Ma, F, Chen, Y-W, Huang, Z-C, Yan, X-P and Wang, J

A novel approach of collision assessment for coastal radar surveillance

A Novel Approach of Collision Assessment for Coastal Radar Surveillance

Feng Ma[1, 2, 4], Yu-wang Chen[2], Zi-chao Huang[1, 3], Xin-ping Yan[1, 3*](xinping_yan@126.com), Jin Wang[4]
Intelligent Transport System Research Center, Wuhan University of Technology, P.R. China[1]
Decision and Cognitive Sciences Research Centre, the University of Manchester, Manchester, M15 6PB, UK[2]
National Engineering Research Centre of Water Transportation Safety (WTS), P. R. China[3]
Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, L3 3AF, UK[4]

Abstract: For coastal radar surveillance, this paper proposes a data-driven approach to estimate a blip's collision probability preliminarily based on two factors: the probability of it being a moving vessel and the collision potential of its position. The first factor can be determined by a Directed Acyclic Graph (DAG), whose nodes represent the blip's characteristics, including the velocity, direction and size. Additionally, the structure and conditional probability tables of the DAG can be learned from verified samples. Subsequently, the obstacles in a waterway can be described as collision potential fields using an Artificial Potential Field model, and the corresponding coefficients can be trained in accordance with the historical vessel distribution. Then, the other factor, the positional collision potential of any position is obtained through overlapping all the collision potential fields. For simplicity, moving speeds of obstacles are considered in this research. Eventually, the two factors are characterised as two pieces of evidence, and the collision probability of a blip is estimated by combining them with Dempster's rule. Through ranking blips on collision probabilities, those that pose high threat to safety can be picked up in advance to remind supervisors. Particularly, a good agreement between the proposed approach and the manual work was found in a preliminary test.

Keywords: Collision Probability; Bayesian Network; Artificial Potential Field; Marine Radar; Nonlinear Optimisation; Dempster's rule

Highlights:

[1] Novel estimation approach of collision probability for radar blips.

[2] Novel method to evaluate the authenticity of a blip using Bayesian Network.

[3] Novel method to evaluate the positional collision potentials using the APF model.

[4] Novel method to obtain the coefficients of potential fields with historical data.

1 Introduction

Marine radar is an active detection tool of coastal surveillance, which does not require replies from supervised vessels. As well as that, it is capable of detecting waterfronts, buoys, and other obstacles. Through marine radar, all the vessels and obstacles are represented as blips on screen with corresponding characteristics, including shapes, velocities, directions and trajectories. In daily managements, these characteristics are used for target extraction and identification. Presently, several other maritime tracking systems have been invented, including the Automatic Identification System (AIS) and maritime satellites. However, the reporting frequency of AIS is too low for real-time tracking (Lin *et al.*, 2007); not many vessels possess satellite transmitters. Therefore, marine radar is still the kernel of a maritime detecting system.

In fact, a considerable proportion of radar blips or objects are caused by noises or stationary objects. In inland waterways or ports, such false or stationary objects are even more than real moving vessels (Ma *et al.*, 2015b). Therefore, supervisors have to identify moving vessels from a plethora of blips manually. However, even if a blip is confirmed to be a real moving vessel, it might not need much attention. For instance, a vessel that is far away from piers, rocks, obstacles, and other vessels is usually safe; in daily management, it does not need much attention. In fact, only a blip that is probably a real moving vessel and is posing a threat to safety needs close inspection (Lin *et al.*, 2007). Particularly, the threat to safety here generally means a potential collision, as the collision avoidance is the main objective of radar surveillance.

Most of radar systems have integrated an Automatic Radar Plotting Aid (ARPA) function to track moving objects. However, the authenticities or collision potentials of targets cannot be obtained by an ARPA function directly. For instance, a late-model coastal surveillance radar system is capable of tracking a 0.5 m^2 target at a distance of 5 miles. However, its ARPA function is not capable of determining whether this 0.5 m^2 target is a real moving vessel, or just a trivial object floating on the water. Presently, the authenticity or collision probability of a target can only be inferred by experienced supervisors. Such manual work might be impractical when there are too many objects in observation. For instance, there are about 20,000 vessels passing through Nantong waterway, Yangtze River, China in one day. Obviously, it is impossible to inspect them one-by-one manually. On the basis of the procedures of manual work, this research aims to develop a data-driven method that helps supervisors identify targets preliminarily so as to enhance their supervision and management efficiency.

It is worth emphasizing that the collision probability in radar surveillance is different from the usual sense. In conventional research, a collision probability is determined by the speed, rotation rate, course, encountered vessels, and environmental factors (Fujii *et al.*, 1974). However, the course and speed measured by radar are not completely credible (IEC 2013; 2014). False alarms might be triggered easily when using them in collision estimation (Ma *et al.*, 2015a). Nevertheless, the positions of targets obtained from radar are comparatively reliable. Therefore, supervisors always take the position as an important factor in the estimation of a blip's collision probability. For example, when a blip or object is located in a dangerous zone, it should attract much attention without regard to whether it is a noise or not. On contrary, if an object is located in open water outside the main channel, which poses limited threat to safety, it might be ignored by supervisors. Particularly, the collision potential of a position is actually determined by surrounding obstacles and environments, including waterfronts, berths, water depths, piers, buoys, shoals and encountered vessels. Apparently, these factors are varying all the time. As a result, to estimate the collision potentials of different positions requires supervisors' experience.

Overall, referring to manual work, there are two major underlying factors in the preliminary identification of a blip that has a high collision probability. The first one is the probability of the blip being a real moving vessel; the other is the corresponding collision potential of its position.

The first factor can be inferred from its characteristics. For instance, a blip that is moving at a usual velocity is likely to be a moving vessel. This inference process is based on the speed of the blip and the experience of the operators. In fact, such experience can be considered as prior information accumulated from a long-time observation. In this light, a probabilistic model might be appropriate in this research (Ranganathan *et al.*, 2004). Among different types of probabilistic models, Bayesian Network (BN) is considered to be efficient and rigorous. Particularly, it is capable of learning structures and the associated coefficients with verified samples under uncertainties (Zhang *et al.*, 2013).

The other factor, or the collision potential of a position, is more complicated. Generally, the term “collision risk” discussed in maritime research is usually considered as the product of a collision probability and the impact of the collision (Williams, 1996). However, the impact involves much detailed information of vessels (Fujii *et al.*, 1974), such as the rudder angle, types of cargo, and the number of people on board the ship. This information is difficult to obtain for radar surveillance. In fact, the primary objective of supervisors in VTS is to avoid all the possible collisions without regarding or weighing the collision consequences. Hence, only the collision probability is investigated in this research.

In relevant research findings, the estimation of the collision probability is generally based on macro perspectives or ship handling. These macro perspectives include waterway design, port engineering and policy-making (Eleye-Datubo *et al.*, 2008). The relevant methods are not capable of describing the successive variation of collision probabilities in microscopic adjacent positions (Dong and Frangopol, 2015). For instance, these methods can be used to estimate the overall collision probability of a bridge zone for setting a speed limit; however, they are not capable of describing the collision probability differences between two points that are 50 meters apart from each other in the bridge zone. In radar surveillance, such a microscopic estimation is essential. Another conventional research perspective of studying the collision probability is for ship handling, which also requires much manoeuvring information of the vessels (Montewka *et al.*, 2012). As described, such information is mostly unknowable for radar surveillance. Therefore, the conventional collision probability estimation methods might not be very suitable for the perspective discussed in this research.

