point 1* and *Point 2* is the trajectory segment with direction change, and each segment contains only trajectory points whose directions have changed or only trajectory points whose directions have not changed. Next, we can analyze each trajectory segment and judge whether it is a trajectory segment with abnormal direction.

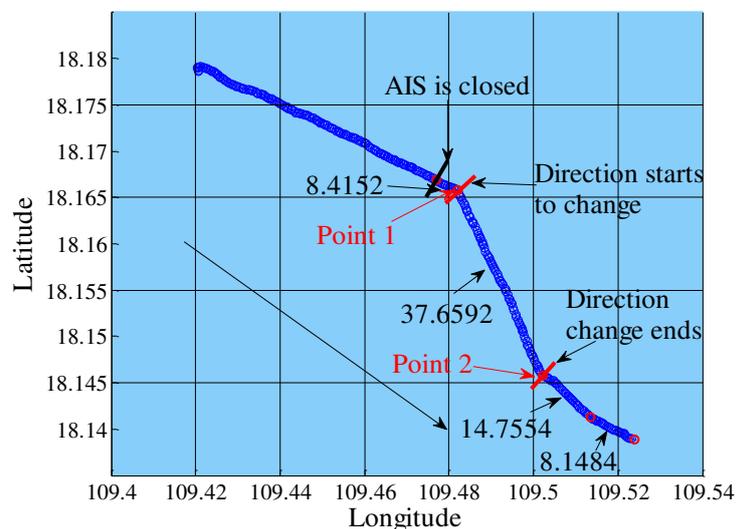


Fig.e 7 An example of a ship with abnormal direction behavior

4.2.2 Detection of Trajectory Segments with Abnormal Direction

When judging whether each trajectory segment has abnormal direction behavior, we need to know whether the direction of each trajectory segment is significantly different from the direction of *rdTrBeforeAC*. Since each trajectory segment contains many points, and each

point has a direction, what we need to compare is the distance between two distributions. The Earth Mover's distance can be used to measure the distance between two distributions. It reflects the minimal amount of work that must be performed to transform one distribution into the other by moving "distribution mass" around (Rubner et al. 1998). The detection process is as follows.

(1) The Earth Mover's distance $wdis$ between the direction of $rdTrBeforeAC$ and the direction of $rdTrAfterAC$ is computed (line 1 in Algorithm 3). (2) After $rdTrAfterAC$ is partitioned, each trajectory segment is represented by $rdTrAfterACSeg_i$, where $i=0,1,2,\dots,n-1$, and n is the number of segments. The set of these trajectory segments is represented by $rdTrAfterACSegs$ (line 2 in Algorithm 3). (3) The Earth Mover's distance $wdisp_i$ between the direction of $rdTrAfterACSeg_i$ and the direction of $rdTrBeforeAC$ is computed. If $wdisp_i > ratio * wdis$, the corresponding trajectory segment is considered as a trajectory segment with abnormal direction (lines 4-9 in Algorithm 3).

In the example in Figure 7, the $wdisp_1$ of the trajectory segment between *Point 1* and *Point 2* is the largest, which is 37.6592. In this example, $wdis=22.8412$, therefore, as long as the appropriate *ratio* is set, the abnormal trajectory segment between *Point 1* and *Point 2* can be correctly detected.

5 Case Study

In the experiment, the real radar trajectory dataset and the simulated radar trajectory dataset are used. The radar trajectory data mainly contains four fields: trajectory identification number, longitude, latitude, and timestamp. The real radar trajectory dataset is collected from the waters near Sanya, China, and this dataset contains 40 radar trajectories. After the trajectory preprocessing process, these radar trajectories contain 6922 radar points. In addition, we also generate a simulated radar trajectory dataset based on real AIS trajectory data, and the specific generation process can refer to (Zhang et al. 2020). The simulated dataset contains 205 radar trajectories. After the trajectory preprocessing process, these radar trajectories contain 47716 radar points.

