A Novel Ship Trajectory Clustering Analysis and Anomaly Detection Method Based on AIS Data

Chuang Zhang[1], Songtao Liu[1], Muzhuang Guo[1**], Yuanchang Liu[2*]
1. Navigation College, Dalian Maritime University, Dalian, China. 2. Department of Mechanical Engineering, University College London, London, UK.

Abstract
The increasing volume of ship traffic has resulted in new challenges for the supervision of maritime safety administration. The conventional manual monitoring approach for maritime traffic is inefficient and lacks specifics, particularly for supervising ships with abnormal trajectories. To address this issue, this study proposes utilizing the minimum description length criterion to extract features from ship trajectory data provided by the automatic identification system (AIS). This approach simplifies the compression of ship trajectories. Additionally, the dynamic time warping trajectory similarity measurement algorithm is employed to optimize the density-based spatial clustering of applications with noise algorithm. This optimization enables the clustering of ship trajectories and the acquisition of normalized ship motion trajectories. Furthermore, a ship trajectory prediction method based on a transformer model is proposed, and the normalized motion trajectory is used as the training set for model training. The trained ship trajectory prediction model is subsequently utilized to predict the target ship trajectory. The AIS ship trajectory data in the vicinity of Yantai Port were used for experimental verification. The results demonstrate the effectiveness of the proposed approach in identifying abnormal ship trajectories.

Keywords: Ship trajectory; Minimum description length; Dynamic Time Warping; Transformer

1. Introduction

The rapid development of the automatic identification system (AIS) has enabled the acquisition of extensive ship motion trajectory data, providing a foundational basis for predicting ship behavior and monitoring abnormal ship behavior. In the realm of maritime research, ship trajectory clustering and anomaly detection algorithms have become focal points, particularly with the development of Internet technology and big data analysis (Yang et al., 2019). The significance of ship trajectory clustering analysis and anomaly detection lies in advancing shipping management intelligence, enhancing navigation safety, and enhancing the efficiency of the shipping industry.

Research of anomaly detection was initially conducted using statistical methods (Kowalska and Peel, 2012), the form of normal trajectory model was a probability model of trajectory point information, which primarily include Kernel Density Estimation (KDE) (Ristic et al., 2008; Dai et al., 2020; Wang et al., 2022), Gaussian Mixture Model (GMM) (Laxhammar et al., 2009), Gaussian Process (GP) (Smith et al., 2012), Hidden Markov Model (HMM) (Shahir et al., 2014), and Bayesian Network (BN) (Mascaro et al., 2014; Zhen et al. 2017; d'Afflisio et al. 2021). The statistical method of anomaly detection uses statistical testing to determine whether the behavior of ship matches a statistical model representing conventional ship behavior. When the matching probability is low, it is considered as abnormal behavior. The disadvantage of statistical method is that the matching accuracy is related to historical data and ignores real-time scene. And this method is difficult to combine multi-source information, expert knowledge, etc., and is not suitable for ship trajectory anomaly detection in the increasingly complex marine traffic circumstance.

In recent years, much effort has been done to improve the performance of the existing algorithms to make them applicable for ship trajectory anomaly detection in the increasingly complex marine traffic circumstance. Clustering, as a tool for big data analysis, is an unsupervised technique that does not depend on any prior knowledge. A trajectory clustering framework based on AIS data was designed to analyze routes, which considered the geographic spatial information and contextual features of ship trajectories, and thereby the density-based clustering algorithm automatically classified different routes (Sheng and Yin, 2018). A new features of local fast ship trajectories method was proposed to search for global and local features of ship trajectories (Tang et al., 2021). Soares et al., (2015) proposed an unsupervised method to segment trajectories without predetermined clear criteria. The DBSCAN algorithm (Ester et al., 1996) is a pioneer technique in the context of density-based clustering. Lei (2016) developed a framework called MT-MAD (maritime trajectory modeling and anomaly detection) to explore frequent movement behaviors and established a single index for combining anomaly scores to determine the suspicious level of each ship's trajectory. The DBSCAN approach is used to cluster course over ground (COG) and speed over ground (SOG) in AIS data, considering both density points and those with similar COGs and SOGs. To handle large datasets, an improved density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm was introduced (Nooshin and Hamid, 2022; Li et al., 2020). Kontopoulos et al. (2021) clustered the track points and extracted the ship's steering points based on the DBSCAN algorithm. DBSCAN algorithm was reused to cluster track lines, in which the Lagrange interpolation algorithm was used to fill the gaps between the steering points. Han et al. (2021) proposed an optimized DBSCAN algorithm for abnormal ship trajectories detection by clustering AIS ship trajectories with the same features, and ultimately validated the performance of this method using data from the Gulf of Mexico. To speed up the computational efficiency of the clustering algorithm. Yang et al. (2022) proposed Density based Trajectory Clustering of Applications with Noise (DBTCAN) algorithm. This method uses Hausdorff distance as a similarity measure to cluster trajectories of different lengths. Therefore, the DBTCAN algorithm can not only recognize noise trajectories but also adaptively select its optimal input parameters, which can be widely applied in ocean research.

Rong et al. (2020) employed the DP algorithm to identify turning points based on ship type, size, final destination, and other maritime traffic patterns. The DBSCAN method clusters turning points, and then combines the density region of these points by KDE. DBSCAN and the Kernel Density Estimation-based Outlier Factor processing algorithm was introduced to calculate the abnormal probability distribution value of trajectory points, eliminating low-probability distribution edge points (Jin et al., 2023). Bai et al. (2023) designed an adaptive threshold fast DBSCAN algorithm for vessel trajectory clustering. The fast DTW algorithm is used to reduce the computational complexity and ensure the accuracy of trajectory similarity, and the DBSCAN parameters are adaptively determined by combining the similarity distribution of trajectories with an improved K-adaptive nearest neighbor. However, the method is distance-based trajectory clustering algorithm and does not consider the speed and acceleration of ship, and thereby attracting the attention of more and more scholars in the field of artificial intelligence technology.

Detecting anomalies in ship trajectories is a complex and challenging task due to the influence of the surrounding environment on the navigation statuses of both the own ship and other ships (Liu et al., 2022). Thus, artificial intelligence technology is increasingly applied in ship trajectory anomaly detection research. Traditional supervised learning techniques were employed to predict arrival time and reorganize data based on spatial grids. Deep learning architectures based on long short-term memory (LSTM) are also explored to address the next position prediction problem (Gözde et al., 2021). Park et al. (2021) employed the spectral clustering method to cluster ship trajectories and established a ship trajectory prediction model using bi-directional LSTM (Bi-LSTM). The support vector machine is employed to predict the ship's next course at the exit of the traffic route in Tokyo Bay based on dynamic historical AIS data. Nevertheless, greater emphasis should be placed on enhancing prediction accuracy by increasing efforts (Nishizaki et al., 2018). Venskus et al. (2017) presented a self-learning adaptive classification method based on the self-organizing map and virtual pheromone for ship trajectory anomaly detection. By utilizing the gradient descent algorithm, the verification dataset is used to calculate the pheromone intensity threshold for trajectory anomaly detection. Venskus et al. (2019) extended their previous work and studied a data batch processing strategy for neural network retraining to detect anomalies in streaming maritime traffic data. The results demonstrate that it is possible to reduce the retraining time while keeping the accuracy relatively unchanged. Mantecón et al. (2019) proposed a supervised deep learning framework for ship anomaly detection, utilizing a convolutional neural network to infer navigation states from the ship trajectory based on AIS information. Zhao and Shi (2019) combined DBSCAN and recurrent neural networks (RNN) to obtain the clustering ship trajectories and prediction of large-scale ship trajectories. Wen et al. (2020) clustered the ship trajectories based on DBSCAN algorithm to identify key regions, and then applied artificial neural networks to learn the relationships between key regions to generate reasonable routes for different ships. However, these algorithms face challenges related to parameter setting, noise recognition, and sensitivity to density distribution in datasets. With the development of deep learning technologies, Li et al. (2023) divided multiple different subsequences by quantifying the similarity of time distribution and matching the time distribution, and then constructed the adaptive transformer model based on transfer learning to conduct the accurate prediction of the future trajectory. However, limited by the long training and learning time of the mentioned models above, the time complexity of parameter calculation is still relatively high and cannot be used for trajectory prediction of a large amount of ship data.

Given this background, analyzing and studying ship trajectory clustering analysis and anomaly detection based on machine learning becomes crucial. The main contributions of this paper are as follows:

(1) The minimum description length (MDL) criterion is designed based on the model description length and the data description length using original AIS ship trajectory data, which ensures to extract ship trajectory features while well maintaining the trajectory features and shape.

(2) The DBSCAN clustering algorithm using the dynamic time warping (DTW) trajectory similarity measurement method is proposed according to the position, speed and course of ship, which essentially solves the problem of slow accuracy and trajectory similarity between ship trajectories and clustering effectiveness.

(3) A ship trajectory prediction model based on the transformer model is constructed for the anomaly detection of ship by using the transformer's self-attention mechanism to capture the long-term dependency relationship between ship trajectory information. The results demonstrate that this method is superior to other algorithms in terms of effectiveness and reliability.

This paper is organized as follows: Section 2 presents the proposed method for ship trajectory clustering analysis and anomaly detection. Section 3 presents the experiments setup include the datasets used in the experiments, evaluation results of the proposed method. At the end of this section, the proposed method is compared with other previous recent studies. Finally, the paper is concluded and some suggestions are given for future studies.

2. Methodology

To address the challenges related to long-term jumps and missing ship trajectories, a method is proposed to divide sub-trajectories and ensure the availability of ship trajectory data. The MDL criterion is utilized to extract ship trajectory feature points and derive ship trajectory features. In water traffic scenarios, the lengths of AIS ship trajectories cannot be made consistent for all trajectories. Therefore, in view of the lack of consideration for water traffic scenario factors when measuring ship trajectory similarity, a ship trajectory feature extension method is introduced, incorporating ship position, speed, and course. This extended feature set is used in conjunction with the DTW trajectory similarity measurement distance and the DBSCAN algorithm to achieve ship trajectory clustering. This approach enhances the accuracy of similarity measurement between ship trajectories and enables the generation of normalized ship motion trajectories. To address ship trajectory anomaly detection, a ship trajectory prediction method based on a transformer model is proposed. The normalized ship motion trajectory obtained from clustering serves as the training data for the network. This enables the transformer model to predict trajectories and serve as a tool for anomaly detection, completing the detection of ship trajectory anomalies. The flowchart depicting ship trajectory clustering and anomaly detection is illustrated in Fig. 1.