Referring to the research conducted in the robot area, the problem can be addressed with an Artificial Potential Field (APF) model, which does not need detailed information of obstacles, and describes the collision probabilities as a continuous function (Volpe and Khosla, 1990). For decades, the APF model has been widely used in robot route planning and manipulation, and it is believed to be efficient and concise.

In summary, this paper aims to propose an intelligent approach to estimate the collision probabilities of radar blips preliminarily using BN and the APF model. It is organised as follows.

Section 2 dedicates to introducing the characteristics of blips and conventional research of collision probability. Section 3 proposes a novel approach to estimate the collision probabilities of blips. In Section 4, a case study is conducted. Section 5 concludes this paper.

2 Literature review

2.1 The uncertainties of marine radar blips

By detecting echo signals which bounce off the surroundings, the coastal surveillance radar can be used to determine the distance, speed, and direction of each moving object in a specific area. The echo signals can be represented as frequency spectrums or blips on a screen. Generally, the blip form is more accessible, which is shown as a radar image. The satellite image and the grey-scale radar image shown in Figure 1 were captured at the same location and surroundings of Yangtze River, Wuhan, China. In the radar image, waterfronts, vessels, buoys, and bridges have been represented as blips at the very beginning of target extraction. The speed, course, and position of targets can be quantified in accordance with the inter-frame differences of corresponding blips. However, radar images or blips are actually not stable. The graphs of blips will be affected by the observation angle and radar resolution notably. Moreover, blips often overlap and connect to each other. Therefore, the direction and speed measured by radar blips are not completely credible (IEC 2013, 2014). In practice, stationary or noise blips might drift like moving vessels; moving vessels approaching to berths might move too slowly, and they look like stationary or noise objects. It is worth noting that each object's speed can be measured with the Doppler velocities too. However, most marine radar systems work on a low Repetition Pulse Frequency (RPF) mode, and the Doppler velocities are ambiguous. Hence, the radar images are used as the major evidence for further identification.

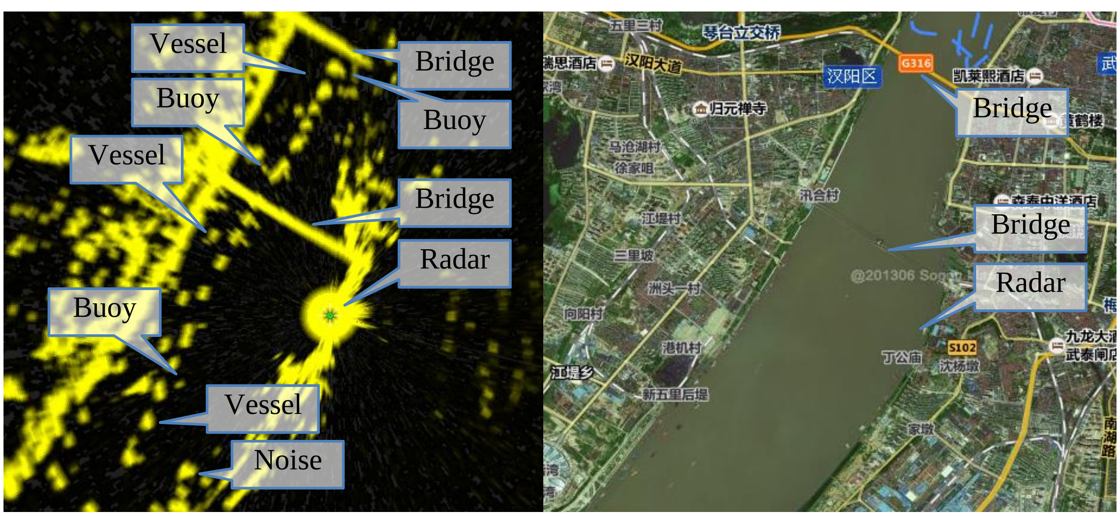

Figure 1 Radar and satellite images of Yangtze River, Wuhan, China

To address the problem of uncertainties described above, radar performance appraisals and improvements have attracted much attention in recent decades (Li *et al.*, 2007; Islam *et al.*, 2012) Many researchers were dedicated to developing a generic filtering algorithm to obtain more accurate trajectories of radar objects (Yoo and Kim, 2008). However, it may be argued that all these filtering algorithms incorporate some assumptions regarding objects' states, which are only applicable in specific conditions.

It is shown in Figure 1 that the marine radar also captured many useless and noise blips, and operators might take them for moving vessels easily. Hence, some intelligent methods have been introduced to distinguish moving vessels from false or stationary objects. For marine radar, Ma *et al.* (2015) proposed a fuzzy k-means (FCM) based classification method to identify the false targets among ARPA targets, and reported the accuracy of 91.0%. Zhou *et al.* (2013) invented a radar target-recognition method based on fuzzy optimal transformation using high-resolution range profiles. Although the existing algorithms are shown to be effective for specific case studies in radar research, they do not constitute a rigorous probabilistic inference process, nor are they proven to be effective in principle or in general. As such, they are of an *ad hoc* nature and might not be as robust as required for real life applications or implementation. In addition to the identification of a blip, operators of radar also need to know the exact probabilities about the blip's states for making appropriate decisions.

2.2 Estimation of collision probabilities

In addition to the authenticity of a blip, its position in a waterway is the other important factor for estimating the corresponding collision probability. In Figure 1, the collision probabilities of the blips near bridges or other channel constructions are obviously higher than the others. To model this phenomenon, the collision probability differences of adjacent positions should be described appropriately.

In fact, the collision probability of a vessel is affected by many factors, including weather, navigators, ship handling, ship condition, encountered vessels and others. Hence, collision probabilities can be modelled from different perspectives (Hänninen and Kujala, 2009) as described in Section 1. The static collision probability model proposed by Fujii *et al.* (1974) is widely used. In such a model, a collision probability is equal to the product of the geometrical probability of a collision course and the causation probability. Obviously, this model is closely related to ship handling. For example, Montewka *et al.* (2010) proposed a new approach for quantifying the geometrical probability to estimate collision probabilities on the basis of maritime and aviation experience. Pedersen (2010) presented a paper to review procedures for reducing the high economic environmental and human costs associated with ship collisions and grounding.

It is worth emphasizing that researchers become increasingly interested in modelling the characteristics of passing vessels with AIS data records since such records are widely believed to be both reliable and objective (Montewka *et al.*, 2010). This research also introduces the AIS records as a fundamental data source in the following discussion.

In summary, the research of collision probability generally starts with a multi-factor qualitative analysis involving ship handling, human factors, and geometrical collision model which are originated from ship domains or minimum distance modelling (Montewka *et al.*, 2012). However, this information is unavailable for coastal radar surveillance, which can only be confirmed with very high frequency (VHF) radio. In daily management, the verification of VHF radio is conducted only when needed; hence, the location of a blip seems to be the only direct and credible evidence for estimating the corresponding collision probability or potential, which is closely related to the dynamic navigation environments of waterways. Any change of berths, piers, buoys and depths might have significant impacts on the distribution of collision probability. Although many researchers have dedicated to proposing methods to model collision risk based on these factors individually (Kujala *et al.*, 2009; Qu *et al.*, 2011), a widely acknowledged and comprehensive modelling method has not been invented yet.