5.1 Experiment on Abnormal Speed Behavior Detection

Due to the lack of trajectory segments with obvious speed changes in the original dataset, in this experiment, we first need to generate corresponding abnormal data. When evaluating the ASBD-RC algorithm, the *FI* score is used, which can take precision and recall into consideration. In order to prove the necessity of the ASBD-RC algorithm, in the experiment, we compare the detection effect of the ASBD-RC algorithm with that of the other two more intuitive algorithms.

5.1.1 Generation of Trajectory Segments with Abnormal Speed Behavior

In order to verify the algorithm more comprehensively, it is necessary to generate trajectory segments with abnormal speed behavior under different combinations of conditions. In this generation process, three conditions are considered. The first condition is the time when AIS is closed. This condition can be expressed as a , $a = n0/N$. Among them, N represents the number of points in the radar trajectory, and $n0$ represents the number of radar points that the ship has passed when the AIS is closed. The larger the value of a , the later the AIS is closed, and vice versa. The second condition is the duration of the trajectory segment

with abnormal speed behavior. This condition can be expressed as b , $b = n1/N$. Among them, N represents the number of points in the radar trajectory, and $n1$ represents the number of points in the trajectory segment with abnormal speed behavior. The larger the value of b , the longer the duration of the trajectory segment with abnormal speed behavior, and vice versa. The third condition is the amount in the change of speed, which can be represented by c . Therefore, the dataset with abnormal speed behavior generated under each combination of conditions can be represented by $abnormalSpeedDatasetR(a,b,c)$ or $abnormalSpeedDatasetS(a,b,c)$. $abnormalSpeedDatasetR(a,b,c)$ represents a dataset generated based on real radar trajectory data, and $abnormalSpeedDatasetS(a,b,c)$ represents a dataset generated based on simulated radar trajectory data.

Figure 8(a) shows 10 normal radar trajectories, and the blue dots represent the radar trajectory points. Figure 8(b) shows the radar trajectories with abnormal speed behavior generated based on the trajectories in Figure 8(a), and the red dots represent the trajectory points where the speed changes. According to the previous definition, the radar trajectories in Figure 8(b) belong to $abnormalSpeedDatasetS(1/2,1/4,3)$. Specifically, for each radar trajectory, the number of points in the radar trajectory is represented by N . The AIS device is turned off when the ship reaches the $(1/2 * N)th$ point of the radar trajectory. When the ship continues to sail to a certain random radar trajectory point, the speed increases by $3 m/s$, and the process lasts for $(1/4 * N)$ trajectory points. It can be seen from Figure 8(b) that the distance between the red dots increases due to the increase of speed.

In the experiment, based on real and simulated radar trajectory data, datasets with abnormal speed behavior under different combinations of conditions are generated.

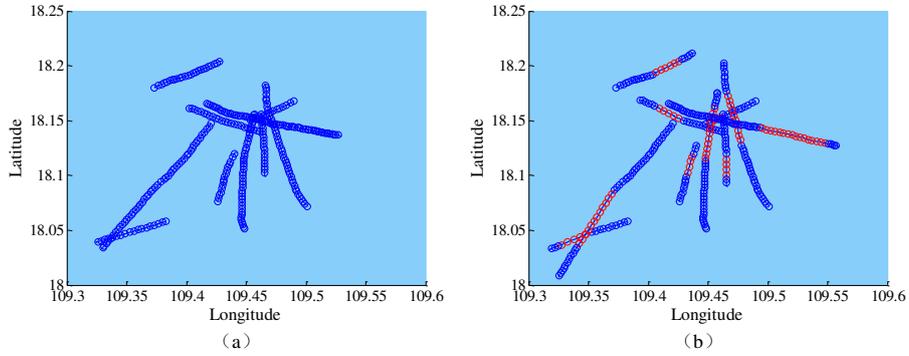


Fig. 8 The normal radar trajectories and the generated radar trajectories with abnormal speed behavior: (a) The normal radar trajectories. (b) The generated radar trajectories with abnormal speed behavior

