Applied Ocean Research 96 (2020) 102049

Contents lists available at ScienceDirect

Applied Ocean Research

journal homepage: www.elsevier.com/locate/apor

Automatic ship route design between two ports: A data-driven method

Yuanqiao Wen[a,b], Zhongyi Sui[c,d], Chunhui Zhou[a,c,d,*], Changshi Xiao[a,c,d], Qianqian Chen[c,d], Dong Han[c,d], Yimeng Zhang[c,d]

[a] National Engineering Research Center for Water Transport Safety, Wuhan, 430063, China
[b] Intelligent Transportation Systems Research Center, Wuhan, 430063, China
[c] School of Navigation, Wuhan University of Technology, Wuhan, 430063, China
[d] Hubei Key Laboratory of Inland Shipping Technology, Wuhan, 430063, China

ARTICLE INFO

Keywords:
Ship route design
Data-driven
Ship dimensions
DBSCAN algorithm
ANN

ABSTRACT

With the forced installation of the ship's automatic identification system (AIS), a large amount of ship trajectory data in the world is generated. These data provide information on latitude, longitude, speed and course, and plenty of materials for maritime pattern extraction and vessel behavior prediction. And how to dig into these AIS data deeply to discover the ship behavior pattern is an important job. There are two key points on the automatic ship route design research: the turning area generation and the turning area linkage. In this paper, we integrate DBSCAN and Artificial Neural Network capable of automatic ship route design based on massive AIS data between certain ports. The main purpose of this study is to recognize the key regions by applying DBSCAN algorithm and then connect these regions automatically by cluster similarity measuring. Then artificial neural network has been used to learn the relationship of turning regions and generate a reasonable route with different ship dimensions. The main achievement of this study have twofold. First, a research framework for automatic generation of ship route is proposed. We can process big AIS data and use them to generate ship route. Second, generation of different routes according to ships of different dimension under the research framework. The method is capable of generating ship route automatically according to different ship dimensions, which has been evaluated on two real routes around the world.

1. Introduction

With the rapid development of water transportation, the navigation environment is becoming more and more complex, and the risk of water traffic is also increasing. This also raises the requirement and challenge for controller and ship operators. Therefore, in order to further ensure the safety of navigation, it is extremely urgent to scientifically understand the transportation system in water area. As in December 2004, the International Maritime Organization (IMO) required all vessels over 299GT to install an automatic identification system (AIS) transponder on board. With the increasing number of ships equipped with AIS, it has become a focus issue to make full use of ship AIS data for maritime traffic research. By obtaining the AIS data of the ship, the intelligent monitoring and forecasting system of maritime traffic can be established, which can effectively monitor the navigable condition of water area, and find the area with high security risk in the navigable water in time. It will help shore-based personnel to better coordinate and manage the navigable water [1].

The trajectory of vessels often has certain regularity. In general, the historical AIS trajectory data of the vessel hide the navigation mode between the two ports. Through the analysis of the historical AIS trajectory data, the behavior pattern of the ship can be extracted, and it plays an important role in the route recommendation, path planing, tracking, early warning and emergency management [2–6]. For the controller, if the ship's route can be extracted based on its voyage information when the ship enters the supervised area, it is of great significance to improve the supervision efficiency of the ship transportation service system and reduce the risk in navigation.

Recently the study of the extraction of ship behavior pattern and the automatic ship route design can be divided into three aspects: trajectory clustering, turning area recognition and automatic routing, machine learning algorithms for training and forecasting. In other words, automatic generation of ship route is not a short-term forecasting, but a long-term forecasting and planning. In this section, the current state of maritime pattern extraction is reviewed, and the emphases in researches are analyzed.

Clustering is one of the important research methods of maritime pattern extraction, which is conducive to obtaining ship routes

* Corresponding author.
E-mail address: chunhui@whut.edu.cn (C. Zhou).

https://doi.org/10.1016/j.apor.2019.102049
Received 26 December 2018; Received in revised form 28 December 2019; Accepted 31 December 2019
0141-1187/

information. This mainly refers to trajectory clustering. Some scholars associate IMO rules with real vessel trajectories based on DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm to extract normal traffic pattern. And they proposed an unsupervised framework of trajectory clustering that is helpful for maritime pattern extraction and abnormal behavior detection [7,8]. Recently, spectral clustering is also a common method. In particular, Principal Component Analysis(PCA) was exploited to find the top *k* principal components for choosing the k clusters. Compared with the traditional spectral clustering algorithm, the new method has superior performance in terms of quantitative and qualitative evaluations [9].

Turning area recognition and automatic routing are a very important aspect and using DBSCAN to cluster points is the most common method. In some studies, the researchers investigate a hybrid approach for unsupervised maritime waypoint discovery, DBSCAN, modified ant colony optimization, and genetic algorithm(GA) have been combined to solve the issues in setting the DBSCAN parameters. Other scientific studies have undertaken similar investigations and many studies have adopted the incremental DBSCAN algorithm, lattice-based DBSCAN algorithm [10–14]. Genetic algorithm also is used to cluster incomplete vessel position data for maritime pattern extraction and route reconstruction, and added a quad tree structure for data processing to increase the speed of calculation [15]. In addition, high winds, ocean current as well as wind-induced ocean waves can affect ship safety, and these factors will also be taken into account in ship routing design [16].

The ship route design by machine learning algorithm is actually a prediction of the ship positions, and then the route planning is based on it. Bradley et al. presented an on-the-fly learning method for probabilistic prediction of future vessel locations using Artificial Neural Networks(ANN) [17]. Therefore, a lot of studies were closely related to this work. In Simsir et al. study, a method was proposed to predict the future coordinates of a manually controlled vessel using Artificial Neural Networks, and a large number of on-line experiments have been done in Istanbul VTS center. The ANN in this study has been trained by using position and speed data and the model can predict ship position accurately. The artificial neural network is designed, trained to predicting the vessel's future behavior, as a cloud based web application, with the ability of overlaying predicted short and long term vessel behavior on an interactive map [18,19]. However, clustering algorithm and neural network were also combined, the historical trajectory was clustered by the K-means clustering algorithm firstly, and then Artificial Neural Network was built with speed, loading capacity, self-weight, maximum power and water level to predict the ship's trajectories [20].

We use a table to summarize the three major aspects in ship route design. As shown in Table 1.

The above studies focus on the ship behavior in an area from the perspective of management, but rarely on certain ports. Despite the availability of ship route generate research, the study of automatic ship route design with different ship dimensions (ship type) is still highly deficient. In this paper, we proposed a method for automatic ship route design based on massive AIS data between two ports. This manuscript is organized as follows: Section 1 provides an introduction to the problem. Section 2 presents the problem background and approach. Section 3 provides the step of data preparation, while Section 4 presents the method to cluster turn points. In Section 5, we describe our approach to generate ship route automatically. And in Section 6, conclusions of our paper are addressed.

2. Research background, research questions and possible methods

2.1. Research background

Before the departure of the vessel, the second officer dose the course design work. With the traditional method, he refers to the route guide and historical route between the departure port and the destination port to set the route of the voyage. The work may take a long time, depending on the complexity of the route. If a method for automatically generated routes can be proposed based on the historical AIS data of the vessels, it can provide reference for the captain and the second officer when designing the route. Before the vessel leaves the port, the second officer only needs to check the rationality of the automatic production route and whether each turning point is in the proper position, without spending a lot of time and effort on the route design.

2.2. Research questions

Most of the current research on automatic generation of ship route is based on AIS data in an area. From the perspective of management, the ship's route is extracted according to the historical trajectory, which is convenient for controllers to manage the traffic. From the perspective of ship officer, they also need a recommended route from departure port to the destination port. Hence, the appropriate method should be proposed for the ship route design automatically.

However, AIS data of different ship dimensions has not been distinguished. These researches did not consider that ships of different dimensions would have different routes. Such research is still in its infancy, few people study the automatic route design with different dimensions between two ports. The actual navigation process of a vessel between two ports is affected by a number of uncertain factors, and is difficult to illustrate with mathematical equations. Moreover, ships of different dimension usually choose different routes. Therefore, this paper puts forward the following research questions.

How to extract marine traffic patterns quickly and effectively?
How to recommend different routes for different ships?

2.3. Possible methods

Many researchers study of marine traffic route use the method of cluster, genetic algorithm, ant colony algorithm and so on. Although there are many kinds of algorithms, different methods are applicable to different scenarios. However, using trajectory clustering algorithm to extract ship behavior patterns is not only computationally expensive but also difficult to extract the navigation of any port in the world. Moreover, the optimization algorithms like genetic algorithm, ant colony algorithm and so on have the disadvantage of complicate and large computation. Furthermore, many scholars use the historical data of ships and other factors to predict the position of ships at the next moment interval and so on, so as to predict the trajectory of ships. The core issue of using neural network to predict ship trajectory is how to

Table 1
Summary of three major aspects in ship route design.