It is worth noting that the collision probability here is not obtained from the frequency analysis of a random process since collision accidents might not happen often actually. Hence, the research on collision probability estimation is usually started with a qualitative analysis of incidents causation (Dong and Frangopol, 2015). It is not illogical to investigate the collision probability in radar surveillance in a potential field. The potential theory might be applicable in this research (Dellacherie and Meyer, 2011).

2.3 Obstacle avoidance modelling with the APF model

The potential theory is originated from mathematical physics. Nowadays, it is also intimately connected with probabilities and the theory of Markov chains (Dellacherie and Meyer, 2011). In many cases, neighbouring objects might attract or repulse each other. The so-called repulsions or attractions among them are actually very difficult to be described or quantified, whilst the distance is the core factor in the attenuation of these forces. By this moment, the potential theory is considered to be attractive for use (Statheros *et al.*, 2008).

In a waterway, a collision probability or a collision potential can be considered as a special "repulsion", which objectively repulses the corresponding vessels away to avoid collision. The closer to obstacles the vessel is, the higher collision potential there should be. The strength of "repulsion" is exactly consistent with the collision potential. When there are in-sufficient records of collision accidents, a collision potential might be quantified by the "repulsions". For instance, it is widely believed that narrow channels between the piers of a bridge are dangerous for passing vessels, or the corresponding collision potentials are high although the accidents that vessels collide with piers are rare. There are very strict regulations for the operators of vessels when crossing piers, including speed limit, no overtaking. These regulations reduce collision accidents objectively. As a result, a collision probability or a collision potential cannot be estimated with a frequency analysis. However, the high collision probabilities or potentials are objective existence, which are changing the behaviours of vessels, making them as far as possible away from the piers. It is not illogical to take the collision potential as "repulsions" that repulse these vessels away from the piers. In the potential theory, those "repulsions" are caused by the corresponding so-called "repulsive potential fields", which are exactly produced by the piers (Volpe and Khosla, 1990).

The phenomenon discussed above is illustrated in Figure 2. In this figure, there are several piers in a waterway. Hundreds of vessels crossed these piers, and vessels' tracks are represented with blue circles and lines. In daily management, these historical records of vessel tracks can be obtained from an AIS database easily. Particularly, these tracks indicate that vessels were obviously willing to take routes which were far away from these piers to lower their collision potentials. On the other hand, such a phenomenon can be regarded as that these vessels were pushed into a narrow channel by some undetectable "repulsions". As shown in Figure 2, these "repulsions" are represented as red arrows. Apparently, the closer to the piers, the greater of the repulsions there would be; the distance is the core factor in the attenuation of the repulsions. As mentioned, the strength of the "repulsions" is consistent with the corresponding collision potential. By analysing the distribution of passing vessels, the corresponding repulsions or repulsive potentials can be quantified. Therefore, the collision potential or probability of a position can be obtained indirectly.

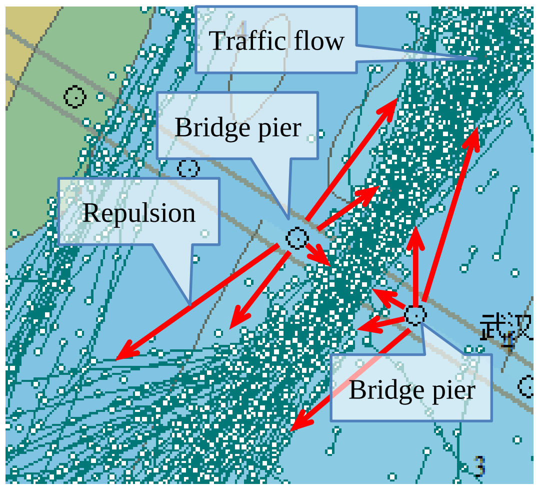

Figure 2 Traffic flow between piers

To describe the ship collision potential as a "force" was firstly proposed by Statheros *et al.* (2008). They used a Virtual Field Force (VFF) to describe the collision potential for collision avoidance in the unmanned surface vessel (USV) research. In fact, similar approaches are common in robot research, and the most frequently used methodology is the Artificial Potential Field (APF). The APF model was invented by Khatib (1986), which was designed for the real-time obstacle avoidance of manipulators and mobile robots (Park *et al.*, 2001). With this model, movements of the robot are governed by potential fields, which are usually composed of two components, attractive potential and repulsive potential fields. An attractive potential field is generally a bowl shape to draw the robot towards the goal. A repulsive potential filed is generally built at the location of an obstacle to push the robot away. As described in Section 1, the collision potentials can be modelled as continuous functions using the APF model. Therefore, the collision potential differences of adjacent positions can be described as the change of the values of these functions.

However, the formulations of the potential fields are different, which are determined by the corresponding scenarios and requirements. In general, several potential functions are frequently used, which are mostly in quadratic and conical forms (Park *et al.*, 2001). The following issue is to determine which potential function is appropriate for modelling collision potential in a waterway. In practice, the shape of the repulsive potential field is very important, and it should be compatible with the influences of corresponding obstacles. In addition, the influence range of the potential field should conform to reality. Hence, the coefficients of the chosen potential function should be assigned very carefully.

Presently, many researchers put much effort to address the problems of local minima and the modelling for arbitrarily shaped obstacles. Research findings that aim to obtain appropriate coefficients of potential field are very limited. Zhang *et al.* (2012) developed an evolved APF method by genetic algorithm, which uses a grid method to generate an obstacle avoidance path to address the local minimum problem. Montiel *et al.* (2015) used a bacterial evolutionary algorithm to address the same issue. Pêtrès *et al.* (2012) proposed an APF-based reactive navigation approach for vessels. In their approach, environment and local constraints are represented as potential fields around the vessels. Moreover, potential fields caused by wind directions and surrounding obstacles will be updated periodically, ensuring an optimal heading for the navigation.

Overall, the APF model is an efficient method for modelling collision potentials in waterway transportation. The problem is how to obtain the appropriate coefficients of potential fields. As described, the distribution of passing vessels might be a good indication (Ma *et al.*, 2015b).

3 A proposed approach

To reduce the burden of VTS supervisors, this research proposes an approach to identify targets that have high collision probabilities from a plethora of radar blips preliminarily. Particularly, this approach consists of two novel methods. The first one is used to estimate the probability of a blip being a true moving vessel using BN. The other novel method is then used to estimate the collision potentials of adjacent positions within the collision potential fields. Eventually, the collision probability of each blip can be considered as the aggregation of authenticity and the corresponding collision potential of its position. For simplicity, only the static obstacles are considered in this research.

3.1 Step 1: The inference process of blips' authenticities using BN

As described, only a small proportion of blips are real moving vessels. In daily management, operators distinguish them from others in accordance with several graphic characteristics, including velocity, course, size, colour, width, and length. Obviously, these factors may be dependent on each other. Therefore, BN is chosen as the basis to establish an identification process whose advantage is that dependencies among all the factors can be modelled appropriately (Zhang *et al.*, 2013). Referring to manual work, three types of evidence are selected in this research: the velocity, motion direction, and blip size, which are presented in Figure 3.

According to the ARPA function requirement IEC 62388 (IEC, 2013; 2014), supervisors are generally able to identify the authenticity of a blip in 30 seconds or 10 continuous frames. Therefore, the velocity and direction characteristics are quantified based on the analysis of 10 frames. In this research, the velocity and motion course are quantified as shown in Figure 3. In the figure, the velocity is equal to the number of units (pixels) that the blip has moved in 10 frames, which is illustrated in sub-figure 3(a). The direction is quantified as the angle between the true north and the motion direction, which is illustrated in sub-figure 3(b). It is worth mentioning that the motion direction values are rounded down to integers.

