



## LJMU Research Online

Chen, C, Chen, XQ, Ma, F, Chen, YW and Wang, J

**Deviation warnings of ferries based on artificial potential field and historical data**

<http://researchonline.ljmu.ac.uk/id/eprint/12042/>

### Article

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Chen, C, Chen, XQ, Ma, F, Chen, YW and Wang, J (2019) Deviation warnings of ferries based on artificial potential field and historical data. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment. ISSN 1475-0902**

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact [researchonline@ljmu.ac.uk](mailto:researchonline@ljmu.ac.uk)

<http://researchonline.ljmu.ac.uk/>



## LJMU Research Online

Chen, C, Chen, XQ, Ma, F, Chen, YW and Wang, J

**Deviation warnings of ferries based on artificial potential field and historical data**

<http://researchonline.ljmu.ac.uk/id/eprint/12042/>

### Article

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Chen, C, Chen, XQ, Ma, F, Chen, YW and Wang, J (2019) Deviation warnings of ferries based on artificial potential field and historical data. Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment. ISSN 1475-0902**

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact [researchonline@ljmu.ac.uk](mailto:researchonline@ljmu.ac.uk)

<http://researchonline.ljmu.ac.uk/>

# Deviation Warnings of Ferries Based on Artificial Potential Field and Historical Data

Chen Chen<sup>1,2</sup>, Xian-Qiao Chen<sup>1,2</sup>, Feng Ma<sup>3\*</sup>, Yu-Wang Chen<sup>4</sup>, Jin Wang<sup>5</sup>

School of Computer Science and Technology, Wuhan University of Technology, Wuhan, China<sup>1</sup>

Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan, China<sup>2</sup>

Intelligent Transportation System Center, Wuhan University of Technology, Wuhan, China<sup>3</sup>

Alliance Manchester Business School, University of Manchester, Manchester, UK<sup>4</sup>

Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK<sup>5</sup>

**Abstract:** Ferries are usually used for transporting passengers and vehicles among docks, and any deviation of the course can lead to serious consequences. Therefore, transportation ferries must be watched closely by local maritime administrators, which involves much manpower. With the use of historical data, this paper proposes an intelligent method of integrating artificial potential field (APF) with Bayesian Network (BN) to trigger deviation warnings for a ferry based on its trajectory, speed and course. More specifically, a repulsive potential field-based model is firstly established to capture a customary waterway of ferries. Subsequently, a method based on non-linear optimisation is introduced to train the coefficients of the proposed repulsive potential field. The deviation of a ferry from the customary route can then be quantified by the potential field. BN is further introduced to trigger deviation warnings in accordance with the distribution of deviation values, speeds and courses. Finally, the proposed approach is validated by the historical data of a chosen ferry on a specific route. The testing results show that the approach is capable of providing deviation warnings for ferries accurately and can offer a practical solution for maritime supervision.

**Keywords:** Route deviation, artificial potential field, non-linear optimisation, Bayesian Network

## 1 Introduction

A ferry is an important water transportation vessel, which is used to carry passengers as well as cargos, including hazardous materials, livestock, vehicles, and even trains. Different from cargo ships, ferries usually commute among wharves or traffic hubs, ranging from a few hundred metres to hundreds of nautical miles. As a special vessel for commuting, a ferry has some unique features as follows. The hull structure and facilities of a ferry are relatively simple, since ferries usually do not have to sail for a long distance, especially the in-land ones are always equipped with a light hull and limited living facilities. However, the cabin and deck of a ferry are generally wide so that it can carry many passengers and much load easily. Ferries are customarily designed as double-ended to make it convenient for passengers and vehicles to get on and off. Furthermore, twin or more full-directional thrusters are equipped to improve the manoeuvrability and controllability of ferries. Some ferries are of a catamaran hull in order to meet the requirements of high speed and seakeeping performance. According to the actual use, ferries can be classified as passenger ferries, car ferries and train ferries. By now, ferries become an integral and essential component of river and island transportation systems.

A ferry sometimes must navigate under poor visibility conditions in order to meet the transportation demand. For instance, truck logistics in Nanjing, Zhenjiang and Yangzhou of Jiangsu

province, China is so heavy that car ferries must operate as usual at night and even in fog conditions. Similarly, in order to meet the requirements of customers, cross-sea ferries take an evening-morning operation which requires them running all the time at night. Moreover, ferry operators have to endure fixed routes and long working hours, hence fatigue in operations is unavoidable. Ferry ships often run between two sides of a strait, which might encounter other vessels passing this strait frequently, leading to high risks. Accidents of ferries have occurred from time to time. For example, ferry Dashun capsized, when it just left Yantai with 10 kilometres away from the shore on November 24 of 1999, and 280 passengers were confirmed dead or missing. In the sinking of South Korea Sewol, 296 lives were lost on April 16, 2014. Although a ferry is generally small in terms of tonnage, it often carries a large number of passengers and vehicles. In the case of collision or overturning, it could easily cause mass human casualties and incur a high negative impact. The safety problem for ferries has attracted much attention from all over the world.

In fact, a ferry is faced with many safety issues, including route deviation, extreme weather and human errors in collision avoidance. In practice, the corresponding maritime administration of ferries usually pays much attention on three factors: deviations, abnormal speeds, and abnormal courses. A ferry is generally equipped with assistant devices including an automatic identification system (AIS), and/or BeiDou Navigation Satellite System (BDS) maritime satellite terminals. Taking Shandong Maritime Safety Administration as an example, it requires all the supervised ferries to be firmly watched by supervisors remotely and manually on top of mandatory navigation facilities and measures in order to ensure safety, especially for those on the Yanda route and China-Japan-Korea routes. However, this tight supervision not only imposes great pressure on human supervisors, and it is also difficult to ensure that the entire process of supervision is effective. In this context, it is critical to develop an intelligent approach to evaluate hazard level for improving the safety of ferry navigation.

There are many factors that could influence the safety of a ferry ship, including weather conditions, human factors and encountering ships. In general, hazard identification is similar for ordinary cargo ships and ferries, but there are particularities for ferries. Firstly, the local maritime administration intends to set a dedicated navigation area or a route for ferries. Meanwhile, ferry companies are also willing to make their ferries running on fix routes for easy management. Therefore, the routes of ferries always follow some principles, specified tracks or guidelines. In other words, any deviation from such principles or specified routes generally means an abnormal situation or a high risk. From the view of local administration, when a ferry behaves abnormally, especially deviating from its specified route, it should be identified immediately with an appropriate mitigating measure taken. However, such customary routes of ferries are unique for each single ferry, how to identify a ferry deviating from its route requires a great deal of knowledge and experience and needs much attention of local supervisors. Moreover, the boundaries of such routes are generally vague. Speeds, courses, and other factors should be taken into consideration for judgement. As a result, reasonable judgments of deviations rely heavily on experienced supervisors. However, the resources of administration are always limited in everyday management. To address this problem, it is necessary to invent artificial intelligence to identify any route deviation automatically, through continuously watching every ferry in its customary route.

