Risk Assessment of Ship Traffic Safety in Port Waters Based on AIS Data

Junting Xiong1, Xin Jin1*, Guangzhi Dong2, Dongliang Gu1, Chengtao Yi1

1Department of Navigation, Dalian Naval Academy, Dalian, Liaoning, 116018, China
2Teaching and Research Guarantee Center, Dalian Naval Academy, Dalian, Liaoning, 116018, China
*email: jerryt@dlmu.edu.cn

Abstract. Maritime safety has always been a matter of great concern to people, and ship traffic safety is one of the most important parts of maritime safety. The port is the starting and ending point of maritime traffic and the connection point of land and sea transportation. Due to the increase in ship traffic, the increase in transport tonnage, and the increase in dangerous goods, the water traffic environment has become more crowded and more dangerous. Under such complex traffic environment conditions, new perspectives should be used to study ship traffic in port waters. Security related issues. Based on the real AIS data of port waters, this paper proposes a new method of ship traffic safety risk assessment by reproducing the trajectory of ships.

1. Introduction

Port waters are a subsystem of the maritime transportation system. How dangerous is the system, how much loss it may cause to society, how much risk is acceptable to people, and how much safety investment is needed to reduce the risk of the system Reduced to safety indicators, these all rely on safety research. The need to qualitatively and quantitatively describe the hazards of the system has given birth to a variety of security risk assessment methods[1]. For example: management methods, multivariate statistical analysis methods, operations research methods, industry organization methods, fuzzy theory, gray theory, neural network methods, computer-aided evaluation and many other evaluation methods. These methods play a corresponding role in their respective fields for specific problems, but there are always problems of versatility and poor compatibility. Therefore, the maritime field lacks scientifically applicable safety evaluation tools for a long time[2].

Based on ship AIS data, this paper uses ArcGIS to draw ship trajectories, and evaluates the traffic safety risks of ships by comparing the number of ship trajectories that constitute the encounter situation with the total number of ship trajectories in the same time under reasonable evaluation indicators. Starting with the data itself, an attempt is made to provide a suitable method for assessing the safety of vessel traffic in port waters.

2. AIS data decoding

The paper uses AIS data of ships in the waters near Luoyuan Bay of Sansha Bay from 0:00:00:00 on April 1st, 2017 to 23:59:59 on April 7th, 2017 as the data source for research and analysis.

Ship AIS data is transmitted in the form of compression encoding, and the information contained in it can be obtained by decoding. There are two transmission modes for AIS information: "clear code"
and "crypto code" [3]. The "clear code" can directly read the information content. Although this encoding method is simple and clear, its information is cumbersome and occupies a large number of characters, which affects the transmission efficiency; The "crypto code" needs to be decoded in accordance with the AIS technical standard ITU-RM.1371 [4] protocol rules to obtain the hidden information Content. Although the parsing process of this encoding method is complicated, it can contain a large amount of information, and the transmission efficiency is very high. Generally, when the amount of AIS information is large, the "crypto code" transmission method is often used.

The paper uses Java programming to decode the AIS secret code, and gets the initial data as shown in Figure 1.

Figure 1. Partially processed AIS data

The paper imports the decoded AIS data into Excel, sorts and adjusts the format to get the result as shown in Figure 2.

Figure 2. AIS data import Excel partial results

3. AIS data preprocessing and statistics

After the decoding process, the basic data of the ship is obtained. The count shows that the static and dynamic information of 1566 ships have been obtained this time. Due to the huge amount of data, it is necessary to preprocess and count these data.

3.1. AIS data preprocessing

In order to make the ship's trajectory in the traffic flow reappearance map identifiable and facilitate subsequent operations, this experiment intercepted the data on April 1st, a total of 221 ships, and the data update interval for the same ship is 2-12 seconds (excluding low speed ships and anchored ships). First use the Excel table tool to eliminate the AIS data (that is, delete the data rows with "#" and blanks, and the data rows with the same information).

After smoothing the sailing trajectory using Java language programming, the ships with the following problems were successfully eliminated: 17 ships with insufficient effective data, 9 ships with loopback problems in the route, 8 ships with the route in ray state, and missing data causing the route to be disconnected 13 Boat. After preprocessing operations, 174 complete navigation trajectory
curves are obtained, and then the ship trajectory is drawn using ArcGIS and superimposed on the map to obtain a rendition of the trajectory of the ship as shown in Figure 3 (the color of the trajectory of the ship generated each time is random, but two adjacent tracks have different colors).