Generally, for a moving vessel, the size of the corresponding blip varies in an appropriate range, which is illustrated in Figures 1 and 3. The size can be considered as how many pixels the corresponding blip is occupying in a radar image after binarization, which is illustrated in sub-figure 3(c).

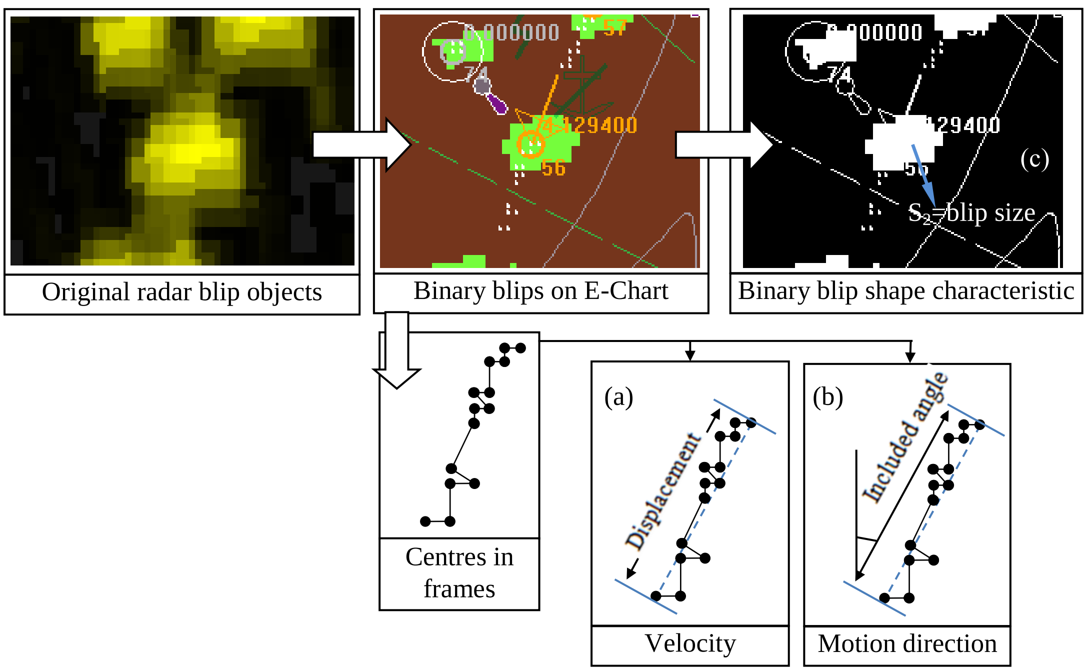

Figure 3 Blip characteristics in frames of radar

Based on the quantified characteristics, the BN-based inference process is conducted as follows: BN is defined by a pair (S, Θ_s), where $S = (\chi, E)$ is a directed acyclic graph (DAG) with a set of nodes χ, and with a set of arcs or nodes $E = \{(X_i, X_j) \mid X_i, X_j \in \chi, X_i \neq X_j\}$ representing probabilistic dependencies among domain variables (Monti and Cooper, 1998). Θ_s represents the parameterization of a probability measure ρ defined over the space of possible instantiations of χ. 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 BNs 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:

$$\rho(\vec{\boldsymbol{x}}) = \rho(x_1, \cdots, x_n) = \prod_{i=1}^{n} \rho(x_i | \boldsymbol{Pai}, \Theta_S) \tag{1}$$

Manual work is capable of identifying the authenticity (A) of a blip with three attributes direction (V), velocity (D), and size (S). Hence, (A), (V), (D), and (S) form a DAG. Subsequently, the structure of the DAG can be learned from verified data samples. Presently, the K2 scoring algorithm is widely accepted for constructing BN from databases or records, proposed by Cooper and Herskovits (1991).

The principle of the K2 scoring algorithm is to assess the appropriateness of a structure based on verified records. Under assumptions associated with lack of missing values and independent coefficients, the K2 scoring algorithm can be further simplified (De Campos and Castellano, 2006). Subsequently, the best scoring structure can be found with a hill-climbing heuristic algorithm. More detailed information about the K2 scoring algorithm can be found in the reference (Cooper and Herskovits, 1991). Presently, the K2 scoring algorithm is fully supported by the software tools of BN, including Netica, Hugin, and the MATLAB bnt toolbox.

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 training data. In this research, the expectation maximization (EM) algorithm is adopted, which is an iterative method to carry out a MLE (Bilmes, 1998). 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 blip being a real moving vessel can be estimated with the new DAG.

3.2 Step 2: The modelling of collision potential field using the APF model

In addition to the authenticity, the collision potential of the position of the studied blip is the other important factor in the estimation of collision probability. The APF model is adopted to describe the collision potential as discussed in Section 2.3. There are many types of APF function, and the Yukawa function is widely used in collision avoidance potential modelling (Volpe and Khosla, 1990), which is presented as,

$$U_{obs,m}(K) = A\frac{e^{-\alpha K}}{K} \tag{2}$$

where $U_{obs,m}$ denotes the avoidance or collision potential value to the m^{th} obstacle. A is a constant, and denotes a maxim value of (collision or avoidance) potential. α is also a constant, and denotes the rate of decay, which is determined by the boundaries of APF. Variable K denotes the pseudo-distance to the m^{th} obstacle, which is different from the actual distance. It is required to take the characteristics of obstacles and the environmental factors into consideration to propose an appropriate formulation of variable K (Volpe and Khosla, 1990), especially in a waterway. Hence, the formulations of variable K for the corresponding obstacles are different, including buoys, piers, rocks, shoals and encountered vessels. For simplicity, only two typical static obstacles (i.e. buoys and piers) are considered.

In Yangtze River, a buoy is generally 1~9 meters long, and a vessel is generally more than 80 meters long. Therefore, a buoy can be considered as a point to a passing vessel. By this moment, an eclipse model or a point model is appropriate, which is defined as follows. Suppose a buoy is located at (x_b, y_b); the pseudo-distance K of the coordinate (x, y) to this buoy is presented as (Volpe and Khosla, 1990),

$$K_b = \sigma_b \cdot \sqrt{(\xi(x - x_b))^2 + (y - y_b)^2} \tag{3}$$

where σ_b denotes a range adjustment coefficient for pseudo-distance. $\xi \in (0,1]$ denotes the ratio between the values of the X and Y axes. Substituting Eq. (3) into Eq. (2), $K = K_b$, the eclipse collision potential equipotential lines are presented in Figure 4, and their centres represent the coordinate of the buoy. In inland rivers, vessels generally sail along the river direction. Therefore, the X axis is set to be parallel to the river direction in this research.

