Design scheme of ship risk prediction model from the perspective of artificial intelligence

Qunsheng Ruan, Fang Luo *, Zhiliang Li, Manying Shi, Zhiling Zhang

School of Information, Mechanical and Electrical Engineering Normal University
name of organization Ningde, China

*Corresponding author e-mail: luofang@ndnu.edu.cn

Abstract. Vessels sailing at sea are full of risks, accidents occur frequently, and it is easy to cause heavy casualties and property losses. To this end, this article focuses on how to build a high-precision and rapid ship risk prediction model, and proposes a design plan for a ship risk prediction model based on big data and artificial intelligence. The plan includes a basic data layer, a technical business layer, and an application layer, with a focus on The technical research scheme of the graph representation model construction of the ship risk big data, the dynamic graph matching algorithm and the ship risk prediction is elaborated in detail, and finally the experimental design method of the scheme is given. The design scheme proposed in this paper can provide reference and inspiration for the research and development of marine risk system.

Keywords: Vessel Risk, Predictive Model, Dynamic Graph.

1. Introduction

The safety of ships at sea has attracted international attention. Ships sailing in the ocean have the characteristics of high mortality and high risk of loss. Although countries have attached great importance to safety production management and adopted a series of safety production prevention and control measures in recent years, the number of accidents, casualties, etc. have not been significantly improved, and the number of marine ship accidents has also increased year by year [1]. In addition, ships often enter sensitive areas such as fishing restricted areas, fishing grounds, and naval military zones, which put pressure on the management of relevant departments and threaten the lives and properties of the vast number of seafarers.

Ship risk prediction and early warning is a relatively complex research topic. It is an algorithm mechanism that detects ship danger in time based on ship navigation dynamics and ocean risk factors, and forms ship risk information in real time. It is aimed at ship prediction or early warning model research The results are beginning to show, and the main ones are: part of the risk prediction system uses the automatic identification system (AIS) [2], which uses the AIS system to implement remote tracking and monitoring of ships, but the AIS-based system has disadvantages such as poor risk prediction and high cost [3, 4]. In order to detect the risk information of ships at sea at an early stage, Xiao et al. [5-7] used distributed computing, machine learning algorithms, deep neural networks, particle swarm algorithms, etc. to make real-time early predictions of ship operations, So as to provide reference information for shipowners to avoid risks. In recent years, with the continuous development of deep
learning algorithms, algorithms such as LSTM, RNN, etc.\cite{8-10} have been applied to ship collision avoidance and risk prediction models, which have achieved good prediction results and have certain application value.

As mentioned above, although many research results have appeared one after another, most of the results are still in the state of academic research, and there is still a gap between the required standards to be applied in practice, and the existing results are only for the risk prediction model. Local technical points lack the overall nature of model research. To this end, we aim at the design of marine ship risk prediction models, with an innovative perspective, practical application-oriented, starting from the overall research of the model, and building the graph representation model in the prediction model, dynamically matching ship characteristics, and risk prediction, etc. The technical research plan makes an in-depth explanation.

2. Model research technical solutions

2.1. Model framework

The ship risk prediction model based on big data and artificial intelligence is a whole. Its research system includes: data preprocessing and modeling data layer, technical business integrating dynamic graph matching technology, machine learning technology, risk prediction technology, etc. Layer, the application layer for ship users and experts' results push and crowdsourcing feedback function. The data layer includes data from different sources. The data components are complex and the data form is multi-form. The data layer provides opaque data services for the technical layer. The algorithm on the technical layer performs intelligent analysis on the data, and the processing results of the algorithm are provided to the application layer. The application layer provides services for the top-level users and provides an interactive interface with expert users. The evaluation opinions are fed back to the technical layer through crowdsourcing feedback methods, and the technical layer adaptively optimizes the algorithms on this layer according to the evaluation results. Its intuitive representation is shown in Figure 1.
2.2. Research route for model key technology

Focusing on the research content and research goals of the model, the following describes the implementation routes of the key technical solutions in the plan, as follows:

1. Construction of graph representation model of ship risk big data

   This research intends to use a combination of machine learning and crowdsourcing technology to construct a conceptual map of ship risk.

   First of all, under the guidance of professional knowledge of marine ship risk, a language model based on neural network is used to calculate the probability model of the probability of generating multi-modal sentences at risk.