5.1.2 Evaluation of the Algorithm

In this paper, the $F1$ score is used to evaluate the ASBD-RC algorithm.

$$F1 = \frac{2 * P * R}{P + R} \quad (4)$$

Among them, P is precision, and $P = TP / (TP + FP)$. R is recall, $R = TP / (TP + FN)$. In the experiment on abnormal speed behavior detection, TP indicates the number of abnormal speed trajectory points detected correctly. FP indicates the number of normal trajectory

points that are incorrectly detected as abnormal speed trajectory points. FN indicates the number of abnormal speed trajectory points that are incorrectly detected as normal trajectory points.

5.1.3 Evaluation Results of the ASBD-RC Algorithm

The evaluation results of the ASBD-RC algorithm are shown in Table 1 and Table 2. Table 1 shows the evaluation results of the ASBD-RC algorithm on the datasets generated based on real radar trajectory data, and Table 2 shows the evaluation results of the ASBD-RC algorithm on the datasets generated based on simulated radar trajectory data.

It can be seen from Table 1 and Table 2 that in most cases, F1 scores are above 0.83, and ASBD-RC algorithm has better performance. Only in Table 1, when $a=1/4$, $b=1/4$, $|c|=2$, the effect of the ASBD-RC algorithm is relatively poor. This is because when the duration of the trajectory segment with speed change is short and the amount in the change of speed is small, the algorithm is more susceptible to the influence of observation noise. Moreover, when AIS is closed early, the number of trajectory points used to set rules is small, which will also affect the detection effect of the algorithm. If the ship's speed changes greatly after the closure of AIS ($|c|$ is large), and the ship sails at the changed speed for a long time (b is large), it indicates that the ship is more likely to shut down the AIS device intentionally (commit illegal activities or be hijacked). This kind of behavior will cause more harm and should be detected. The evaluation results show that when $|c|$ increases or b increases, the detection effect of the algorithm is better, which shows that the ASBD-RC algorithm is suitable for detecting more harmful abnormal speed behavior, and has application value.

Table 1. The evaluation results of the ASBD-RC algorithm on the datasets generated based on real radar trajectory data.

The change of speed (c)		-4	-3	-2	2	3	4
a=1/4, b=1/4	Precision	0.8925	0.9169	0.7519	0.7242	0.8107	0.8868
	Recall	0.8486	0.8222	0.8192	0.8575	0.9032	0.9397
	F1 score	0.8700	0.8670	0.7841	0.7852	0.8544	0.9125
a=1/4, b=2/4	Precision	0.9272	0.9423	0.8695	0.8893	0.9283	0.9572
	Recall	0.9532	0.8974	0.8966	0.9342	0.9749	0.9708
	F1 score	0.9400	0.9193	0.8828	0.9112	0.9510	0.9640
a=2/4, b=1/4	Precision	0.9359	0.9049	0.8170	0.7806	0.9137	0.9270
	Recall	0.9116	0.8939	0.8462	0.9043	0.9174	0.9676
	F1 score	0.9236	0.8994	0.8313	0.8379	0.9155	0.9469

Table 2 The evaluation results of the ASBD-RC algorithm on the datasets generated based on simulated radar trajectory data

The change of speed (c)		-4	-3	-2	2	3	4
a=1/4,	Precision	0.8501	0.8093	0.7969	0.8371	0.8732	0.9018

b=1/4	Recall	0.9160	0.9339	0.8710	0.8675	0.9309	0.9542
	F1 score	0.8818	0.8672	0.8323	0.8520	0.9011	0.9273
	Precision	0.9419	0.9262	0.8967	0.9415	0.9557	0.9593
a=1/4, b=2/4	Recall	0.9673	0.9759	0.9587	0.8794	0.9644	0.9917
	F1 score	0.9544	0.9504	0.9267	0.9094	0.9600	0.9752
	Precision	0.8973	0.8658	0.8581	0.8913	0.9130	0.9300
a=2/4, b=1/4	Recall	0.9536	0.9669	0.9164	0.8736	0.9652	0.9737
	F1 score	0.9246	0.9136	0.8863	0.8824	0.9384	0.9513