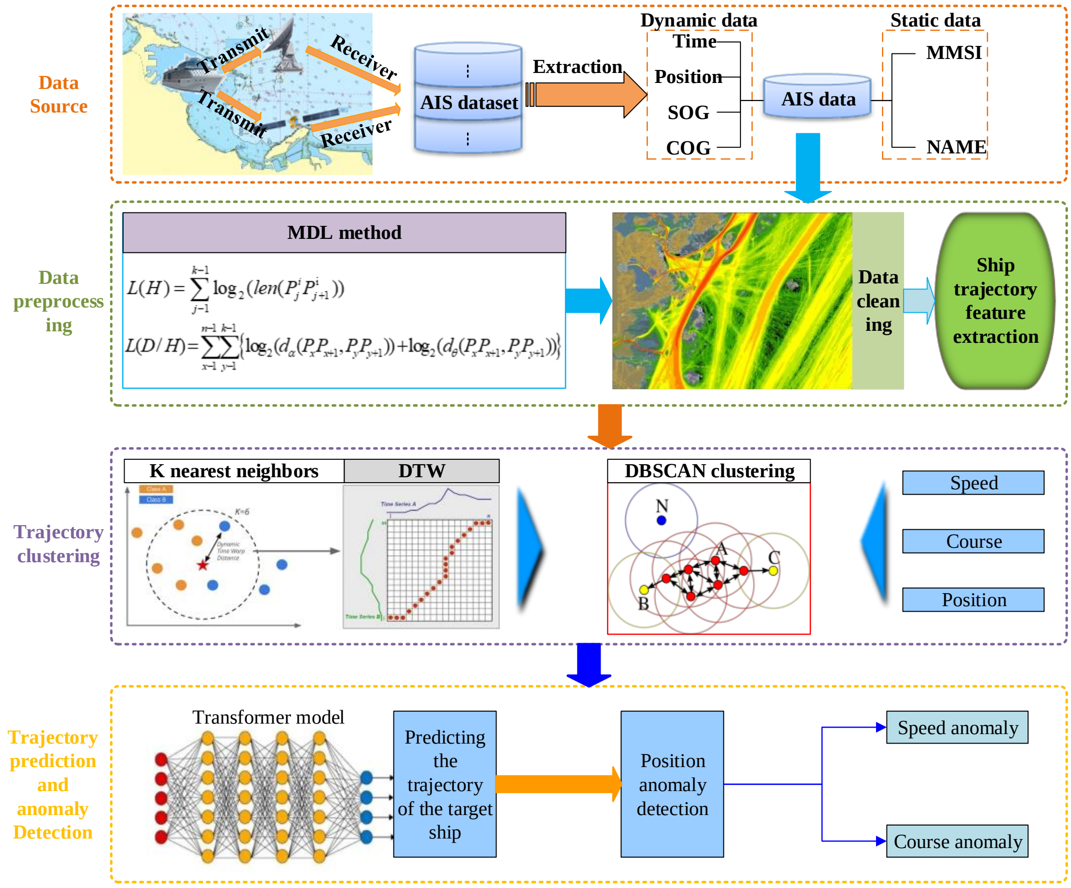

Fig. 1. Flowchart of ship trajectory clustering and anomaly detection.

2.1 Feature Extraction of AIS Data

The ship trajectory data consists of a large number of AIS data points. Meanwhile, due to the large tonnage of the ship, slow sailing speed, and gradual changes in motion status, a large overlap of ship attribute information exists between trajectory points in a short period of time. Certain trajectory points even possess identical position information, resulting in more redundancy in AIS data. In addition, there are a large number of outliers in the original AIS trajectory data. To enhance data mining efficiency, it is essential to ensure the accuracy, completeness, and conciseness of ship trajectory information. Therefore, preprocessing operations are conducted to extract ship trajectory feature points from AIS data before performing clustering analysis. The MDL criterion is chosen as the method for extracting trajectory feature points. This criterion allows for compressing ship trajectories and reducing data storage and transmission costs, thus facilitating data visualization and analysis. The fundamental idea of MDL is to encode and compress a given set of instance data using a specific model to save storage space. The compressed data, along with the model used, can be saved for future correct recovery of the instance data. The total data length to be saved corresponds to the sum total of the length of the encoded and compressed instance data and the length required to store the model. This combined length is known as the total description length.

The MDL principle consists of two components: (1) the model description length ($L(H)$), which represents the length needed to describe the model H and can be interpreted as the complexity of the model itself. Generally, models with lower complexity feature shorter description lengths. (2) The data description length ($L(D/H)$), which denotes the length required to describe the data D under the given model H. It can be interpreted as the amount of information necessary to encode the data D using the model H. Typically, if the model H effectively explains the data D, the description length of data D will be shorter under the given model conditions. As shown in Fig. 2, for a trajectory $\{P_1, P_2, P_3, P_4, P_5\}$ in a string of trajectory data, in the MDL approach, each point in the trajectory is traversed to calculate the MDL under a compression model (MDL_{par}) and the MDL under a non-compression model (MDL_{nopar}) for that point. If $MDL_{par}>MDL_{nopar}$, the point is considered a feature point. Assuming a trajectory sequence $T=\{P_1, P_2, ..., P_n\}$ is divided into multiple segments, the set of feature points is denoted as $P=\{P_{c1}, P_{c2}, P_{c3}, ..., P_{ck}\}$. The specific calculation formula for MDL is as follows:

$$L(H)=\sum_{j-1}^{k-1}\log_2(len(P_j^i P_{j+1}^i)) \tag{1}$$

$$L(D/H)=\sum_{x-1}^{n-1}\sum_{y-1}^{k-1}\left\{\log_2(d_\alpha(P_xP_{x+1},P_yP_{y+1}))+\log_2(d_\theta(P_xP_{x+1},P_yP_{y+1}))\right\} \tag{2}$$

where $Len(P)$ represents the length of sequence P, and d_α and d_θ are the vertical distance and angular distance between sequences, respectively.

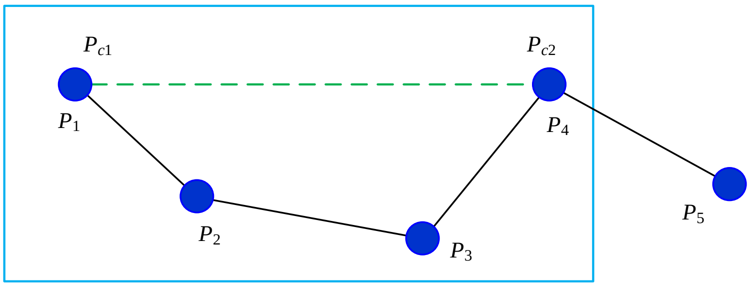

Fig. 2 Instance diagram of MDL.

According to the above formula,

$$L(H)=\log_2(len(P_1P_4)) \tag{3}$$

$$\begin{aligned} L(D/H)=&\log_2(d_\perp(P_1P_4,P_1P_2)+d_\perp(P_1P_4,P_2P_3)+d_\perp(P_1P_4,P_3P_4))+\\ &\log_2(d_\theta(P_1P_4,P_1P_2)+d_\theta(P_1P_4,P_2P_3)+d_\theta(P_1P_4,P_3P_4)) \end{aligned} \tag{4}$$

The specific process of MDL principle compression trajectory is as follows:

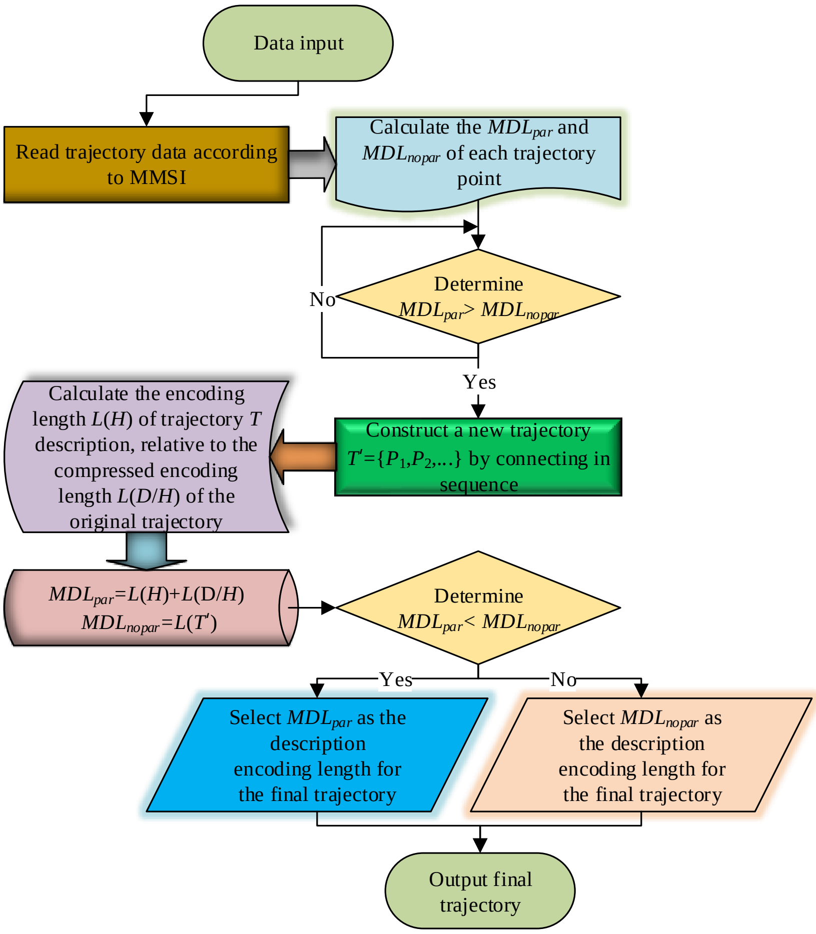

Fig. 3 Flowchart of MDL compression trajectory.

2.2 Cluster analysis of ship trajectories based on improved DTW distance algorithm

Trajectory data obtained from the AIS is a type of spatiotemporal data that includes information such as position and time. Trajectory pattern mining and trajectory classification are essential components of trajectory data research (Wang et al., 2023). Trajectory pattern mining involves grouping similar trajectories together to analyze group behavior or periodic behavior of moving objects. On the other hand, trajectory classification entails establishing a model to analyze the similarity between trajectories and partitioning them into different states. When assessing trajectory similarity, it is crucial to select an appropriate trajectory similarity measurement algorithm. To determine the similarity between different trajectories, an improved DTW distance algorithm is proposed, and the DBSCAN method is employed for clustering.

