Abnormal Detection Method of Transship Based on Marine Target Spatio-Temporal Data

Wen Ying[1,2], Mingwang Ou[2], Qichun Liang[2], Zhixia Yang[2], Man Zhao[1]*

[1] *China University of Geosciences, China*
[2] *CETC Ocean Corporation, China*
yingwen@cetc.com.cn, oumingwang@cetc.com.cn, liangqichun1@cetc.com.cn, yangzhixia@cetc.com.cn, zhaoman@cug.edu.cn

Abstract

Marine target situational awareness is developing towards intelligence. The practical application of marine law enforcement, route control, fishery supervision and other fields makes marine ship track mining and anomaly analysis a hot direction of technology development. At present, most of the research focuses on the feature recognition of single target or single target signal, while the research on correlation anomaly detection based on multi target or multi target signal is relatively less. In this paper, based on the large spatio-temporal data of ocean targets such as radar and AIS track, the corresponding recognition algorithm is designed for the abnormal transport behavior. The real-time batch detection of illegal transshipment is realized, which verifies the effectiveness of the algorithm and shows good performance in practical application scenarios.

Keywords: Marine target spatio-temporal data, Ship track mining, Ship behavior identification, Abnormal transshipment detection algorithm

1 Introduction

Marine target monitoring plays an important role in the field of marine safety control [1]. At present, the construction scale of information network for marine target detection is increasing, and the construction and deployment of observation stations are showing an obvious rasterization trend [2]. In addition, maritime satellites are widely used to receive AIS signals from marine vessels, and the volume of marine target monitoring data is rapidly expanding [3]. The improvement of system and data scale brings a series of new technical challenges, among which the most critical problem is to solve the blind detection and low application efficiency caused by insufficient application of intelligent technology. Ship behavior anomaly detection based on target track is the core step to realize intelligent marine target monitoring. By timely finding a relatively small amount of suspected abnormal behaviors in the track data of millions or even tens of millions of levels of marine targets, it is helpful to focus on the actual problems occurring on the sea, and provide scientific analysis and guidance for continuous key observation or regulatory action.

2 Related Work

At present, many sea-related research institutions and teams have carried out a series of research work on this issue, and have obviously formed two research directions, namely, the detection of abnormal behavior of ships and the analysis of specific behavior. In the aspect of abnormal behavior detection, the current main detection aspects include the abnormal behavior detection model based on Naive Fourier algorithm and clustering model based on DBSCAN algorithm. Among them, the model based on Naive Fourier Algorithm has been proved to have good detection capability in some ports, sea entrances, and other local waters with relatively fixed characteristic patterns (Castaldo [4], Salma [5]), but in the practical application level, the model needs to be reconstructed and trained according to the actual situation of different waters, with relatively weak universality. Zhao-kun Wei et al. [6] have designed a kind of Naive Fourier detection method that does not rely on track information, but only uses ship AIS static information and dynamic information to monitor ship abnormal behavior. Still, the application field is limited to limited waters. DBSCAN algorithm is a typical data space clustering algorithm. This kind of algorithm is mainly used in maritime route supervision application scenarios, such as Carsten H. Botts [7], S. C. Jia et al. [8], using DBSCAN algorithm to identify maritime and inland water routes and ships deviating from the route or C. Wang [9] using DBSCAN clustering to identify abnormal behaviors inside and outside the port. With the development of relevant research, some combination algorithms based on DBSCAN have also emerged. For example, Y. M. Wang [10] proposed a hybrid clustering algorithm that integrates K-means and DBSCAN algorithms. After K-means algorithm is used for preliminary clustering to extract characteristic attributes, DBSCAN algorithm is used for clustering to improve anomaly detection efficiency and accuracy. Z. R. Wang et al. [11] proposed a combinatorial algorithm that integrates DBSCAN and iForest, combining speed and ship density. The distribution rule and the change of clustering effect are selected to detect the abnormal behavior of ships.

*Corresponding Author: Man Zhao; E-mail: zhaoman@cug.edu.cn
DOI: 10.53106/160792642023092405010

Other studies focus on the analysis and detection of ship-specific behavior by focusing on the motion characteristics of a particular class of ships, or ship-specific behavior or derived effects for analysis, mining, and prediction [12]. Kamran F. Toloue et al. [13], Y. Wang et al. [14], Erico N. de Souza et al. [15], etc. identify the operation status of fishing boats through VMS data and AIS data for detection, W. W. Zhao et al. [16] identify the ship segments through ship AIS data, and Y. Q. Wen et al. [17] analyze the relevant ship route voyage based on static AIS data and other shipping data, so as to predict the global crude oil trade trend.