5.1.4 Comparison with Other Algorithms

In this part, we compare the detection effect of the ASBD-RC algorithm with that of the other two algorithms, which are the rule-based algorithm and the extended rule-based algorithm.

The rule-based algorithm is to detect trajectory points with abnormal speed according to the set rules. The specific process is the same as the first step in the ASBD-RC algorithm. That is, if $meanSpeedA$ is greater than $meanSpeedB$, the points in $rdTrAfterAC$ whose speeds are greater than $(meanSpeedB+ratio*stdSpeedB)$ are regarded as trajectory points with abnormal speed; Otherwise, the points in $rdTrAfterAC$ whose speeds are less than $(meanSpeedB-ratio*stdSpeedB)$ are regarded as trajectory points with abnormal speed.

The extended rule-based algorithm can be divided into two steps. The first step is the same as the rule-based algorithm. The second step is to use the “extend” method to connect the detected points into abnormal trajectory segments, that is, if $P_{i(k)}$, $P_{i(l)}$ are two adjacent abnormal trajectory points detected, and $|l-k| \leq numThreshold$, the trajectory segment between $P_{i(k)}$ and $P_{i(l)}$ is regarded as an abnormal trajectory segment.

The detection results of the three algorithms are shown in Table 3. The numbers in Table 3 are the F1 scores of the algorithms on datasets with abnormal speed behavior generated based on real data. As can be seen from Table 3, compared with the rule-based algorithm which regards the trajectory points whose speeds deviate from the mean value as abnormal, the ASBD-RC algorithm and the extended rule-based algorithm are better, which indicates that it is beneficial to connect the abnormal trajectory points for abnormal speed behavior detection. And in most cases, compared with the extended rule-based algorithm, the ASBD-RC algorithm has better performance and better stability, which shows that after detecting trajectory points with abnormal speed behavior, the DBSCAN algorithm can more correctly connect them into abnormal trajectory segments.

Table 3 Comparison of the ASBD-RC algorithm with other algorithms (The numbers in the table are the F1 scores of the algorithms on real radar trajectory data)

The change of speed (c)		-4	-3	-2	2	3	4
a=1/4, b=1/4	ASBD-RC	0.8700	0.8670	0.7841	0.7852	0.8544	0.9125
	Extended rule-based	0.8558	0.8614	0.7918	0.6913	0.7738	0.8473
	Rule-based	0.8070	0.7771	0.6900	0.6537	0.7353	0.8065

	ASBD-RC	0.9400	0.9193	0.8828	0.9112	0.9510	0.9640
a=1/4, b=2/4	Extended rule-based	0.9305	0.9195	0.8953	0.8740	0.9266	0.9433
	Rule-based	0.9012	0.8613	0.8078	0.8601	0.9107	0.9299
	ASBD-RC	0.9236	0.8994	0.8313	0.8379	0.9155	0.9469
a=2/4, b=1/4	Extended rule-based	0.9171	0.8907	0.8462	0.7930	0.8785	0.9271
	Rule-based	0.8746	0.8315	0.7481	0.7660	0.8528	0.8982

5.2 Experiment on Abnormal Direction Behavior Detection

Similar to the experiment on abnormal speed behavior detection, in this experiment, we also need to generate trajectory segments with abnormal direction behavior. The *FI* score is also used to evaluate the ADBD-PE algorithm. Similarly, in order to prove the necessity of the ADBD-PE algorithm, we will also compare the detection effect of the ADBD-PE algorithm with that of the other two algorithms.