Suppose we have two trajectory sequences, $Traj_1=[p_1, p_2,..., p_n]$ and $Traj_2=[q_1, q_2,..., q_m]$, where m and n represent the number of trajectory points in $Traj_1$ and $Traj_2$, respectively. The DTW distance between $Traj_1$ and $Traj_2$ can be calculated using Equation (5), where $dist(p, q)$ refers to the Euclidean distance between trajectory points p and q. The $Rest(Traj_1)$ and $Rest(Traj_2)$ terms indicate the remaining trajectory segments of $Traj_1$ and $Traj_2$, respectively, after removing the first trajectory point. As illustrated in Fig. 4, the minimum distance between points on the trajectory sequence is calculated to determine the matching combination of trajectory points. During this process, certain trajectory points (such as q_2 and p_3) may be reused. Fig. 5 depicts the combination of all matching trajectory points in the form of a matrix. The shortest Euclidean distance matrix between each point of the two trajectory sequences is computed to determine the shortest matching path from the upper-left corner to the lower-right corner of the matrix. The sum of the weights of the connections on the optimal path represents the DTW distance.

$$DTW(T_1,T_2)=\begin{cases} 0 & m=n=0 \\ \infty & m=0 or n=0 \\ dist(a_1,b_1)+\min\begin{cases} DTW(Rest(T_1),Rest(T_2)) \\ DTW(Rest(A),B) \\ DTW(A,Rest(B)) \end{cases} & \end{cases} \tag{5}$$

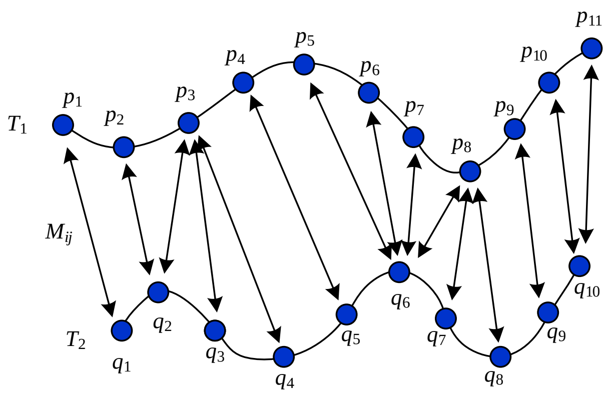

Fig. 4 Schematic of DTW calculation.

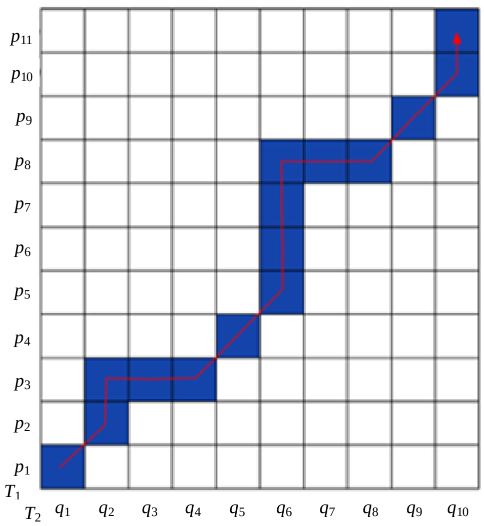

Fig. 5 Diagram of DTW time-warping distance.

The calculation process of the DTW algorithm is as follows:

1) For two trajectories with assumed lengths of n and m, they are represented as point sequences of $Traj_1$=[p_1, p_2,..., p_n] and $Traj_2$=[q_1, q_2,..., q_m], respectively.

2) An $n \times m$ distance matrix D is defined, where $D(i, j)$ represents the distance between point p_i and point q_j.

3) An $n \times m$ cumulative distance matrix C is initialized, where $C(i, j)$ represents the minimum distance between the partial trajectories from point p_1 to point p_i and from point q_1 to point q_j.

4) The first line of C and the first line D are initialized, such that $C(1,1) = D(1,1)$, $C(i, 1) = \infty$, and $C(1, j) = \infty$, where the range of i is $[2, n]$ and the range of j is $[2, m]$.

5) In $i \in [2, n]$, $j \in [2, m]$, the value of $C(i, j)$ is calculated. Among them, $D(i, j)$ represents the distance between point p_i and point q_j, whereas $\min\{C(i-1, j), C(i, j-1), C(i-1, j-1)\}$ represents the minimum value of three adjacent cells.

6) The final DTW distance is $C(n, m)$.

Because ship AIS data contains dynamic trajectory information, including position, course, and speed, and considering that the course and speed in the trajectory information will also have a certain impact on ship trajectory similarity, the AIS data is expanded to include position, speed, and course. This expansion is achieved by using a feature vector expansion method, which extends the feature vectors of trajectory points from two-dimensional vectors (longitude and latitude) to four-dimensional vectors (longitude, latitude, course, and speed). This comprehensive consideration of similarity allows for a more accurate calculation of distances between trajectory points. Assuming X_T represents the ship's navigation characteristic data at time T, the expanded feature matrix can be expressed as follows:

$$X_T = [x_t, y_t, s_t, c_t] \tag{6}$$

Each row of the expanded feature matrix represents the feature vector of a trajectory point, where x_t and y_t represent the longitude and latitude, respectively. s_t indicates the speed, and c_t indicates the course. DBSCAN is a density-based clustering algorithm that automatically discovers clusters of any shape in data and identifies noise points. In this study, the improved DTW distance algorithm is utilized as a replacement for the ε-field in DBSCAN clustering.

2.3. Ship trajectory prediction and anomaly detection

A ship trajectory prediction model based on the transformer architecture is developed using the normal trajectory model obtained from DBSCAN clustering as the foundation for data. Simultaneously, the criteria for trajectory anomaly detection are translated into feature deviation values, encompassing position, speed, and course. Real-time ship trajectory anomalies are detected by assessing the predictability of normal trajectories and the unpredictability of abnormal trajectories.

2.3.1 Transformer model

The proposed transformer model in this study incorporates several key components, including position coding, a multihead attention mechanism, residual connections, and normalization (Zhang et al, 2023).

(1) Position coding

The transformer architecture, being a parallel input model that lacks the sequential iteration advantage of RNN structures, necessitates the introduction of position coding to incorporate positional information. The input position information is expressed as follows:

$$\begin{cases} PC_{(pos,2i)} = \sin(pos / 10000^{2i/d_{\mathrm{model}}}) \\ PC_{(pos,2i+1)} = \cos(pos / 10000^{2i/d_{\mathrm{model}}}) \end{cases} \tag{7}$$

where *pos* is the position of sample, d_{model} is the total latitude of the input feature, i is the latitude label of the input feature, and the value range of $i \in [0, d_{\text{model}}/2]$.

(2) Multihead attention mechanism

The main component of the transformer model is built using a multihead attention mechanism, which comprises the self-attention mechanism. The self-attention mechanism operates on input sequences by attending to different positions within the same sequence. In Fig. 6, the self-attention mechanism is exemplified using a_2. In this mechanism, three parameters are generated: q, k, and v. For each q, calculations are performed with k from a_1, a_3, and a_4 to produce attention scores. These scores are multiplied by their respective input v, resulting in four extracted vectors. These vectors are then added together to obtain the output, b_2, produced by the self-attention model for a_2.

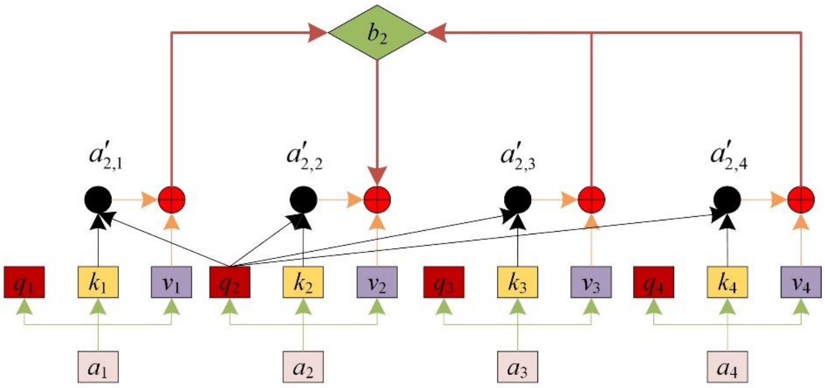

Fig. 6 Diagram of b_2 calculation.

$$b_2 = \sum_i a'_{2,i} v^i \tag{8}$$

$$\begin{cases} Q = XW^Q \\ K = XW^K \\ V = XW^V \end{cases} \tag{9}$$

The attention score matrix (Fig. 7) composed of α is calculated using Q and K. Subsequently, the final output matrix of attention layer comprising b is calculated using V and the attention score matrix. The output result is recorded as $Attention(Q, K, V)$:

$$Attention(Q,K,V) = softmax(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}})V \tag{10}$$

where Q is the query matrix, K is the key matrix, V is the value matrix, and X is the input matrix; W, Q, K, and V are linear transformation weight matrices, and *softmax* is used to calculate the weight. d_k is the dimension of input data.

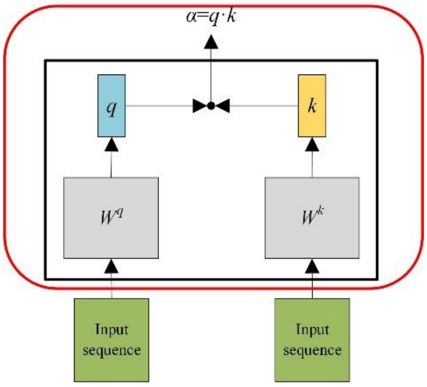

Fig. 7 Calculation process of attention score.

The calculation of the sub-attention mechanism involves applying linear transformations to the *K*, *Q*, and *V* matrices. These linear transformations serve to capture different aspects of the input and enable the model to focus on multiple levels of information. After each self-attention operation, the outputs from multiple attention heads are spliced together to form the final output. This allows the model to integrate information and enhance its overall performance. The multiple-head attention mechanism, denoted as *MultiHeads*, can be represented as follows:

$$MultiHeads(Q,K,V) = Concat(head_1,\ldots,head_h)W^o \tag{11}$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{12}$$

where W_i^Q, W_i^K, and W_i^V are the linear transformation weight matrices of the *i*-th head of *Q*, *K*, and *V*, respectively, and W^o is the linear transformation weight matrix after the multihead attention matrix is concatenated, *Concat* is a feature concatenation function.

(3) Residual connection

The submodule of the transformer model primarily consists of a multihead attention mechanism layer and a feedforward layer. Between these layers and the input layer, as well as between the attention layer and the feedforward layer, there exist residual connections and data normalization. Residual connections allow for the connection of input and output data, addressing the issues of gradient vanishing and weight matrix degradation. The dimensions of the input and output data from the attention layer are consistent, facilitating the residual connection. The residual connection formula for the multihead attention mechanism (*H*) and the residual connection formula of feedforward (*H'*) are as follows:

$$H' = X_{input} + Attention(Q,K,V) \tag{13}$$

$$H = H' + Feedforward(H') \tag{14}$$

where X_{input} is input sequence.

2.3.2 Ship trajectory prediction model based on Transformer model

In order to detect ship trajectory anomalies and construct a transformer model for predicting ship trajectory data, a threshold-based detection method is proposed. The approach involves defining trajectory clusters obtained through clustering as normalized ship motion trajectories within a selected water area. These trajectories are then used as training data to train the constructed transformer model, enabling the transformer to learn the normal ship motion trajectory model and predict future trajectory points using the historical trajectory point data of the target ship. The transformer network, composed of attention mechanisms, features excellent parallel capability and omits the processing steps of previous historical experience. The specific process steps are illustrated in Fig. 8 and can be summarized as follows:

(1) Dataset construction: The dataset consists of ship trajectory data following clustering analysis. Trajectories that can participate in clustering are selected as normalized motion trajectories within the specified area to establish the dataset. The dataset is divided into a training set and a test set, with an 8:2 ratio based on the ship's MMSI number. Each ship's trajectory data includes longitude, latitude, speed, and course, and a single trajectory's feature value at time t is denoted as $X_t = [x_t, y_t, s_t, c_t]$.