The above research focuses on the detection and recognition of single target or single target signal features [18], and the interaction between multiple ships or the association relationship between multiple ship signals in the actual application scenario can reflect some special ship activity patterns, which often contains higher information value than the single target analysis. But the current research on the detection of multi-target or multi-target signal association anomaly is relatively less. From T. Zhang et al [19], based on the association analysis of multiple AIS target data, a MMSI deception detection algorithm is proposed to identify the tampering of maritime mobile service identification behavior. The theoretical validity of the algorithm is verified by simulation data.

In this paper, common multi-target/multi-target signal anomaly detection problems, namely abnormal transshipment detection, are identified and tested in real time and in batches, and shows good performance in the actual application scenario. Section 3 give the analysis of the anomaly transshipment behavior and the detection strategy and implementation method of detecting the anomaly transshipment behavior; Section 4 gives some test results and application cases to verify the effectiveness of the method; Section 5 is the summary and prospect of the paper.

3 Strategy and Implementation

3.1 Behavior Pattern Analysis

Transshipment operations usually refers to a change of the goods between ship and ship. In most case, it belongs to the normal behavior in the process of shipping. It is beneficial to reduce operation links, reduce transport costs, and speed up transport, but most of transshipment operating in the port, wharf, anchorage or on inland waters involves a ship on a sea voyage, where lightering operation behavior happening largely means goods smuggling and other illegal activity. The operation mode of anomaly transshipment can be divided into four modes: direct berthing, astern replenishment, longitudinal replenishment and boat transferring. Figure 1 shows the samples of these four different modes.

Direct berthing is also known as STS transferring, which is the most common mode of transshipment, generally by one of the ships in advance with fender and floating, and another ship drew up to it. This way of transshipment is mainly used for bulk cargo transport. Parallel replenishment refers to two ships near each other within a certain distance, to keep the same speed in parallel, through the rope or tube to transport the goods, which is commonly used in delivering liquid goods, such as refined petroleum. A similar pattern called astern replenishment refers to two ship travelling front and back with similar speed and keeping a close distance. The boat transferring mode is mostly used for the transfer of high-value goods, such as luxury goods and drugs, which two ships approach a certain distance, and one of them will put down small boats ferrying people and goods to the other ship.

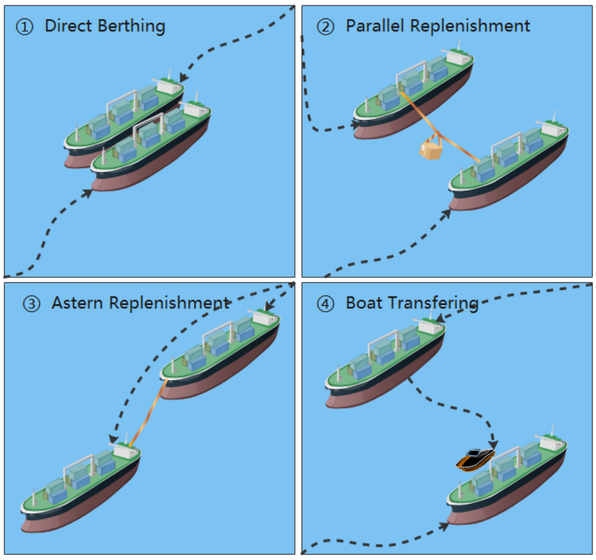


Figure 1. Four different mode of illegal transshipment

Combined with the analysis of the above operation modes, it can be found that illegal transshipment operations at sea generally have the following common behavior characteristics:

(1) The operation sites are generally selected in the sea 15km away from the shore, mostly concentrated in areas 5-15km away from the main sea lanes and near the maritime border.

(2) The operation time is generally selected in the unconventional period, and the probability of night operation is large.

(3) Before the transshipment operation, the sailing track of the relevant ships is most likely to come from different directions. During the transshipment process, the sailing track of the ships is mostly in the same direction or reverse direction. After the completion of the operation, it is most likely to go to different directions.

(4) Before the transshipment operation, there is an obvious speed decline process, which will keep low speed or even float for a period of time. After the operation, the speed will obviously rise.

Apart from these common characteristics shown above, there are other characteristics for the four different operation mode that can be summarized as shown in Table 1.

Table 1. Behavioral characteristic of four transshipment operating modes

Modes	Distance	Heading	Speed	Duration
Direct berthing	Normally ≤0.1km	Close to 0°/180°	Close to 0m/s	Normally ≥5min
Parallel replenishment	Normally ≤0.3km	Close to 0°	≤2 knots with same speed	Normally ≥5min
Astern replenishment	Normally ≤0.5km	Close to 0°	≤2 knots with same speed	Normally ≥5min
Boat transferring	Normally ≤1km	-	Close 0m/s	Normally ≥5min

3.2 Detection Algorithm Design

Based on the analysis of the behavior mode above, it can be seen that in the process of transferring vessels, the distance must be relatively close within the same time break, and the speed of the two vessels is less than a certain threshold and lasts for a certain period of time. Therefore, corresponding detection model can be constructed based on the above characteristics. The specific processing steps are shown in Figure 2.

