A DATA-DRIVEN FRAMEWORK FOR EXTRACTING GLOBAL MARITIME SHIPPING NETWORKS BY MACHINE LEARNING

Lei LIU[1], Ryuichi SHIBASAKI[2] and Yue HU[3]

[1]School of Engineering, the University of Tokyo
(7 Chome-3-1 Hongo, Bunkyo City, 113-8654, Tokyo)
E-mail: lei1992@seu.edu.cn
[2]Professor, School of Engineering, the University of Tokyo
(7 Chome-3-1 Hongo, Bunkyo City, 113-8654, Tokyo)
E-mail: shibasaki@sys.t.u-tokyo.ac.jp (Corresponding Author)
[3]College of Transportation Engineering, Tongji University
(No.4800, Caoan Road, Jiading District, 201804, Shanghai)
E-mail: 1910934@tongji.edu.cn

This paper proposes a framework for extracting global maritime shipping network using automatic identification system data based on machine learning methods. The framework consists of berthing area identification, trajectory selection and separation, waypoint area identification, edge generation, and network construction. To realize route planning under the consideration of historical navigation information on the network, a route planning method based on a probability-directed graph model is proposed. The real-world global AIS data of bulk carriers in 2018 is used to prove the framework. Results show that the framework is effective in extracting maritime shipping networks over global waters and makes the estimated networks applicable for analyzing the navigation speed on edges and the magnitude of the flows between nodes. Additionally, combined with the estimated shipping networks, generated routes considering node connection probabilities match well to the observed trajectories in most cases.

Key Words : *maritime shipping network, automatic identification system (AIS), shipping route, DBSCAN, A* algorithm, waypoint, directed graph*

1. Introduction

Maritime shipping network is highly important for ship routing and scheduling and flexibility analysis of the shipping system[1]. Given that the current network extraction studies mainly focus on local waters due to the limited applicability of the approaches and the difficulties to harmonize parameters in large waters[2-4], this paper proposes a framework for extracting global maritime shipping traffic networks using automatic identification system (AIS) data based on machine learning methods. In order to verify the maritime shipping networks and explore the application of the networks on ship routing, a probability-based route planning method combined with A* algorithm is proposed after transforming the estimated shipping network into a probability-directed graph.

2 Conceptual framework

Fig.1 shows the complete framework on network extraction and route planning. The left part displays the flow of maritime shipping network extraction, whereas the right part shows the route recommendation and planning based on the left results.

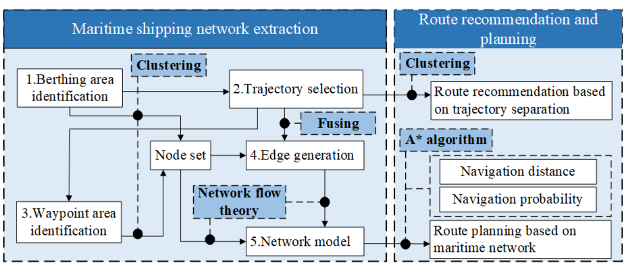

Fig.1 The framework of network extraction and route planning

The core steps in the framework are as follows:

- The berthing ship points are identified according to the navigation status in AIS data and the berthing area can be obtained through a clustering algorithm as the first-class nodes;

- To ensure the integrity of the extracted shipping network and reduce computation, a trajectory hybrid separation[5] is used to find the trajectories between all pairs of berthing areas obtained in the first step to get valid ship trajectories;
- As the identification elements of the second-class nodes, waypoints are where a ship changes its current motion behavior, mainly involving speed and course variation. Different from berthing points, waypoint identification requires more complex filtering conditions. After determining waypoint candidates, the clustering algorithm combined with instantaneous and continuous variation features gets waypoint areas as the second-class nodes;
- A segmented track between adjacent nodes can be taken as the edge by identifying all berthing and waypoint points in a trajectory that belong to the berthing and waypoint areas, and the trajectory fusion is performed to get the unique edge when there are multiple pairs of segmented tracks between adjacent nodes[6];
- With the complete node set (berthing and waypoint areas) and the arcs generated from ship trajectories, a directed maritime shipping network is constructed using network flow theory.