5.2.1 Generation of Trajectory Segments with Abnormal Direction Behavior

In order to verify the ADBD-PE algorithm more comprehensively, it is necessary to generate trajectory segments with abnormal direction under different combinations of conditions. Similar to the abnormal speed behavior detection experiment, three conditions should be considered in this experiment. The first two conditions are the same as the previous experiment, that is, the time when AIS is closed and the duration of trajectory segment with abnormal direction, which are represented by a and b respectively. The third condition is the amount in the change of direction. It is represented by c , and its unit is degree. A positive value indicates a clockwise change, and a negative value indicates a counterclockwise change. Therefore, the dataset with abnormal direction behavior generated under each combination of conditions can be represented by $abnormalDirectionDatasetR(a,b,c)$ or $abnormalDirectionDatasetS(a,b,c)$. $abnormalDirectionDatasetR(a,b,c)$ represents a dataset generated based on real radar trajectory data, and $abnormalDirectionDatasetS(a,b,c)$ represents a dataset generated based on simulated radar trajectory data.

Figure 9 shows 10 radar trajectories with abnormal direction behavior generated based on the trajectories in Figure 8(a). The radar trajectories in Figure 9 belong to $abnormalDirectionDatasetS(1/2,1/4,40)$, that is, the AIS device is turned off when the ship reaches the $(1/2*N)th$ point of the radar trajectory. When the ship continues to sail to a certain random radar trajectory point, its sailing direction changes 40 degrees clockwise, and the process lasts for $(1/4*N)$ trajectory points, as shown by the red dots in Figure 9. Finally, the ship recovers the original direction, and N is the number of points in the radar trajectory.

In the experiment, based on the real and simulated radar trajectory data, datasets with abnormal direction behavior under different combinations of conditions are generated.

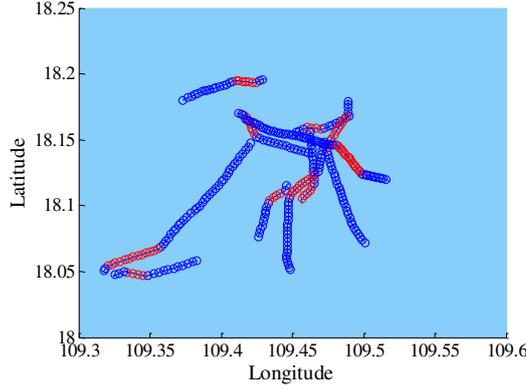


Fig. 9 The generated radar trajectories with abnormal direction behavior

5.2.2 Evaluation of the Algorithm

In this paper, the $F1$ score is also used to evaluate the ADBD-PE algorithm. In the experiment on abnormal direction behavior detection, TP indicates the number of abnormal direction trajectory points detected correctly. FP indicates the number of normal trajectory points that are incorrectly detected as abnormal direction trajectory points. FN indicates the number of abnormal direction trajectory points that are incorrectly detected as normal trajectory points.

5.2.3 Evaluation Results of the ADBD-PE Algorithm

The evaluation results of the ADBD-PE algorithm are shown in Table 4 and Table 5. Table 4 shows the evaluation results of the algorithm on the datasets generated based on real radar trajectory data, and Table 5 shows the evaluation results of the algorithm on the datasets generated based on simulated radar trajectory data.

As can be seen from Table 4 and Table 5, the change trend of the $F1$ score of the ADBD-PE algorithm on different datasets is similar to that of the ASBD-RC algorithm. When a , b , and $|c|$ are small, the ADBD-PE algorithm is more vulnerable to noise, and the detection effect of the algorithm is relatively poor. When the amount in the change of direction increases ($|c|$ increases) or the duration of the trajectory segment with abnormal direction increases (b increases), the detection effect of the ADBD-PE algorithm becomes better. It also shows that the ADBD-PE algorithm is suitable for detecting more harmful abnormal direction behavior.

