Probability Analysis of Ship Collision and Grounding in Inland Waterway Based on Big Data Analysis

Yuheng Tang, Yuankang Mao, Mengjie Wu, Te Shi and Cunlong Fan

1School of Energy and Power Engineering, Wuhan University of Technology, Wuhan, Hubei, 430063, China
2School of Computer and Science and Technology, Wuhan University of Technology, Wuhan, Hubei, 430070, China
3Intelligent Transportation System Research Center, Wuhan University of Technology, Wuhan, Hubei, 430063, China
4National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, Hubei, 430063, China

*Corresponding author’s e-mail: yuheng.tang@whut.edu.cn

Abstract. With the rapid development of technology, mining valuable information from massive data has become an important part of the community. However, there are many deficiencies in the probability analysis of collision and grounding of ships in inland waterways, such as the insufficient use of data. The main purpose of this paper is to analyze the probability of collision and grounding of ships in inland waterways.

Initially, big data analysis technology is used to preprocess the massive shipping data, and then the probability of the processed data is calculated by establishing the probability calculation model of collision and grounding. The results show that in the channel intersection area with poor shipping conditions, the increase of ship traffic volume will lead to the increase of ship collision probability; in the case of low visibility and long driving time, the probability of ship grounding will increase.

1. Introduction
With the rapid development of inland navigation, more and more large-scale ships appear on the surface of the inland river. Some river mouths and river corners often suffer from ship congestion. The existence of ship congestion will increase the probability of ship collision and grounding to a large extent. Once the ship collision occurs, it will bring great casualties and property losses, and lead to environmental pollution. Therefore, the probability analysis of ship collision and grounding in inland waterway is of great practical significance.

In the aspect of ship collision and grounding, many experts and scholars have studied it and achieved great results. In [1], aiming at the evaluation and verification of ship collision damage, the author adopts experimental and numerical simulation methods to study the depth model scale and full-scale collision test, and verifies the ability and accuracy of the proposed analysis method. In [2], in order to determine the absorbed energy and the degree of collision damage in ship collision analysis, the author verifies the experimental results from the public domain. In [3], the author proposes a data-driven method, which can initially estimate the collision probability of the pen tip: it is the probability of the moving ship and the possibility of its position collision. In [4], the author makes a systematic
analysis of ship accidents in the past ten years to evaluate the current safety level of most ship types in the world merchant fleet. In [5], the author studies the current ship risk model of ship collision, ship structure collision and landing, explores the availability of the model for the marine autonomous surface ships, and on this basis, develops the ship risk model standards for the autonomous ships. In [6], by investigating the possibility of propagating damage on the damaged ship due to the low cycle and high stress fatigue process caused by wave load, the author extended the structural safety assessment after conventional accidents. In [7], the author constructs a Bayesian network structure to analyze ship collision accidents in the downstream of Yangtze River, and validates the model by using the existing methods. The proposed consequence estimation model considering emergency management is of great help to the study of ship collision probability.

With the further development of information technology, the era of big data has come. More and more industries use big data analysis to analyze and develop the industry reasonably. In [8], the author proposes to apply the big data analysis in agricultural practice to solve various practical problems related to agriculture. In [9], the author considers applying big data analysis to government data analysis to analyze the data set published by the government on the portal, to promote government innovation and improve the transparency and accountability of the government. In [10], the author applies big data analysis to the risk assessment of railway failure, analyzes a type of rail surface defect by using big data analysis method, to assess the risk of rail failure, to ensure the safety of railway transportation. Big data analysis technology has unique advantages in dealing with massive data. For the current situation that inland river shipping is so developed, the shipping data generated every day is huge. It is necessary and feasible to apply big data analysis technology to the probability analysis of ship collision and grounding in inland waterway.

With the continuous development of inland waterway shipping, it is the surge of shipping data. For a large number of shipping data, this paper puts forward the analysis of collision and grounding of inland waterway ships based on big data analysis, in order to find out the causes and probability of collision and grounding of inland waterway ships, so as to provide effective suggestions for ensuring the safety of inland navigation.

2. Methods

2.1 Collision and grounding of ships
Ship grounding refers to the accident of stopping or damaging the ship when the ship stays in the shallow water, while ship collision refers to the direct or indirect collision between the ship and other ships or objects in the process of running. Ship grounding is generally divided into two categories: maneuverable grounding and drift grounding. For the probability analysis of ship collision and grounding, it is necessary to analyze the causes of ship collision and grounding. The causes of ship collision and grounding are mainly divided into two aspects: subjective factors and objective factors. The subjective reasons mainly lie in the quality of the crew, the preparation of the voyage and the operation of the crew. The objective reasons mainly include the conditions of the inland river and the ship itself.