Artificial potential field (APF) has been widely applied in fields of a lane keeping system of intelligent vehicles, obstacle avoidance path planning of robots and road risk awareness. To make the APF model applicable to the identification of customary routes for ferries, this research analyses the historical trajectories distribution of specific ferries over a period of time. The coefficients of APF can be obtained by building optimisation models to capture the actual distribution. Subsequently, navigation areas or customary routes of these ferries can be learned. The warnings of route deviation can be comprehensively identified by a Bayesian Network-based method. The proposed method is validated on a ferry Bohaiyinzhu, which travels from Lvshun to Penglai (the Yanda route).

This research aims to propose an intelligent approach to trigger deviation warnings for ferries on the basis of AIS history data and the APF model. The rest of this paper is organised as follows. The literature of vessel risk assessment and real-time warning, route deviation, the APF model application in lane departure and ship navigation are briefly reviewed in Section 2. A novel approach is proposed to assess hazards for ferries in Section 3. The proposed approach is validated through a case study in Section 4. Section 5 concludes this paper.

## 2 Literature review

The research dedicated to analyse route deviation or other risks on ferries is very limited, as ferries are generally considered as normal ships in references. Ship risk assessment refers to risk assessment of vessel accidents under different navigation environments. Formal safety assessment (FSA) was put forward by researchers in the UK, which was also approved by the International Maritime Organization (IMO) as a standard and systematic process for accessing maritime safety in 2002 [1]. Since then FSA has been widely used as a tool to establish an evaluation framework of ship navigation risk, and it provides not only qualitative information but also quantitative results. Lois *et al.* [2] applied the FSA framework in a critical evaluation of cruise passenger vessels, and they conducted a test case study to demonstrate the feasibility of the proposed approach. Wang *et al.* [3] discussed how a formal safety assessment methodology was applied to containership operations. Montewka *et al.* [4] introduced a qualitative scoring system on risk for FSA and showed its applicability on an exemplary risk model for a RoPax ship.

In recent years, researchers have introduced many other new methods for water transportation safety evaluation. Hu *et al.* [5] presented a risk-assessment approach based on fuzzy functions for the assessment of pilotage safety in Shanghai Harbour. Mentés *et al.* [6] proposed an FSA and fuzzy-set-based theory approach to evaluate risks of cargo ships at coasts and open seas of Turkey. You *et al.* [7] presented an approach to assess the risk and sustainability associated with ship collision accidents, which was illustrated on a maritime transportation system in the Delaware River area. Wang *et al.* [8] proposed a decision-making method based on subjective analysis of evidential reasoning (ER) and fuzzy set modelling when there was a large amount of uncertainty. Montewka *et al.* [9] introduced a systematic, transferable and proactive framework based on BNs and a set of analytical methods for the estimation of the risk model parameters, which had been applied to a maritime transportation system operating in the Gulf of Finland and the results have good agreement with the available records. In

summary, these researchers have proposed a set of methods on risk assessment for general purposes. However, these methods are not designed for triggering a real-time warning for vessels.

In fact, previous research demonstrated that there is a high correlation between real-time risk warning of a specific vessel and its navigation trajectory. Sutulo *et al.* [10] showed a kinematic mathematical model for short-term path prediction in manoeuvring simulation. Sang *et al.* [11,12] provided a novel method for restoring ship trajectory and proposed a real-time ship safety early warning method. This method collects AIS information to calculate ship dynamic information of the distance to closest point of approach (DCPA) and the time to closest point of approach (TCPA) and predicts ship trajectory to obtain the real-time collision risk. The timeliness and reliability of this method were verified by means of an accident case of Wuhan Yangtze River Bridge. According to historical data, Gan *et al.* [13] predicted the long-term trajectories of ship with satisfactory accuracy using K-Means clustering and Artificial Neural Network (ANN) models. Chen *et al.* [14] developed a cross-disciplinary application of ecological methods in habitat use of the wild animal principle to extract fairways boundaries of ships crossing Taiwan Strait.

Nowadays, a majority of large-tonnage ships are well equipped with an electronic chart display and information system (ECDIS), which allows users to set a "deviation limit" and a "deviation angle" according to their coordinates and heading. In the navigation, if the vessel deviates from its route more than the "deviation limit" or the "deviation angle", such an ECDIS might trigger a warning. However, such deviation limits are generally given by users without a rational and scientific basis. In addition, such a mechanism does not take the customary routes into account, and it was not intelligent enough in supporting decision-making and early warning.

In practice, the route deviation of ships is very similar to the lane deviation of intelligent vehicles. The lane departure system of intelligent vehicles is mainly based on the relative distance of vehicle-road or the expected path deviation feedback. However, this method cannot be transplanted to the application of ferries directly, since waterways are usually much wider than vehicle lanes. APF proposed by Khatib *et al.* [15] has been widely used in the path planning system of robots and intelligent vehicles. In practice, the APF-based method has been proved to be capable of planning a path efficiently [16]. In general, APF consists of attractive fields and repulsive fields. A destination in APF is usually described as a source of attractive potential field (PF), which attracts an agent to approach such a destination. On contrary, an obstacle in PF is usually described as a source of repulsive field located at its position so that the obstacle might repulse agents away [17]. Wolf *et al.* [18] presented a set of PF components, including lane potential, road potential, car potential and velocity potential. This model was proved to be valid on a multi-lane, populated high-way where intelligent vehicles could be automated in lane-keeping or lane-changing in different scenarios. Rasekhipour *et al.* [19] considered non-crossable obstacles and crossable obstacles where a road environment APF was built closely to the actual situation. In the research, a minimum value of PF can be generated at the central region of the lane. Homoplastically, a maximum value is decided on the boundary of the road. In practice, the coefficients of PF functions are determined by the position and dynamic status of the corresponding

vehicle. By applying the APF model, boundaries of lane can always be found automatically, hence the autonomous vehicle cannot overstep the lane boundary under the influence of the PF function.

In fact, APF was widely used in the field of ship movement because of its advantages in complex movement. Ma *et al.* [20] made effective use of repulsion PF to describe navigation risk distribution and obtained coefficients of the repulsion field based on historical trajectories and a nonlinear optimisation method. His research reveals the correlations among APF, ship behaviours and risk potential distributions [21]. Rong *et al.* [22] evaluated near ship-ship collision situations in the Tagus River Estuary using a simulation model based on the APF method of ship navigation in restricted waters. Chen *et al.* [23] designed an improved APF method to address the problem of collision avoidance for unmanned ships. Yogang Singh *et al.* [24] elaborated the use of APF in path planning of an autonomous surface vehicle (ASV) in a real time marine environment in Portsmouth Harbour.

As elaborated previously, the distance or deviation to the customary waterways or routes is not the only factor in the determination of deviation warnings. Course, speed and many other related factors are also important. Hence, a decision should be made with a combination of these factors. In such a case, Bayesian Network [25], a commonly used machine learning method can be a practical option to make inference based on dependant or independent factors.

This research puts forward the APF model of the Yanda Route to simulate that the ship is traveling in a customary route. By means of the ferry historical trajectory coming from AIS, the coefficients of APF can be learned. The following step is to establish a model to describe the customary route of a ferry based on APF and Bayesian Network. Then, warnings can be triggered automatically for all ferries sailing on this waterway.