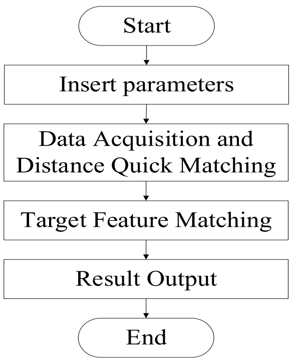


Figure 2. Execution steps of transshipment detection

3.2.1 Insert Parameters

According to the actual demand of transshipment detection scenarios, it is required that the operation model can effectively respond to two working modes: region detection and specific object association detection. Region detection refers to the selection of the maritime area and time period to be detected, and the ergodic matching of all maritime target trajectories in this area and time period, and then the output of all target groups with suspected transshipment behavior, as well as the related attribute information of the target groups. The specific target association detection, on the other hand, refers to the matching of other maritime targets related to a specific target in the spatial and temporal region, and the output of the target group with suspected transshipment behavior with the specific target. Therefore, the input parameters required by the transshipment model mainly includes time interval (start time and end time), space interval (upper left coordinate and lower right coordinate), and specific target information (MMSI or the unique identification number of fusion target data).

3.2.2 Data Acquisition and Quick Matching

Combining data acquisition with fast matching process is due to these two processes have an integral nature in the implementation level. Though the input conditions can easily retrieve and fetch the relevant target data in the space-time region, in the case of large region and time span, the volume of retrieved data and the calculation of correlation-matching will increase dramatically, thus to avoid massive data traversal calculation, a three-dimensional Geohash-like algorithm is used there, for achieving fast matching by the Euclidean distance of the target trajectory in the space-time region. Using this method, the obviously irrelevant target matching groups are filtered and computational costs of further matching are reduced. The specific processing steps are as follows:

Firstly, according to the input conditions, the target data matching the spatial and temporal interval were retrieved from the database and transformed to the following format. The trajectory of a certain object in a certain time slot with an unique identifier A can be represented as:

$$Tr[A] = \{PA_{(0)}, PA_{(1)}, PA_{(2)}, \ ..., PA_{(n)}\}. \quad \textbf{(1)}$$

Where $PA_{(n)}$ is the nth trajectory point in $Tr[A]$, which can be represent as a six element groups:

$$PA_{(n)} = \{ID_A, t_{(n)}, lon_{(n)}, lat_{(n)}, sog_{(n)}, cog_{(n)}\}. \quad \textbf{(2)}$$

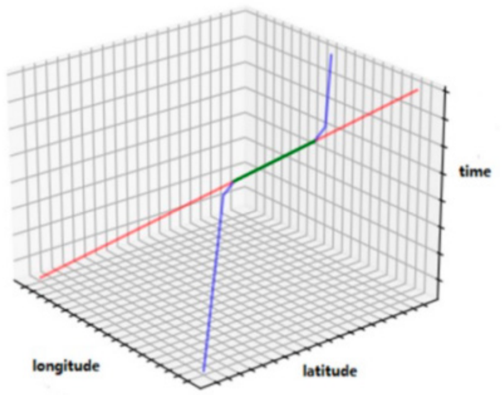
Where ID is the unique identifier of the trajectory, it could be a MMSI in AIS trajectory, a bench number for a radar trajectory, or a fusion ID if a radar and AIS data fusion process was previously performed, $t_{(n)}$, $lon_{(n)}$, $lat_{(n)}$, $sog_{(n)}$, $cog_{(n)}$, each represent the time stamp, the longitude, the latitude, the speed over ground, and the course over ground of the nth trajectory point. For the conversion of longitude and latitude coordinate system of data, WGS84 is used uniformly to convert the coordinate system of AIS, radar and other data sources to WSG84; Clean the dirty data, clean and filter the invalid data and null values in the data, and enhance the accuracy and reliability of the algorithm model. Different data update frequencies are inconsistent, and different target update frequencies of the same data are also inconsistent. During the re execution of the algorithm, the time of the data field is discretized to keep the target time of different update frequencies consistent. Thus, the input matrix has a format as follows (Table 2):