   Then, the tagging algorithm is used to obtain the word segmentation result, and then the supervised machine learning method is used to calculate the weight of the word related to the ship's risk, and under the set threshold condition, the core word (entity relationship) in the sentence of the ship information is extracted.

   Then, in order to judge the validity of the extracted entity relationship, it needs to be further handed over to the crowdsourcing expert system, which completes the entity link between the entity relationship and the entity objects extracted from the ship risk dictionary machine.

   Finally, build a conceptual feature library of ship data.

   The visual representation of the technical route for constructing the conceptual model of ship risk data is shown in Figure 2.

2) Ship dynamic feature map model: The establishment of an accurate data model is the basis of the entire research, and the latest ship feature map model representation method is adopted. Specifically, ship characteristics can be represented by a directed graph as \( G=(V, E) \). \( V \) represents a collection of nodes, and a node in \( V \) represents an entity object. Based on the above conceptual diagram, if there is a relationship between two entity objects, add a directed edge to represent it.

   Figure 3 is an example of a ship risk characteristic map. One of the difficulties in constructing feature maps is how to determine the specific connections between entities. Concept maps generally only
contain basic hierarchical relationships such as risk concepts and the synonymy and subordination between instances, and rarely involve more in-depth relationship mining of different entity types. Therefore, traditional entity relationship mining algorithms can be used to conduct in-depth analysis on specific ship-related data to obtain possible entity association tags. In order to improve the accuracy of machine learning, the crowdsourcing principle of the expert question answering system is further used to process and optimize the machine learning results, so as to obtain an accurate ship feature map model. In addition, ship risk data is constantly evolving over time. In order to accurately describe this dynamic change attribute of data, a dynamic graph representation method is needed. Specifically, the data of the same ship at different times can be expressed as a dynamic graph sequence structure based on the ship characteristic graph.

2) Multi-modal data fusion: Based on the ship dynamic feature map structure, it is further expanded to support multi-modal data fusion. Since the graph model itself can well describe the complex relationships of various entities in the actual world, each node in the graph can represent an entity, and the entity itself can be data with any structure. Therefore, if the data of each modal is represented as a node in the graph, the edges in the graph can be used to represent the relationship between multi-modal data. For example, in the historical feature map structure of a ship, there is a corresponding on-site navigation image, then the feature and image can be represented as two nodes in the graph and connected by an edge to indicate that there is a connection between the two. By establishing the connection between multi-modal data in this way, risk data fusion can be well realized. The difficulty that needs to be solved here is how to establish the specific connection between multi-modal data. For multi-modal data with ambiguous semantic relations, a crowdsourcing question and answer system can also be used to establish the connection between the data. First, use the expert question-and-answer system to mark the potentially connected multi-modal data; secondly, based on the expert's mark results, design the divergent marked results as crowdsourced questions and submit them to the system for processing.

![Fig. 3 Example of ship risk characteristic map](image_url)

2. Dynamic graph matching algorithm

It is proposed to adopt a distributed large-scale dynamic graph pruning strategy, and use the two-way simulation matching method based on graph simulation matching to prune the data graph. The two-way simulation requires that all child nodes of the current node conform to the binary relationship. The parent...
node must also conform to the binary relationship. When the data graph is updated, the data graph can be pruned through the idea of binary simulation, a large data graph can be pruned into a relatively small data graph, and the data graph can be maintained continuously. Only during the matching process You need to perform incremental matching on the small images. The research technical route of dynamic graph matching algorithm is shown in Figure 4.

**Fig. 4** The technical route of the dynamic graph matching method under the binary simulation idea

3. Vessel risk prediction

The research route of this model is divided into two sections: offline processing and online processing. The two-stage approach can help ensure the timeliness of algorithm processing under big data.

Offline processing stage: A role mining algorithm is proposed to find the roles in the ship-state-risk progression matrix. It is planned to use the random walk model to measure the strength of the relationship between the ship, role, and state indicators, for each syndrome indicator Assign weights to each role in and build a set of roles.

Online processing stage: Use the dynamic graph matching algorithm mentioned in the above research to find similar ships and establish a role-based similarity model for them. Finally, comprehensively use the ship-item-scoring matrix and the role-based similarity graph network to predict risk results, and accept crowdsourced feedback to continuously improve and optimize the diagnosis and prediction model.