Method	Advantage	Disadvantage
Clustering algorithm	It is very effective for extracting marine traffic patterns.	There exists problem of big calculation and the recommended route cannot be given.
Turning area recognition and connection	This method is simple calculation	It is difficult to select turning points in turning area and it is difficult to prove the recommended routes are reasonable
Machine learning algorithm	This method has the function of self-learning and it can map any complex nonlinear relationship.	A large number of samples are needed for training. The reasoning process cannot be explained.

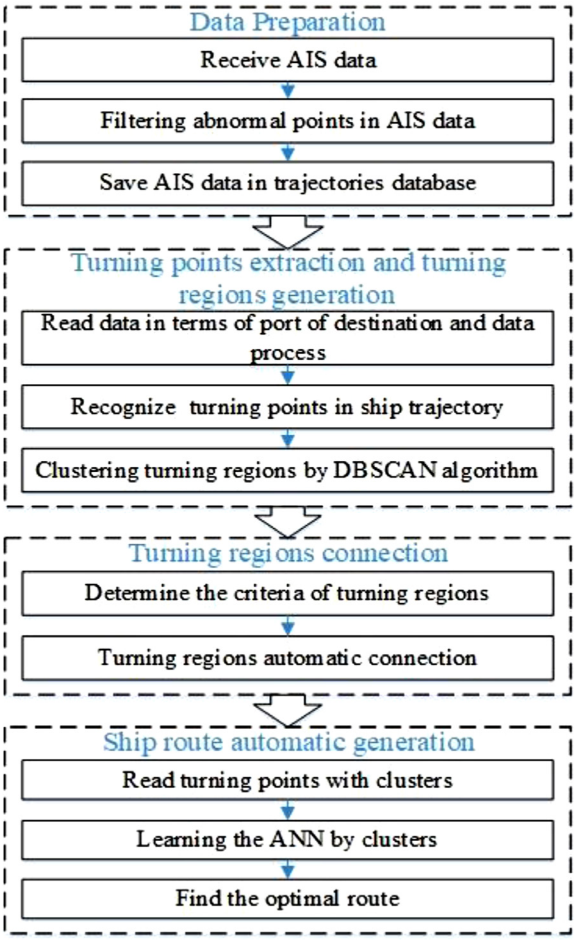

Fig. 1. The flow diagram of algorithm.

transform time series into supervised learning, however, for the generation of routes, the calculation with the method is undoubtedly enormous.

In this paper, we proposed a method of automatic ship route design based on DBSCAN algorithm and back propagation neural network algorithm. The historical AIS data between the departure port and the destination port are extracted by the voyage information of the vessel, and the DBSCAN algorithm is used to cluster the ship turning area. However, the position of ship turning in turning region is random, which is also affected by many factors, such as ship dimensions, weather, ship collision, etc. So in this paper, the longitude, latitude, speed, heading, length, width and draft of the ship in turning area are used as the input and the longitude, latitude, length, width and draft used as the labels of the three-layer model to train the neural network. The learning ability of the neural network is used to "learn" the law of vessel movement, and then the key positions of the vessel are predicted, thus the ship route of different ship dimension is generated. This approach not only avoids the complex modeling process, but also ensures that the established model is in line with the actual system. The algorithm flow diagram of this paper is shown in Fig. 1. As illustrated in Fig. 1, the algorithm architecture can be divided into four essential parts: Data preparation; Turning points extraction and turning regions generation; Turning regions connection; Ship route automatic generation.

3. Data preparation

Automatic ship route design requires proper AIS data and the quality of data will have a great impact on the clustering results and the neural networks. AIS data are extracted from the database according to departure port, destination port and ship dimensions. The data preparation involves two aspects: data preprocessing and data selection.

AIS data are received as a series of messages following a non-standard pattern of irregular time intervals and there are many bad data in AIS data. So the preprocessing is to clean the data and remove errors, for example range anomaly, velocity anomaly and course anomaly [21]. The data is then classified and stored in the database according to the voyage information and ship dimensions (Table 2). As shown in Table 2, a small portion of AIS data in database includes departure, destination, time, latitude, longitude, speed, course, length and width.

The data stored in the database cannot be used directly, and data selection is also needed. The amount of data stored in the database is huge, so the amount of computation required is enormous. Therefore, Laplacian Eigenmaps [22] and kernel density estimation [23] are used to process the AIS data, and the trajectories with high similarity are selected as the experimental data. And a large number of trajectories will make the behavor characteristics of ships more ambiguous. Therefore, the purpose of this step is to make the ship's behavioral characteristics more obvious.

Using Laplacian Eigenmaps, the high-dimensional trajectory data is reduced to one dimension, and each trajectory data is mapped to a point in the high-dimensional space, so that it becomes a vertex v_i in the graph Ψ.

In this paper, DTW distance is used to measure the similarity of trajectories [24–26].

$$DTW(i, j) = dis(x, y) + \min\{DTW(x-1, j), DTW(x, y-1), DTW(x-1, y-1)\} \tag{1}$$

In Eq. (1): $DTW(i, j)$is the DTW distance between trajectory i and trajectory j. $dis(x, y)$is the Euclidean distance between point x and point y.

The similarity $S_{ij}(DTW(i, j))$ between the trajectories is assigned to the edges e_{ij} connecting the vertices $v_i$and v_j, thereby obtaining an undirected weighted graph $\Psi(V, E)$ based on the similarity of the ship trajectory. It can be represented by join matrix$A = (a_{ij})_{m\times m}$.

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & 1 & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & & \ddots & & & \vdots \\ a_{i1} & & & 1 & & a_{im} \\ \vdots & & & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & 1 \end{bmatrix} \tag{2}$$

In Eq. (2): $a_{ij} = a_{ji}, \forall\ i = j, a_{ij} = 1$

Table 2
Preprocessed AIS data.

Departure	Destination	Time	Latitude	Longitude	Speed	Course	Length	Width
Jeddah	Singapore	t0	39.2551	19.89	21.1	153	299	45
Jeddah	Singapore	t1	39.2897	19.8216	21.4	154	330	42
Jeddah	Singapore	t2	39.3353	19.7364	21.2	152	330	43
Jeddah	Singapore	t3	39.3883	19.6405	20.4	151	335	45
Jeddah	Singapore	t4	39.4434	19.5425	20	150	336	48

Constructing a Laplacian similarity matrix:

$$L = D^{-1}(D - A)D^{-1} \tag{3}$$

In Eq. (3): $D$is a Diagonal matrix, $D_{ii} = \sum_{j=1}^{n} a_{ij}$

Calculate the eigenvector corresponding to the second small eigenvalue λ_k of matrix L.

$$q = (q_1, q_2, \cdots, q_i, \cdots, q_m)^T$$

The q^T represents the characteristics of trajectory data, and it is dimensionless quantity.

Kernel density estimation is a nonparametric density estimation method. Assume that 3 objects are given in 2 dimensional space, $D = \{q_1, q_2, \cdots, q_m\} \subset F^d(F^d$ is a collection of objects), and the kernel density estimation of object $q(q \in F^d)$ is defined as

$$\hat{f}(q) = \frac{1}{m\eta}\sum_{i=1}^{m} K\left(\frac{q - q_i}{\eta}\right) \tag{4}$$

In Eq. (4): $K(\frac{q - q_i}{\eta})$ is kernel function; η is the bandwidth of smooth parameter, $\eta = 1.05\sigma m^{-1/5}$, σ is standard deviation of samples.

Because Gauss kernel function has better smoothness, in this paper, we use the standard Gauss function as kernel function. As shown in Eq. (5).

$$K\left(\frac{q - q_i}{\eta}\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(q - q_i)^2}{2\eta^2}\right) \tag{5}$$

Therefore, the density of trajectories can be estimated as Eq. (6):

$$f(T) = \frac{1}{\sqrt{2\pi}\, m\eta}\sum_{i=1}^{m}\exp\left(-\frac{(q - q_i)}{2\eta^2}\right) \tag{6}$$

Laplacian Eigenmaps is graph-based dimensionality reduction algorithm. It expects the points related to each other (the points connected in the graph) to be as close as possible in the reduced dimension space, so that the original data structure can be maintained after dimensionality reduction. In this paper, Laplacian Eigenmaps is applied to reduce the dimension of high dimensional trajectory data to 1D. And then these one-dimensional data are processed by KDE and the standard Gauss function is used as the kernel function. According to the KDE results, the density data of the top 70% are used as experimental samples (As shown in Fig. 2). If there are multiple routes between the two ports, they will be processed separately.