Table 4 The evaluation results of the ADBD-PE algorithm on the datasets generated based on real radar trajectory data

The change of direction (c)		-60	-40	-20	20	40	60
a=1/4, b=1/4	Precision	0.8898	0.9079	0.7781	0.6960	0.8077	0.9131
	Recall	0.9682	0.8775	0.7718	0.7857	0.9469	0.9890
	F1 score	0.9273	0.8924	0.7749	0.7381	0.8718	0.9496
a=1/4, b=2/4	Precision	0.9724	0.9556	0.9141	0.8731	0.9300	0.9546
	Recall	0.9771	0.8948	0.8020	0.8751	0.9593	0.9671

a=2/4, b=1/4	F1 score	0.9747	0.9242	0.8544	0.8741	0.9444	0.9608
	Precision	0.9713	0.8729	0.7122	0.7691	0.8859	0.9546
	Recall	0.9606	0.9647	0.7838	0.9503	0.9960	0.9965
	F1 score	0.9659	0.9165	0.7463	0.8501	0.9377	0.9751

Table 5 The evaluation results of the ADBD-PE algorithm on the datasets generated based on simulated radar trajectory data

The change of direction (c)		-60	-40	-20	20	40	60
a=1/4, b=1/4	Precision	0.9192	0.8572	0.5632	0.6000	0.7896	0.8896
	Recall	0.9310	0.8550	0.8075	0.7772	0.8565	0.9353
	F1 score	0.9251	0.8561	0.6636	0.6772	0.8217	0.9119
a=1/4, b=2/4	Precision	0.9634	0.9344	0.8397	0.8439	0.9299	0.9538
	Recall	0.9410	0.8932	0.7497	0.8197	0.9106	0.9439
	F1 score	0.9521	0.9134	0.7922	0.8316	0.9202	0.9488
a=2/4, b=1/4	Precision	0.9287	0.8864	0.7110	0.6950	0.8561	0.9127
	Recall	0.9489	0.9104	0.8589	0.8576	0.9253	0.9616
	F1 score	0.9387	0.8982	0.7780	0.7678	0.8894	0.9365

5.2.4 Comparison with Other Algorithms

In this part, we compare the detection effect of the ADBD-PE algorithm with that of the other two algorithms, which are the rule-based algorithm and the ADBD-PM algorithm.

The rule-based algorithm is to detect trajectory points with abnormal direction according to the set rules. First, the mean value and standard deviation of the direction of the points in $rdTrBeforeAC$ need to be computed, and they are named $meanDirectionB$ and $stdDirectionB$ respectively. Then, the mean value of the direction of the points in $rdTrAfterAC$ is computed, which is named $meanDirectionA$. If $meanDirectionA$ is greater than $meanDirectionB$, the points in $rdTrAfterAC$ whose directions are greater than $(meanDirectionB + ratio * stdDirectionB)$ are regarded as trajectory points with abnormal direction; Otherwise, the points in $rdTrAfterAC$ whose directions are less than $(meanDirectionB - ratio * stdDirectionB)$ are regarded as trajectory points with abnormal direction.

The ADBD-PM algorithm can be divided into two steps. The first step is to use the Douglas-Peucker algorithm to partition $rdTrAfterAC$. Each trajectory segment is represented by $rdTrAfterACSeg_i$, where $i=0,1,2,\dots,n-1$, and n is the number of segments. The second step is similar to the second step of the ADBD-PE algorithm, except that when measuring the difference between the directions of two trajectory segments, the mean value is used instead of the Earth Mover's distance. Specifically, when judging whether $rdTrAfterACSeg_i$ has abnormal direction behavior, first, the mean value of the directions of the points in $rdTrAfterACSeg_i$ is computed and recorded as $meanDirectionSeg_i$. If $|meanDirectionSeg_i - meanDirectionB| > ratio * |meanDirectionA - meanDirectionB|$, it is considered that $rdTrAfterACSeg_i$ has abnormal direction behavior.

