Research on Ship Track and Navigation Behavior Characteristics Based on Deep Learning

Yue Yang, Tianyi Liu and Xiaolong Wu
Dalian Naval Academy, Liaoning, China

Keywords: Neural Network, Deep Learning, Ship Track Characteristics, AIS System.

Abstract: Along with the expansion of China's Marine interests and the improvement of comprehensive national strength, Marine navigation safety and ship abnormal navigation behaviour problems are increasingly acute. In this paper, a deep learning CALS algorithm based on neural network is proposed to analyse more than AIS ship sailing track data in the Chinese sea area in 2021, and build ship sailing track prediction and early warning models. By exploring the navigation characteristics of various types of ships in the Chinese sea area and the characteristics of navigation behaviour at different time and space scales, this paper is to solve the problem of abnormal trajectory detection and early warning of ships, and provide information support for safe navigation and military mission decision-making.

1 INTRODUCTION

With the advent of the era of intelligent big data, the statistical analysis and mining technology of big data play an important role. With the continuous advancement of Chinese-style modernization in the process of building a maritime power and the deepening application of Automatic Identification System (AIS) in maritime safety and communication between ships and ports, ships and ships, etc., The historical accumulation of ship track time series data in the China Sea area has exploded, which provides strong data support for mining and analyzing the distribution characteristics of various ship tracks and the characteristics of navigation behavior at different time and space scales.

With the continuous progress of China's comprehensive national strength and the continuous expansion of overseas interests, the mission of the People's Navy has gradually changed from offshore defense to far sea defense. In order to better explore the navigation characteristics of different types of ships and navigation behavior characteristics at different time and space scales in the sea of China, and make use of the above characteristics to support military mission decision-making and ensure navigation safety, this paper starts from the analysis of ship navigation trajectory and ship behavior. It uses AIS data of all ships in the sea of China in 2021. Based on LSTM long and short-term memory network machine learning algorithm, the innovative neural network structure uses the real-time sailing trajectory data of nearby ships to analyze the abnormal sailing position, tracking anomaly, heading anomaly, speed anomaly and other abnormal situations. The new algorithm desire to screen out the ships that exceed the historical path range and may have suspicious sailing behaviors. So, it can provide early warning for military mission decision-making and safe navigation.

2 DATA PROCESSING

In this paper, about 3×10^9 AIS data of 13 types of vessels, such as fishing vessels and pilot vessels, under 11 conditions such as sailing and losing control in the Chinese sea area in 2021 are collected. After decoding the original AIS message information, two tables of Pos and ShipInfo are obtained, corresponding to ship dynamic information and static information respectively. The generation of ship navigation trajectory adopts the method of connecting ship dynamic information in AIS message information to a single ship according to the time series, and records the real-time transmission of navigation-related information.

Yang, Y., Liu, T. and Wu, X.
Research on Ship Track and Navigation Behavior Characteristics Based on Deep Learning.
DOI: 10.5220/0012284200003807

In *Proceedings of the 2nd International Seminar on Artificial Intelligence, Networking and Information Technology (ANIT 2023)*, pages 370-374
ISBN: 978-989-758-677-4

2.1 Data Cleaning and Preprocessing

According to the data format requirements of AIS, data that is not within a reasonable range in ship dynamic information is excluded. The constraints are shown in Table 1.

Table 1. AIS message data format table.

message	data within
MMSI	[200000000,900000000]
Lat	[-181000000,181000000]
Lon	[-91000000,91000000]
Sog	[0,52576]
Cog	[0,35990]
Hdg	[0,35990]
Rot	[-1200,1200]

In this paper, the single sailing track of a ship is the basic research unit. So it is necessary to divide the original track data. Based on information such as time interval of adjacent track points, maximum ship speed, longitude and latitude of track points, flight segments are divided automatically. The specific algorithm steps are as follows:

Step1. The AIS message information after excluding the cross-border data is sorted according to the ship MMSI number and the priority of the data update time.