### **3 A proposed approach**

#### **3.1 The area and object under study**

The Yanda waterway or route is an important waterway which connects Liaoning province with Shandong province, China, and is known as the "golden waterway" of Bohai Gulf. The shipping activity of this route is very busy, which is being closely watched by the local maritime administration. Presently, the entire Yanda route has been covered by the shore-based AIS base stations, hence the AIS data from the ships in Bohai Gulf can be obtained easily. Presently, there are 19 roll-on-roll-off passenger ships, which sail in the roundtrips from Yantai to Dalian, Penglai to Lvshun, and Longkou to Lvshun. In accordance with the AIS records, from July 1st to July 31st of 2018, there were 6 ferries sailing on the Yanda route. Their Marine Mobile Service Identity (MMSI) numbers are 412328370, 413409000, 413408000, 412330020, 414096000 and 414095000. The detailed information of these ships is shown in Table 1. According to the AIS data from the local maritime administration, the historical coordinates and speeds of the six ships can be extracted. Then, the average speeds can be calculated, and meanwhile the trajectories can be obtained. Subsequently, the ship flow heat map of the Yanda route is made, as shown in Fig. 1. The bar in this figure represents the correlations between the colour and the ship density.

Table 1 Traffic Characteristics of Ferries on Yanda

MMSI	Ship name	Ship length (m)	Ship width (m)	Max speed (knot)	Average speed (knot)	route
412328370	Bohaiyinzhu	161.20	24.80	18.5	13.7	Lvshun-Penglai
413409000	Bohaizhenzhu	163.95	25.00	18.5	14.1	Lvshun-Penglai
413408000	Bohaiyuzhu	163.95	25.00	18.5	14.2	Lvshun-Longkou
412330020	Bohaibaozhu	163.95	25.00	18.5	14.0	Dalian-Yantai
414096000	Bohaicuizhu	178.80	28.00	20	14.2	Dalian-Yantai
414095000	Bohajingzhu	178.80	28.00	20	14.4	Dalian-Yantai

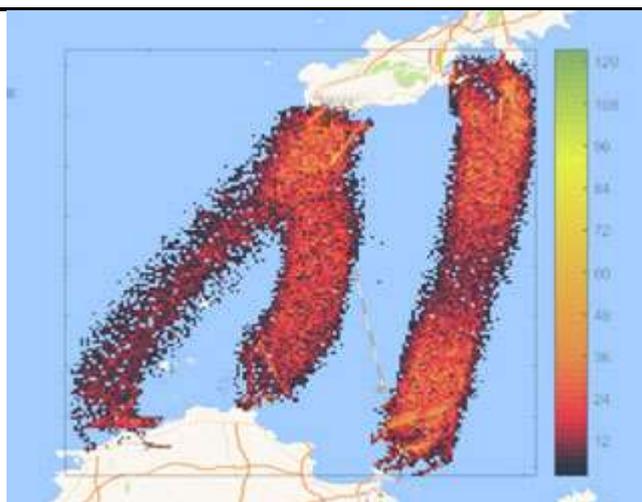


Figure1 The heat map of the ferry density distribution of the Yanda route

It can be inferred that several customary routes of ferries can be easily identified. In order to address the problem of recognizing route deviations, this research chooses the “Bohaiyinzhu” ferry as the study object. Bohaiyinzhu is a typical roll-on roll-off passenger ship with 162.2 m long and 24.8 m wide, which is capable of carrying 200 vehicles and 1,128 passengers simultaneously. This ferry is equipped with a satellite positioning system (GPS) and an AIS terminal. Therefore, by extracting the AIS historical data of Bohaiyinzhu, the historical trajectory of this ferry ship can be obtained.

Moreover, it clearly shows in Fig. 1 that the ferries are more likely to sail at the centre of the customary routes or waterways. On the other hand, such a phenomenon can be regarded as that this ferry ship was pushed into narrow routes by some undetectable “repulsions”. Apparently, the closer to the boundaries, the greater the repulsion a ferry ship gets; the distance is the core factor in the attenuation of the repulsions. The strength of the “repulsions” can be considered as consistent with the corresponding hazard level, since ships are always intended to sail in safe routes. By analysing the distribution of passing vessels, the corresponding repulsions or repulsive potentials can be quantified. Therefore, the hazard level of a position can be obtained indirectly. By this meaning, the intelligence of identifying route deviation is presented as follows.

### 3.2 The modelling of channel potential field using the APF model

As mentioned previously, a vehicle usually travels within a lane. In the corresponding research, such boundaries or constraint relationships are often described with the model of PF. Similarly, ferries always travel along a specific or customary route. The only difference is that the route of a ferry is much wider than a lane of an automobile. In this occasion, a lane or a channel PF model can be used to describe relationships between a vessel and its customary route. It is easy to find a vessel PF value according to its position in an APF model, then it is possible to estimate a hazard distribution based on its position when the coefficients are known. For simplicity, only the repulsive PF is used in this research.

The channel PF is introduced to build a model of a customary route, which can explain why a ferry is willing to keep in its own route. Following such a principle, the route boundaries generate a repulsive potential or a repulsive force to the corresponding ferry. If the ferry is close to boundaries, the potential field might push it back to the centre of the route.

To formulate such an APF model, a coordinate system is established at the first step. Let the x-axis align with the heading of the ferry in its customary route and let us choose the origin point such that the right boundary is centred at  $y = 0$ , as shown in Fig. 2. The ferry was considered as a point in this coordinate frame, and its coordinate is presented as  $(x, y)$ .

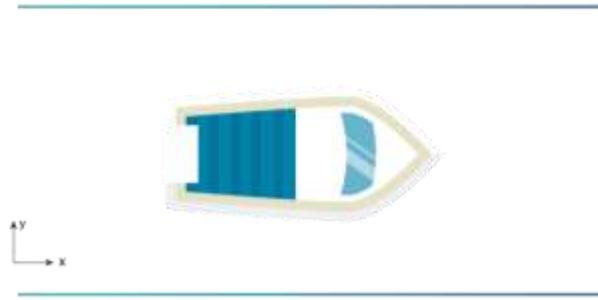


Figure 2 Channel Setup: ferry motion coordinate system

The channel PF prevents the ferry from leaving the channel by becoming infinite at the channel boundaries. Facing the centre of the channel, the channel potential field is gradually reduced to zero. According to the experience of APF in relevant research on robots [26], a repulsive PF of each boundary of the customary route is presented:

$$U_{channel,j} = \frac{1}{2}\eta \left( \frac{1}{y-y_j} \right)^2, \quad (1)$$

where  $\eta$  is a scaling factor and  $y_j$  is the  $j^{th}$  channel boundary coordinate,  $j \in \{1,2\}$ .

The curve of channel edge potentials with typical coefficients is illustrated in Fig. 3. With this model, the safety level or hazard level can be inferred.