Table 2. The input matrix of transshipment detection

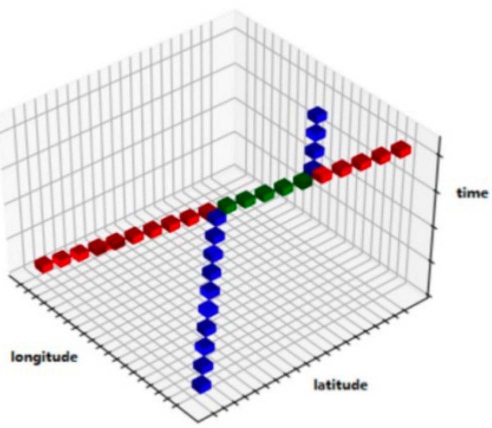
ID	Time	Lon	Lat	Sog	Cog
ID_A	$t_{(0)}$	$lon_{(0)}$	$lat_{(0)}$	$sog_{(0)}$	$cog_{(0)}$
ID_A	$t_{(1)}$	$lon_{(1)}$	$lat_{(1)}$	$sog_{(1)}$	$cog_{(1)}$
ID_A	$t_{(2)}$	$lon_{(2)}$	$lat_{(2)}$	$sog_{(2)}$	$cog_{(2)}$
		...			
ID_A	$t_{(n)}$	$lon_{(n)}$	$lat_{(n)}$	$sog_{(n)}$	$cog_{(n)}$
ID_B	$t_{(0)}$	$lon_{(0)}$	$lat_{(0)}$	$sog_{(0)}$	$cog_{(0)}$
ID_B	$t_{(1)}$	$lon_{(1)}$	$lat_{(1)}$	$sog_{(1)}$	$cog_{(1)}$
ID_B	$t_{(2)}$	$lon_{(2)}$	$lat_{(2)}$	$sog_{(2)}$	$cog_{(2)}$
		...			
ID_B	$t_{(n)}$	$lon_{(n)}$	$lat_{(n)}$	$sog_{(n)}$	$cog_{(n)}$
		...			
ID_N	$t_{(n)}$	$lon_{(n)}$	$lat_{(n)}$	$sog_{(n)}$	$cog_{(n)}$

Secondly, the triplet of time, longitude, and latitude is used to construct a three-dimensional spatial-temporal coordinate space, and each target trajectory in the input matrix is depicted and discrete processing with isometric grouping. Then, by collecting the target track fragments falling in the same group, the fast matching relationship of target track can be obtained. According to the results, the target track fragments that have not been effectively correlated can be deleted in the input matrix, which can effectively realize the filtering of irrelevant data. Figure 3(a) shows the trajectory of two targets in the spatial-temporal coordinate, Figure 3(b) shows the result of isometric grouping and the overlapping grouping area of two target trajectories, and the pseudocode of the algorithm is shown in Algorithm1.

Algorithm 1 firstly gets the input matrix of target trajectory and the limit parameters (including the time and space interval, segmentation precision) used to generate an equidistant discrete coordinate map (line 1 in Algorithm 1). Then, each trajectory dot in the input matrix is printed into the coordinate space (line 2-4 in Algorithm 1). After completing the above steps, the multiple target trajectory points falling on the same unit are selected, forming a result matrix, which has a similar structure as the input matrix. Finally, the trajectory dot in the result matrix is sorted by trajectory and time sequence (line 9 in Algorithm 1), forming the output.



(a) The trajectory of two targets



(b) The result of isometric grouping

Figure 3. Fast matching by the euclidean distance of the target trajectory in the space-time region

Algorithm 1. Fast matching by the euclidean distance of the target trajectory

Input: Input Matrix of Target Trajectory, Tr_{in}; Time interval, *timeInterval*[t_{start}, t_{end}];
Space interval
spaceInterval[lon_{start}, lon_{end}, lat_{start}, lat_{end}]; Isometric Grouping Threshold,
threshold[*timePrecision*, *spacePrecision*].
Output: Filtered Output Matrix of Target Trajectory, Tr_{out}.

```
1: coordinateMap[cell,dot] <- createCoordinateMap(
     timeInterval, spaceInterval, threshold)
2: for each target trajectory in Tr_in do
3:     coordinateMap <- updateCoordinateMap(Tr_in)
4: end for
5: for each cell in coordinateMap do
6:     if the count of trajectory dot falls in the cell is greater than
       1 then
7:         Tr_out <- addMatchDot(dot)
8: end for
9: Tr_out <- sortingAsTrajectory()
10:return T_rout
```

3.2.3 Target Feature Matching

In the step of data collection and quick matching, the target data in the space-time area is retrieved and quickly filtered based on the target distance characteristics, effectively eliminating the obviously irrelevant target groups. In the step of target feature matching, the potential lightering operation behavior is determined by further analyzing the multi-dimensional feature elements of the target groups. The processing logic of target feature matching is to match the distance characteristics, speed characteristics, and heading characteristics of each time segment of target groups that have been determined to have potential relationships in the process of data acquisition and rapid matching on each time segment to form a "suspected docking point", and then match the duration of the suspected docking of this target group to finally form the docking detection results.

3.2.3.1 Distance Feature Matching

Distance feature matching includes two processes: relative distance calculation and trend analysis. It can be seen from the analysis of offshore docking behavior mode that maintaining a relatively close relative distance between barges is a necessary condition for offshore docking activities. From this distance feature matching, the longitude and latitude information of the current time segment target group is first converted to obtain the relative distance of the target. Then, the distance change trend feature can be obtained by analyzing the relative distance of the previous time clip. If the distance of the current time clip target group is less than the established relative distance threshold (set as 1km by default), and the distance change trend shows that two targets are close to or maintain the distance, then add a "distance feature consistent" label to the track points of the time clip target group.