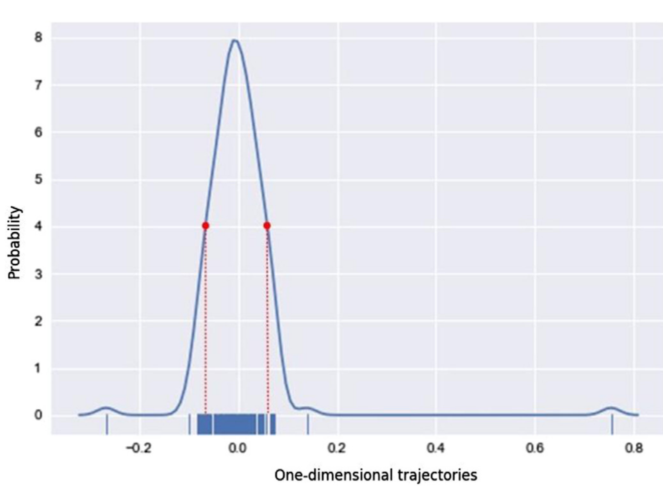

Fig. 2. Schematic of ship trajectory identification based on KDE.

4. Clustering turning regions by DBSCAN algorithm

4.1. Recognize turning points in ship trajectory

The input of the automatic ship routing model is a time series consisting of the key positions of the ship. It is important to identify the turning point from the original AIS trajectory data for automatic ship route production.

The trajectory of a vessel is made up of AIS data with time series, and each of which records the time, latitude, longitude, speed, and course of the ship's location. The AIS trajectory can be represented as $P = \{n_1, n_2, \cdots n_{t-1}, n_t, n_{t+1}\cdots\}$, $n = [\varphi, \lambda, \nu, c]$, φ, λ, ν and c represent latitude, longitude, speed and course, respectively. When the ship in the steering operation, the turning rate will increase, and the heading will change significantly after the steering. Therefore, the turning points of the vessel can be detected from the historical trajectory where the heading angle difference condition occurs. As shown in Fig. 3.

The heading angle difference between the two points in the AIS data can be expressed as$\Delta\theta = c_{t+1} - c_t$. If $\Delta\theta$ is greater than the set threshold, and then n_t can be considered as a vessel turning point. Furthermore, inspired by article [14], we assume two key parameters, $Course_{start}$ and$Course_{end}$. The $Course_{start}$is the start course of the turning point, taking the direction of the turning point n_t as the starting course of the turning point. And $Course_{end}$ is the end course of the turning point and it is the course after the end of a turning movement, taking the heading of a point after the turning point as the end of the course. If one movement is over, the ship will continue to turn direction, the $Course_{end}$ will be considered as the start course of the next turning point(As shown in Fig. 3(a)). If the turn is over, the ship will no longer change its course greatly, that can be considered as the end of the turning action (As shown in Fig. 3(b)).

This paper selects ship AIS data for the Jeddah-Singapore and Shanghai-Shenzhen as experimental data. According to the ship voyage information, the historical AIS data was screened from the AIS database, and statistical analysis of the heading angle difference between the two sequential points on each route is made, as shown in Fig. 4.

It can be seen from Fig.4 that the turning angle of the ship is mainly concentrated in the interval of 0–5°, and the proportion of the angular difference exceeding 5° is relatively small. Therefore, the setting of the threshold is 5°, and the point with angle changes more than 5° is identified as the turning point.

Nevertheless, it is not difficult to see from Fig.4 that there is a situation where the angle difference is greater than 30° or even 45°, which may be caused by the existence of errors in the data or the absence of AIS data. In the 4.2 section, these special points are treated as noise points, so it doesn't effect the clustering result.

4.2. Cluster turning regions using DBSCAN

DBSCAN divides all the data points in the dataset into 3 categories: Core point, Border Point and Noise Point as shown in Fig. 5. The criteria for which type of point are mainly determined by the two parameters of the DBSCAN algorithm: *Eps* and *MinPts*. *Eps* is the domain-wide size of the object, and *MinPts* is the set threshold. Whether the number of points in the *Eps* of the object p is greater than the threshold (*MinPts*) determines the p is a core point or a non-core point (boundary point and noise point) [27].

- Core Point: A point contains more than *MinPts* points in the radius *Eps* neighborhood, point C as shown in Fig. 5.
- Border Point: A point contains less than *MinPts* points in the radius *Eps* neighborhood, but the point is in the neighborhood of the core point. Point B as shown in Fig. 5.
- Noise Point: If the point is neither the core nor the border point, it is the noise point. Point A as shown in Fig. 5.

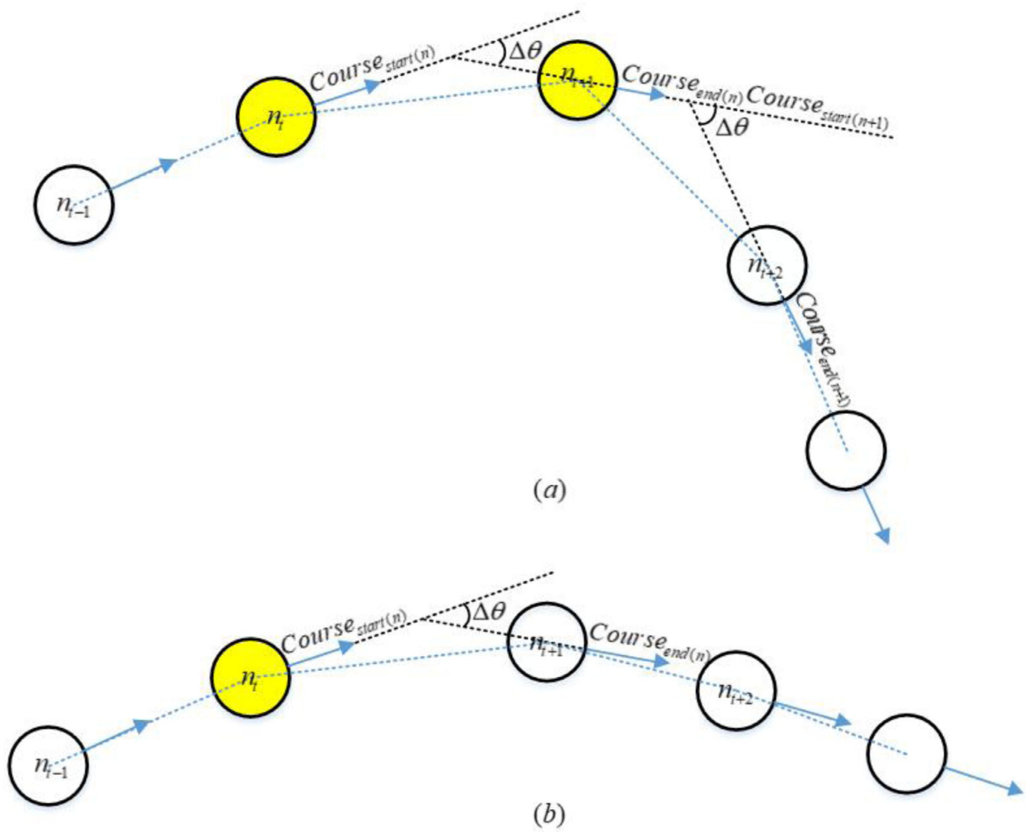

Fig. 3. Recognize turning points.

DBSCAN finds clusters by judging the connection of *Eps* neighborhoods of each point. If the number of points in the *Eps* neighborhood of object p exceeds *MinPts*, then the *p* is the core object, and the new cluster is determined by this. Hence, the DBSCAN algorithm detects the directly density-reachable objects of the core object p one by one, and classifies the retrieved objects into the same cluster as the core object p, which may lead to the merging of several density-reachable clusters in this process. If all points have been retrieved and classified, then the clustering process is finished.

If all ship track points are clustered using DBSCAN clustering algorithm, some high density areas can be identified. But a large number of points will be clustered into one class if *Eps* is too large (as shown in Fig. 6(a)). If the *Eps* is too small, the high density areas cannot be found (as shown in Fig. 6(b)).

In this paper, all the turning points in the AIS data are identified and the traditional DBSCAN algorithm is used to cluster the turning points with relatively close distance to obtain the turning area. The DBSCAN algorithm can reduce the influence of noise points on the clustering results, and the important areas of arbitrary shape can be found, which is similar to the shape of the ship's turning area in the actual navigation situation.