2.2 Big data analysis technology
The most basic task of big data analysis is to use big data analysis technology to mine potentially valuable information from a variety of, daily surge, including a wealth of massive data. On this basis, effectively use the information as an important basis for decision-making. The big data analysis process is mainly divided into five steps:

(1) Define analysis objectives
Big data analysis first defines the analysis objectives, which are mainly to clarify the main problems of data analysis and the analysis objectives to be achieved by data analysis. In short, it is to define the problem to be studied.

(2) Data acquisition
If you want big data analysis to be able to mine valuable potential information for the project, the first thing is to follow the three principles of full, detailed and accurate data collection. Data collection refers to the collection of data as comprehensive as possible, comprehensive data can better find the correlation between data in data analysis; Detail refers to the collection of data dimensions as many as possible, more and more detailed data dimensions, which can make the data analysis model more accurate and reasonable; Quasi means to ensure the accuracy of data as much as possible during data collection. The results of false data after data analysis must be inaccurate, which will have a misleading impact on decision-making.

(3) Data exploration

In order to ensure the accuracy of the final model of big data analysis, the key lies in the exploration of the original data. Data exploration includes outlier analysis, missing value analysis, correlation analysis and periodic analysis. The main content of data exploration is to judge whether a data set can meet the analysis objectives of the expected definition, what are the internal laws, trends and attributes of the data, whether there are the same attributes among the data sets and whether there are data states not considered.

(4) Data preprocessing

On the one hand, it improves the quality of data, on the other hand, it can better use big data analysis technology and analysis tools. Data preprocessing mainly includes data cleaning, data integration, data change and data specification.

(5) Modeling analysis

Modeling analysis mainly includes four steps: model establishment, model training, model verification and model prediction. Model building is not accomplished overnight. It needs to be verified repeatedly according to the actual situation and select the most suitable model for analysis.

2.3 Probability analysis of ship collision and grounding in inland waterway based on big data analysis

With the continuous development of inland navigation, the number of ships on the inland river is increasing rapidly, which may lead to the occurrence of ship collision and grounding events. The traditional analysis method in the analysis of the probability of ship collision and grounding on the inland waterway, for the massive shipping data processing it can’t do the reasonable mining as far as possible, because the traditional analysis method can’t extract the valuable potential information from the massive shipping information. Therefore, this paper introduces the big data analysis technology into the probability analysis of ship collision and grounding in inland waterway. Using the advantages of big data analysis technology to quickly and accurately mine valuable information, we can process a large number of shipping data, and analyse the probability of ship collision and grounding.

IWRAP is a model based on AIS historical data to calculate the probability of ship collision, which can be used to calculate and analyse different kinds of ship collision probability. The probability of collision of ships in the route is determined by the number of potential collision accidents and the probability of cause:

\[ P = 1 - \exp(-N_a \times P_c \times \Delta t) \]  

Where: \( P \) is the probability of ship collision; \( P_c \) is the probability of ship collision; \( \Delta t \) is the unit time to analyse the probability of ship collision; \( N_a \) is the number of accident risks, obtained according to different collision types between ships.

For the calculation of ship grounding probability, the dynamic draught and other factors are mainly considered, so Monte Carlo method is used to establish the calculation model of ship grounding probability. Considering the response of the ship to the wave, the change of draft caused by the heeling, the sinking and the bow of the ship in shallow water and the distribution characteristic function of the channel bottom, a series of random numbers are generated by Monte Carlo method, and the probability \( p (Z < 0) \) of the ship grounding in the given channel limited depth is calculated according to the following formula:

\[ Z = d + HTIDE - (D + R + B_f + H + S) \]
Among them, d is the channel depth; HTIDE is the tide height; D is the draft of the design ship type; R is the vertical response of the ship to the waves; Bf is the bottom of the channel; H is the draft change caused by the heeling; S is the squatting caused by the shallow water.

3. Data
In this paper, the lower reaches of the Yangtze River selected as the research object of the inland waterway. The lower reaches of the Yangtze River are winding, fast and dense, and there are many ferries crossing the river. The industry along the river is developed, and there are various accident places. The Yangtze River channel is mainly divided into three types: large ship channel, small ship channel and no navigation area, and the small ship channel is on the outside, followed by the large channel, and the no navigation area is in the middle, which is symmetrical distribution. For the analysis of ship collision probability, the main choice is the intersection of the Yangtze River and the Beijing Hangzhou canal, where ships meet more and the channel conditions are more complex. On this basis, we consider the probability of ship collision with the increase of ship traffic. For the analysis of the probability of ship grounding, it is mainly based on the situations of ship grounding on the channel of Jiangsu section of the Yangtze River.