The detection results of the three algorithms are shown in Table 6. The numbers in Table 6 are the F1 scores of the algorithms on datasets with abnormal direction behavior generated based on real radar trajectory data. In essence, the rule-based algorithm regards the trajectory points whose directions deviate from the mean value as abnormal. It can be seen from Table 6 that both the ADBD-PE algorithm and the ADBD-PM algorithm are superior to the rule-based algorithm, which indicates that it is necessary to partition the trajectory in abnormal direction behavior detection. At the same time, in most cases, the ADBD-PE algorithm is better than the ADBD-PM algorithm, which shows that it is more accurate to use the Earth Mover's distance to measure the difference between the directions of two trajectory segments.

Table 6 Comparison of the ADBD-PE algorithm with other algorithms (The numbers in the table are the F1 scores of the algorithms on real radar trajectory data)

The change of direction		-60	-40	-20	20	40	60
(c)							
a=1/4, b=1/4	ADBD-PE	0.9273	0.8924	0.7749	0.7381	0.8718	0.9496
	ADBD-PM	0.8665	0.8388	0.6779	0.6709	0.7891	0.7726
	Rule-based	0.7402	0.6694	0.5785	0.5590	0.5941	0.7158
a=1/4, b=2/4	ADBD-PE	0.9747	0.9242	0.8544	0.8741	0.9444	0.9608
	ADBD-PM	0.9613	0.9508	0.8119	0.8819	0.9366	0.8881
	Rule-based	0.9160	0.9129	0.8063	0.7859	0.8583	0.8301
a=2/4, b=1/4	ADBD-PE	0.9659	0.9165	0.7463	0.8501	0.9377	0.9751
	ADBD-PM	0.9310	0.8891	0.7087	0.8130	0.8932	0.8997
	Rule-based	0.9026	0.7902	0.6686	0.7017	0.7923	0.7935

6 Conclusion

This paper mainly studies the detection of abnormal ship behavior after the closure of AIS. In the detection process, the ship behavior before the closure of AIS is considered as normal. We mainly detect whether the behavior of the ship after the closure of AIS is significantly different from that before the closure of AIS. Abnormal behavior can be divided into abnormal speed behavior and abnormal direction behavior. The ASBD-RC algorithm is proposed to detect abnormal speed behavior. In the ASBD-RC algorithm, firstly, the radar trajectory points whose speeds deviate from the mean value of normal speed are detected, and then the DBSCAN algorithm is used to connect the detected points into trajectory segments with abnormal speed. The ADBD-PE algorithm is proposed to detect abnormal direction behavior. In the ADBD-PE algorithm, firstly, the radar trajectory after the closure of AIS is partitioned, and then based on the Earth Mover's distance, whether each trajectory segment has abnormal direction behavior is judged. Finally, based on the real and simulated radar trajectory data, we generate abnormal behavior datasets under different combinations of conditions. These datasets are used to evaluate the two abnormal behavior detection algorithms. The experimental results show that the algorithms have good performance on most datasets, and are suitable for detecting more harmful abnormal behaviors. At the same

time, in order to verify the necessity of the two abnormal behavior detection algorithms, we also compare the two algorithms with other algorithms. The comparison results show that the proposed algorithms have better and more stable detection results.