Step2. Traverse the track points according to the order of the ship MMSI number, until the track point X_i is found, and the time interval between it and the adjacent track point X_{i-1} is greater than the set threshold T. And its sailing speed is 0, or the sailing state is anchored or berthed and the sailing speed is close to 0, the track point X_i is marked as the track division point.

Step3. Repeat Step2 until the end of the current traversed ship MMSI number track sequence. The track data corresponding to the current ship MMSI number is cut according to the marked track division points, and the start and end point of the track are determined according to the time sequence. The segment data is identified and presented.

Step4. Switch to the next ship MMSI number, and repeat Step2 and Step3 until all ship MMSI numbers are traversed. And the ship track data set is finally generated.

The data sample after cleaning is shown in Table 2.

Table 2. Sample data after cleaning.

MMSI	Lat	Lon	Sog	Cog	Hdg
200018895	22.00362	113.136257	1955	25640	25400
200018895	22.002265	113.130978	2418	33680	33600
200018895	22.007665	113.12822	2624	34370	34300
...	...	...	...	...	...
200018895	22.016293	113.108553	3550	32480	31700
200018895	22.022635	113.105698	3704	34250	34200
200018895	22.029445	113.103275	3807	34410	34500

2.2 Data Normalization

In order to improve model training efficiency and prediction accuracy, eliminate prediction errors caused by dimensional disunity, maximum-minimum normalization is adopted to normalize track data. The value of processed data is between [0,1]. The normalization formula for mapping the actual value ε_i of the attribute ε to ε_i' in the interval [0,1] is as follows.

$$\varepsilon_i' = \frac{\varepsilon_i - \min_\varepsilon}{\max_\varepsilon - \min_\varepsilon} \tag{1}$$

In (1) $\min_\varepsilon$ represents the minimum value of the attribute ε and $\max_\varepsilon$ represents the maximum value of the attribute ε.

Normalization processing can effectively eliminate the problem that the feature weights caused by different dimensions are not in line with the reality. At the same time, it can amplify the feature differences, reduce the data scale and the training amount of the model. In the follow training part, in order to better visualize the results or compare the errors, it is often necessary to carry out anti-normalization operations, and the formula is as follows.

$$\varepsilon_i = \varepsilon_i' \left(\max_\varepsilon - \min_\varepsilon\right) + \min_\varepsilon \tag{2}$$

3 SELECTION OF FEATURE PARAMETER

In AIS message information, there are seven track-related parameters. According to the requirements of the algorithm, five parameters such as longitude (Lon, degree), latitude (Lat, degree), speed to ground (Sog, mm/sed), course to the ground (Cog, 1/100 degree) and ship heading (Hdg, 1/100 degree) were selected. And the consistency of the five parameters was tested through Kendall's consistency test. The analysis process was shown in Figure 1.

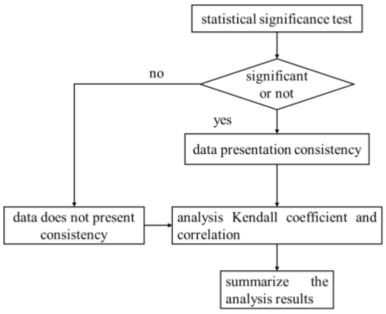

Figure 1. Kendall’s consistency checking process.

The analysis results are shown in Table 3.

Table1. Kendall’s analysis results.

name	rank mean value	Kendall’s coefficient	χ^2	P value
Lat	1.038	0.875	34978.682	0.000
Lon	2.047			
Sog	3.154			
Cog	4.628			
Hdg	4.134			
Significance level is 1%.				

Table 3 shows the results of Kendall's model test, including rank mean value, Kendall coordination coefficient W, χ^2 value and significance P value. The results of Kendall coefficient consistency test show that the significance P value of the overall data is 0.000, which presents significance at the level 1% and rejects the null hypothesis, so the data presents consistency. Meanwhile, the Kendall’s coordination coefficient value of the model is 0.875, which is within the interval of [0.8,1.0]. Therefore, the degree of correlation is almost identical. Based on the test results, the rationality of the selection of track related feature parameters is verified.