The target longitude and latitude position and distance uses the Haversine formula to calculate the distance between two points on the arc surface. The specific calculation method is as follows:

$$haversine(\frac{d}{R}) = haversine(\Delta lat) + \cos(lat1)\cos(lat2)haversine(\Delta lon). \tag{3}$$

$$haversine(\theta) = \sin^2\frac{\theta}{2} = \frac{(1-\cos\theta)}{2}. \tag{4}$$

Where R is the radius of the earth, and the average value is 6371 km. Lat1 represents the latitude of point 1, lat2 represents the latitude of point 2, lon represents the longitude difference between two points, and lat represents the latitude difference between two points. Before calculation, conversion between angle and radian is required.

The decision logic for distance change trend is as follows:

$$\Delta Trd_{t_1} = \begin{cases} \text{"}approach\text{"} \ \ if \ \ d_{t_1} - d_{t_0} > D_{th2} \\ \text{"}maintain\text{"} \ \ if \ \ \left|d_{t_1} - d_{t_0}\right| \le D_{th2} \\ \text{"}away\text{"} \ \ if \ \ \left|d_{t_1} - d_{t_0}\right| > -D_{th2} \end{cases}. \tag{5}$$

Where D_{th} is the sensitivity threshold. When the distance between two time segment targets changes within the range, it can be regarded as a normal disturbance.

3.2.3.2 Speed Feature Matching

Speed feature matching is aimed at the obvious feature that ships control ship speed in the process of barge passing at sea. Determine the speed level of target groups in the current time segment. If the speed of target groups in the same time segment is less than the speed threshold time segment, add a "speed feature consistent" label to the track points of target groups in the time segment. The decision logic is as follows:

$$C_V = \text{"}ture\text{"} \ \ if \ \ sog_{PA_{t_1}} \le V_{th1} \ \ \&\& \ sog_{PB_{t_1}} \le V_{th1}. \tag{6}$$

3.2.3.3 Course Feature Matching

It can be seen from the above analysis of docking behavior mode that the relationship characteristics between ship course and distance, speed, and other characteristics are different in the process of docking at sea. Considering the complexity of the actual navigation of ships at sea, the relationship characteristics of course are not necessary to determine the behavior of docking at sea, but when the distance, speed, and duration characteristics fully point to the behavior of docking at sea, the characteristics of the course relationship can further enhance the accuracy of the determination, and to a certain extent, can provide effective support for the specific methods used to distinguish the connection and the study and judgment of the connection intention. In the three connection modes of direct berthing, parallel replenishment, and longitudinal replenishment, the two ships should ensure the same or completely opposite bow directions as far as possible. Even if they are transported by small ships, the probability of keeping parallel static between the two ships is greater than the probability of random direction. Therefore, the course relationship characteristics can be effectively obtained by calculating the interpolation of the course angle between the two ships. The specific calculation method is as follows:

First, coordinate conversion of ship course information, as shown in the formula:

$$\theta_{P_{t_1}} = \begin{cases} \theta_{P_{t_1}} & if \ \ \theta_{P_{t_1}} \le 180^\circ \\ \theta_{P_{t_1}} - 360 & if \ \ \theta_{P_{t_1}} > 180^\circ \end{cases}. \tag{7}$$

Then, calculate the heading difference of the target group in the current time segment, as shown in the formula:

$$\Delta\theta_{PABt_1} = \left|\theta_{PA_{t_1}} - \theta_{PB_{t_1}}\right|. \tag{8}$$

Finally, if the course difference is less than the course threshold, add a "Course Feature Compliance" label to the track points of the time segment target group, which means that the track points of the time segment target are more likely to be connected at sea.

3.2.3.4 Duration Matching

As the sea connection will be accompanied by the transfer of the cargo, its obvious distance, speed, and heading characteristics will last for a period of time, so after the matching of distance, speed and heading is completed, the duration of the segments of the target grouping that meet the relevant characteristics will be matched, The main matching method is to take the first (pair) track point in the target grouping time sequence with the labels of "distance feature matching" and "speed feature matching" as the starting time segment, and then successively search whether the target grouping in the subsequent time segment also has the above labels, and record the duration until the track point appearing in two consecutive time segments has no label end, so as to determine the duration of the record, When the duration reaches or exceeds the duration threshold, record the target ID, start time, and end time of the group respectively. At the same time, if the "Course feature compliance" mark of the track point in this period is greater than the threshold value, add the confidence level of the current segment as "high".

Algorithm 2. Target feature matching

Input: Filtered Input Matrix of Target Trajectory, Tr_{in}; Distance Threshold, D_{th} ; Sog Threshold, V_{th} ; Cog Threshold, θ_{th} ; Time Threshold, T_{th} .

Output: The Output Matrix of Transshipment Detection, Tr_{out} .