4. Results
Result1: Probability analysis of ship collision
According to the navigation situation of ships at the intersection of the Yangtze River and the Beijing Hangzhou canal, the collision probability of ships in the Yangtze River channel (channel 1) and the Beijing Hangzhou canal channel (channel 2) changes with the increase of the number of ships. Figure 1 shows the ship probability of the two channels with the increase of ship traffic.

![Figure 1. Collision probability with increasing ship throughput.](image)

As can be seen from Figure 1, with the increase of ship traffic volume, the collision probability of both channels increases in a straight line, but compared with the Yangtze River channel, the collision probability of the Beijing Hangzhou canal channel increases by a larger margin. At the same time, the probability of collision increases with the increase of the angle of intersection. Especially in the Beijing Hangzhou canal channel, the probability of collision is about 1.7%.

Result2: Probability analysis of ship grounding
According to the channel conditions of Jiangsu section of the Yangtze River, the probability of ship grounding under different visibility and driving time is considered. Table 1 shows the probability distribution of ship grounding in different visibility and driving time.
Table 1. Probability of ship grounding under different visibility and driving time.

| Visibility (Good--Bad) | Driving Time /h |
|------------------------|-----------------|
|                        | 2               |
|                        | 3               |
|                        | 4               |
|                        | 5               |
|                        | 6               |
| 4                      | 0.0004          |
|                        | 0.0013          |
|                        | 0.0020          |
|                        | 0.0025          |
|                        | 0.0029          |
| 5                      | 0.0010          |
|                        | 0.0016          |
|                        | 0.0024          |
|                        | 0.0030          |
|                        | 0.0035          |
| 6                      | 0.0012          |
|                        | 0.0020          |
|                        | 0.0028          |
|                        | 0.0034          |
|                        | 0.0041          |
| 7                      | 0.0015          |
|                        | 0.0024          |
|                        | 0.0032          |
|                        | 0.0040          |
|                        | 0.0047          |
| 8                      | 0.0018          |
|                        | 0.0026          |
|                        | 0.0037          |
|                        | 0.0044          |
|                        | 0.0052          |

It can be seen from table 1 that under the condition of good visibility, the probability of ship grounding is relatively small. With the continuous decline of visibility, the probability of ship grounding is increasing. At the same time, the ship's driving time also leads to the ship's grounding to a certain extent, which shows that the driving time is too long and the driver's fatigue is also one of the reasons for the ship's grounding. In the case of the worst visibility, if the driving time is too long, the ship will be stranded.

5. Conclusions
In the era of information explosion, data is a very important resource. How to mine valuable information from massive information is very important. In this paper, the big data analysis technology is introduced into the study of the probability of collision and grounding of inland waterway ships, and the advantages of big data analysis technology are fully used to analyze the various factors that may affect the probability of collision and grounding of inland waterway ships. The research method adopted in this paper can also be applied to other inland waterways, which is of great practical significance for specifying ship navigation rules, assessing ship navigation risks and reducing the occurrence of ship collision and grounding events.

Acknowledgments
The research presented in this paper was sponsored by Independent Innovation Research Fund Project of Wuhan University of Technology (2019-ND-B1-07).

References
[1] Zhang, S., Villavicencio, R., Zhu, L., & Pedersen, P. T. (2019). Ship collision damage assessment and validation with experiments and numerical simulations. Marine Structures, 63, 239-256.
[2] Zhang, S., & Pedersen, P. T. (2017). A method for ship collision damage and energy absorption analysis and its validation. Ships and Offshore Structures, 12(sup1), S11-S20.
[3] Liu, B., Garbatov, Y., Zhu, L., & Soares, C. G. (2018). Numerical assessment of the structural crashworthiness of corroded ship hulls in stranding. Ocean Engineering, 170, 276-285.
[4] Eliopoulou, E., Papanikolaou, A., & Voulgarellis, M. (2016). Statistical analysis of ship accidents and review of safety level. Safety science, 85, 282-292.
[5] Thieme, C. A., Utne, I. B., & Haugen, S. (2018). Assessing ship risk model applicability to Marine Autonomous Surface Ships. Ocean Engineering, 165, 140-154.
[6] Gledić, I., Parunov, J., Prebeg, P., & Ćorak, M. (2019). Low-cycle fatigue of ship hull damaged in collision. Engineering Failure Analysis, 96, 436-454.
[7] Wu B., Tian H., Yan X., Soares, C. G. (2019), A probabilistic consequence estimation model for collision accidents in the downstream of Yangtze River using Bayesian Networks. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 1748006X-1982570X.
[8] Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture, 143, 23-37.
[9] Hardy, K., & Maurushat, A. (2017). Opening up government data for Big Data analysis and
public benefit. Computer law & security review, 33(1), 30-37.
[10] Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., ... & De Schutter, B. (2017). A big data analysis approach for rail failure risk assessment. Risk analysis, 37(8), 1495-1507.