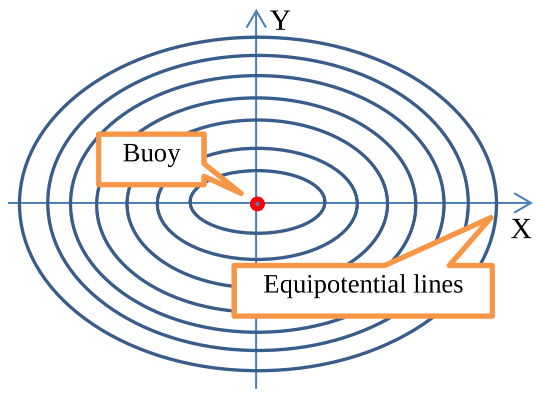

Figure 4 A buoy (ellipse) repulsive APF with equipotential lines

Different from a buoy, the pier of the Wuhan Yangtze River Bridge is 60 meters long and 5 meters wide. Hence, the shape and dimensions of a bridge pier should not be neglected, and it is not appropriate to take it as a point. Therefore, a rectangle model is adopted. Its pseudo-distance K is presented as follows (Volpe and Khosla, 1990),

$$K_p = \sigma_p \cdot \min(\sqrt{(x - x_p')^2 + (y - y_p')^2}) \tag{4}$$

where $(x_p', y_p') = \begin{cases} |x_p' - x_p| < l, y_p' = y_p \pm w \\ |y_p' - y_p| < w, x_p' = x_p \pm l \end{cases}$, (x_p, y_p) denotes the centre of the pier, l denotes the length of the pier in the X axis, w denotes the length of the pier in the Y axis. σ_p is an adjustment coefficient of the bridge pier pseudo-distance. Substituting Eq. (4) into Eq. (2), $K = K_p$, the rectangle equipotential lines are presented in Figure 5. It is worth mentioning that the *potential edge rectangle* in the centre represents the maximum value of collision potentials. Generally, the *potential edge rectangle* is larger than the actual geometrical dimensions of the corresponding pier. The reason lies in that operators should keep their vessels away from the piers at a considerable distance to ensure safety (Fujii *et al.*, 1974). In this figure, the dimensions of the pier are marked as a red dotted rectangle in the centre. The X axis here is also set to be parallel to the river direction in this research.

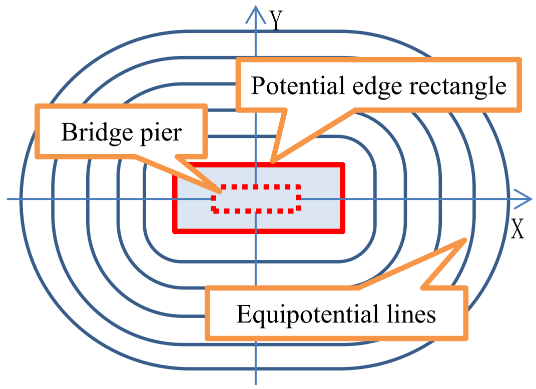

Figure 5 A pier (rectangle) repulsive potential field with equipotential lines

Using the methods and models discussed above, all the piers and buoys can be modelled as sources of collision (repulsive) potential fields, which pose threats to passing vessels. Moreover, in any place of the waterway, the corresponding collision potential can be considered as the combination of the different collision potential fields, which can be obtained with Eqs. (2), (3) and (4).

3.3 Step 3: A nonlinear optimisation of the coefficients of potential fields

The prominent problem of the proposed avoidance or collision potential model is that all the coefficients are unknown. In former research, these coefficients are generally assigned based on experience or some assumptions (Bing *et al.*, 2011). This research aims to propose a novel method to address this problem with any available data.

As described previously, the distribution of collision potential can be inferred based on the behaviours of a large amount of passing vessels, since vessels always take the routes that pose low threat to their safety. The lower collision potential is, the more vessels there should be. The IMO (International Maritime Organization) requires every single vessel to be equipped with an AIS terminal for remote monitoring. As described in Section 1, the AIS is not very suitable for real-time tracking since its reporting frequency is too low. Nevertheless, the positions from AIS are credible, which are obtained from a GPS sensor. Therefore, it is possible to find out the characteristics of vessel distribution in a waterway accurately based on sufficient AIS records.

In a relatively close or isolated scenario, when the collision avoidance is the major concern for ship manoeuvring and the obstacles are known and relatively stationary, the appropriate coefficients of collision potential fields should make the distribution of collision potentials consistent with the distribution of passing vessels in AIS records. In this light, the coefficients can be obtained in a nonlinear constraint optimisation model as follows.

Suppose there are m piers and n buoys in a relatively close area of a waterway. The coordinates of the piers are $\{(x_1^p, y_1^p), \cdots, (x_m^p, y_m^p)\}$, and the coordinates of the buoys are $\{(x_1^b, y_1^b), \cdots, (x_n^b, y_n^b)\}$. Based on the formulations of Section 3.2, the combined collision potential of the position (x, y) caused by these piers and buoys is presented as,

$$P(x,y,\overrightarrow{para})=\sum_{i=1}^{m}A_p\frac{\exp(-a_p\cdot K_p(x,y,x_i^p,y_i^p,\xi,\sigma_p))}{K_p(x,y,x_i^p,y_i^p,\xi,\sigma_p)}+\sum_{i=1}^{n}A_b\frac{\exp(-a_b\cdot K_b(x,y,x_i^b,y_i^b,\xi,\sigma_b))}{K_b(x,y,x_i^b,y_i^b,\xi,\sigma_b)} \quad (5)$$

where $\overrightarrow{para}=\{a_p,\xi,\sigma_p,a_b,w,l,\sigma_b\}$ denotes all the undetermined coefficients of Eqs. (2), (3) and (4); the functions $K_p(\cdot)$ and $K_b(\cdot)$ are used to calculate the pseudo-distances to buoys and piers, and their formulations are given in Eqs. (3) and (4). It is worth mentioning that A_p and A_b denote the maximum values of the collision potentials caused by a pier and a buoy. For simplicity, they can be considered as equal, or $A_p=A_b=1$.

As discussed, the collision potential distribution caused by obstacles should conform to the real distribution of passing vessels. Suppose a cross profile of a major channel contains *L* discrete statistical points or sections $\{(x_1,y_1),\cdots,(x_L,y_L)\}$. There is a point (x_k,y_k) on this cross profile, 1≤ *k*≤*L*. Its corresponding collision potential is presented as $P(x_k,y_k,\overrightarrow{para})$, given by Eq. (5). Therefore, the collision potentials for all the *L* points of this profile can be presented as $\{P(x_1,y_1,\overrightarrow{para}),\cdots,P(x_L,y_L,\overrightarrow{para})\}$, and the maximum and minimum collision potentials of the *L* points are presented as $P_{\max}=\max[P(x_1,y_1,\overrightarrow{para}),\cdots,P(x_L,y_L,\overrightarrow{para})]$, $P_{\min}=\min[P(x_1,y_1,\overrightarrow{para}),\cdots,P(x_L,y_L,\overrightarrow{para})]$. Therefore, the normalised collision potential of the point (x_k,y_k) is presented as,

$$P_{normal}(x_k,y_k)=[P(x_k,y_k,\overrightarrow{para})-P_{\min}]/(P_{\max}-P_{\min}) \quad (6)$$

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

$$\overrightarrow{P^*}=\{1-P_{normal}(x_1,y_1),\cdots,1-P_{normal}(x_L,y_L)\} \quad (7)$$

Suppose the distribution (densities) of the passing vessels of the *L* points is denoted as a vector $\vec{d}=\{d_1,\cdots d_L\}$, and the maximum and minimum passing vessel numbers of the *L* points are presented as $d_{\max}=\max(d_1,\cdots,d_L)$, $d_{\min}=\min(d_1,\cdots,d_L)$. Hence, the normalised distribution of vessels on the *L* points is presented as,

$$\overrightarrow{d^*}=\{(d_1-d_{\min})/(d_{\max}-d_{\min}),\cdots,(d_L-d_{\min})/(d_{\max}-d_{\min})\} \quad (8)$$

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

$$\begin{aligned}&\overrightarrow{para}=\{a_p,\xi,\sigma_p,a_b,w,l,\sigma_b\}=\\&\arg\min_{feasiable}\sum_{i=1}^{L}\left|[1-P_{normal}(x_i,y_i)]-(d_i-d_{\min})/(d_{\max}-d_{\min})\right|\end{aligned} \quad (9)$$

Since Eq. (9) is continuously differentiable, the gradient function of Eq. (9) can be obtained easily. Therefore, the appropriate $\overrightarrow{para}$ can be obtained with the '*fmincon*' function of MATLAB (Liu *et al.*, 2003). Then the collision potential of each point in a waterway can be obtained as the combination of all the collision potential fields, given by Eq. (5).