(2) Data formatting: To simplify the data, the time interval between trajectory data points is adjusted to 1 min and the data is normalized. The deviation standardization method is employed for normalization, and the transformer model is used to normalize the data obtained from ship trajectory feature prediction. The standardization formula for deviation is as follows:

$$X^* = \frac{X - Min}{Max - Min} \tag{15}$$

where Max and Min are the maximum value and minimum value in the sample data, respectively. X is the original training data, and X^* is the normalized data.

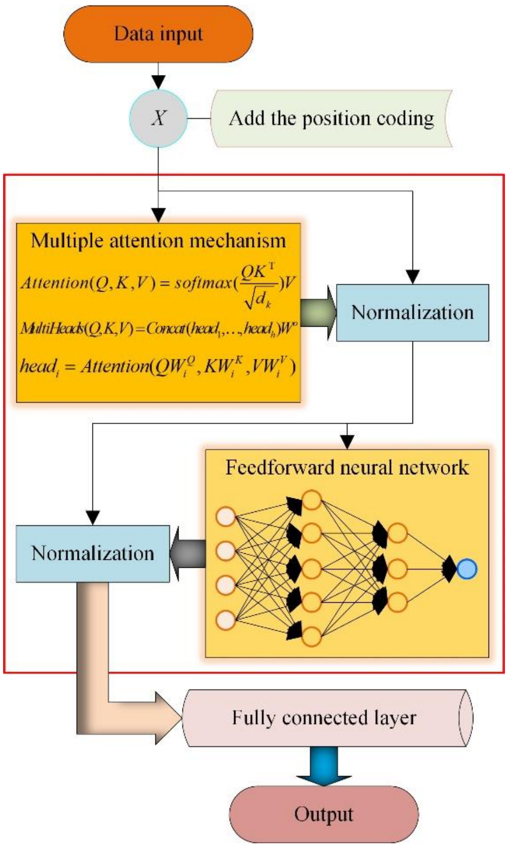

Fig. 8 Trajectory prediction process based on Transformer model.

(3) Step size and output: In order to achieve multivariate prediction with the transformer model, the original data is first transformed into a supervised learning dataset. The dataset is then converted into trajectory feature data $\{X_t, X_{t+1},..., X_{t+n}\}$ to serve as the input for the transformer model. Herein,

the input consists of the trajectory feature values of 10 consecutive time points, whereas the output corresponds to the trajectory feature values of the 11th time point. This setup allows the model to be trained on the input–output pairs and predict future ship trajectories.

When constructing the transformer model, the original encoder and decoder of the transformer is not used, and the decoder is replaced with a fully connected layer. The general process for prediction is as follows:

1) Input encoding: The input sequence is passed through an encoder, generating a series of encoding vectors. Each encoding vector encapsulates a portion of information from the input sequence.

2) Initialization of the output: A special starting symbol is added to the output sequence, and it is converted into a vector representation using an embedding layer. This vector serves as the output for the first time step.

3) Prediction of the output sequence: Starting from the first time step, the output vector and the encoding vector from the current time step are fed into the fully connected layer. The fully connected layer processes these vectors and generates the output vector for the next time step. This process is repeated until the length of the output sequence reaches the specified maximum value or the model produces a special ending symbol.

4) Generation of final prediction: Each output vector in the output sequence is converted into corresponding markers to obtain the predicted output sequence.

3. Experiment and discussion

The source of the data come provided by a professional data manufacturer of China (The time coverage is from January to June in 2022), and the format of data is shown in Table 1. The focus of this study is the sea area near the Port of Yantai in China. To facilitate the study, AIS data from a rectangular water area within the Yantai sea area was extracted. The latitude and longitude ranges selected for this area are as follows: (121.37°–121.53°E), (37.53°–37.67°N). The extracted AIS data represents the original trajectory map within this specified water area. Fig. 9 displays the extracted original trajectory map, providing a visual representation of the ship trajectories observed in the study area.

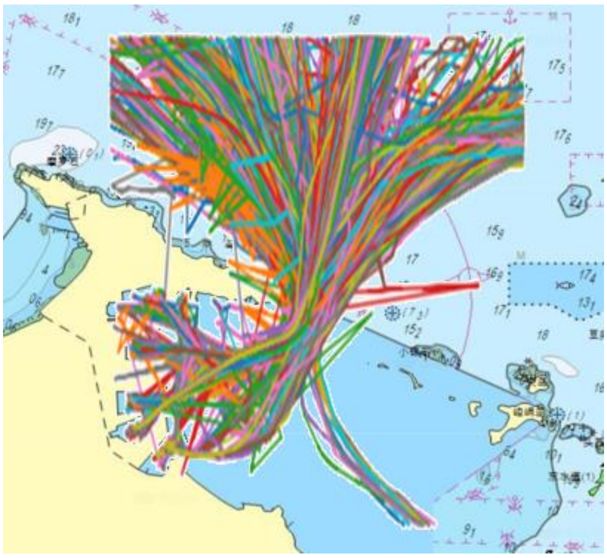

Fig. 9 Schematic of the original trajectory in the region.

Table 1 Sample table of partial ship trajectory data

Time	MMSI	longitude（°）	latitude（°）	speed（kn）	course（°）
6/11 16:01	413556760	121.390627	37.566217	8.1	82.3
6/11 16:01	413301080	121.430653	37.654277	9.8	347.3
6/11 16:01	636017220	121.401017	37.591267	9.9	172
6/11 16:01	413020510	121.442738	37.587440	9.2	222.1
6/11 16:01	414760000	121.441682	37.587040	9.3	224.8
6/11 16:01	413556760	121.391107	37.566290	8.1	81.3
6/11 16:01	413020510	121.442338	37.587088	9.2	222.1
6/11 16:01	413020540	121.439147	37.581850	4.1	16.4
6/11 16:01	414760000	121.441253	37.586712	9.3	226.4

3.1 AIS data acquisition and preprocessing

To address information errors in AIS data, this research utilizes interpolation and sub-trajectory partitioning methods to handle cases of missing speed, course, and trajectory points. Speed and course errors are directly supplemented using interpolation. However, for cases of missing trajectory points, it is necessary to distinguish between trajectory drift and trajectory point missing. Although the reasons for these abnormalities are similar, the difference lies in the duration of missing trajectory points. To differentiate between them, a clear threshold of 10 min is employed. If the time interval between adjacent points in a ship's trajectory with the same MMSI is less than 10 min, interpolation is used to supplement the missing points. However, if the interval exceeds 10 min, the trajectory is divided into two different sub-trajectories. It should be noted that sub-trajectory division may result in some sub-trajectories containing only single points or a small number of trajectory points. As these sub-trajectories do not effectively reflect the motion characteristics of the ship, only ship motion trajectories containing more than 100 data points are considered for further analysis in this research. The flowchart illustrating the process of sub-trajectory division is presented in Fig. 10, and Fig. 11 displays the trajectory image after the data cleaning procedure.

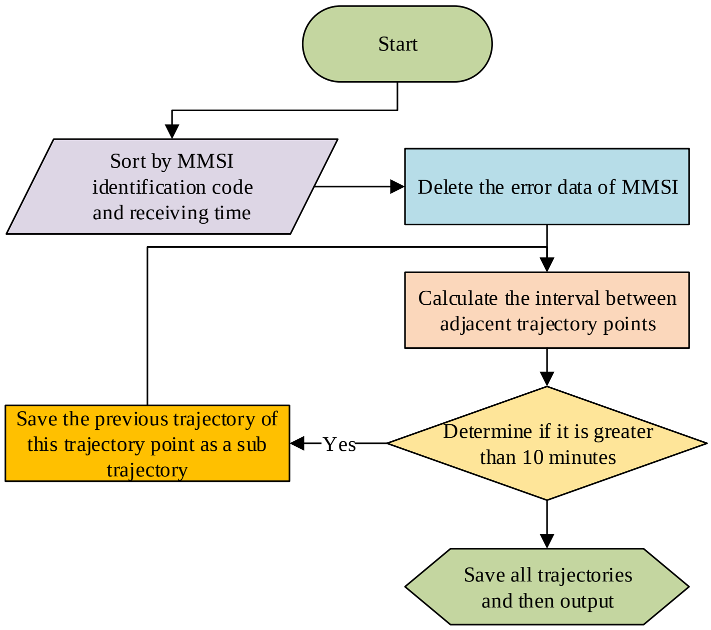

Fig. 10 Flowchart of sub-trajectory division.

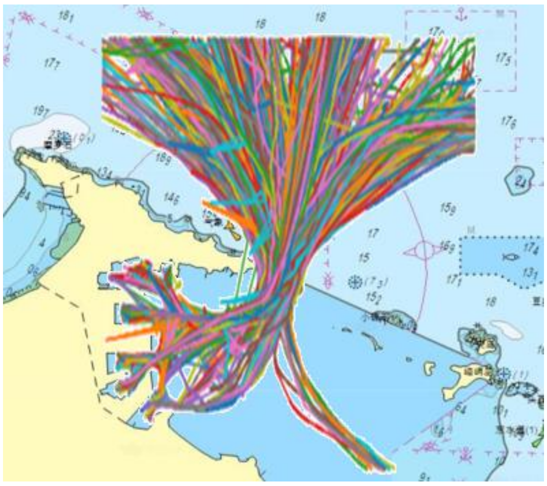

Fig. 11 Trajectory after data cleaning.

To demonstrate the effect of MDL trajectory compression, two ships with MMSI 997760305 and 413556260 were selected for compression. A comparison between the uncompressed and compressed trajectories of these ships is depicted in Fig. 12 and 13, respectively.

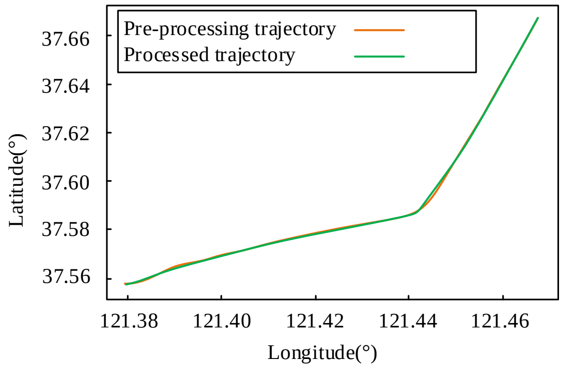

Fig. 12 Comparison between uncompressed and compressed trajectory of MMSI 997760305.

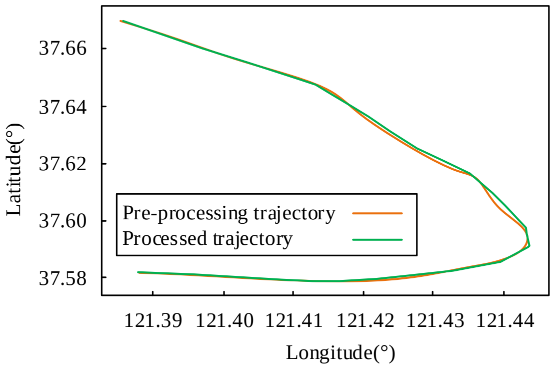

Fig. 13 Comparison between uncompressed and compressed trajectory of MMSI 413556260.