Normally all data are generated by ships of the same scale are on the same route. Therefore, the same behavior patterns hidden in these turning points and the turning areas of the ship can be more accurately identified through the DBSCAN clustering algorithm. Clustering algorithm was implemented by Scikit-learn's python library. The library requires the input of a distance matrix showing pairwise distances between all turning points in the experimental dataset. In this paper, a self-adaptive algorithm was adopted to determine the parameters in DBSCAN by analyzing of statistical characteristics of data sets [28].

1 Calculate the distance distribution matrix of sample data.
2 The distance distribution matrix is fitted by *Inverse Gaussian* to determine the *Eps* parameter.
3 The appropriate *minPts* are selected by analyzing the distribution characteristics of the number of noise points.

This method can realize the automation of clustering process without manual input of parameters. The clustering results are shown in Fig. 7. Due to the different sailing modes and busyness of each route, different parameters should be selected for AIS data in different regions. In the actual application process, the officer can modify the parameters according to experience to achieve the optimal effect.

4.3. Automatic linkage of turning cluster

The clustering results obtained by DBSCAN are numbered consecutively in Arabic numerals. For example, 0–11 represents 12 different clustering results. It is generally known that the numbering of the clustering results is randomly generated, and the order of the numbers is not the order from the departure port to the destination. Hence, how to automatically connect the turning regions obtained by clustering in the correct order is a key issue in the process of automatically generating routes. The turning regions are automatically connected in the correct way to lay the foundation for the training of the neural network model. Normally, the distance from the turning point to the departure port on the route is gradually increasing. Occasionally in some special waters, it turns the opposite concept.

In view of the above situation, this paper introduces the concept of *turning region course difference*, and proposes an optimal turning region matching algorithm. From the starting point, it searches for the best matching key region and automatically connects the turning region.

In Fig. 8, the turning region obtained by clustering is represented by*Cluster*, and each *Cluster*is composed of several turning points whose center coordinates are obtained by averaged the coordinates of the turning points. The start course of the turning region is defined as

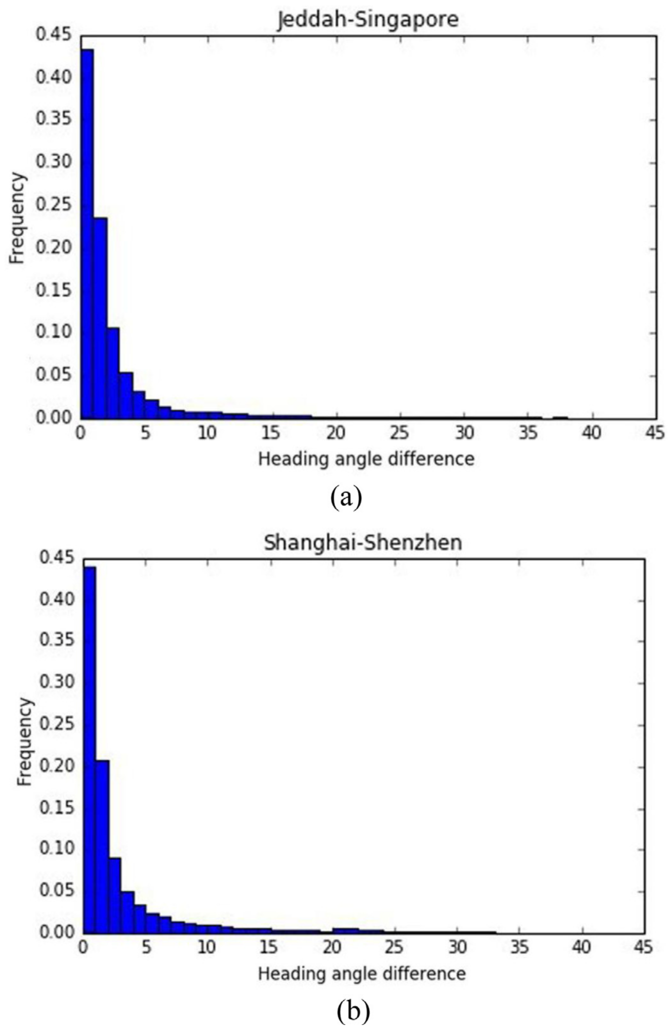

Fig. 4. A statistical analysis of the heading angle difference.

$Cluster_{start}$, which is averaged from all the $Course_{start}$ of turning points in$Cluster$. The end course of the turning region is defined as $Cluster_{end}$, which is averaging from all the $Course_{end}$ of turning points in $Cluster$. As shown in Eqs. (7) and (8).

$$Cluster_{start(a)} = \frac{\sum_{i=1}^{n_a} Course_{start(i)}}{n_a} \tag{7}$$

$$Cluster_{end(a)} = \frac{\sum_{i=1}^{n_a} Course_{end(i)}}{n_a} \tag{8}$$

The difference of steering angle betweenturning regions is defined as the difference between $Cluster_{start}$ and $Cluster_{end}$ in two different turning regions. The distance between the turning regions is calculated by the Euclidean distance between the center points of the region. As shown in Eq. (9) and Eq. (10), $c_{ij}$and c_{ji} indicate different meanings.

$$c_{ij} = Cluster_{end(i)} - Cluster_{start(j)} \tag{9}$$

$$c_{ji} = Cluster_{end(j)} - Cluster_{start(i)} \tag{10}$$

Construct distance matrix E and angle matrix F, e_{ij} indicate the distance between $Cluster_i$ and $Cluster_j$,f_{ij} indicate the difference of steering angle between $Cluster_i$ and $Cluster_j$($f_{ij} = c_{ij}, f_{ji} = c_{ji}$).The Arabic numeral 1 represents the departure port and k represents the destination port.

$$E = \begin{bmatrix} 0 & e_{12} & e_{13} & \cdots & e_{1k} \\ e_{21} & 0 & e_{23} & \cdots & e_{2k} \\ e_{31} & e_{32} & 0 & \cdots & e_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e_{k1} & e_{k2} & e_{k3} & \cdots & 0 \end{bmatrix} \tag{11}$$

$$F = \begin{bmatrix} 0 & f_{12} & f_{13} & \cdots & f_{1k} \\ f_{21} & 0 & f_{23} & \cdots & f_{2k} \\ f_{31} & f_{32} & 0 & \cdots & f_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & f_{k3} & \cdots & 0 \end{bmatrix} \tag{12}$$

Normalize the distance matrix E and the angle matrix F to obtain the matrices G and H:

$$G = \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1k} \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2k} \\ g_{31} & g_{32} & g_{33} & \cdots & g_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{k1} & g_{k2} & g_{k3} & \cdots & g_{kk} \end{bmatrix} \tag{13}$$

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1k} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2k} \\ h_{31} & h_{32} & h_{33} & \cdots & h_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{k1} & h_{k2} & h_{k3} & \cdots & h_{kk} \end{bmatrix} \tag{14}$$

$$g_{ij} = \frac{e_{ij} - \min(e_{ij})}{\max(e_{ij}) - \min(e_{ij})} \quad (i, j = 1, 2, \cdots, k) \tag{15}$$

$$h_{ij} = \frac{f_{ij} - \min(f_{ij})}{\max(f_{ij}) - \min(f_{ij})} \quad (i, j = 1, 2, \cdots, k) \tag{16}$$

The shorter distance and the smaller course difference between the two turning region, the more similar between the two region. Considering the two factors such as the distance between and the course difference between the turning regions, the similarity between the turning regions is measured. The similarity matrix S is shown in the Eq. (17).

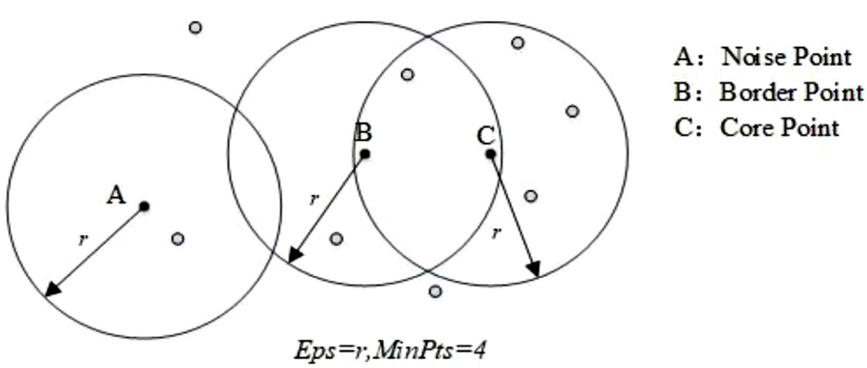

Fig. 5. A schematic diagram of the type of point in DBSCAN.