3.4 Step 4: The combination of the two factors using Dempster's rule

A blip's probability of being a real moving vessel and the collision potential of its position can be obtained with Steps 3 and 4. The next issue is to estimate the collision probability based on these two factors, which can be considered as two pieces of evidence. Particularly, they are based on the AIS and radar blips obtained in the same location. Hence, they are not independent in a strict sense. However, it is difficult to quantify their dependencies. Considering the contribution in the risk recognition of manual operation, the two pieces of evidence can be regarded as being approximately independent of each other for simplicity. Hence, in this research, Dempster's rule is applicable in the evidence combination (Li and Pang, 2013), which is given below. In the future research, methods such a Belief Rule Base (BRB) approach may be introduced to address this problem in more detail.

Suppose $\Theta = \{\theta_0, \theta_1\}$ is a set of mutually exclusive and collectively exhaustive propositions for the collision probability estimation of a blip. θ_0 is the *Collision* state, denoting a situation that the corresponding blip will collide with an obstacle; θ_1 is the *Non-collision* state, denoting a situation that the corresponding blip will not collide with any obstacle. Let Ø represent the empty set. In practice, the Unknown state θ_2 can be represented by the frame of discernment Θ itself, and it means the state that is neither θ_0 nor θ_1. Thus, the power set of Θ consists of 4 subsets of Θ, and is denoted by 2^Θ or $P(\Theta)$, as follows:

$$P(\Theta) = \{\emptyset, \theta_0, \theta_1, \theta_2\} \tag{10}$$

A *Basic Probability Assignment* (*bpa*) is a function *p*: $2^\Theta \to [0,1]$ that satisfies,

$$p(\emptyset) = 0, \sum_{\theta \subseteq \Theta} p(\theta) = 1 \tag{11}$$

where the basic probability $p(\theta)$ is assigned exactly to a proposition θ and not to any smaller subset of θ. Then, the two factors discussed previously can be transformed to two pieces of evidence as follows.

There is a blip located at the position (x_k, y_k), and its probability of being a real moving vessel is estimated as *p* based on Section 3.1. Apparently, only a real moving vessel might collide with an obstacle. Hence, based on the authenticity of the blip only, the basic probabilities about the $\theta_0, \theta_1, \theta_2$ states can be obtained as follows, or a piece of evidence can be constructed,

$$e_1: \{p(\theta_0), p(\theta_1), p(\theta_2)\} = \{p, (1-p), 0\} \tag{12}$$

In the area under investigation, there are M individual points $\{(x_1, y_1), (x_2, y_2), \cdots, (x_M, y_M)\}$. Their collision potentials to obstacles are presented as $\{P(x_1, y_1, \overrightarrow{para}), \cdots, P(x_M, y_M, \overrightarrow{para})\}$ based on Eq. (5), where $\overrightarrow{para}$ is obtained with the method presented in Section 3.3. The maximum collision potential is presented as

$P_{\max}' = \max[P(x_1, y_1, \overrightarrow{para}), \cdots, P(x_M, y_M, \overrightarrow{para})]$; the minimum collision potential is presented as $P_{\min}' = \min[P(x_1, y_1, \overrightarrow{para}), \cdots, P(x_M, y_M, \overrightarrow{para})]$. Therefore, the normalised collision potential of position (x_k, y_k) is presented as,

$$P'_{normal}(x_k, y_k) = [P(x_k, y_k, \overrightarrow{para}) - P'_{\min}] / (P'_{\max} - P'_{\min}) \tag{13}$$

Based on the collision potential of the blip's position only, the basic probabilities about the $\theta_0, \theta_1, \theta_2$ states can be obtained as follows, or the piece of evidence is constructed as,

$$e_2: \{p(\theta_0), p(\theta_1), p(\theta_2)\} = \{P'_{normal}(x_k, y_k), 1 - P'_{normal}(x_k, y_k), 0\} \tag{14}$$

Dempster's rule can be used to combine the two pieces of evidence, which is presented as follows:

$$m(\theta) = [m_1 \oplus m_2] = \begin{cases} 0 & \theta = \emptyset \\ \frac{\sum_{B \cap C = \theta} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \phi} m_1(B) m_2(C)} & \theta \neq \emptyset \end{cases} \tag{15}$$

where θ is a proposition that can be any subset of a set of hypotheses; $m(\theta)$ is the basic probability for θ; $m_1(B)$ is the basic probability for proposition B from the first piece of evidence; $m_2(C)$ is the basic probability for proposition C from the second piece of evidence; lastly, $\emptyset$ is the empty set. Therefore, the basic probability about the *Collision* state θ_0, or the collision probability of the blip based on the two pieces of evidence is presented as:

$$p(\theta_0) = P'_{normal}(x_k, y_k) \times p / \{1 - P'_{normal}(x_k, y_k) \times (1 - p) - [1 - P'_{normal}(x_k, y_k)] \times p\} \tag{16}$$

4 A case study

4.1 Experimental platforms

To validate the proposed approach, an experiment was conducted. The experimental platform was placed on a wharf boat, which was 1.5 kilometres upstream of the Wuhan Yangtze Bridge of the Yangtze River Wuhan waterway. The testing radar was FURUNO FAR 2127S, working on X-band (9GHZ). In Figure 6, the left-hand side presents the location of the radar and the scan area, and the right-hand side presents the radar antenna. In this experiment, the radar intermediate-frequency signal was fetched and converted to grey-scale images using an S3C-3000 radar processor. One of the images is presented on the left hand side of Figure 1.

Figure 6 An experimental platform of marine radar in Yangtze River, Wuhan, China

The experiment lasted from 09:00 to 10:55 on the 17th April 2015. In total, 173 targets were captured, including 119 vessels and 54 stationary targets or noises. In the experiment, all the targets were verified manually. It is noted that many observations or blips were indeed from the same target since the radar scanned the area once per 2.4 seconds. In total, 15,286 individual observations (blips) have been captured. In these observations (blips), 11,958 observations are from moving vessels and 3,328 observations are from noises or stationary targets. In the following research, all the stationary and noise targets are treated as noise samples for simplicity.

Particularly, the verified samples are divided into two parts randomly. The first half is used to obtain the structure and CPTs of BN as discussed in Section 3.1, and the second half is used for identification validation.

Meanwhile, an AIS receiver was placed in the same area, which received 2,300,000 AIS messages from 15th March to 12th April 2015. Particularly, all the AIS messages are obtained from the same area as that of the blip recognition. These records will be used for training the coefficients of collision potential fields as described in Section 3.3.