Furthermore, after applying the MDL criterion to extract feature points and compress all ship trajectories in the sea area, the overall results are presented in Fig. 14. Both individual trajectory data and the entire region's ship trajectory data exhibit a simplified and adjusted representation, while preserving the original characteristics of the ship trajectories.

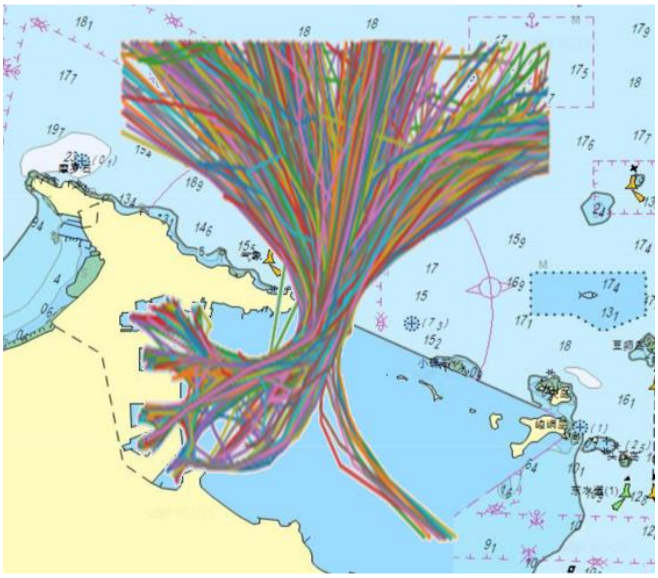

Fig. 14 Schematic of trajectory after MDL criterion in Yantai Port.

3.2 Ship clustering analysis

In terms of measuring trajectory similarity, the Frechet distance, DTW distance, fast DTW distance (Bai et al., 2023), and the improved DTW distance algorithm (after expanding trajectory similarity) are employed. Using ship trajectory data from the Port of Yantai, the DBSCAN density clustering algorithm is utilized for clustering.

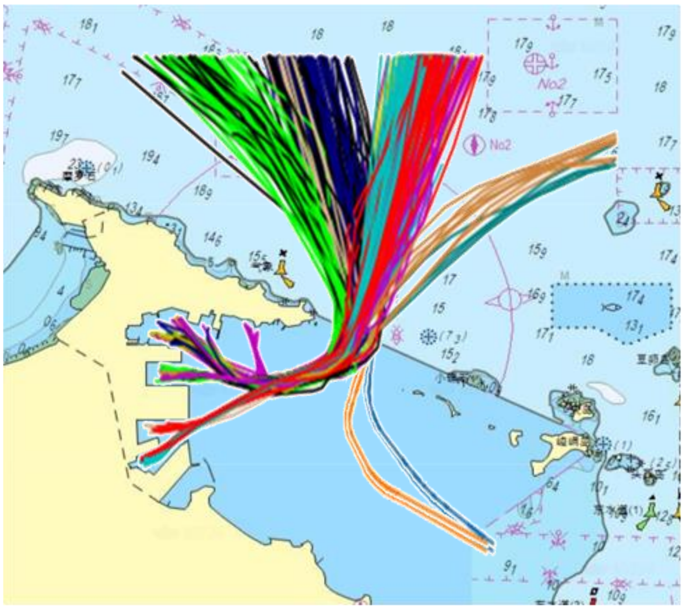

Fig. 15 Diagram of the clustering effect based on Frechet distance.

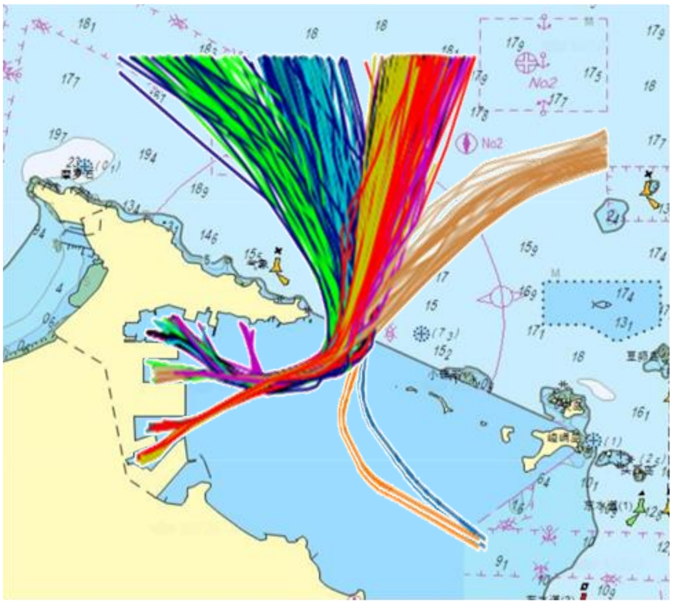

Fig. 16 Diagram of clustering effect based on DTW distance.

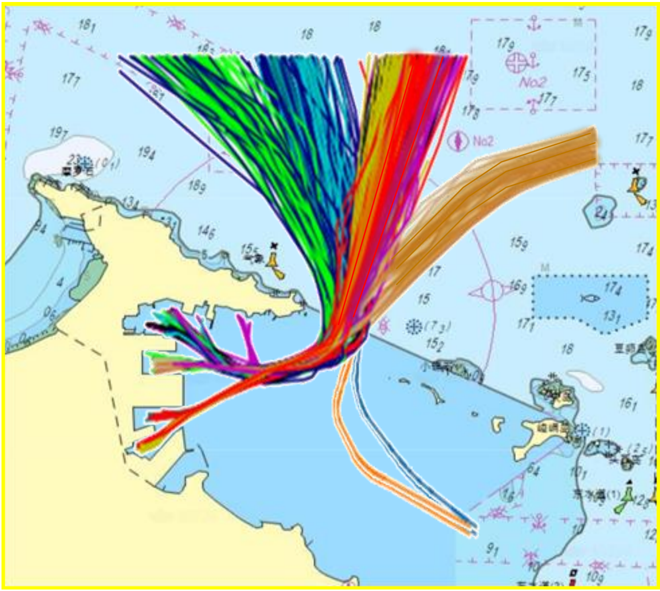

Fig. 17 Schematic based on improved fast DTW distance.

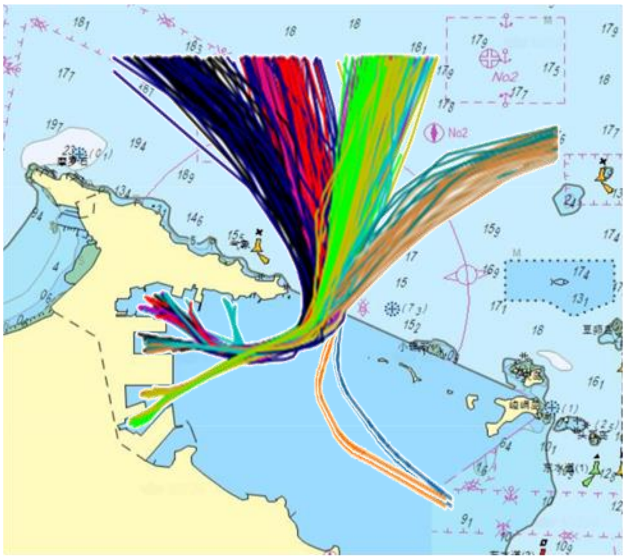

Fig. 18 Schematic based on improved DTW distance.

Table 2 Clustering number of different distance formulas

Distance similarity formula	Clustering number
Frechet distance	11
DTW	12
Fast DTW	12
Improved DTW	13

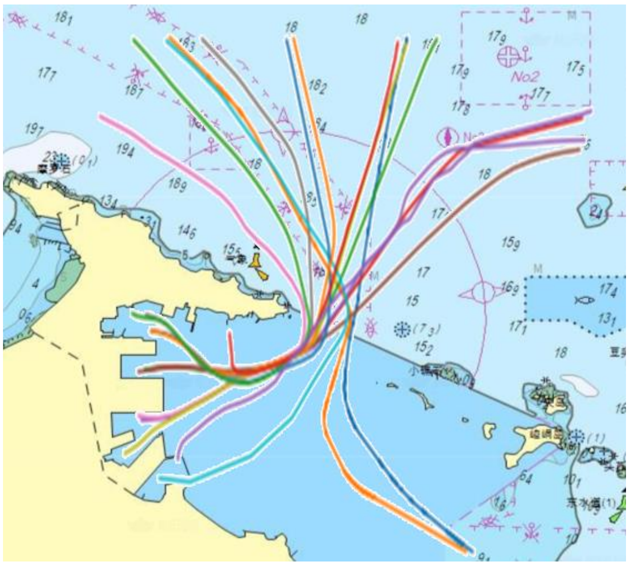

Fig. 19 Main shipping trajectory in the region.

The DBSCAN algorithm was employed to cluster ship trajectories based on the four trajectory similarity measurement methods. Each trajectory segment was considered as a core, and surrounding objects were traversed to obtain trajectory class clusters. The resulting clustering is depicted in the graph, where each color represents a different type of trajectory cluster. The clustering effect based on the four trajectory similarity measurement methods was evaluated, and the optimal clustering effect is observed. An analysis of the ship trajectories within the selected range revealed a total of 15 main ship navigation trajectories in the area (as shown in Fig. 19). Table 2 provides a summary of the clustering results using the DBSCAN algorithm based on different similarity measurement methods. The proposed improved DTW distance yielded 13 clusters of ship trajectories near the Port of Yantai, closely aligning with the primary ship movement patterns. The DBSCAN algorithm based on the Frechet distance function resulted in 11 clusters. Both fast DTW and DTW yielded 12 clusters of ship trajectories near Yantai Port, mainly because the purpose of fast DTW was to reduce overfitting and improve computing speed. It can be observed that the DTW distance-based clustering outperformed the Frechet distance-based clustering within the research sea area. The DBSCAN algorithm utilizing the improved DTW distance effectively clustered ship trajectories into different clusters based on the shape characteristics of the trajectory space and the position characteristics of the start and end points. This improved accuracy in clustering algorithms is particularly valuable in areas with cross routes and high traffic intensity. The experiment demonstrated the accuracy and effectiveness of the ship normal trajectory model established using this method.

3.3 Abnormal detection of ship trajectory

The primary ship trajectory features include position, speed, and course, play a crucial role in detecting abnormal ship trajectories. The evaluation of deviation values in these features allows for the detection of trajectories that deviate from the normal trajectory model.

The position of a ship serves as the most intuitive indicator of its trajectory status and is the most important evaluation factor in detecting abnormal ship behavior. The detection process involves ascertaining abnormalities in the behavior of the ship's position; in the presence of such abnormalities, the ship's course and speed are further examined for abnormalities. The abnormality with respect to the position of a ship is determined based on the AIS data transmitted by the ship by comparing the ship's position with the predicted position and evaluating the distance deviation. Considering various factors such as meteorological and sea conditions, ship maneuverability, and collision avoidance behavior, it is challenging for ships to precisely adhere to their intended route while sailing at sea. Consequently, two deviation warning thresholds have been established during the detection of positional abnormalities: a low threshold of 150 m and a high threshold of 250 m. Abnormal situations surpassing the high threshold trigger an alert. As depicted in Fig. 20, the target ship's trajectory is predicted using its current and historical trajectory points. Assuming that the next trajectory point after X3 is X4, two warning circles are defined based on the predicted point X4. Point X5 lies outside the warning circle due to abnormal ship course, while the abnormality of point X6 can be attributed to abnormal speed. According to this study, if we set the abnormal threshold for speed and course at 0.03 times of the predicted speed of 10 kn, then a normal trajectory would fall within the range of 9.7–10.3 kn for the actual speed.