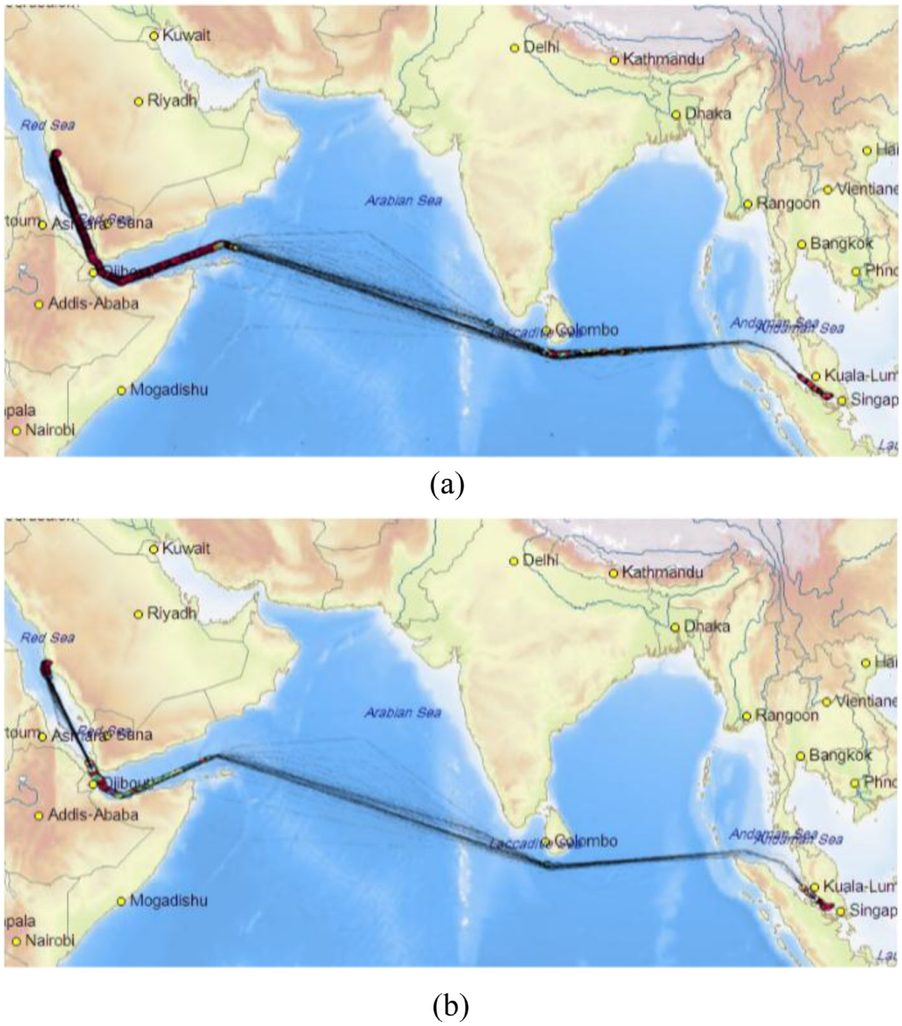

Fig. 6. Clustering results of different parameters.

$$S = \alpha G + (1 - \alpha)H \tag{17}$$

In Eq. (17), α is a weight parameter depends on actual condition. In this paper, we consider G and H as equally important, so α is 0.5.

As shown in Fig. 9, the algorithm starts from the location of the departure port to look for the best turning region in turn. When the last turning region is matched, the whole process ends. Since the location of the departure port and the destination port are known, start matching from the departure port, and the destination port does not participate in the actual calculation. In Fig. 10, the data in the red rectangular box are involved in the calculation actually. The first match is the process of finding the minimum value in the yellow area.

For example, the turning points of Jeddah-Singapore routes are clustered and 12 turning regions are obtained. Set the starting course to 225° and automatically connect the turning region. The automatic connection result is shown in Fig. 11.

5. Ship route automatic generation

5.1. Back propagation neural network model

The BP neural network refers to a multi-layer feed forward neural network based on error back propagation, which is under supervision. It is able to approximate any nonlinear mapping with arbitrary accuracy and it can learn and adapt to position information. It also has a distributed structure for data storage and processing. The neural network has the better features of tolerance, learning and robustness. The experimental results show that neural network is adaptive in complex questions.

After completing a forward propagation, the error of the actual output value can be obtained from the teacher signal, and then the back propagation algorithm is used to obtain the output error of each layer, and the connection weight of each neuron is corrected. When the sum of square of the error of the neural network output layer is less than the specified error, the training is completed, and the weight and deviation of the network are saved. A diagram of neural network model as shown in Fig. 12.

5.2. Activation function

The most commonly used activation function in traditional neural networks is the Sigmoid function, and the sigmoid function is regarded as the core of the neural network. From the perspective of mathematics, the nonlinear Sigmoid function has a larger signal gain to the central region, less signal gain on both sides, and a good effect on the feature space mapping of the signal. The nonlinear combination of the weighted input is used to produce the nonlinear decision boundaries. From the perspective of neuroscience, the central region resembles the excitatory state of neurons. The bilateral regions resemble the inhibitory states of neurons. Therefore, in neural network training, key features can be pushed to the central region, and non-key features can be pushed to both sides [29,30].

In this paper, we use the Relu function as the activation function

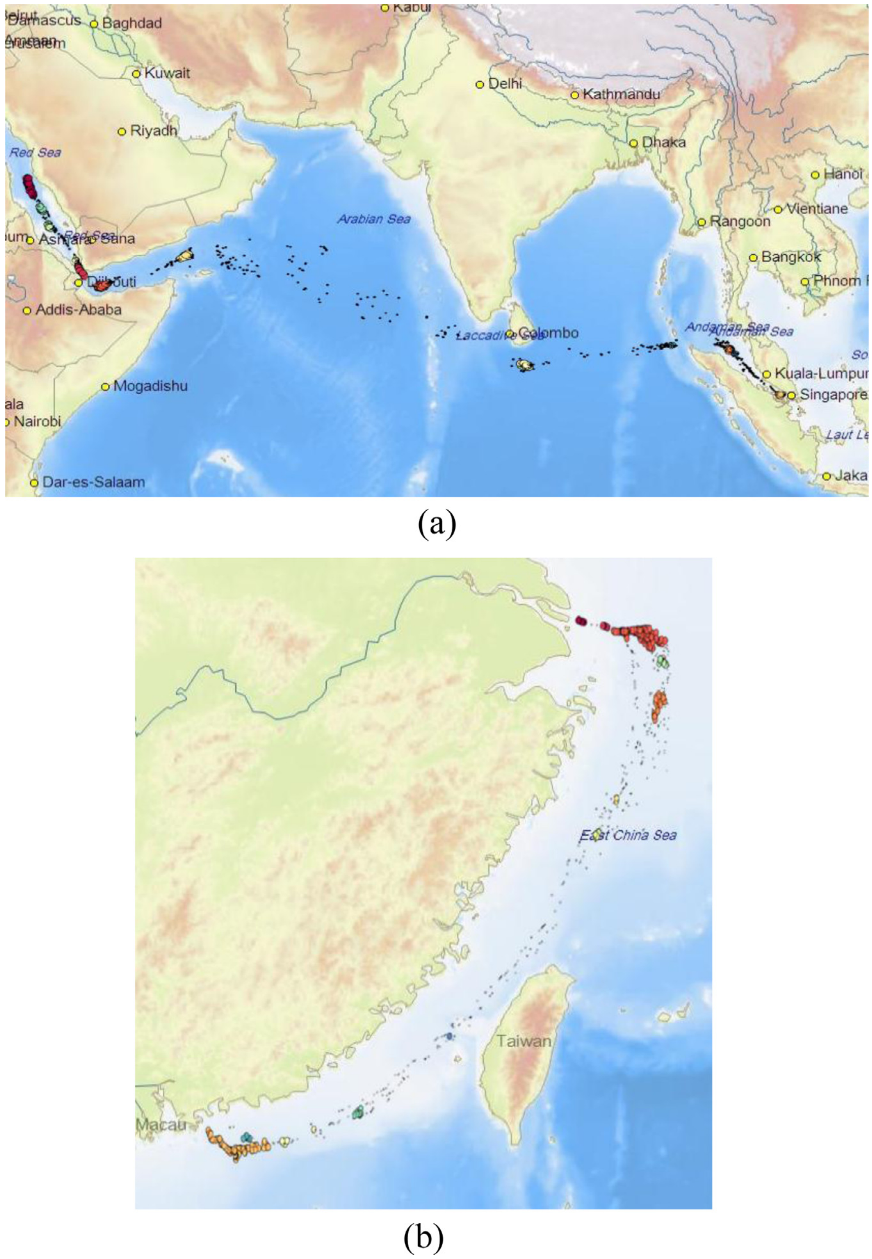

Fig. 7. Using DBSCAN to cluster turning points.

and it is widely used in deep learning in recent years. Compared with the Sigmoid function, the advantages of the Relu function are simple calculation, simple derivative, fast convergence, unilateral suppression and relatively wide excitatory boundary. But the network is very fragile in training, and it is easy to see that many neurons have a 0 value, which causes the training to stop. Therefore, a small learning rate is generally set to avoid this situation.