4 CONSTRUCTION OF CALS ALGORITHM MODEL

This paper starts from long short-term memory LSTM network machine learning algorithm, the basic idea is as follows. CNN-Attention-LSTM neural network structure was used to train the ship trajectory prediction model. Real-time AIS data was used to predict the ship trajectory. Sliding window was used to model the distribution of the difference sequence. And the difference between the predicted value and the actual value was converted into a unified anomaly score. According to the anomaly score, we can determine whether the trajectory is abnormal and determine the type of abnormal trajectory. The process is shown in Figure 2.

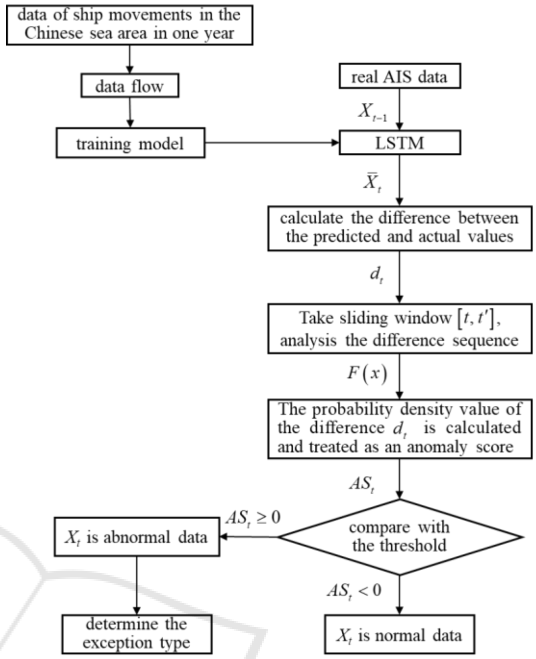

Figure 2. Flow chart of CALS algorithm.

In this paper, the CNN model is combined with LSTM to extend the application scenario of the algorithm, so that it can support long input sequences, which can be read by the CNN model as block or subsequence information. The working mode of CNN-LSTM is that the sub-sequence information is passed into the CNN model as input information, and then LSTM summarizes and processes the sub-sequence information before output. The LSTM based on one-dimensional convolutional neural network enhances the processing power of LSTM convective data and improves the learning efficiency. At the same time, considering the AIS data set has significant time characteristics, time attention mechanism is introduced on the basis of CNN-LSTM. So that the neural network can make more reasonable use of information sources and better explore data characteristics when making predictions. The structural relationship of the three is shown in Figure 3.

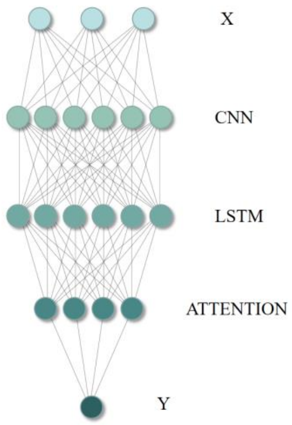

Figure 3. CNN-Attention-LSTM diagram.

5 TEST RESULTS AND ANALYSIS

5.1 Reading and Processing of Real-Time AIS Data

This method is implemented based on python language. Taking a total of 1,048,576 AIS message information of ships in the sea area of China in 2021 as an example, after data cleaning, model construction, anomaly detection, type identification and other steps, the method simulates real-time AIS data push through Socket technology to realize real-time detection of ships with abnormal track.

In order to reduce repeated reading of the model, reduce performance waste, and improve prediction speed, this paper uses Socket technology to create threads to monitor real-time data, form resident services, and use sockets to pass parameters. The specific algorithm steps are as follows.

Step1: Listen for client connection requests, and use sockets to pass data to threads after successfully establishing a connection.

Step2: Start the thread, process the data and return the result to the client.

Step3: Close the connection, the thread suspends, and continue listening for client connection requests.

The specific implementation effect is shown in Figure 4.