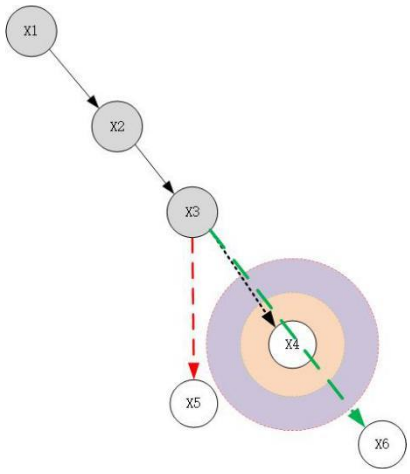

Fig. 20 Schematic of abnormal trajectory.

The flowchart illustrating the process of ship trajectory anomaly detection is presented in Fig. 21. By combining the historical trajectory of the ship with the transform model trained based on historical ship trajectories within the region to obtain the predicted trajectory of the target ship, the position of a ship trajectory status is detected for abnormal ship behavior, in the presence of such abnormalities, the ship's course and speed are further examined for abnormalities. To evaluate the effectiveness of the experiment, two abnormal trajectories featuring abnormal steering and acceleration are selected. Fig. 22 and 23 depict the instances of abnormal steering and abnormal acceleration during ship sailing.

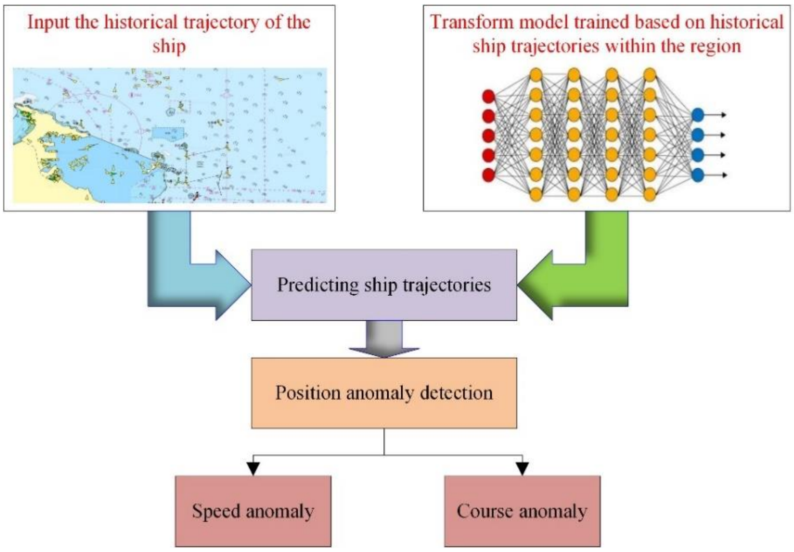

Fig. 21 Anomaly detection process of ship trajectory.

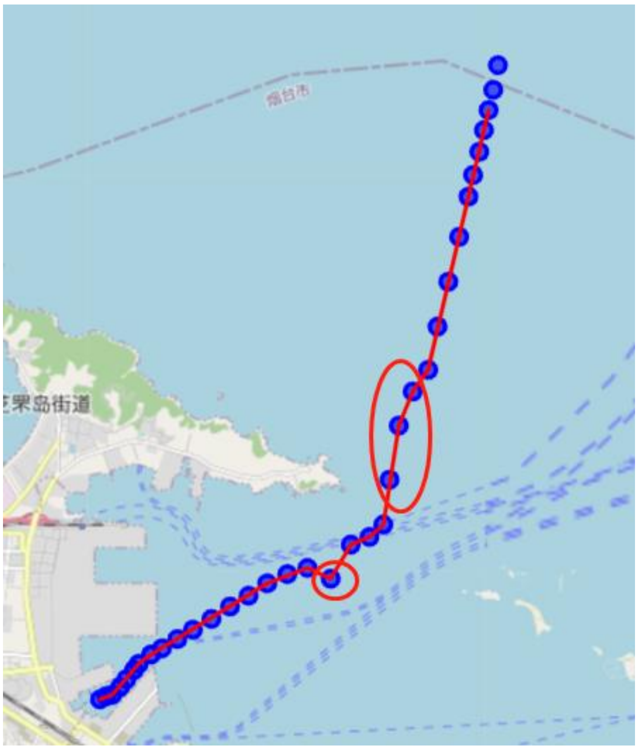

Fig. 22 Schematic of abnormal steering trajectory (trajectory 1#).

Fig. 23 Schematic of abnormal acceleration trajectory (trajectory 2#).

Fig. 24 illustrates the detected abnormal trajectory points during abnormal ship turning, which are marked with yellow circles as warning indicators. To assess the effectiveness of the detection, the differences between the real trajectory and the predicted trajectory are displayed in Fig. 25, 26, and 27. In Table 3, it can be observed that four abnormal trajectory points are successfully detected during the abnormal ship turning. However, according to the set course anomaly threshold, trajectory number 15 should also be detected as abnormal (assuming the starting point of the input trajectory is identified as 0 in the anomaly detection); however, it is not detected.

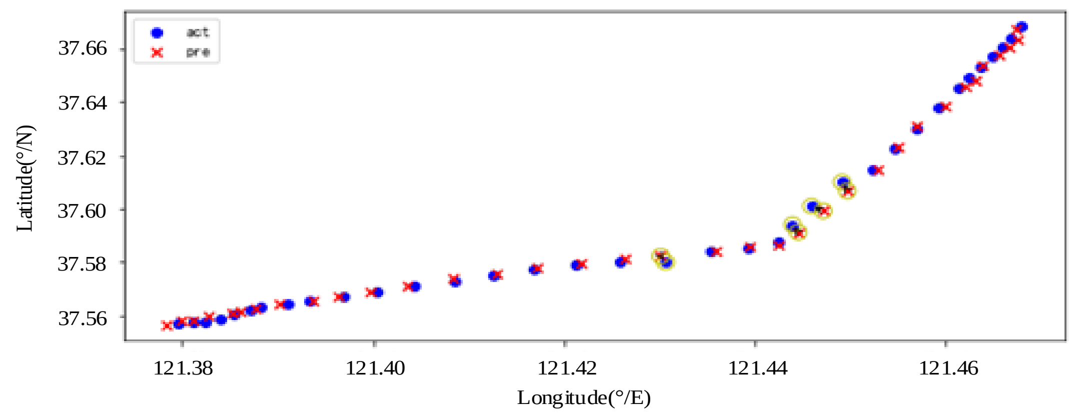

Fig. 24 Abnormal detection diagram of ship steering.

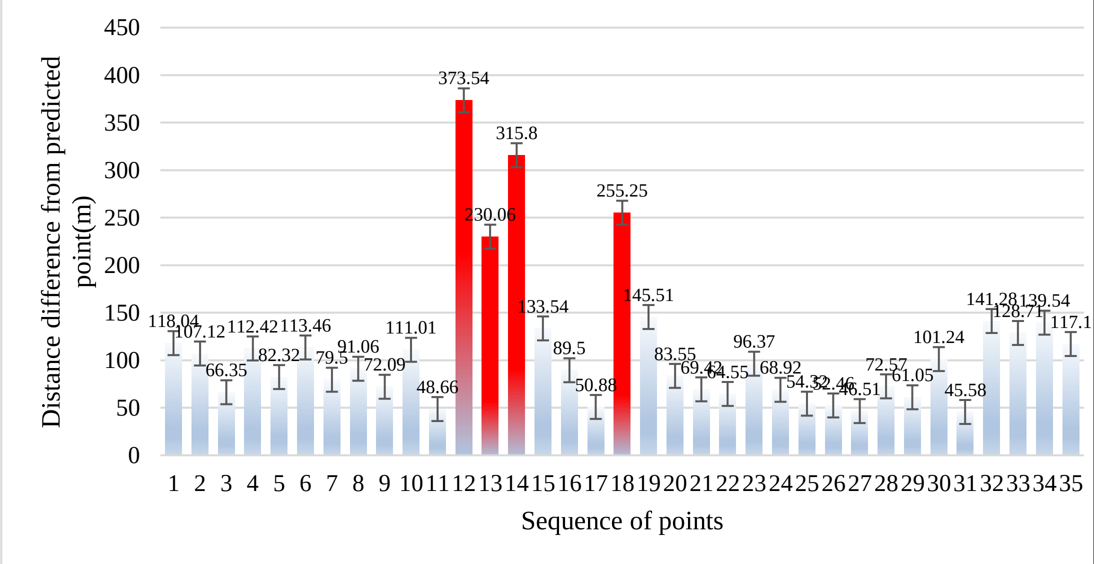

Fig. 25 Schematic of the distance difference between the real trajectory and the predicted trajectory of ship steering.

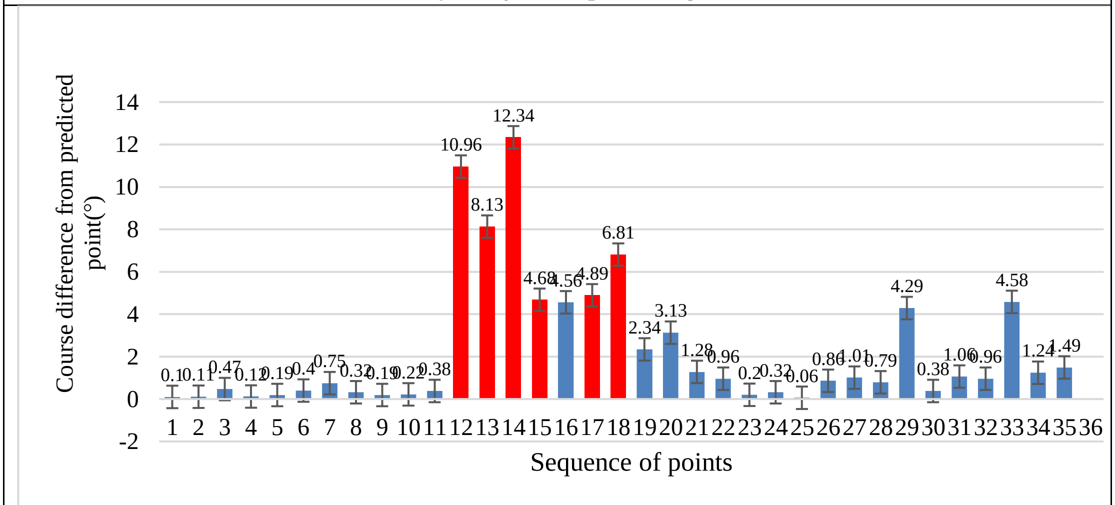

Fig. 26 Schematic of the course difference between the real trajectory and the predicted trajectory of ship steering.