5.3. Data standardization

Since the units of longitude, latitude, speed, and heading are different, the magnitude difference is also relatively large between the four parameters, and therefore, directly inputting raw data into a neural network for training could degrade the performance and convergence of the network. So, the input data must be standardized before training the neural network and after the prediction result is obtained, the data is deformalized to be the prediction result.

The z-score standardization method is used to standardize data based on the mean and standard deviation of the raw data. The processed data conforms to the standard normal distribution with a mean of 0 and a standard deviation of 1. As shown in Eq. (18):

$$x^* = \frac{x - \mu}{\sigma} \tag{18}$$

In Eq. (18):x^* is standardized data, μ is the mean value of sample data, σ is the standard deviation of sample data.

5.4. ANN training

After data filtering, turning point recognition, turning region clustering and automatic connection, the ship route generation model was created by the three-layer neural network structure with a sufficient number of neurons in the hidden layer. Turning points data like latitude and longitude, speed, course, length, width and draft of the ship are stored in database. All stored data can be used to decide model structure and train ANN. The trained ANN has been used to predict next

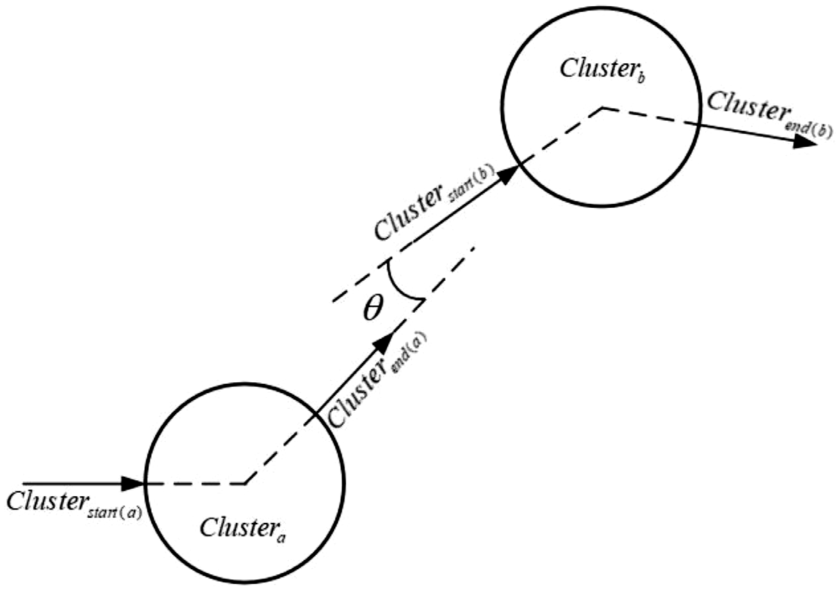

Fig. 8. Searching for the optimal matched turning region.

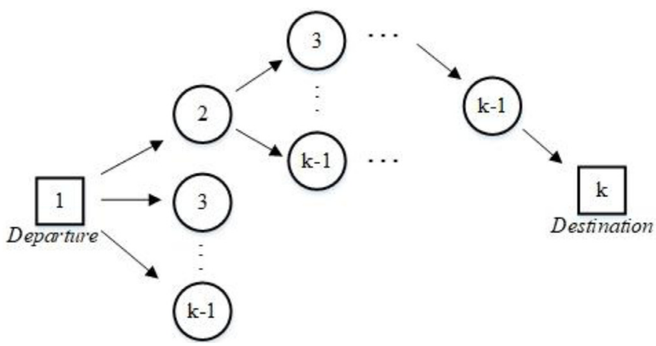

Fig. 9. Match the best turning region.

	1	2	3	···	k-1	k
1	S_{11}	S_{12}	S_{13}	···	S_{1k-1}	S_{1k}
2	S_{21}	S_{22}	S_{23}	···	S_{2k-1}	S_{2k}
3	S_{31}	S_{32}	S_{33}	···	S_{3k-1}	S_{3k}
⋮	⋮	⋮	⋮	⋱	⋮	⋮
k-1	S_{k-11}	S_{k-12}	S_{k-13}	···	S_{k-1k-1}	S_{k-1k}
k	S_{k1}	S_{k2}	S_{k3}	···	S_{kk-1}	S_{kk}

Fig. 10. Turning region similarity matrix.

position of turning points. When training a neural network, the training data and labels are determined by different ship dimensions. The input variables and labels of the neural network are shown in Table 3 and 4.

The number of input neurons is determined by seven factors (latitude, longitude, speed, course, length, width and draft), and the number of labels was set by the five variables (latitude, longitude, length, width and draft).

In artificial neural network, to determine the number of hidden layer nodes is a very important and complex problem and at present, there is no standard method to determine the number of nodes in the hidden layer. If the number of hidden neurons is too small, the performance of the artificial neural network may be poor or unable to be trained, and sufficient number of connection weights may not be generated to satisfy the artificial neural network's learning of the sample. If too many hidden neurons are selected, although the systematic error of the artificial neural network can be made smaller, at the same time, the artificial neural network training time is increased, and it is easy to fall into local minimum and fail to achieve the best result.

For how to determine the number of nodes in the hidden layer, this paper refers to the empirical equation as follows:

$$Z = \sqrt{x + y} + \mathrm{r} \tag{19}$$

In Eq. (19), Z is the number of nodes in the hidden layer, x is the number of nodes in the input layer, y is the number of nodes in the output layer, r is a constant in [1,10].

In the neural network architecture, the numbers of hidden neurons and hidden layers have been found by trial. The back propagation neural network model designed in this paper has 4 input nodes and 2 output nodes. After several experiments, it can be found that the performance of the artificial neural network is the best when the number of hidden layer nodes are 8.

After clustering by the DBSCAN algorithm, multiple turning areas are obtained. The method proposed in this paper is not simply predicting a turning point, but generating a route. When the artificial neural network algorithm is adopted to transform the turning point prediction problem into supervised learning, the turning point data of the next turning region need to be used as the previous neural network model's label. As shown in Fig. 13, the ANN is used as many times as it needed.

The turning regions are empirical by clustering a large amount of historical AIS data. In other words, the probability of the ship turning in turning region cluster by DBSCAN is very high. However, the position

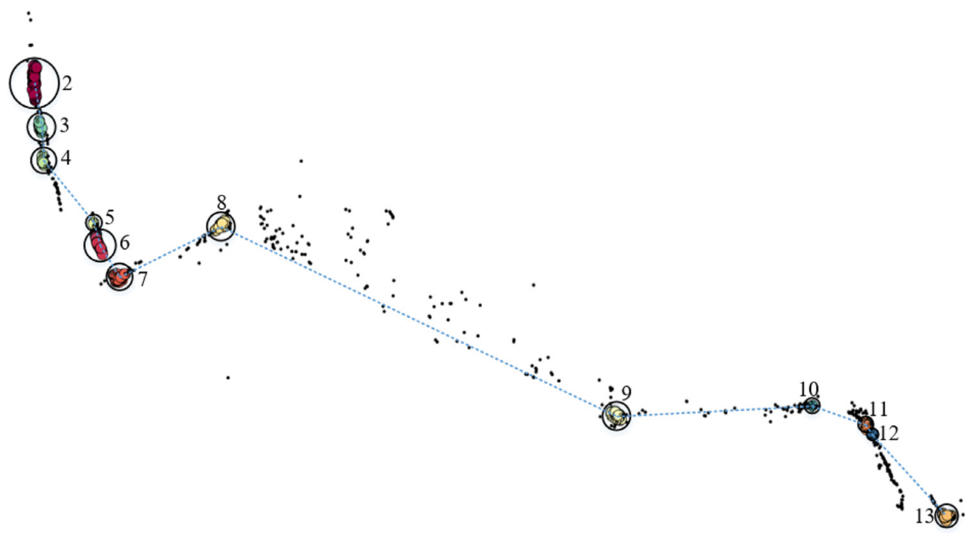

Fig. 11. Automatic linkage of turning clusters.