Create Target Array: targetArray
timeCellRaw <- groupByTimeCell(Tr_{in})
timeCell <- clearTimeCellRaw(timeCellRaw)
for each time cell in timeCell **do**
  if calDistance($d_{PA_{t1}}$, $d_{PB_{t1}}$, D_{th}) **then**
    distLabel <- addDistanceLabel(Tr_{in})
  if calSog($sog_{PA_{t1}}$, $sog_{PB_{t1}}$, V_{th}) **then**
    sogLabel <- addSogLabel(Tr_{in})
  if calCog($\theta_{PA_{t1}}$, $\theta_{PB_{t1}}$, θ_{th}) **then**
    cogLabel <- addCogLabel(Tr_{in})
end for
if calDuration(distLabel, sogLabel, cogLabel, T_{th}) **then**
  targetArray <- addTarget(Tr_{in})
Tr_{out} <- arrayToMatrix(targetArray)
return Tr_{out}

(4) Results output

Since the full track data of the target has been stored in the database, there is no need to repeatedly output the track point data of the offshore docking target group. According to the duration matching result, the offshore docking detection result data can be quickly obtained, and the format is shown in the following table (Table 3):

Table 3. The output matrix of transshipment detection

ID1	ID2	Duration	StartTime	Stop Time	Confidence
ID_A	ID_B	T_{PAB}	$t_{(sPAB)}$	$t_{(ePAR)}$	High (e.g.)
ID_X	ID_Y	T_{PXY}	$t_{(sPXY)}$	$t_{(ePXY)}$	Normal (e.g.)
			…		

4 Verifications

Different from the verification of detection methods for fishing behavior of fishing vessels, the effectiveness of detection algorithms can be verified through the work logs obtained by fishing vessel management companies or regulatory agencies. The verification of the effectiveness of detection algorithms for offshore docking under actual application scenarios is faced with practical difficulties. Because research institutions generally lack effective monitoring and confirmation means for the actual situation of marine activities, and due to privacy protection and other factors, there are relatively few public materials that can be used as the basis for verification. At the same time, there is a lack of data sets on barge passing ships at sea, and the model algorithm cannot be scored through a large number of data sets. The accuracy can only be verified through the deduction of actual cases of maritime law enforcement departments. According to this situation, this study searched for several cases of illegal trading at sea published by relevant law enforcement agencies in recent years, and traced them through historical AIS data retrieval according to the location and approximate time of the incident, so as to verify the effectiveness of this detection method.

Case 1 is shown in Figure 4. In 2020, according to the intelligence information provided by maritime law enforcement agencies, illegal reselling of refined oil products by oil tankers will occur in China's sea area A, which will occur between 22:00 on Day 1 and 3:00 in the morning on Day 2. Therefore, according to the characteristics of the reselling of refined oil products on the sea, real-time monitoring of the sea area will be carried out using ship barge detection algorithms, and a number of ship barge incidents will be found in the sea area around 2:00 in the morning on Day 2. Through the identification screening, speed analysis and track tracing of all barge passing ships in this period, the suspicious ship x and ship y were locked. Both ships are Chinese nationals, and ship x is a large oil tanker. When sailing out of the offshore drilling platform, it suddenly decelerated when passing the sea area A. The barge passing behavior occurred with fishing boat y waiting for a long time. One hour later, the two ships left one after another and headed for nearby ports a and b.

As a result, the law enforcement agencies intercepted ships x and y respectively at the port to search, and found that there were 60 tons of oil products on ship y, all of which were from ship x. Therefore, it can be determined that the two ships had lightering behavior at sea, and carried out illegal product oil trading.

Case 2 is shown in Figure 5. In 2021, according to the intelligence information provided by maritime law enforcement agencies, there will be many illegal product oil smuggling incidents abroad in the product oil smuggling high incidence area B on Day D1; Therefore, the barge detection algorithm was used to conduct real-time monitoring on sea area B, and three barge passing incidents were detected at 3-5 a.m. on Day 1. After the identification and historical behavior analysis of the incident ships, it was found that the barge "mother ships" were all large oil tankers in China. They started from different drilling platforms and began to slow down after entering sea area B, and had barge passing behaviors with small fishing boats in China and other countries. The barge transit time was more than 2 hours. After the small boat left, the "mother ship" slowly sailed in the sea area.

As result, the maritime law enforcement agencies dispatched law enforcement forces and successfully seized three ships suspected of smuggling refined oil products in the sea area, with a total of 30 people involved. The people involved were from China, Taiwan, and the Philippines. The preliminary investigation showed that the smuggling of refined oil products involved more than 8,000 tons. This case was defined as a case of Chinese oil tankers smuggling refined oil products overseas, which can prove that the three ships actually had multiple lightering behaviors in the sea area.