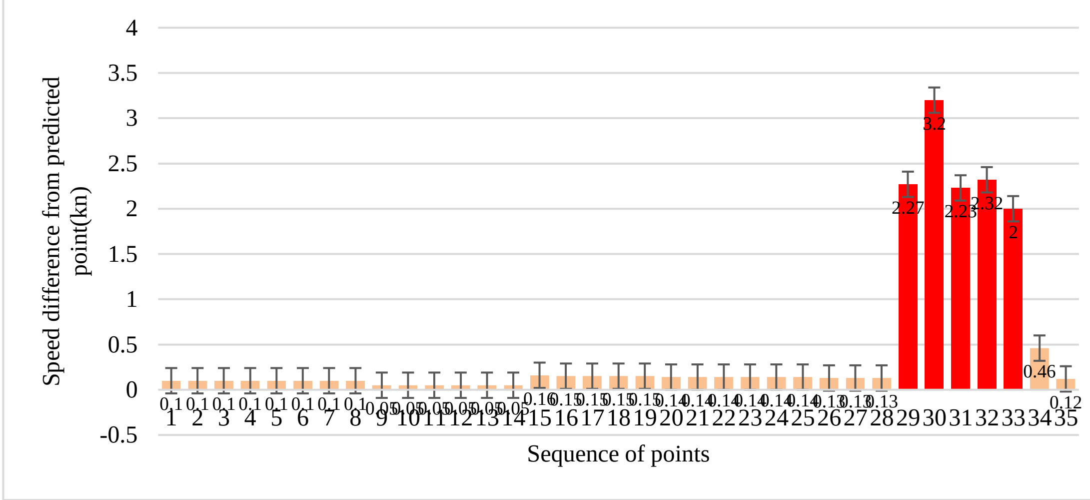

Fig. 27 Schematic of the speed difference between the real trajectory and the predicted trajectory of ship steering.

Table 3 Detection results when the ship turns abnormally.

The serial number of the trajectory	Distance difference from predicted position	Course	Speed	Detection result
12	373.54	214.96	16.9	Warning
13	230.06	202.91	16.1	Warning
14	315.80	189.32	15.2	Warning
15	133.54	190.83	14.8	No warning
17	255.25	202.09	14.4	Warning

To analyze the second trajectory with the acceleration anomaly, the proposed anomaly detection method from this research is employed, resulting in the outcomes displayed in Fig. 28 to 31. Fig. 28 demonstrates the detection of abnormal trajectory points—marked with yellow circles—as a warning for abnormal ship acceleration. In order to assess the effectiveness of the detection, three discrepancies between the real trajectory and the predicted trajectory are presented in Fig. 29, 30, and 31. Table 4 illustrates that five abnormal trajectory points are successfully detected during the abnormal ship turning. However, the speed anomaly with a serial number of 34 is not detected, even though it exhibits abnormal behavior.

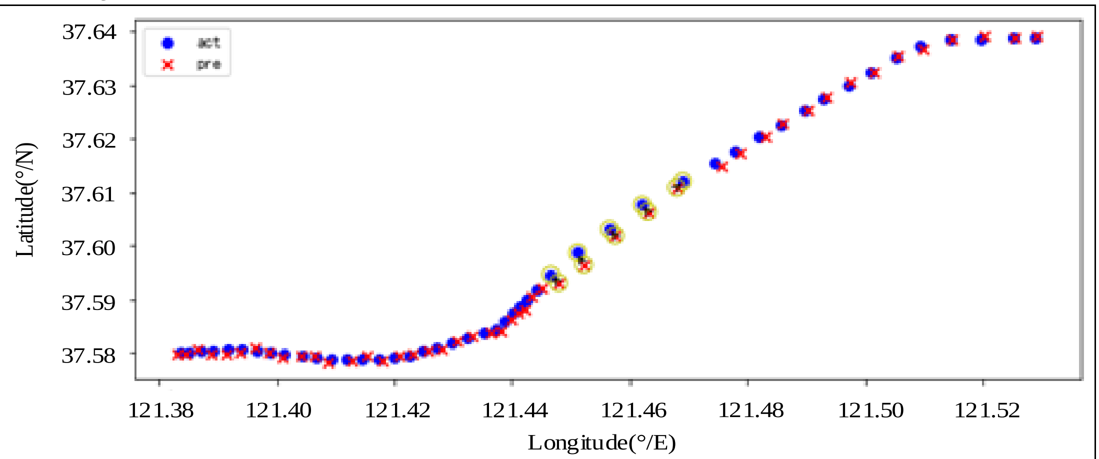

Fig. 28 Abnormal detection diagram of ship acceleration.

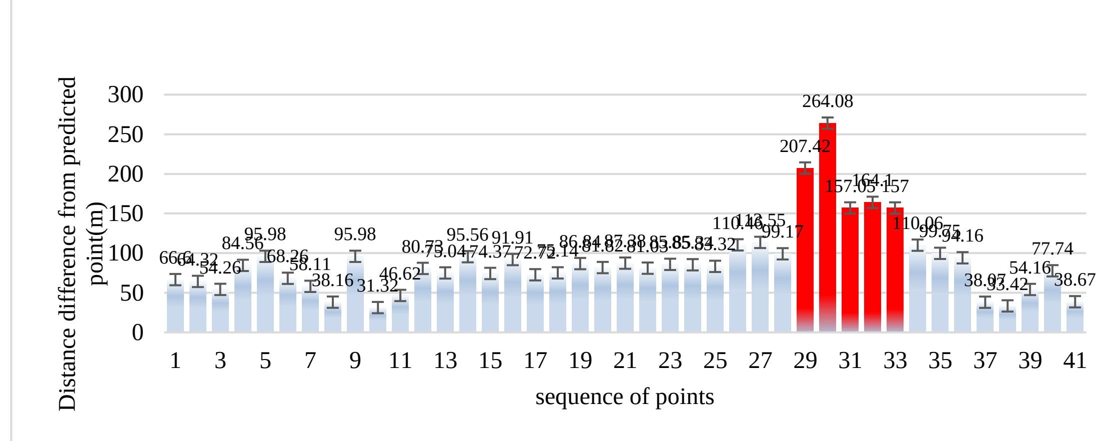

Fig. 29 Schematic of the distance difference between the real trajectory and the predicted trajectory of ship acceleration.

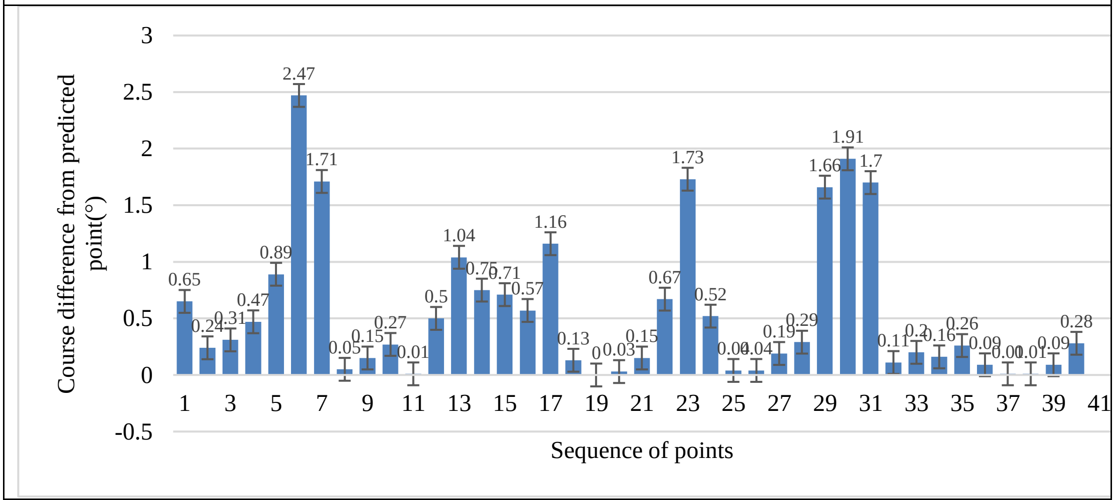

Fig. 30 Schematic of the course difference between the real trajectory and the predicted trajectory of ship acceleration.

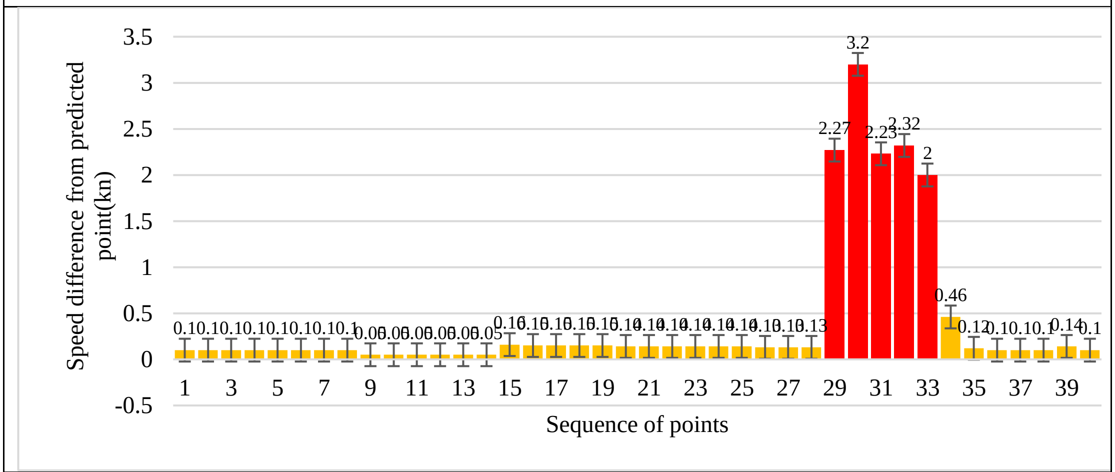

Fig. 31 Schematic of the speed difference between the real trajectory and the predicted trajectory of ship acceleration.

Table 4 Detection results when the ship accelerates abnormally.

The serial number of the trajectory	Distance difference from predicted position	Course	Speed	Detection result

29	207.42	35.5	11.1	Warning
30	264.08	41.5	13.3	Warning
31	157.08	44.1	14.6	Warning
32	164.09	43.8	15.6	Warning
33	157.00	51.1	14.9	Warning
34	110.06	50.9	13.3	No warning

The detection accuracy of the ship trajectory anomaly detection model based on transformer prediction can be calculated using the formula: $DR = 1 - (N/M \times 100\%)$, where M represents the total number of abnormal trajectory points and N represents the number of undetected abnormal trajectory points in the anomaly detection results. For ship steering anomaly trajectories, the detection accuracy is determined to be 80%, indicating that 80% of the abnormal trajectory points were correctly identified by the model. Similarly, for ship acceleration anomaly trajectories, the detection accuracy is 83.3%, indicating that 83.3% of the abnormal trajectory points were accurately detected by the model. These high accuracy rates within the water area demonstrate the effectiveness of the model in assisting maritime safety administrations in identifying abnormal ship behavior.