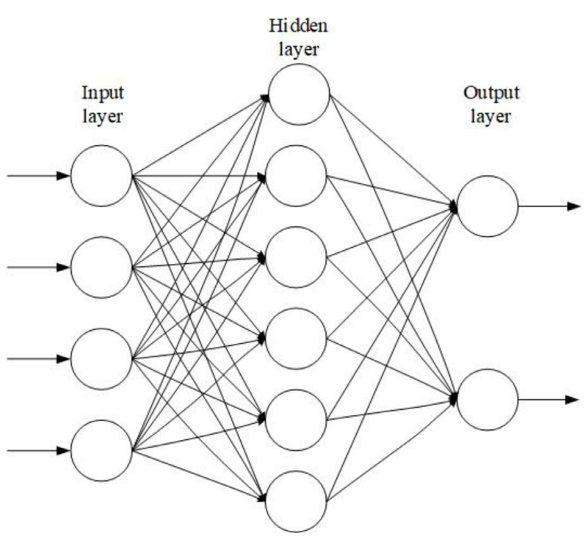

Fig. 12. Artificial Neural network model.

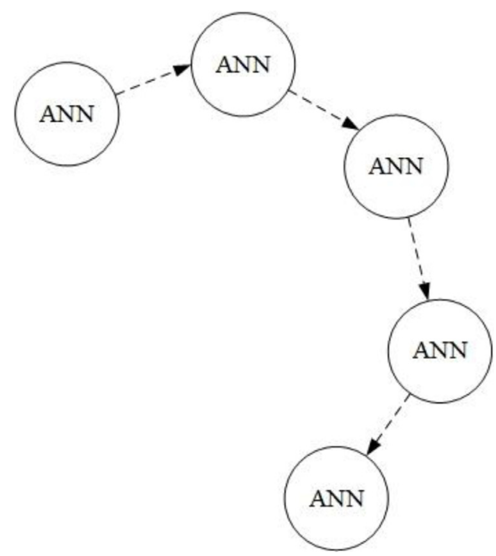

Fig. 13. Multiple neural network models.

Table 3
Input variables of neural network model.

Input	Variable name
Lat0	Latitude of ship in turning point
Lon0	Longitude of ship in turning point
Speed0	Speed of ship in turning point
Course0	Course of ship in turning point
Length0	Length of ship in turning point
Width0	Width of ship in turning point
Draft0	Draft of ship in turning point

Table 4
Labels of neural network model.

Output	Variable name
Lat1	Latitude of ship in next turning point
Lon1	Longitude of ship in next turning point
Length1	Length of ship in next turning point
Width1	Width of ship in next turning point
Draft1	Draft of ship in next turning point

of ship turning in turning region is random, which is also affected by many factors, such as weather, ship collision avoidance, etc. Therefore, a random matching method is adopted to determining the label of each artificial neural network.

5.5. Results and comparative analysis

Such proposed model can be used for any specific vessel to generate her route and the performance has been found to be successful in the turning regions. Traditionally, all data are used to extract one of the most popular route, however, ship of different dimensions often plan different routes. Therefore, four ships of different dimension are selected to validate the model proposed in this paper. Parameters of four ships are shown in the Table 5 and 6.

Table 5
Parameters of two ships in Shanghai-Shenzhen route.

	MMSI	Call sign	Length	Width	Draft
Ship1	209539000	5BFN4	139	23	7.3
Ship2	413171000	BPKD	335	43	12.3

Table 6
Parameters of two ships in Jeddah-Singapore route.

	MMSI	Call sign	Length	Width	Draft
Ship3	636016046	D5ED4	169	27	9.9
Ship4	477314700	VROL7	368	51	15.7

As shown in Fig. 14, the chart depicts validation of the method proposed in this paper for two routes around the world. The (a), (b) represent the Jeddah-Singapore, Shanghai-Shenzhen routes respectively. Blue line depicts the real route from departure to destination while red line depicts the automatic ship route design result. In Fig. 14(a), the red line depicts the automatic route with ship1, and the yellow line depicts the automatic route with ship2. In Fig. 14(b), the red line depicts the automatic route with ship3, and the yellow line depicts the automatic route with ship4. From the experimental results, ships of different dimensions have different modes of navigation.

In previous studies, after recognition of the turning regions, the center of the area are usually connected as recommended routes. A reasonable result should be similar to AIS data. In order to verify the rationality of the method in this paper, DTW algorithm is used to measure the similarity between generated trajectories and AIS data. The average DTW distance between the route generated by connecting central points and the real trajectory is calculated. Meanwhile, the average DTW distance between route generated by connecting central points and real trajectories is also calculated. As shown in Table 7. The results show that the route obtained by the neural network is closer to the real situation. Moreover, the disadvantage of the traditional method is that the route cannot be determined according to different ship dimension.

The above experiments show that it is necessary to recommend routes according to ship dimension. Different ships have different routes. Recommending routes according to the characteristics of different ships will improve the safety of navigation. Furthermore, recommended routes will be updated when the database changes

6. Conclusions

This paper proposes a method for automatic ship route design to generate the vessel route based on massive AIS data that occur in a predictable pattern (between two ports). The main purpose of this study is to recognize the key regions by applying DBSCAN algorithm and then connect these regions automatically by cluster similarity measuring.

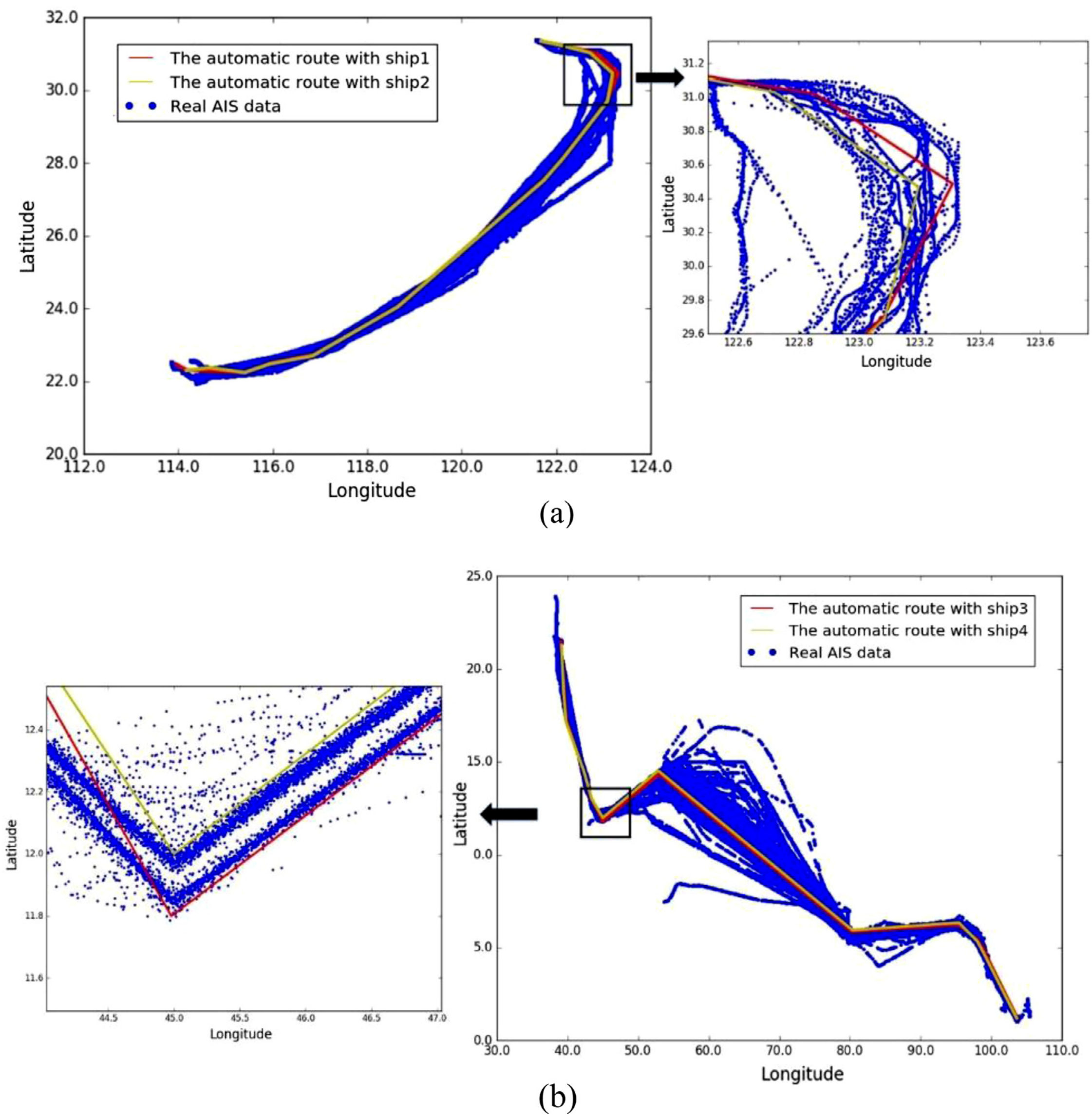

Fig. 14. Real AIS data and automatic route of different vessel.