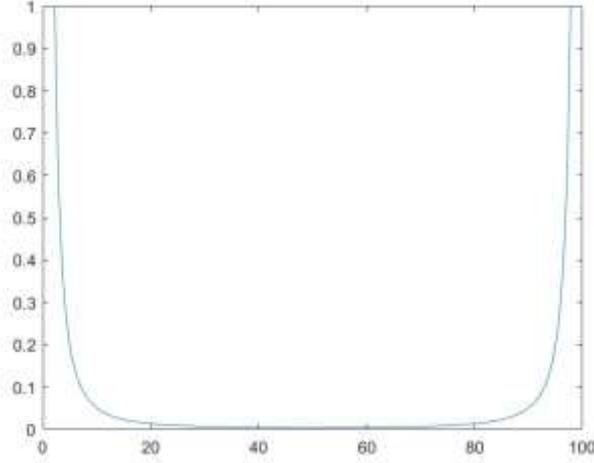


Figure 3 Edge potential fields with typical coefficients

### 3.3 A nonlinear optimisation of the coefficients of PF

The following question is to obtain the coefficients of PF. As discussed previously, the left and right boundaries of the route are considered as producing direct repulsive forces to the ferry, given by Eq. (1). In this view, the closer to the centre of this channel or this route, the smaller the corresponding repulsive, as illustrated in Fig. 3. As elaborated previously, the distribution of historical ferry tracks can be considered as another representation of the hazard level.

By selecting a profile in the customary route of the ship, the navigation direction of the ship in which the  $x$ , coordinate of all points on this profile is the same. For any point of the profile, the potential field value is only related to  $y$ , which is represented as:

$$P(y, \overrightarrow{para}) = \frac{1}{2}\eta \frac{1}{(y-y_1)^2} + \frac{1}{2}\eta \frac{1}{(y-y_2)^2} \quad (2)$$

where  $\overrightarrow{para} = \{\eta, y_1, y_2\}$  denotes all the undetermined coefficients of Eqs. (2);  $y_1$  and  $y_2$  denote  $y$  axis coordinates of the left and right sides of this channel profile.

Suppose a cross profile contains  $L$  discrete statistical points or sections  $\{(x, y_1), \dots, (x, y_L)\}$ . There is a point  $(x, y_k)$  on this cross profile,  $1 \leq k \leq L$ . Its corresponding potential field is represented by  $P(y_k, \overrightarrow{para})$ . Therefore, the potentials for all  $L$  points of this profile can be represented as  $\{P(y_1, \overrightarrow{para}), \dots, P(y_L, \overrightarrow{para})\}$ , and the maximum and minimum potentials of the  $L$  points are represented as  $P_{max}[P_1 \overrightarrow{para} \dots P_L \overrightarrow{para}]_{max}$ ,  $P_{min}[P_1 \overrightarrow{para} \dots P_L \overrightarrow{para}]_{min}$ . Therefore, the normalised potential of the point  $(x, y_k)$  is presented as:

$$P_{normal}(y_k) = [P(y_k, \overrightarrow{para}) - P_{min}] / (P_{max} - P_{min}) \quad (3)$$

Hence,  $[1 - P_{normal}(y_k)]$  can be regarded as a normalised safety degree of point  $(x, y_k)$  on this profile. The normalised safety degree distribution of the  $L$  points can be presented as:

$$\overrightarrow{P^*} = \{1 - P_{normal}(y_1), \dots, 1 - P_{normal}(y_L)\} \quad (4)$$

The distribution of the passing vessels of the  $L$  points can be denoted as a vector  $\vec{d} = \{d_1, \dots, d_L\}$ , and the maximum and minimum passing vessel numbers of the  $L$  points as  $d_{max}$  and  $d_{min}$ , respectively. Hence, the normalised distribution of vessels on the  $L$  points is presented as:

$$\vec{d}^* = \{d_1 - d_{min}, \dots, d_L - d_{min}\} / (d_{max} - d_{min}) \quad (5)$$

It was discussed that the appropriate coefficients of the potential fields  $\overline{para}$  should make the deviation between  $\overline{d}^*$  and  $\overline{P}^*$  minimum. Therefore, the coefficients can be obtained with a nonlinear optimisation model, which is presented as,

$$\overline{para} = \{\eta, y_1, y_2\} = \underset{feasible}{argmin} \sum_{i=1}^L |[1 - P_{normal}(y_i)] - (d_i - d_{min}) / (d_{max} - d_{min})| \quad (6)$$

Such an optimisation can be implemented by the *fmincon* function of MATLAB 2018 [21]. With all these equations and historical data, the safety degree can be quantified and normalised.

### 3.3 Trigger a warning based on Bayesian Networks

Based on the quantified characteristics of a ferry ship, a BN-based inference process is conducted as follows. BN is defined by a pair  $(S, \Theta_S)$ , where  $S = (\chi, E)$  is a directed acyclic graph (DAG) with a set of nodes  $\chi$ , and with a set of arcs or nodes  $E = \{(X_i, X_j) | X_i, X_j \in \chi, X_i \neq X_j\}$  representing the probabilistic dependencies among domain variables [26].  $\Theta_S$  represents the parameterization of a probability measure  $\rho$  defined over the space of possible instantiations of  $\chi$ . Given a node  $X_i \in \chi$ , **Pai** is used to denote the set of parents of  $X_i$  in  $S$ . The essential property of BN is summarized by the Markov property, which asserts that each variable is independent of its non-descendants given its parents. The application of the chain rule, together with the Markov property, yields the following factorization of the joint probability of any particular instantiation  $\vec{x}$  of all  $n$  variables:

$$\rho(\vec{x}) = \rho(x_1, \dots, x_n) = \prod_{i=1}^n \rho(x_i | \mathbf{Pai}_i, \Theta_S) \quad (7)$$

In practice, manual work gives a warning (W) based on the deviation to customary routes (D), the speed (V), the course deviation (C), and other factors. The deviation to the customary route can be given from Section 3.2. The course deviation can be considered as the included angle between the customary route and the course of this ferry ship. Hence, (W), (D), (V) and (C) form a DAG (Directed Acyclic Graph). Subsequently, the undetermined structure of the DAG can be learned from historical data and verified data samples. Presently, the K2 scoring algorithm is widely accepted for constructing BN from databases or records, proposed by Cooper and Herskovits [27], which is fully supported by the software tools of BN, including Netica, Hugin, and the MATLAB bnt toolbox. **The MATLAB bnt toolbox is an open source software program, which is released on Google Open Source projects. Moreover, it is a popular and widely used BN-based inference tool.**

When the structure is determined, the conditional probability tables (CPTs) of the DAG can be learned from verified samples too. Usually, a maximum likelihood estimation (MLE) is used to implement CPTs estimation when given a training data. In this research, the expectation maximization (EM) algorithm is adopted, which is an iterative method to carry out an MLE [28]. Such a process is also supported by the software tools described above. Hence, the details of the EM algorithm will not be given here. Lastly, the probability of a ship should be warned can be estimated with the new DAG.

## 4. A case study

### 4.1 The distribution of vessels of Yanda

As discussed in Section 3, a ferry traveling in the customary route is affected by the PF of the route, and the route distribution should follow the routine that there are more ships in the middle of the route

and fewer ships near the boundary of the route. The distribution of Bohaiyinzhu route, which is relatively isolated on the vast sea, should follow this rule.