3 Results and conclusions

The real-world global AIS data of bulk carriers in 2018 is used to extract the maritime shipping networks to prove the framework. There were total 5,457 nodes in the estimated global maritime shipping networks, including 2,769 berthing areas and 2,688 waypoint areas, as shown in **Fig.2**.

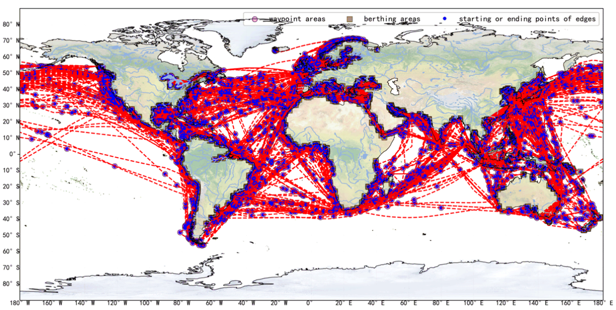

Fig.2 The estimated maritime dry bulk shipping network

The clustering parameters for berthing area identification and waypoint area identification revealed a similar inter-point distance of berthing points and a varied inter-point distance of waypoints in different waters. The results also suggested that the framework is effective in extracting maritime shipping networks over large waters and makes the estimated shipping networks applicable for analyzing the navigation speed on edges and the magnitude of the flows between nodes.

The comparison of planned routes on the estimated shipping networks and the grid reference system showed that the routes generated on the former were smoother and more similar to the observed trajectories. Regarding the probability-based route planning method, distance-based planning was more stable than probability-based planning in the estimated shipping networks while probability-based planning was more applicable to local areas in the grid reference system. In addition, route planning based on the estimated shipping networks can save more time than that based on the grid reference system when routes get longer.

This study focused on the extraction of maritime shipping networks in global waters and optimizes the conditions for waypoint identification for the extension of study waters. However, it still took the estimated and unique recognition thresholds, and thus could not avoid the problem that the recognition thresholds were difficult to standardize for different ships in different scenarios. One possible alternative is to automatically identify whether a vessel changes its navigation status based on its continuous AIS points, which may be achieved by constructing multi-dimensional features of ship's continuous points combined with classification algorithms. Additionally, after transforming the shipping network into a probability-directed graph by counting the frequency of connections between nodes, the planned routes with the highest probability on this basis show some unreasonable outcomes, especially for large waters. The most likely reason is that the way of calculating the probability of node connections is too single while node connections are actually related to the origin, destination, ship size, weather conditions, etc. Consequently, multiple factors need to be fully considered when implementing route planning considering connection probability.

Overall, the proposed framework and methods may help provide a comprehensive framework to obtain and analyze maritime shipping traffic networks, and enrich route planning methods by considering historical navigation patterns. Future work can not only focus on solving the above-mentioned problems to improve the methodology of shipping network extraction but can also follow current research for a deeper analysis of the estimated networks.

REFERENCES

1) Varlamis, I., et al., Building navigation networks from multi-vessel trajectory data. *GeoInformatica*, 2020. 25(1): p. 69-97.
2) Shukai, Z., et al., Data-driven based automatic maritime routing from massive AIS trajectories in the face of disparity. *Ocean Engineering*, 2018. 155: p. 240-250.
3) Naus, K., Drafting Route Plan Templates for Ships on the Basis of AIS Historical Data. *Journal of Navigation*, 2019. 73(3): p. 726-745.
4) Yan, Z., et al., Exploring AIS data for intelligent maritime routes extraction. *Applied Ocean Research*, 2020b. 101.
5) Liu, L., et al., A Hybrid-Clustering Model of Ship Trajectories for Maritime Traffic Patterns Analysis in Port Area. *Journal of Marine Science and Engineering*, 2022. 10(3).
6) Rong, H., A.P. Teixeira, and C. Guedes Soares, Ship trajectory uncertainty prediction based on a Gaussian Process model. *Ocean Engineering*, 2019. 182: p. 499-511.