In future work, we will study the detection of abnormal ship behavior in more complex environments. For example, when a ship has obvious changes in direction or speed before the AIS is closed, how to detect the abnormal behavior of the ship after the closure of AIS is a problem worthy of study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- Arguedas VF, Pallotta G, Vespe M (2018) Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring. *IEEE Transactions on Intelligent Transportation Systems* 19(3):722-732.
- Brun L, Saggese A, Vento M (2014) Dynamic scene understanding for behavior analysis based on string kernels. *IEEE Transactions on Circuits and Systems for Video Technology* 24(10):1669-1681.
- Castaldo F, Palmieri FAN, Regazzoni CS (2016) Bayesian analysis of behaviors and interactions for situation awareness in transportation systems. *IEEE Transactions on Intelligent Transportation Systems* 17(2):313-322.
- Douglas DH, Peucker TK (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization* 10(2):112-122.
- Ester M, Krieger HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc. 1996 Int. Conf. Knowledge Discovery and Data Mining (KDD '96)*, Portland, Oregon, USA, pp. 226-231.
- Fu P, Wang H, Liu K, Hu X, Zhang H (2017) Finding abnormal vessel trajectories using feature learning. *IEEE Access* 5:7898-7909.
- Habtemariam BK, Tharmarasa R, Meger E, Kirubarajan T (2012) Measurement level AIS/radar fusion for maritime surveillance. *Proceedings of the Signal and Data Processing of Small Targets 2012* 8393: 83930I.
- Huang L, Wen YQ, Guo W, Zhu XY, Zhou CH, Zhang F, Zhu M (2020) Mobility pattern analysis of ship trajectories based on semantic transformation and topic model. *Ocean Engineering* 201:107092.
- Kazemi S, Abghari S, Lavesson N, Johnson H, Ryman P (2013) Open data for anomaly detection in maritime surveillance. *Expert Systems With Applications* 40(14):5719-5729.
- Laxhammar R, Falkman G (2014) Online learning and sequential anomaly detection in trajectories. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36(6):1158-1173.
- Lei PR (2016) A framework for anomaly detection in maritime trajectory behavior. *Knowledge and Information Systems* 47(1):189-214.
- Mascaro S, Nicholson AE, Korb KB (2014) Anomaly detection in vessel tracks using Bayesian networks. *International Journal of Approximate Reasoning* 55(1):84-98.

- Mazzarella F, Vespe M, Alessandrini A, Tarchi D, Aulicino G, Vollero A (2017) A novel anomaly detection approach to identify intentional AIS on-off switching. *Expert Systems With Applications*. 78:110-123, 2017.
- Piciarelli C, Micheloni C, Foresti GL (2008) Trajectory-based anomalous event detection. *IEEE Transactions on Circuits and Systems for Video Technology* 18(11):1544-1554.
- Riveiro M, Pallotta G, Vespe M (2018) Maritime anomaly detection: A review. *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery* 8(4):e1266.
- Rong H, Teixeira AP, Guedes Soares C (2020) Data mining approach to shipping route characterization and anomaly detection based on AIS data. *Ocean Engineering* 198:106936.
- Rubner Y, Tomasi C, Guibas LJ (1998) A metric for distributions with applications to image databases. *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, Bombay, India, pp. 59-66.
- Sidibé A, Shu G (2017) Study of automatic anomalous behaviour detection techniques for maritime vessels. *Journal of Navigation* 70(4):847-858.
- Tu E, Zhang G, Rachmawati L, Rajabally E, Huang GB (2018) Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology. *IEEE Transactions on Intelligent Transportation Systems* 19(5):1559-1582.
- Wan Z, Chen J, Makhloufi AE, Sperling D, Chen Y (2016) Four routes to better maritime governance. *Nature* 540(7631):27-29.
- Wang F, Lei Y, Liu Z, Wang X, Ji S, Tung AKH (2020) Fast and parameter-light rare behavior detection in maritime trajectories. *Information Processing and Management* 57(5):102268.
- Yan Z, Xiao Y, Cheng L, He R, Ruan X, Zhou X., Li M, Bin R (2020) Exploring AIS data for intelligent maritime routes extraction. *Applied Ocean Research* 101:102271.
- Zhang T, Zhao S, Cheng B, Chen J (2020) Detection of AIS closing behavior and MMSI spoofing behavior of ships based on spatiotemporal data. *Remote Sensing* 12(4):702.