This research collected AIS data from the AIS equipment on the Bohaiyinzhu from 0:00:16 on July 1, 2017 to 2:5:50 on July 22, 2017. After eliminating error messages, a total of 117,115 records were obtained. A software program was developed on the VC++ platform, which is capable of parsing the collected AIS information from database. By drawing the historical coordinates of Bohaiyinzhu on the electronic chart, two customary routes can be easily identified in Fig. 4.



Figure 4 The trajectory of Bohaiyinzhu

Based on the approach elaborated in Section 3.3, the research chooses several profiles in Fig. 4 to obtain appropriate coefficients of the PF model of the Yanda Route. Profile *K1* lies at the range of [121.05,121.1] longitude and [38.3,38.4] latitude, containing 3,537 trajectory points. Profile *K2* lies at the range of [120.88, 120.98] longitude, [37.94, 37.96] latitude, containing 1,123 trajectory points. Profile *K3* lies at the range of [120.90, 120.97] longitude and [38.54, 38.56] latitude, containing 1,135 trajectory points.

Subsequently, the densities of each profile can be further processed by the Origin software. The frequency or the **ship distribution maps** of these three profiles can be obtained respectively, as shown in Fig. 5 and Fig. 7, where the X-axis represents the longitudes, and the Y-axis the densities. Taking profile *K1* as an example, the data of profile *K1* processed by the Origin software was divided into 17 groups. The Origin software also produces the position of the centre point of each group, the quantity of points in each group, the frequency of each group, *etc.* The results are shown in Table 2.

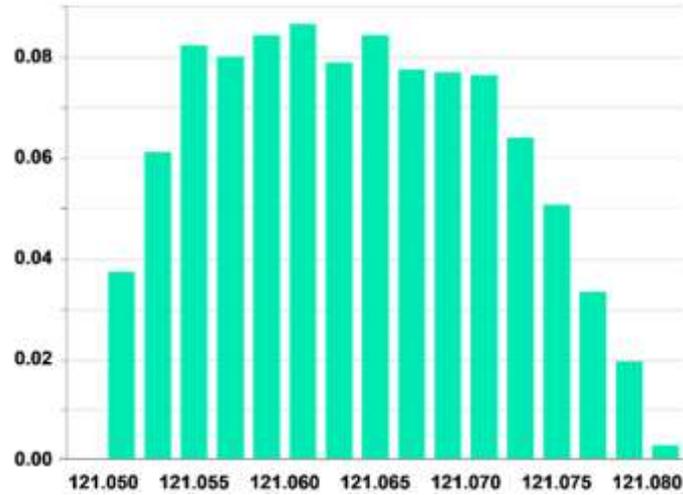


Figure 5 The ship distribution of profile *K1*

Table 2 The frequency outputs of profile *K1* data

No	Bin Center	Bin End	Count	Relative Frequency	Normalized
1	121.049	121.05	0	0	0
2	121.051	121.052	133	0.0376	0.43322
3	121.053	121.054	217	0.06135	0.70684
4	121.055	121.056	292	0.08256	0.95114
5	121.057	121.058	284	0.08029	0.92508
6	121.059	121.06	299	0.08453	0.97394
7	121.061	121.062	307	0.0868	1
8	121.063	121.064	280	0.07916	0.91205
9	121.065	121.066	299	0.08453	0.97394
10	121.067	121.068	275	0.07775	0.89577
11	121.069	121.07	273	0.07718	0.88925
12	121.071	121.072	271	0.07662	0.88274
13	121.073	121.074	227	0.06418	0.73941
14	121.075	121.076	180	0.05089	0.58632
15	121.077	121.078	119	0.03364	0.38762
16	121.079	121.08	70	0.01979	0.22801
17	121.081	121.082	11	0.00311	0.03583

Then, the distribution of vessels on profile *K1* can be normalised with Eq. (5) and presented in Fig. 6, where the X-axis represents longitude, and the Y-axis the normalised distribution.

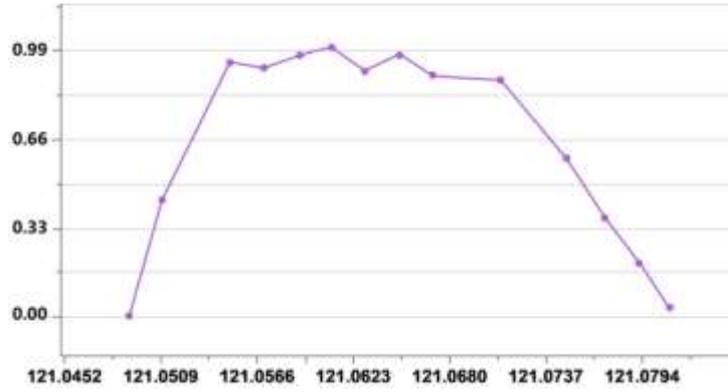


Figure 6 The normalised distribution of ships on profile *K1*

Similarly, we can process the data of profile *K2* and profile *K3* in the same way, and their distribution charts and normalized distribution as shown in Fig. 7 and Fig. 8.