The research technical route of the aforementioned machine-intelligent ship risk prediction model is shown in Figure 5. An example of a network structure based on role-based similarity is shown in Figure 6.
2.3. Model experiment design scheme

1. Experimental tools and environment: It is recommended to use Python and VC++ language tools to implement the algorithms and models in the subject; for graph data storage and visualization, use Neo4j, D3.js and other graph data management systems and graph data visualization tools. Algorithm experiments can be performed on GPU services equipped with high-performance computing.

2. Algorithm comparison and evaluation: In terms of the accuracy and performance of the algorithm, the proposed method of this article is compared with the traditional collaborative filtering algorithm, single-layer DKT (Deep Knowledge Tracing) model and two-layer DKT model algorithm and other high-level algorithms. The algorithm is continuously optimized according to the comparison results; the current best algorithm can be used as the comparison algorithm of the map data matching algorithm in this scheme. Part of the data representation model can be displayed through visualization tools, and the industry-recognized test benchmarks are used to verify the calculation accuracy and generalization ability of the risk prediction model proposed in this article.

3. Conclusion

From the perspective of big data and a new generation of information technology, this paper uses dynamic graph technology and the ship-project-scoring matrix method to make an innovative explanation of the technical design of ship risk prediction models. The design scheme is practical and feasible, and can be used for ship risk prediction systems. R&D provides better technical guidance and valuable reference. At the same time, the solution proposed in this article is only a macro-level technical solution, and many details are not explained. Therefore, in the actual development of the model, the design and application of related algorithms need to be further subdivided. In the future, on this basis, we will develop an application model for this plan to provide support for the maritime department's decision-making on ship risk management and control.

Acknowledgments

This work is supported by Natural Science Foundation of Fujian Province of China (No.2020J01435, No.2019J01846). Science and Technology Guiding Project of Fujian Province, China (No.2019Y0046). Industry University Research Collaboration Project between Ningde city and Xiamen University (No.2020C001). Ningde Normal University for service local action special plan (No.2020ZX505).
References

[1] J.H. Chen, W.T. Bian, W. Zheng, et al. Factor assessment of marine casualties caused by total loss [J]. INTERNATIONAL Journal of Disaster Risk Reduction, 2020, 47:1-10.

[2] C.T. Wei, H.C. Kie, Y.L. Wu. The Automated Risk Estimation for the Navigation of Autonomous Ships by Learning with Navigation Feature[J]. International Journal of Computational Method, 2021, 18(3):1211-1220.

[3] J.K. Kim. Semi-continuous spatial statistical analysis using ais data for vessel traffic flow characteristics in fairway[J]. Journal of Marine Science And Engineering , 2021, 9(4):1-20.

[4] M.H. Liang, R.W. Liu, S.C. Li. An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation[J]. Ocean Engineering, 2021, 225:1-18.

[5] Z. Xiao, L.Y. Zhang, X.J. Fu, et al. Concurrent processing cluster design to empower simultaneous prediction for hundreds of vessels' trajectories in near real-time[J]. IEEE Transactions On Systems Man Cybernetics-Systems, 2021, 51(3):1830-1843.

[6] T. Chai, X. Han. A study on ship collision conflict prediction in the Taiwan Strait using the EMD-based LSSVM method[J]. PoS one, 2021, 16(5):e0250948.

[7] Y.S. Zheng, X.G. Zhang, Z.J. Shang, et al. A decision-making method for ship collision avoidance based on improved cultural particle swarm[J]. Journal of Advanced transportation, 2021, 2021:1-31.

[8] C. Zhang, D. Zhang, M.Y. Zhang, et al. An integrated risk assessment model for safe arctic navigation [J]. Transportation Research Part A-Policy and Practice, 2020, 142:101-114.

[9] M. Jie, W.K. Li, C.F. Jia. Risk prediction for ship encounter situation awareness using long short-term memory based deep learning on intership behaviors [J]. Journal of Advanced transportation, 2020, 2020:1-15.

[10] Y.M. Su, J.F. Lin, D.G. Zhao, et al. Real-Time Prediction of Large-Scale Ship Model Vertical Acceleration Based on Recurrent Neural Network [J]. Journal of Marine Science and Engineering, 2020, 8(10): 1-11.