For the above two test trajectories, the proposed method is used to predict the ship trajectory compared with the traditional LSTM network, and the detailed results of trajectory prediction are shown in Table 5 and Table 6 respectively. It can be seen that the proposed method has the better predictive performance (i.e., the minor error) in two test trajectories on three evaluation indexes (position, course and speed). Among the 6 results of trajectory 1# adopting proposed method, the Mean error and Maximum error value has increased by 14.05%, 24.84%, 51.04%, 13.1%, 52.78% and 5.9% in position, course and speed, respectively. Similarity, for trajectory 2#, the Mean error and Maximum error value has increased by 15.62%, 32.2%, 30.38%, 18.21%, 61.3% and 4.48% in position, course and speed, respectively.

Table 5 Prediction error based on transformer model

Trajectory features	trajectory 1#		trajectory 2#	
	Mean error	Maximum error	Mean error	Maximum error
Position	112.56	373.54	89.7	264.08
Course	2.11	12.34	0.55	2.47
Speed	0.34	3.2	0.41	3.2

Table 6 Prediction error based on LSTM model

Trajectory features	trajectory 1#		trajectory 2#	
	Mean error	Maximum error	Mean error	Maximum error
Position	130.96	497.0	106.3	389.45
Course	4.31	14.2	0.79	3.02
Speed	0.72	3.4	1.06	3.35

4. Conclusion

In this research, a ship trajectory anomaly detection model is developed, taking into account the characteristics of AIS trajectory data in maritime traffic research. The proposed model includes an improved DTW algorithm that considers local trend characteristics and ship motion information, enhancing the accuracy of trajectory similarity measurement. Additionally, a transformer trajectory prediction model is constructed, incorporating a threshold-based anomaly detection method. The transformer model is trained using normalized motion trajectories obtained through DBSCAN cluster analysis, enabling it to predict ship trajectories based on historical data. The effectiveness of the proposed method is demonstrated through AIS data from the Port of Yantai, highlighting its high accuracy in detecting abnormal ship trajectories. Future research endeavors may involve considering the impact of meteorological environments, as such information can influence ship clustering and anomaly analysis. In addition, enhancing the adaptability of method parameters is one of the important research interests in the future.

Disclosure statement

No potential conflict of interest was reported by the authors.

Acknowledgements

This work was supported by the Applied Basic Research Plan of Liaoning Province in 2022 (2022JH2/101300265); the Fundamental Research Funds for the Dalian maritime university (017231034); 2023 DMU navigation college first-class interdisciplinary research project (2023JXB14).

References

Bai, Xiangen, Xie, Zhexin, Xu Xiaofeng, Xiao, Yingjie, 2023. An adaptive threshold fast DBSCAN algorithm with preserved trajectory feature points for vessel trajectory clustering. Ocean Engineering. 280, 114930.

Dahlbom, Anders, and Lars, Niklasson, 2007. Trajectory clustering for coastal surveillance. 2007 10th international Conference on Information Fusion. IEEE, 1-8.

Dai, Jianhua, Liu, Ye, Chen, Jiaolong, Liu, Xiaofeng, 2020. Fast feature selection for interval-valued data through kernel density estimation entropy. International Journal of Machine Learning and Cybernetics. 11(12), 2607-2624.

D'Afflisio, Enrica, Braca, Paolo, Willett, Peter, 2021. Malicious AIS spoofing and abnormal stealth deviations: a comprehensive statistical framework for maritime anomaly detection. IEEE Transactions on Aerospace Electronic Systems. 57(4), 2093-2108.

Ester, Martin, Kriegel, Hans-Peter, Sander, Jorg, Xu, Xiaowei, 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Kdd. Vol. 96. No. 34. pp. 226-231.

Gözde, Boztepe Karataş, Pinar, Karagoz, Orhan, Ayran, 2021. Trajectory pattern extraction and anomaly detection for maritime vessels. Internet of Things. 16, 100436.

Han, Xuyang, Armenakis, Costas, Jadidi, Mojgan, 2021. Modeling vessel behaviours by clustering AIS data using optimized DBSCAN. Sustainability. 13(15), 8162.

Jin, Chen, Hao, Chen, Quan, Chen, Xin, Song, Wang, Hongdong, 2023. Vessel sailing route extraction and analysis from satellite-based AIS data using density clustering and probability algorithms. Ocean Engineering. 280, 114627.

Kontopoulos, Ioannis., Varlamis, Iraklis, Tserpes, Konstantinos, 2021. A distributed framework for extracting maritime traffic patterns. International Journal of Geographical Information Science. 35(4), 767-792.

Kowalska, Kira, Peel, Leto, 2012. Maritime anomaly detection using Gaussian process active learning. In: 2012 15th International Conference on Information Fusion, pp. 1164-1171.

Laxhammar, Rikard, Falkman, Göran, Sviestins, Egils, 2009. Anomaly detection in sea traffic-a comparison of the Gaussian mixture model and the kernel density estimator. In: 2009 12th International Conference on Information Fusion, pp. 756-763.

Lei, Po Ruey, 2016. A framework for anomaly detection in maritime trajectory behavior. Knowledge and Information Systems. 47(1), 189-214.

Li, Kexin, Guo, Jian, Li, Ranchong, Wang, Yujun, Li, Zongming, Miu, Kun, Chen, Hui. 2023. The

Abnormal Detection Method of Ship Trajectory with Adaptive Transformer Model Based on Migration Learning. In: 2023 4th International Conference on Spatial Data and Intelligence, pp. 204-220.
Li, Huanhuan, Liu, Liu, Jingxian, Yang, Zaili, Liu, Ryanwen, Wu, Kefeng, Wan, Yuan, 2020. Adaptively constrained dynamic time warping for time series classification and clustering. Information Sciences. 534, 97-116.
Liu, Ryanwen, Liang, Maohan, Nie, Jiangtian, Yuan, Yanli, Xiong, Zehui, Yu, Han, Guizani, Nadra, 2022. STMGCN: Mobile edge computing-oempowered vessel trajectory prediction using spatio-temporal multigraph convolutional network. IEEE Transactions on Industrial Informatics. 18(11), 7977-7987.
Mantecón, Tomas, Casals, David, Navarro-Corcuera, Juanjosé, del-Blanco, Carlos. R., Jaureguizar, Fernando. 2019. Deep learning to enhance maritime situation awareness. In 2019 20th International radar symposium. IEEE, 1-8.
Mascaro, Steven, Nicholso, Ann, Korb, Kevin, 2014. Anomaly detection in vessel tracks using Bayesian networks. International Journal of Approximate Reasoning. 55(1), 84-98.
Nooshin, Hanafi, Hamid, Saadatfar, 2022. A fast DBSCAN algorithm for big data based on efficient density calculation. Expert Systems with Applications. 203, 117501.
Park, Jinwan, Jeong, Ju-kwon, Park, Youngsoo, 2021. Ship trajectory prediction based on Bi-LSTM using spectral-clustered AIS data. Journal of Marine Science and Engineering. 9(9), 1037.
Ristic, Bronko, La Scala, B., Morelande, M., Gordon, N., 2008. Statistical analysis of motion patterns in AIS data: anomaly detection and motion prediction. In: 2008 11th International Conference on Information Fusion, pp. 1-7.
Rong, Hao, Teixeira, A. P., Soares, Carlos Guedes, 2020. Data mining approach to shipping route characterization and anomaly detection based on AIS data. Ocean Engineering. 198, 106936.
Shahir, Hamed Yaghoubi, Glässer, Uwe, Nalbandyan, Narek, 2014. Maritime situation analysis: a multi-vessel interaction and anomaly detection framework. In: 2014 IEEE Joint Intelligence and Security Informatics Conference, pp. 192-199.
Sheng, Pan, Yin, Jingbo, 2018. Extracting shipping route patterns by trajectory clustering model based on automatic identification system data. Sustainability 10(7), 2327.
Smith, Mark, Reece, Steven, Roberts, Stphen, Rezek, Iead., 2012. Online maritime abnormality detection using Gaussian processes and extreme value theory. In: 2012 IEEE 12th International Conference on Data Mining, pp. 645-654.
Soares Júnior, A., Moreno, Moreno, B. N., Times, V. C., Matwin, S., Formiga Cabral, L. D. A., 2015. GRASP-UTS: an algorithm for unsupervised trajectory segmentation. International Journal of Geographical Information Science, 29(1-2), 46-68.
Tang, Chunhua, Chen, Meiyue, Zhao, Jiahuan, Liu, Tao, Liu, Kang, Yan, Huaran, Xiao, Yingjie, 2021. A novel ship trajectory clustering method for finding overall and local features of ship trajectories. Ocean Engineering. 241, 110108.
Venskus, Julius, Treigys, Povilas, Bernatavičienė, Jolita, Medvedev, Viktor, Voznak, Miroslav, Kurmis, Mindaugas, Bulbenkiene, Violeta, 2017. Integration of a self-organizing map and a virtual pheromone for real-time abnormal movement detection in marine traffic. Informatica. 28(2), 359-374.
Venskus, Julius, Treigys, Povilas, Bernatavičienė, Jolita, Tamulevičius, Gintautas, Medvedev, Viktor, 2019. Real-time maritime traffic anomaly detection based on sensors and history data embedding. Sensors. 19(17), 3782.

Wang, Jiecheng, Liu, Yantong, Chang, Jincai, 2022. An improved model for kernel density estimation based on quadtree and quasi-interpolation. Mathematics. 10(14), 2402.
Wang, Siwen, Li, Ying, Hu, Xing, 2023. A novel method for ship trajectory prediction in complex scenarios based on spatio-temporal features extraction of AIS data. Ocean Engineering. 281, 114846.
Wen, Yuanqiao, Sui, Zhongyi, Zhou, Chunhui, Xiao, Changshi, Chen, Qianqian, Dong, Han, Zhang, Yimeng, 2020. Automatic ship route design between two ports: a data-driven method. Applied Ocean Research. 96, 102049.
Yang, Dong, Wu, Lingxiao, Wang, Shuaian, Jia, Haiying, Li, Kevin X, 2019. How big data enriches maritime research–a critical review of Automatic Identification System (AIS) data applications. Transport Reviews. 39(6), 755-773.
Yang, Jiaxuan, Liu, Yuan, Ma, Lingqi, Ji, Chengtao, 2022. Maritime traffic flow clustering analysis by density based trajectory clustering with noise. Ocean Engineering. 249, 111001.
Zhang, Yuqi, Li, Chaofeng, Shang, Shaopeng, Chen, Xinqiang, 2023. SwinSeg: Swin transformer and MLP hybrid network for ship segmentation in maritime surveillance system. Ocean Engineering. 281, 114885.
Zhao, Liangbin, Shi, Guoyou, 2019. Maritime anomaly detection using density-based clustering and recurrent neural network. Journal of Navigation. 72(4), 894-916.
Zhen, Rong, Jin, Yongxing, Hu, Qinyou, Shao, Zheping, & Nikitakos, Nikitas, 2017. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and naïve bayes classifier. Journal of Navigation. 70(3), 1-23.