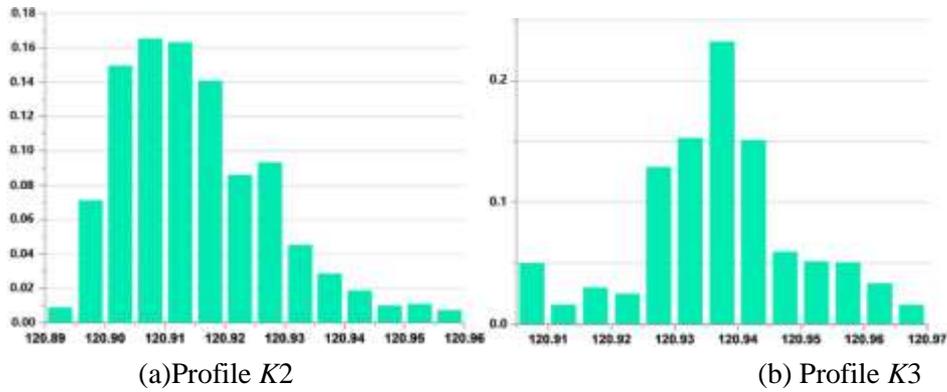


Figure 7 The ships distribution

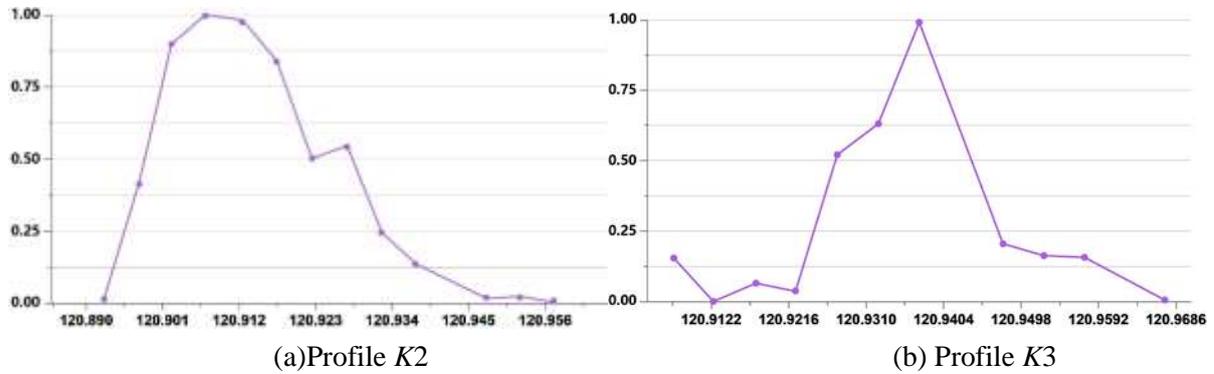


Figure 8 The normalised distribution of ships

#### 4.2 The training of the artificial potential coefficients

The next step is to obtain the coefficients of these potential fields. Apparently, the coefficients should make the potentials consistent with the distribution of vessels. Therefore, the coefficients can be obtained in a nonlinear optimisation model, as Eq. (6). Taking profile *K1* as an example, 17 bin centre points were chosen randomly as elaborated in Section 4.1. In this occasion,  $i = 17$ , the coefficients are obtained as  $\overrightarrow{para} = \{\eta, y_1, y_2\} = \{132.2629, 121.0280, 121.0998\}$  using the ‘*fmincon*’ function of MATLAB 2018.

With these coefficients, the potential field of all points on profile *K1* can be obtained by Eq. (2), and the normalised “potential distribution” (hazard level) and normalised “safety distribution” are presented in Fig. 9 and Fig. 10, which are defined in Eq. (3) and Eq. (4).

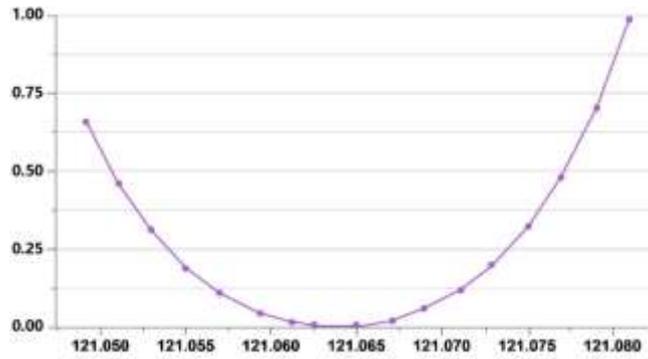


Figure 9 The normalised distribution of potential filed on profile *K1* (hazard level)

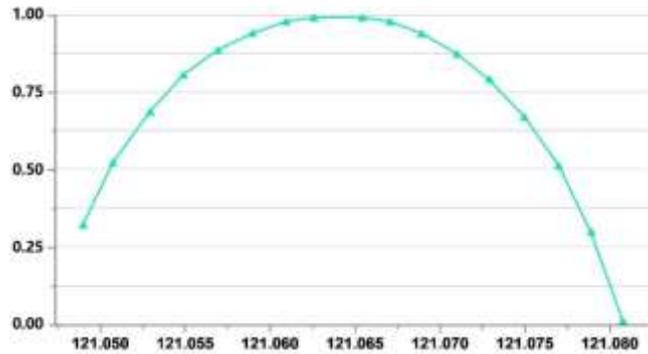


Figure 10 The normalised distribution of safety degree on profile *K1*

By comparing Fig. 6 and Fig. 10, a good agreement can be found. In other words, the distribution of potentials is consistency with the distribution of ships on profile *K1*.

Moreover, the artificial potential coefficients of profile *K2* and profile *K3* can be obtained in the same way, the outputs are  $\{3.6911, 120.8700, 120.9799\}$  and  $\{3.0342, 120.8760, 120.9990\}$ , respectively. Based on Eq. (3), the normalised potential distributions of profile *K2* and profile *K3* are presented in Fig. 11. Based on Eq. (4), the normalised safety degrees of profile *K2* and profile *K3* are presented in Fig. 12. Obviously, a high agreement can also be found between Fig. 8 and Fig. 12.

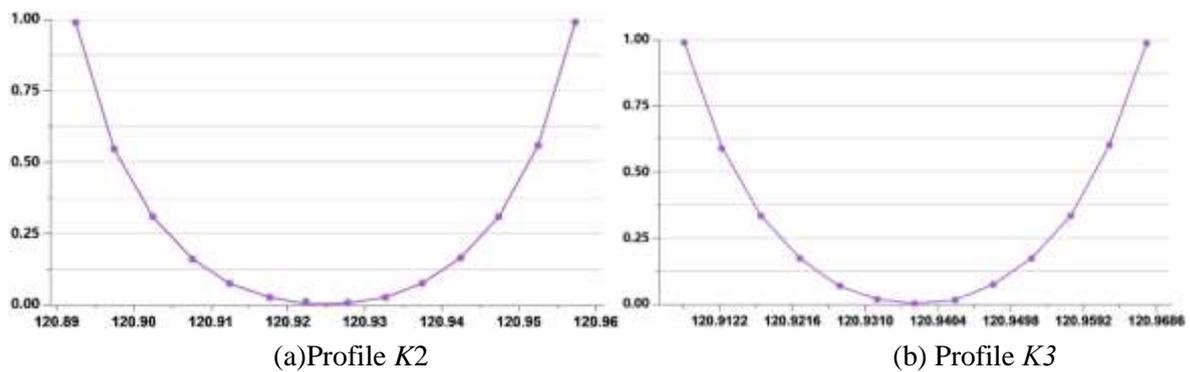


Figure 11 The normalised distribution of potential filed (hazard level)

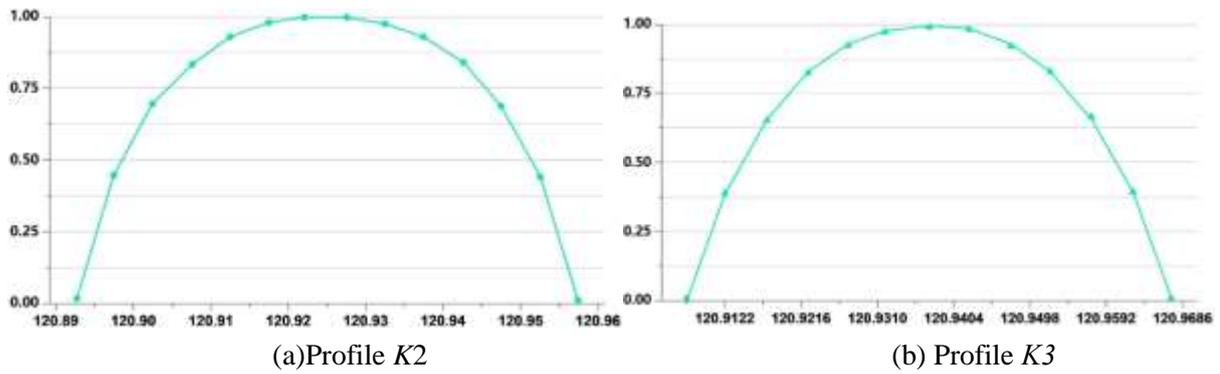


Figure 12 The normalised distribution of safety degree

Basing on the obtained distribution information of ships, a heatmap of the safety degree of the entire Yanda route can be obtained in Fig. 13.