Figure 3. The preprocessed ship trajectory reproduction map

3.2. AIS statistics
This paper calculates the length and type of 1566 ships, superimposes them with the pre-processed ship information, and makes an Excel table. According to the table information, we draw the ship types and ship lengths in the waters near Luoyuan Bay in Sansha Bay on April 1st, 2017. The distribution is as follows:

3.2.1. Ship type distribution.
The statistical results show that the highest proportion of cargo ships is 54% (including 31% of general cargo ships and 23% of bulk cargo ships), followed by container ships and oil tankers each accounting for 18%. And 11%, passenger ships, fishing boats and other types of ships accounted for a relatively small number. Since the maneuverability of different types of ships is very different, such as container ships and fishing boats, more than 90% of the ships are motorized ships from the distribution situation, and the number of fishing ships and ships with limited maneuverability is very small. Therefore, from the perspective of simplifying the collision avoidance situation, all non-motorized ships are excluded, and this article only focuses on motorized ships.

3.2.2. Ship length distribution.
The length of the ship directly reflects the tonnage of the ship to a certain extent, and the maneuverability of the ship is related to the tonnage of the ship. Generally speaking, the maneuverability of large ships will be slightly worse. Due to the influence of wind currents, small ships are generally not easy to maintain direction and speed. Most of the ships are between 50 m-150 m in length, and there are relatively few small ships below 50 m and large ships above 150 m. Therefore, in order to simplify the collision avoidance situation, we exclude ships less than 50 m and greater than 150 m in length. The paper assumes that the maneuverability of motor ships is basically the same.

4. Reappearance of ship encounter trajectory
Before recreating the trajectory of the ship that constitutes the encounter situation, the information of 174 ships needs to be imported into the Mysql database, and then the ship trajectory screening is performed. The entire screening process is carried out on the Hadoop platform. The screening index is two ships. The purpose of the distance and relative position is to filter out the ship data that constitute an encounter situation at the same time and generate avoidance maneuvers (two ships
avoiding). For the distance between the two ships and the relative azimuth, use the mid-latitu-
de algorithm to solve:

\[
\begin{align*}
D\phi &= \phi_2 - \phi_1 = s \cdot \cos C \\
Dep &= s \cdot \sin C \\
\varphi_m &= \frac{\varphi_1 + \varphi_2}{2} \\
D\lambda &= \lambda_2 - \lambda_1 = Dep \cdot \sec \varphi_m
\end{align*}
\]

(1)

In formula (1), is the difference in latitude, is the difference in longitude, is the east-west distance,
is the median latitude, and is the desired distance between the two ships and the relative bearing.

Use Hadoop to analyze the position data of 174 ships. According to the topography of Luoyuan
Bay in Sansha Bay, it can be determined that the researched waters are open waters. Therefore, this
article sets the following ship domain models[7] as constraints for trajectory screening, and the range of
model domains is set as shown in Figure 4.

Figure 4. Trajectory screening constraints

The trajectory screening is carried out according to the above-mentioned ship field. When the ship
meets the distance requirement for action within the relative position, it will be recorded, which proves
that there has been an encounter situation. The screening result is shown in Figure 5.

Figure 5. Hadoop screening results

After being screened by Hadoop, 48 complete trajectories (that is, 24 encounter situations) are
obtained. As shown in Figure 6, each trajectory is plotted with the principle of adjacent and different
colors. In this day, each ship Encounter situations exist, and evasive manipulations have taken place in
all encounter situations. Figure 3 and Figure 6 are the final renderings of the ship's trajectory to be
compared.

![Figure 6. Reproduction diagram of ship's trajectory that constitutes the encounter situation](image)

### 5. Ship traffic safety risk assessment

According to the relevant research of relevant experts and scholars in my country, the evaluation indices and evaluation process are both operability and practical, avoiding too many factors to consider causing the main factors to be inconspicuous and affecting the rationality of the evaluation results\(^8\), select eight factors including visibility, gale, channel width, traffic special points, obstruction distribution, traffic intensity, navigation aid conditions, and marine traffic accidents are used as basic evaluation indices\(^9\). And based on historical experience data, five levels are divided into Table 1, namely, level I, level II, level III, level IV, and level V, respectively representing safe, relatively safe, generally safe, unsafe, and very unsafe.