4.2 Step 1: Authenticity inference of blips

To implement the proposed approach in this research, a software program is developed and shown in Figure 7. As shown in this figure, radar images have already been overlapped on the S57 (A map format defined by the IMO) electronic chart of the waterway. Three typical verified objects were notified as the red rectangles, and the enlarged images are also shown in Figure 7. They are buoys No.27, No.55, and a moving vessel No.15. The white circles and orange circles are the objects' labels. The centres of the objects are also marked accordingly. Especially, the white dots are the former centres of the object. Intuitively, the moving vessel objects are different from noises in terms of the attributes of the velocity, course, and graphic shape.

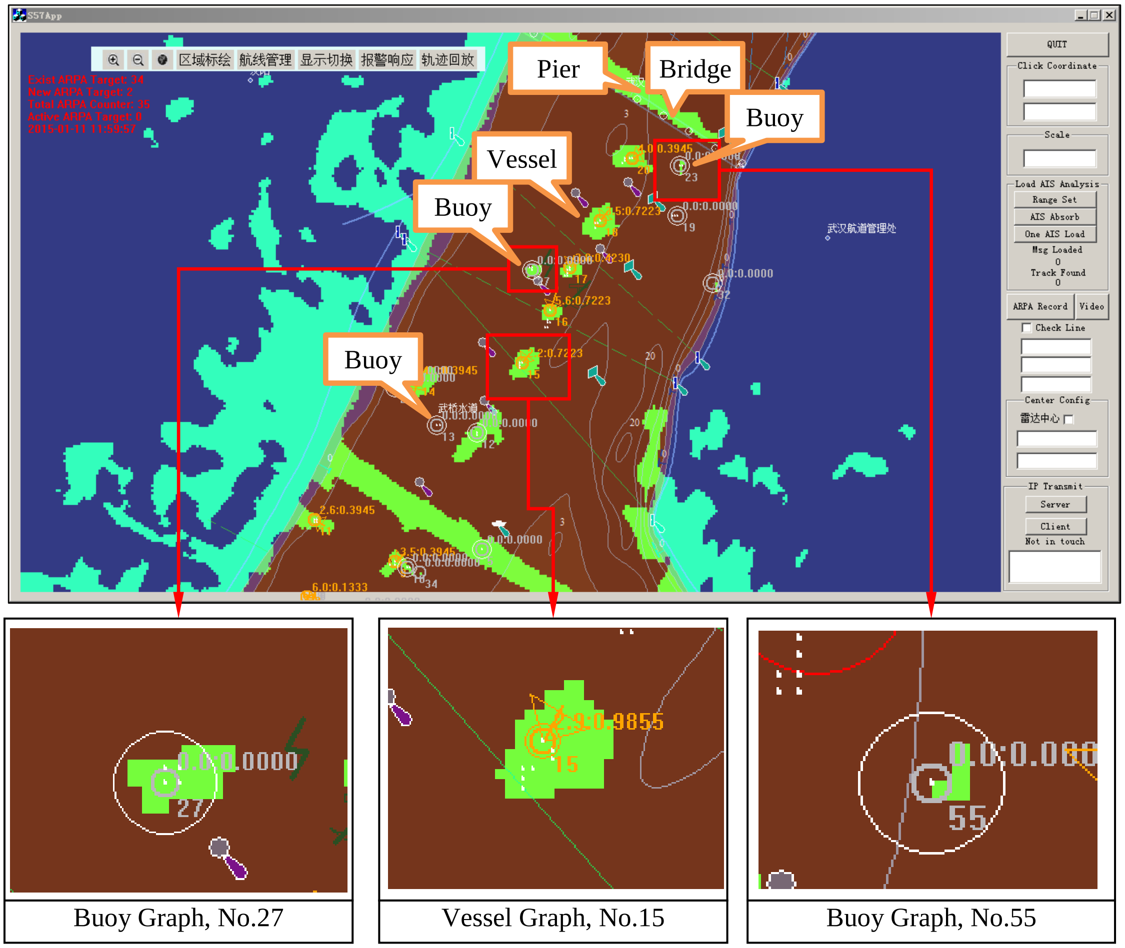

Figure 7 The experimental software application program based on VC++

Using the methods proposed in Figure 3 and Section 3.1, these characteristics are quantified in the software program. All the blips in sequential images have been transformed to verified records that are presented in a text form with discrete values. A typical record is presented in Figure 8. The record contains several fields, which are separated by commas and represent different types of discrete attribute values. In this way, the course (direction), velocity, and size are all stored in one record. Moreover, the verified vessel and noise records are saved separately.

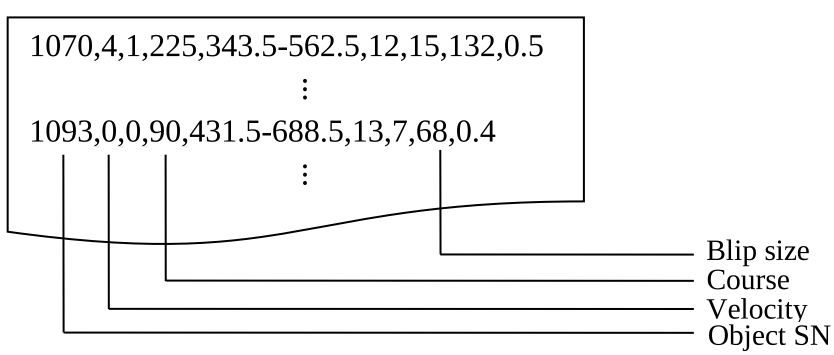

Figure 8 Text record definitions

- *Velocity* (D)

In this research, the authenticity of a blip being a real moving vessel is denoted as two states: $\mathbf{A}_1$ (Noise, a noise or stationary object), $\mathbf{A}_2$ (Vessel, a moving vessel). The first half of the verified samples include 7,643 quantified records, and 20 of them are presented in Appendices A.1. The vessel and noise velocity distributions are presented in Figure 9, where the X axis represents the observation values and the Y axis represents the frequencies. It is clear that the moving vessels are more likely to move at the velocity of 5 to 17 units (pixels) per 10 frames. However, the noise blips are more likely to move at the velocity of lower than 4 units per 10 frames. In this figure, there are 25 original observation values. The full interval should be discretised to sub-intervals to decrease the complexity of the DAG (Monti and Cooper, 1998). In general, a smaller interval in discretisation makes the model closer to reality. However, smaller intervals will increase the complexity, especially when modelling joint probabilities in a BN. Based on the method proposed by Ma *et al.* (2015b), the full interval can be discretised to 4 sub-intervals or states $\{\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3, \mathbf{D}_4\}$ = {{0, 1, 2, 3, 4, 5}, {6, 7, 8, 9, 10, 11, 12}, {13, 14, 15}, {16, 17, 18, 19, 20, 21, 22, 23, 24, 25}}.