Figure 13 The heat map of potential of whole Yanda route

### 4.3 The warning with this approach

Based on Section 4.2, a normalised safety degree or potential coursed by deviation can be obtained directly on any position in the Yanda route, even there is no ferry ship that has ever sailed on this position before. Subsequently, as elaborated previously, whether to trigger a deviation warning or not also relies on the corresponding speed, the course and other factors. It is worth noting that the course here stands for the included angle between the heading and the corresponding customary route. Based on the model proposed in Section 3.3, an initial DAG can be presented as the left side of Fig. 14.

At the very beginning, the correlations among the nodes of this figure are unknown; the corresponding conditional probability tables (CPTs) also remain to be determined. As discussed previously, Bayesian Network is a typical supervised learning method, which requires much labelled

data for training. In other words, warning samples and normal samples should be given and labelled in advance. In this occasion, historical data with warning records is required. However, it is difficult to obtain the real warning records in the daily management of the Yanda route. Therefore, to validate the proposed approach, the mentioned 117,115 records were labelled manually by experienced and competent operators who were invited from local administrations. In such records, 2,512 ones are considered as risky. Then, the input samples are equally divided into the training ones and the validation ones.

Moreover, the nodes or the attributes of a BN-based model should be discretized before training. Taking the speed as an example, the ferry's possible speed ranges from 0.0 to 20.0 knots, which means that there are 200 possible values. To make this model easy to be trained, such values should be discretized into several intervals in order to avoid problems associated with an unacceptably high amount of combinations of data for use. With the help of the method proposed in [22], the discretized nodes are shown as Table 3.

Table 3 Discretization of the attributes (nodes)

Attributes (Nodes)	Intervals of nodes					
	V1	V2		V3		V4
Speed (knots)	0~8.1	8.2 ~13.0		13.1~18.0		$\geq$ 18.1
Course (included angle with the customary route, degree)	C1	C2	C3	C4	C5	C6
	0~5	6~12	13~21	22~30	31~45	$\geq$ 45
Hazard level caused by Deviation from the customary routes	D1	D2	D3	D4	D5	
	$\geq$ 0.8	0.6~0.7	0.4~0.5	0.3~0.4	0~0.2	

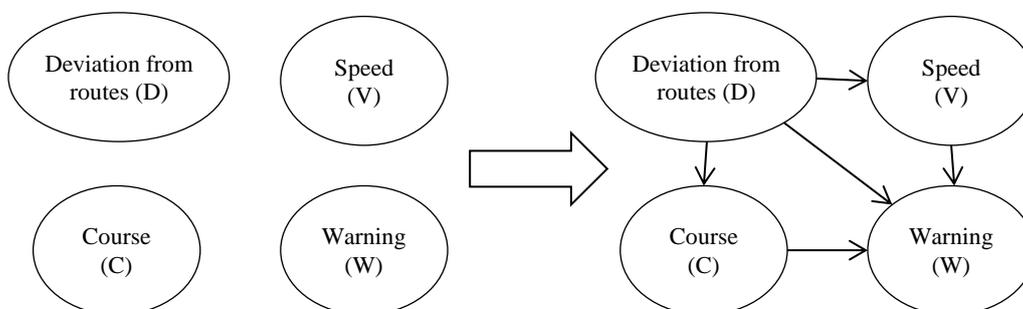


Figure 14 the DAG of warning

With the help of the K2 algorithm and the bnt toolbox of MATLAB 2018b, the formulated DAG is updated based on the training ones, and the updated DAG is presented on the right side of Fig.

The CPTs of the updated DAG can be estimated with the MLE method available in the bnt toolbox. The results are shown in Tables 5-7 in the appendix. Surprisingly, the nodes of D, V, C are not independent basing on the updated DAG. The deviation to the customary routes would take great influence on the speed and the course directly. This seems reasonable, considering that when a ferry deviated from its customary routes, it might change the course and speed to return its routes.

At last, the updated DAG that is shown on the right side of Fig. 14. Few examples based on this DAG are given in the appendix. Moreover, this DAG is tested on the validation samples, and compared with the manual labels. Eventually, the confusion matrix is presented in Table 4 when 50% or 0.50 is set as the threshold in triggering deviation warnings. It can be concluded that the proposed approach is capable of simulating the manual work efficiently.

Table 4 Confusion matrix

	Warning	No warning	Accuracy
Normal samples	511	55,534	99.01%
Warning samples	2,111	101	95.21%

The outstanding advantage of this approach is not only the efficiency, but also the generalizing ability. In this case study, such a route is about 300 nautical miles, and the training samples only include 58,557 records. With the help of this approach, practical artificial intelligence of giving deviation warnings to ferries can be formulated based on samples whose size is not unnecessarily large.

## 5 Conclusions

This research proposed an intelligent approach, which is capable of simulating the manual work, giving deviation warnings to ferry ships when sailing in their customary routes. The customary route can be learned from historical data automatically, meanwhile the deviation from the customary route can be quantified as a normalised deviation hazard level by the proposed APF model. Subsequently, a warning can be given by a Bayesian Network-based model in accordance with the deviation, speed, course and sufficient training. In the field testing, the proposed approach provided a promising result in replicating manual work. In the following research, the intelligent approach of integrating APF with BN should be improved further through incorporating Deep Learning and other unsupervised machine learning methods in order to make the whole process automatically.

## Acknowledgements

The first author is financed by the China Scholarship Council under Grant 201706950028. This research is partially supported by EU H2020 RISE 2016 RESET – 730888.

## Appendix

Table 5 The Conditional Probability Table of D

	D1	D2	D3	D4	D5
Hazard level caused by Deviation from the customary routes (D)	0.83	0.05	0.03	0.03	0.05

Table 6 The Conditional Probability Table of V

	D1				D2				D3				D4				D5			
	V1	V2	V3	V4																

Speed (V)	0.10	0.13	0.12	0.65	0.10	0.13	0.21	0.56	0.12	0.15	0.25	0.37	0.34	0.23	0.21	0.21	0.25	0.28	0.21	0.25
-----------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Table 7 The Conditional Probability Table of C

	D1				D2				D3				D4				D5			
	C1	C2	C3	C4																
Course (C)	0.87	0.04	0.04	0.05	0.80	0.14	0.02	0.04	0.56	0.34	0.04	0.06	0.34	0.44	0.14	0.08	0.45	0.14	0.28	0.13

Table 7-1 The Conditional Probability Table of W

	V1																			
	D1				D2				D3				D4				D5			
	C1	C2	C3	C4																
Warning (W1)	0.03	0.04	0.04	0.05	0.10	0.14	0.21	0.34	0.22	0.43	0.36	0.38	0.45	0.44	0.42	0.61	0.95	0.96	0.96	0.96
Warning (W2)	0.97	0.96	0.96	0.95	0.9	0.86	0.79	0.66	0.78	0.57	0.64	0.62	0.55	0.56	0.58	0.39	0.05	0.04	0.04	0.04

Table 7-2 The Conditional Probability Table of W

	V2																			
	D1				D2				D3				D4				D5			
	C1	C2	C3	C4																
Warning (W1)	0.03	0.04	0.04	0.05	0.10	0.14	0.21	0.34	0.22	0.43	0.36	0.38	0.45	0.44	0.42	0.61	0.96	0.96	0.96	0.96
Warning (W2)	0.97	0.96	0.96	0.95	0.9	0.86	0.79	0.66	0.78	0.57	0.64	0.62	0.55	0.56	0.58	0.39	0.04	0.04	0.04	0.04