| Classification                  | I    | II   | III  | IV   | V    |
|---------------------------------|------|------|------|------|------|
| Visibility (I\(_1\))           | <15  | 15—20| 21—30| 31—45| >45  |
| Gale (I\(_2\))                 | <30  | 30—45| 46—75| 76—120| >120 |
| Channel width (I\(_3\))        | <0.3 | 0.3—0.4| 0.5—0.6| 0.7—0.8| >0.8 |
| Traffic special points (I\(_4\))| <0.1 | 0.1—0.2| 0.3—0.4| 0.5—0.6| >0.6 |
| Obstruction distribution (I\(_5\))| <0.3 | 0.3—0.5| 0.6—0.8| 0.9—1.1| >1.1 |
| Traffic intensity (I\(_6\))    | <15  | 15—30| 31—60| 61—120| >120 |
| Navigation aid conditions (I\(_7\))| <1   | 1—3  | 4—6  | 7—9  | >9   |
| Marine traffic accidents (I\(_8\))| <10 | 11—20| 21—30| 31—40| >40  |
| Comprehensive risk assessment (I) | <1.0 | 1.0—2.9| 3.0—4.9| 5.0—6.9| >=7  |

To assess the risk of ship traffic safety, not only the comparison of the ship’s trajectory should be
considered, but also the environmental conditions of the port waters on the day of data collection should be taken into account. The following formula is obtained from a comprehensive analysis:

\[
I = \sum_{j=1}^{8} \omega_j I_j + k \frac{L_2}{L_1}
\]  

In formula (2), \(\omega_j\) is the weighting coefficient of various environmental factors on April 1st, \(k\) is the weighting coefficient of ship trajectory comparison, \(L_1\) is the total number of ship trajectories on April 1st, and \(L_2\) is the composition of the encounter situation. The number of trajectories, after substituting the corresponding data, can get a comprehensive risk assessment \(I\) of 1.3, which belongs to level II, which is relatively safe.

6. Conclusions

The risk assessment method used in this paper has done qualitative and quantitative research on the risk assessment of ship traffic safety in port waters. Starting from empirical methods, the selected representative evaluation indicators are used to qualitatively analyze the environmental conditions of the waters; starting from the data itself, using the ship's trajectory is compared, and the ship's traffic conditions are quantitatively analyzed. Finally, the risk assessment results are comprehensively given. Generally speaking, the assessment results are more scientific, reasonable and accurate. Neither the water environment factors nor the ship traffic factors are artificially added, but the research object exists objectively. Therefore, the analysis results or the design plan are more in line with objective reality and more reasonable optimization; it can fully consider the intermediary transition of things Nature, from this intermediary transition process, a floating selection of a certain state can give a series of analysis results and design schemes under different safety levels, which provides favorable conditions for the decision-making and optimization of engineering and technical personnel in the field of navigation.

References

[1] Tang, X.H., Fan, Y.T., Cai, C.Q. (2008) Research on grid analysis and prediction of marine traffic risk. Journal of Applied Basic Science and Engineering, (03): 425-435.
[2] Chen, Z.O. (2015) Tianjin Port Hong Kong main channel navigation traffic organization scheduling modeling research. China Maritime, 5: 46-49.
[3] Geng, L.H., Zheng, Z.Y., Qi, L. (2017) Method for extracting information from ship encounters in AIS data. Chinese Journal of Science and Technology, 12(7): 802-805.
[4] Liu, H., Li, W.F. (2016) Automatic identification of maritime traffic flow area based on AIS data. China Navigation, 4: 88-90.
[5] Xiao, F., Han, L., Gulijk, C.V., et al. (2015) Comparison study on AIS data of ship traffic behavior. Ocean Engineering, 95(3):84-93.
[6] Mohamed, R., Aly, H., Youssef, M. (2017) Accurate real-time map matching for challenging environments. IEEE Trans. on Intelligent Transportation System, 18(4): 847-857.
[7] Hansen, M.G., Jensen, T.K., LEHN-SCHIOLEN, T., et al. (2013) Empirical ship domain based on AIS data. Journal of Navigation, 66: 931-940.
[8] Xu, G.B., Xuan, S.Y., You, Q.H. (2012) Application of fuzzy set pair model based on entropy weight in navigation risk assessment of port waters. Journal of Shanghai Maritime University, (01): 7-11.
[9] Guo, C., Fang, Q.G., Hu, S.P., et al. (2012) Consistency analysis model of risk assessment conclusions for port waters. Journal of Shanghai Maritime University, (01): 16-19.