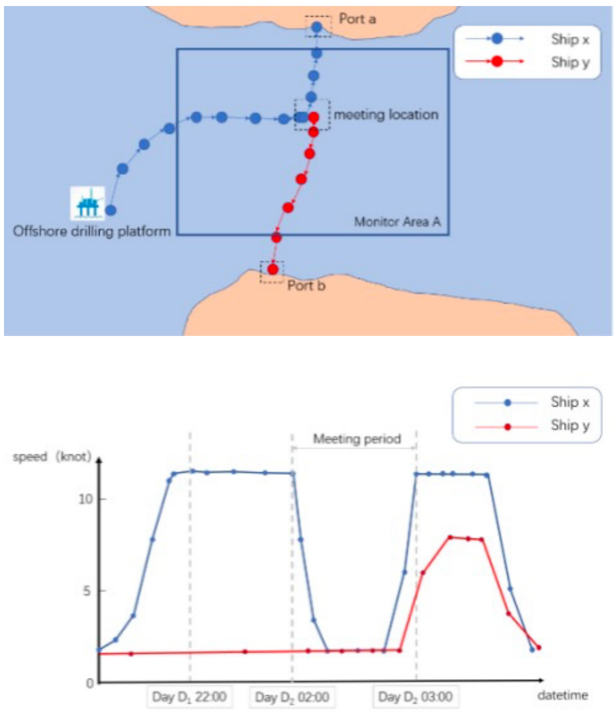


Figure 4. Illustration of case I

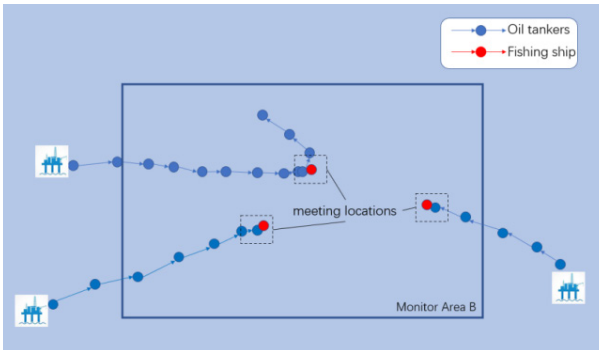


Figure 5. Illustration of case II

5 Conclusion

In this paper, based on the space-time big data of marine targets such as radar and AIS track, the corresponding recognition and detection algorithm is designed for the common multi-target/multi-target signal anomaly detection problem of ship sea connection. Integrating Algorithm 1 and Algorithm 2, the time complexity of the algorithm is O (n). At present, the time complexity of the algorithm matching for the lightering model of maritime targets is generally O (n * n) while the one-to-one matching of targets on the sea surface is carried out by using the exhaustive method, which is not suitable for large-scale data processing scenarios due to its low execution. The algorithm is improved based on the principle of hash, which improves the efficiency of the algorithm and reduces the number of matching between targets. At the same time, it enhances the reliability and reusability of data, and is suitable for real-time streaming computing and large-scale offline computing. The real-time and batch detection of transmission and identity tampering is realized, and the effectiveness of the algorithm is basically verified, which shows good performance in practical application scenarios. In the process of verifying the detection method of maritime connection behavior, it is found that the current maritime criminal activities also pay more and more attention to the camouflage and concealment in the information domain. In the actual maritime activities, they show that they have arrived at the operation area, closed communication equipment such as AIS and GPS, and even tampered with the information of active communication equipment. In the future research, the team will consider combining maritime behavior detection with trajectory prediction methods, carry out the research on "Stealth barge passing detection method", deal with the situation that ships close or tamper with information in illegal maritime activities, and verify its effectiveness in practical application.

References

[1] Z. Zhang, D. Li, M. Zhao, Y. Yao, S. Y. Lee, Parallel Planning of Marine Observation Tasks Based on Threading Building Blocks, *International Journal of Performability Engineering*, Vol. 17, No. 9, pp. 756-765, September, 2021.

[2] Z. Chen, R. Zhai, D. Li, J. Bian, S. Zhang, S. Xu, Performance Evaluation of a Tactical Data-link System Based on MSK and 16QAM, *IEEE Access*, Vol. 9, pp. 84316-84326, June, 2021.

[3] Q. He, Y. Tian, D. Li, W. Liu, M. Jian, Satellite Imaging Task Planning Using Particle Swarm Optimization and Tabu Search, *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, Hainan, China, 2021, pp. 589-595.

[4] F. Castaldo, F. A. Palmieri, V. Bastani, L. Marcenaro, C. Regazzoni, Abnormal Vessel Behavior Detection in Port Areas Based on Dynamic Bayesian Networks, *17th International Conference on Information Fusion (FUSION)*, Salamanca, Spain, 2014, pp. 1-7.

[5] S. Zouaoui, V. Roy, N. Maïzi, Behavior Analysis Modulus for harbor security, *2012 Oceans*, Hampton Roads, VA, USA, 2012, pp. 1-9.

[6] Z. Wei, X. Xie, W. Pan, R. Zhao, Ship Abnormal Behavior Detection Based on Naive Bayes, *Journal of Transportation Systems Engineering and Information Technology*, Vol. 17, No. 6, pp. 147-154, December, 2017.