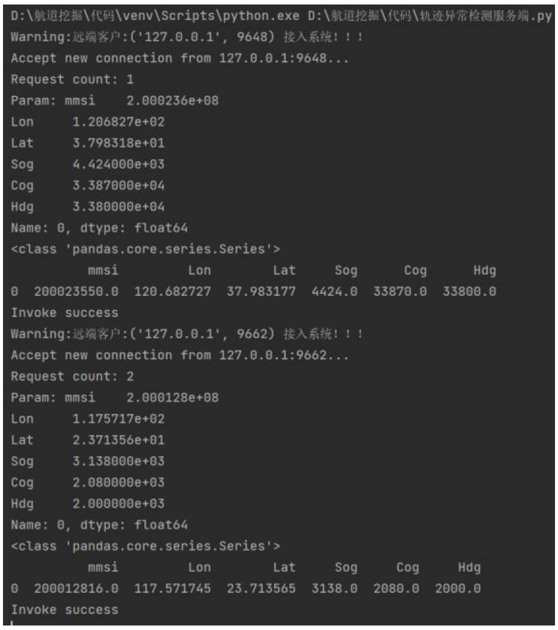

Figure 4. Socket monitors real-time data effects.

5.2 Prediction Model Evaluation

Two commonly used standards of absolute mean error (MAE) and root mean square error (RMSE) are selected for the evaluation indicators of the prediction model. They can measure the deviation between the observed value and the true value and reflect the actual situation of the predicted value error. The smaller is the value, the higher is the accuracy of the model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2} \quad (3)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left|\tilde{y}_i - y_i\right| \quad (4)$$

The error values of the three prediction models are shown in Table 4.

Table 4. Error values of the three prediction models.

Model	CNN-Attention-LSTM	CNN-LSTM	LSTM
RMSE	0.021401891	0.027483279	0.033624116
MSE	0.017074395	0.022646103	0.031701226

It is not difficult to see from the Table 4 that CNN-Attention-LSTM model has a great reduction in absolute mean error and root mean square error compared with CNN-LSTM model and LSTM model. It can be seen that the CNN-Attention-LSTM model used in this paper has high accuracy.

6 CONCLUSION

From the perspective of Marine security, the research on ship track and navigation behavior characteristics in China's sea area conducted in this paper has performed well in the actual test. It reached the expected goal, and can be applied in the field of ship trajectory anomaly warning, providing technical support for commanders and combatants to judge and make decisions on the enemy situation and the safe navigation of Chinese ships.

ACKNOWLEDGMENTS

This work was financially supported by Dalian Naval Academy research and innovation team fund DJYKYKT2021-018， and student research fund DJYKYKT2022-003.

REFERENCES

Wenyao Ma. Consistency Detection of ship abnormal behavior [D]. *Dalian Maritime University*, 2018.

Bochen Yang. Research and application of ship trajectory analysis based on AIS [D]. *University of Electronic Science and Technology of China*, 2018.

Shuang Li. Fishing boats suspicious behavior detection based on AIS [D]. *Dalian Ocean University*, 2022. DOI:10.27821/d.cnki.gdlhy.2022.000214.

Jianguo Kong, Yabin Li. Research on track prediction based on CNN-LSTM-attention model [J]. *Aeronautical Computing Technology*, 2019, 53(01):1-5.

Chengyong Liu, Wenjie Qiao. Research on ship track prediction model based on LSTM and attention mechanism [J]. *Navigation of China*, 2021, 44(04):94-100+106.

Yuling Jiang, Zhennan Xiong, Jihong Tang. Ship trajectory clustering algorithm based on trajectory segment DBSCAN [J]. *Navigation of China*, 2019, 42(03):1-5.

Jingyi Xie. Research on ship track prediction method based on navigation strategy learning [D]. *University of Chinese Academy of Sciences* (national center for space science), 2021. DOI:10.27562/d.cnki.gkyyz.2021.000030.

Yingjie Yan. Transformer state anomaly detection based on sliding window and clustering algorithm [J]. *High Voltage Technology*, 2016, 42(12): 4020-4025. DOI: 10.13336/j.1003-6520.hve.20161128041.