Table 7
Comparison of two methods.

Route	Average DTW distance between automatic generated route and real trajectories	Average DTW distance between route generated by connecting central points and real trajectories
The automatic route with ship1	97644.32 km	115426.56 km
The automatic route with ship2	9856.15 km	115426.56 km
The automatic route with ship3	12374.76 km	142045.62 km
The automatic route with ship4	11879.52 km	142045.62 km

Then artificial neural network has been used to learn the relationship of turning regions according to different ship dimensions and generate a reasonable route. The method shows great potential in learning and understanding vessel behavior patterns and is capable of generating vessel route successfully, which has been evaluated on Jeddah-Singapore, Shanghai-Shenzhen routes respectively. The route obtained by the neural network is closer to the real situation, the different routes according to different ships will be design automatically.

The experimental results validate the effectiveness of the proposed method. However, because the route generation is based on the ship's historical trajectory, and the approach of approximation and fitting is adopted, thus the potential risk shuold not be ignored. Before sailing at sea, the captain and the second officer should fulfill their obligations and do a good job in safety inspection. On the one hand, under different loading conditions, the same kind of ships need different water depth for their navigation. If the support of electronic chart and the constraint of water depth can be added into the automatic generation process of the route, the better result will be obtained. On the other hand, the navigation environment, including the channel and navigation rules, may change dynamically, so the route obtained by using historical data may be possible out of date. Therefore, it is necessary to use the latest data in the automatic route generation. It is necessary to get the latest navigation environment information in time and add it to the route generation algorithm as a semantic constraint.

CRediT authorship contribution statement

Yuanqiao Wen: Conceptualization, Methodology, Supervision. **Zhongyi Sui:** Data curation, Visualization, Writing - original draft. **Chunhui Zhou:** Conceptualization, Supervision, Writing - review & editing. **Changshi Xiao:** Supervision. **Qianqian Chen:** Software, Validation. **Dong Han:** Validation. **Yimeng Zhang:** Investigation.

Conflict of Competing Interest

None.

Acknowledgments

The authors would like to thank the anonymous reviewers and editors for their constructive comments, which is very helpful to improve the paper. This work was supported by Double First-rate Project of WUT, the National Key R&D Program of China through Grant No. 2018YFC1407405 and the National Key R&D Program of China through Grant No. 2018YFC0213900 and the National Science Foundation of China (NSFC) through Grant No. 51679180.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.apor.2019.102049.

References

[1] Y Wen, Y Huang, C Zhou, J Yang, C Xiao, X Wu, Modelling of marine traffic flow complexity, Ocean Eng 104 (2015) 500–510.
[2] H Shen, H Hashimoto, A Matsuda, et al., Automatic collision avoidance of multiple ships based on deep Q-learning, Appl. Ocean Res. 86 (2019) 268–288.
[3] Song, Liu, Bucknall. Smooth A* algorithm for practical unmanned surface vehicle path planning, 2019, 83, 9-20.
[4] P.A.M. Silveira, A.P. Teixeira, C. Guedes Soares, Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal, J. Navigat. 66 (6) (2013) 879–898.
[5] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, G.B. Huang, Exploiting AIS data for intelligent maritime navigation: a comprehensive survey from data to methodology, IEEE Trans. Intell. Transp. Syst. 19 (5) (2017) 1559–1582.
[6] S.A. Breithaupt, A. Copping, J. Tagestad, J. Whiting, Maritime route delineation using AIS data from the Atlantic coast of the US, J. Navigat. 70 (2) (2017) 379–394.
[7] B. Liu, E Souza, S Matwin, M. Sydow, Knowledge-based clustering of ship trajectories using density-based approach, Proc. IEEE International Conference on Big Data, 2015, pp. 603–608.
[8] F Sun, Y Deng, F Deng, Q Zhu, H Chu, Unsupervised maritime traffic pattern extraction from spatio-temporal data, Proc. International Conference on Natural Computation, 2015, pp. 1218–1223.
[9] H Li, J Liu, W Liu R, N Xiong, K Wu, T Kim, A Dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis, Sensors 17 (8) (2017) 1792.
[10] A Dobrkovic, E Iacob M, V Hillegersberg J, Using machine learning for unsupervised maritime waypoint discovery from streaming AIS data, Proc. The 15th International Conference on Knowledge Technologies and Data-driven Business, 2015, pp. 1–8.
[11] P R Lei, T H Tsai, W C Peng, Discovering maritime traffic route from AIS network, Proc. Network Operations and Management Symposium, 2016, pp. 1–6.
[12] G Pallotta, M Vespe, K Bryan, Vessel pattern knowledge discovery from AIS data: a framework for anomaly detection and route prediction, Entropy 15 (6) (2013) 2218–2245.
[13] Z Xiao, L Ponnambalam, X Fu, W Zhang, Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors, IEEE Trans. Intell. Transp. Syst. (99) (2017) 1–13.
[14] S Zhang, G Shi, Z Liu, Z Zhao, Z Wu, Data-driven based automatic maritime routing from massive AIS trajectories in the face of disparity, Ocean Eng. 155 (2018) 240–250.
[15] A Dobrkovic, M E Iacob, J V Hillegersberg, Maritime pattern extraction from AIS data using a genetic algorithm, Proc. IEEE International Conference on Data Science and Advanced Analytics, 2016, pp. 642–651.
[16] Chen, Shigeaki, Kenji, Study on a numerical navigation system in the East China Sea, Appl. Ocean Res. 52 (2015) 257–266.
[17] J Rhodes B, A Bomberger N, M Zandipour, Probabilistic associative learning of vessel motion patterns at multiple spatial scales for maritime situation awareness, Proc. International Conference on Information Fusion, 2007, pp. 1–8.
[18] U Simsir, S Ertugrul, Prediction of manually controlled vessels' position and course navigating in narrow waterways using Artificial Neural Networks, Appl. Soft. Comput. 9 (4) (2009) 1217–1224.
[19] D Zissis, K Xidias E, D Lekkas, A cloud based architecture capable of perceiving and predicting multiple vessel behaviour, Appl. Soft. Comput. 35 (2015) 652–661.
[20] S Gan, S Liang, K Li, J Deng, T Cheng, Ship trajectory prediction for intelligent traffic management using clustering and ANN, Proc. International Conference on Control, 2016, pp. 1–6.
[21] L.Z. Sang, A. Wall, Z. Mao, X.P. Yan, J. Wang, A novel method for restoring the trajectory of the inland waterway ship by using AIS data, Ocean Eng. 110 (2015) 183–194.
[22] M Belkin, P Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, Advances in NISP 14, MIT Press, Cambridge, MA, 2001.
[23] S J Sheather, M C Jones, A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation, J. Roy. Stat. Soc. 53 (3) (1991) 683–690.
[24] E Keogh, C A Ratanamahatana, Exact indexing of dynamic time warping, Knowl. Inf. Syst. 7 (3) (2005) 358–386.
[25] J Taylor, X Zhou, N M Rouphail, R J Porter, Method for investigating intradriver heterogeneity using vehicle trajectory data: a Dynamic Time Warping approach, Transport. Res. B 73 (2015) 59-80. 19.
[26] A Vakanski, I Mantegh, A Irish, F Sharifi, Trajectory learning for robot programming by demonstration using hidden markov model and dynamic time warping, IEEE Trans. Syst. Man. Cybern. Part B Cybern. 42 (4) (2012) 1039–1052.
[27] M Ester, P Kriegel H, X Xu, A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise, Proc. International Conference on Knowledge Discovery and Data Mining, 1993, pp. 226–231.
[28] LN Xia, JW Jing, SA-DBSCAN: A self-adaptive density-based clustering algorithm, J. Graduate School Chin. Acad. Sci. 26 (4) (2009) 530–538.
[29] I Kaastra, M Boyd, Designing a neural network for forecasting financial and economic time series, Neurocomputing 10 (3) (1996) 215–236.
[30] G Zhang, B E Patuwo, M Y Hu, Forecasting with artificial neural networks: The state of the art, Int. J. Forecast. 14 (1) (1993) 35–62.