[7] C. H. Botts, A Novel Metric for Detecting Anomalous Ship Behavior Using a Variation of the Dbscan Clustering Algorithm, *SN Computer Science*, Vol. 2, No. 5, Article No. 412, September, 2021.

[8] S. Jia, F. Yang, Intelligent Recognition Method for Abnormal Behavior of Ships in Port Waters Under Complex Scenes, *Journal of Ship science and technology*, Vol. 40, No. 24, pp. 7-9, December, 2018.

[9] C. Wang, *Research on Berthing Behavior of Vessels Based on AIS Data*, Master's Thesis, Dalian Maritime University, Liaoning, China, 2020.

[10] Y. M. Wang, *Detection and Early Warning Method of Ship Abnormal Behavior Based on Massive AIS Data*, Ph. D. Thesis, Dalian Maritime University, Liaoning, China, 2020.

[11] Z. Wang, K. Zhao, C. Cai, M. Ding, P. Wang, Analysis of Vessel Abnormal Behavior Based on DBSCAN and IForest Algorithms, *Journal of Ship Electronic Engineering*, Vol. 41, No. 4, pp. 89-94, April, 2021.

[12] Y. Zheng, D. Li, L. Wang, M. Zhao, W. Ying, Robustness of the Planning Algorithm for Ocean Observation Tasks, *International Journal of Performability Engineering*, Vol. 16, No. 4, pp. 629-638, April, 2020.

[13] K. F. Toloue, M. V. Jahan, Anomalous Behavior Detection of Marine Vessels Based on Hidden Markov Model, *6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, Kerman, Iran, 2018, pp. 10-12.

[14] Y. Wang, Y. B. Wang, J. Zheng, Analyses of Trawling Track and Fishing Activity Based on the Data of Vessel Monitoring System (VMS): A Case Study of the Single Otter Trawl Vessels in the Zhoushan Fishing Ground, *Journal of Ocean University of China*, Vol. 14, No. 1, pp. 89-96, February, 2015.

[15] E. N. de Souza, K. Boerder, S. Matwin, B. Worm, Improving Fishing Pattern Detection from Satellite AIS

Using Data Mining and Machine Learning, *Journal of PLoS ONE,* Vol. 11, No. 7, Article No. e0158248, July, 2016.

[16] W. Zhao, Z. Hu, C. Wei, Ship Sailing Period Identification Based on AIS, *Journal of Computer Applications and Software*, Vol. 35, No. 10, pp. 111-116, October, 2018.

[17] Y. Q. Wen, Y. M. Zhang, L. Huang, C. Zhou, C. Xiao, F. Zhang, Mechanism of Ship Behavior Dynamic Reasoning Based on Semantics, *Journal of Navigation of China,* Vol. 42, No. 3, pp. 34-39. 2019.

[18] S. Jindal, M. Sachdeva, A. K. S. Kushwaha, Human Activity Recognition Using Ensemble Convolutional Neural Networks and Long Short-Term Memory, *International Journal of Performability Engineering*, Vol. 18, No. 9, pp. 660-667, September, 2022.

[19] T. Zhang, S. Zhao, B. Cheng, J. L. Chen, Detection of AIS Closing Behavior and MMSI Spoofing Behavior of Ships Based on Spatiotemporal Data, *Journal of Remote Sensing*, Vol. 12, No. 4, Article No. 702, February, 2020.

Biographies

Wen Ying graduated from Imperial College London with a master's degree in optics and photonics. He is currently pursuing a PhD degree in Geoscience information in China's University of Geosciences (Wuhan). He works in CETC Ocean Corporation as the director of the system integration department. His current research direction covers Ocean Big Data Applications, intelligent operation, maintenance and control of sea-related system and equipment.

Mingwang Ou received his B.S. degree in software engineering in 2016 in the South of China's University and M.S. degree in software engineering in 2019 in Hainan University. He is a software and data mining engineer in CETC Ocean Corporation, China. He current research direction covers Ocean Big Data Applications, Artificial Intelligence Algorithm, Ocean Data Analysis and Mining, computer vision and Software R&D.

Qichun Liang received his M.S. degree in GIS in 2017 in University of Chinese Academy of Sciences. He is a Vice General Manager in Research Institute Co., Ltd, China. His current research interests include Ocean big data and Ocean remote sensing.

Zhixia Yang received her M.S. degree in Electronic Engineering in 2017 in China's University of Mining and Technology. She is a system integration engineer in CETC Ocean Corporation. Her current research interests include ocean data analysis and intelligent control system.

Man Zhao received her B.S. degree, Master degree in Computer Science and PhD in Geoscience information. She is an associate professor in the School of Computer, China's Univ. of Geosciences (Wuhan), Hubei, China. Her current research interests include Intelligent decision-making and planning.