Table 7-3 The Conditional Probability Table of W

	V3																			
	D1				D2				D3				D4				D5			
	C1	C2	C3	C4																
Warning (W1)	0.03	0.04	0.04	0.05	0.12	0.42	0.45	0.45	0.33	0.33	0.56	0.56	0.45	0.45	0.62	0.66	0.96	0.96	0.96	0.96
Warning (W2)	0.97	0.96	0.96	0.95	0.88	0.58	0.55	0.55	0.67	0.67	0.44	0.44	0.55	0.55	0.38	0.34	0.04	0.04	0.04	0.04

Table 7-4 The Conditional Probability Table of W

	V4																			
	D1				D2				D3				D4				D5			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
Warning (W1)	0.03	0.04	0.04	0.05	0.22	0.22	0.35	0.35	0.22	0.22	<b>0.36</b>	0.36	0.45	0.45	0.42	0.42	0.82	0.82	0.82	0.82
Warning (W2)	0.97	0.96	0.96	0.95	0.78	0.78	0.65	0.65	0.78	0.78	<b>0.64</b>	0.64	0.55	0.55	0.58	0.58	0.18	0.18	0.18	0.18

Based on the updated DAG shown on the right side of Fig. 14 and the CPTs, the deviation warning can be triggered automatically. An example is given here. A ferry is sailing near the customary route, whose hazard level evaluated by the APF model is 0.5, speed is 18.1 knots, the included angle between its heading and the customary route is 18 degrees. Based on Table 3, the corresponding status can be assigned to D3, V4, C3. In this occasion, the inference probabilities of W1

warning) is 0.36 based on Table 7-4. The vessel was considered to be safe by this approach, even it was sailing at a high speed.

## References

- [1] Revised guidelines for formal safety assessment (FSA) for the use in the rule-making process, MSC-MEPC.2/CIRC.12.[S]. London: MSC, 2013.
- [2] Lois P, Wang J, Wall A, *et al.* Formal safety assessment of cruise ships.[J]. *Tourism Management*, 2004, 25(1):93-109.
- [3] Wang J, Foinikis P. Formal safety assessment of containerships[J]. *Marine Policy*, 2001, 25(2):143-157.
- [4] Montewka J, Goerlandt F, Kujala P. On a systematic perspective on risk for formal safety assessment (FSA)[J]. *Reliability Engineering & System Safety*, 2014, 127:77-85.
- [5] Hu S, Fang Q, Xia H, *et al.* Formal safety assessment based on relative risks model in ship navigation[J]. *Reliability Engineering & System Safety*, 2017, 92(3):369-377.
- [6] Montes A, Akyildiz H, Yetkin M, *et al.* A FSA based fuzzy DEMATEL approach for risk assessment of cargo ships at coasts and open seas of Turkey[J]. *Safety science*, 2015, 79: 1-10.
- [7] Akyildiz H, Montes A. An integrated risk assessment based on uncertainty analysis for cargo Cessel safety[J]. *Safety science*, 2017, 92: 34-43.
- [8] Wang J. A subjective modelling tool applied to formal ship safety assessment[J]. *Ocean Engineering*, 2000, 27(10):1019-1035.
- [9] Montewka J, Ehlers S, Goerlandt F, *et al.* A framework for risk assessment for maritime transportation systems—A case study for open sea collisions involving RoPax vessels[J]. *Reliability Engineering & System Safety*, 2014, 124: 142-157.
- [10] Sutulo S, Moreira L, Soares C G. Mathematical models for ship path prediction in manoeuvring simulation systems[J]. *Ocean Engineering*, 2002, 29(1):1-19.
- [11] Sang L, Wall A, Mao Z, *et al.* A novel method for restoring the trajectory of the inland waterway ship by using AIS data[J]. *Ocean Engineering*, 2015, 110:183-194.
- [12] Sang L, Yan X, Wall A, *et al.* CPA calculation method based on AIS position prediction[J]. *The Journal of Navigation*, 2016, 69(6):1409-1426.
- [13] Gan S, Liang S, Li K, *et al.* Ship trajectory prediction for intelligent traffic management using clustering and ANN[C]. *Ukacc International Conference on Control*. IEEE, 2016.
- [14] Chen J, Lu F, Peng G. A quantitative approach for delineating principal fairways of ship passages through a strait[J]. *Ocean Engineering*, 2015,103:188-197.
- [15] Khatib O. Real-Time Obstacle Avoidance for Manipulators and Mobile Robots[J]. *International Journal of Robotics Research*, 1986, 5(5):500-505.
- [16] Park, M G, Jeon, J H, Lee, M C. Obstacle avoidance for mobile robots using artificial potential field approach with simulated annealing[C]. In *Proceedings of IEEE International Symposium on Industrial Electronics*, 2001, 3, 1530-1535.
- [17] Barraquand J, Latombe J C. Robot motion planning: a distributed representation approach. *International Journal of Robotics Research*, 1991, 10(6):628-649.
- [18] Wolf M T, Burdick J W. Artificial potential functions for highway driving with collision avoidance[C]. *IEEE International Conference on Robotics and Automation*. IEEE, 2008:3731-3736.
- [19] Rasekhipour Y, Khajepour A, Chen S K, *et al.* A Potential Field-Based Model Predictive Path-Planning Controller for Autonomous Road Vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(5):1255-1267.

- [20]Ma F, Chen Y W, Huang Z C, *et al.* A novel approach of collision assessment for coastal radar surveillance[J]. *Reliability Engineering & System Safety*, 2016, 155:179-195.
- [21] Ma F, Chen Y W, Yan X P, *et al.* A novel marine radar targets extraction approach based on sequential images and Bayesian Network[J]. *Ocean Engineering*, 2016, 120:64-77.
- [22]Rong H, Teixeira A, Soares C G. Evaluation of near-collisions in the Tagus River Estuary using a marine traffic simulation model[J]. *Zeszyty Naukowe/Akademia Morska w Szczecinie*, 2015,43 (115): 68-78.
- [23]Chen Y, Li T. Collision avoidance of unmanned ships based on artificial potential field[C]. *Chinese Automation Congress (CAC)*, 2017. IEEE, 2017: 4437-4440.
- [24]Singh Y, Sharma S, Sutton R, *et al.* Path Planning of an Autonomous Surface Vehicle based on Artificial Potential Fields in a Real Time Marine Environment[J]. 2017.
- [25]Trucco P, Vagno E, Ruggeri F, *et al.* A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation[J]. *Reliability Engineering & System Safety*, 2008, 93(6):845-856.
- [26]Monti, S. & Cooper, G. F. (1998). A multi-variate discretization method for learning Bayesian networks from mixed data. In *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, 404-413.
- [27]Cooper, G. F. & Herskovits, E. (1991). A Bayesian method for constructing Bayesian belief networks from databases. In *Proceedings of the Seventh conference on Uncertainty in Artificial Intelligence*, 199186-94.
- [29]Bilmes, J. A. (1998). A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute*, 4(510), 126.