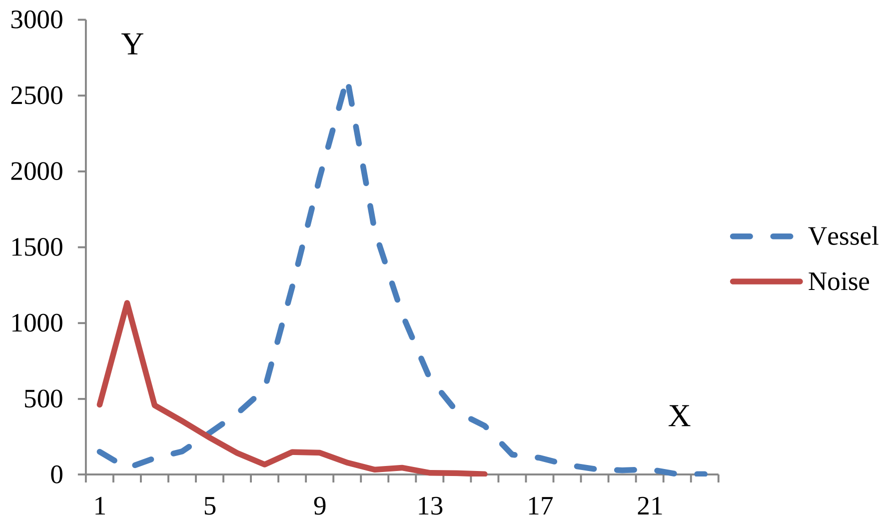

Figure 9 Speed distributions of moving vessel and noise targets

- *Course* (V)

The motion direction distributions in 10 frames of vessel and noise blips are presented in Figure 10, where the X axis represents the course values and the Y axis represents the frequencies. The differences between vessels and noises are distinctive in the distributions. Following the same procedures for modelling the velocity (D) node, the full interval of direction values should be discretised to 5 sub-intervals or states $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{V}_5\}$ = {{0, ⋯,13}, {14, ⋯,43}, {44, ⋯,190}, {191, ⋯,228}, {229, ⋯,359}}.

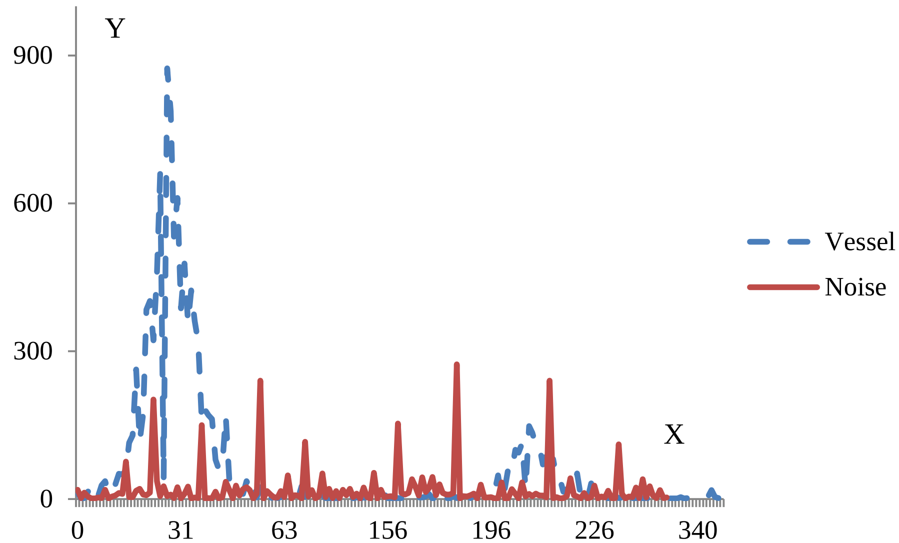

Figure 10 Motion course distributions of moving vessel and noise targets

- *Size* (S)

The size distributions of vessel and noise blips are presented in Figure 11, where the X axis represents the size values and the Y axis represents frequencies. Following the same procedures for modelling the velocity (D) node, the full interval should be discretised to 5 sub-intervals or states $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4, \mathbf{S}_5\} = \{\{11, \cdots, 13\}, \{14, \cdots, 43\}, \{44, \cdots, 190\}, \{191, \cdots, 228\}, \{229, \cdots, 322\}\}$.

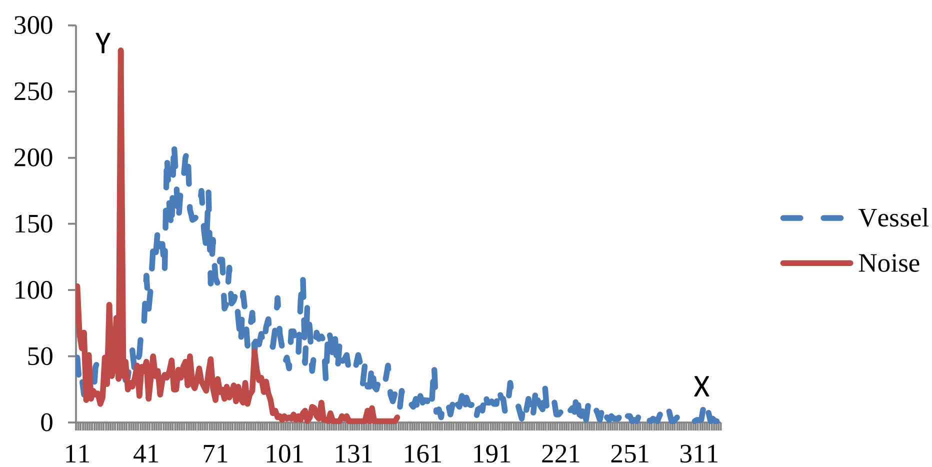

Figure 11 Blip size distributions of moving vessel and noise targets

- *BN inference and the result validation*

Subsequently, the DAG structure can be updated with the "learn_struct_k2" function in MATLAB bnt tool box based on the first half verified samples. The updated structure is shown in Figure 12. The output of "learn_struct_k2" function is presented in Appendices A.2. According to the new DAG, the velocity (D) exerts effects on direction (V) and size (S) too. Obviously, these nodes or attributes are not independent of each other.

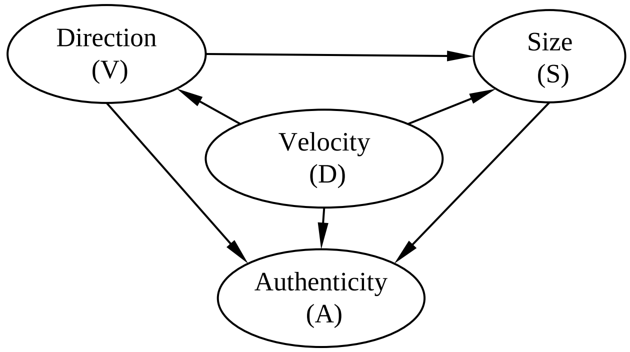

Figure 12 The DAG for authenticity recognition

The first half of the verified samples can also be used for learning the CPTs with a 'learn_param' function in MATLAB bnt tool box, which was described in Section 3.1. Eventually, the new DAG and CPTs will be used to estimate the probability of a blip being a true moving vessel in observation. The detailed CPTs are presented in TABLE III~VI of Appendices A.3.

Subsequently, the second half of the verified samples are used for validation. In practice, a final decision has to be made based on the probability. Referring to manual work, 50% is an intuitive and reasonable threshold for use. If the reasoning probability of a blip being a moving vessel is larger than 50%, the blip (observation) is considered as a true moving vessel. Otherwise, it can be considered as a noise or stationary object.

Table I Results of the analysis of the verified samples using the developed model

	Total	Correct identification	In-correct identification	Accuracy
Noises or stationary object	1,648	1,369	279	83.07%
Moving vessels	5,995	5,631	364	93.93%
Total	7,643	7,000	643	91.59%

Table I shows the results obtained from the developed model. As shown in Table I, it can be seen that there are 5,995 verified observations of being moving vessels and 1,648 verified observations of being noises or stationary objects in the analysis. The developed model produced 1,369 correct identifications out of 1,648 observations from noises or stationary objects, leading to the recognition accuracy of 83.07%. As for the 5,995 verified observations of being from moving vessels, the model produces the recognition accuracy of 93.93%. In total, the global accuracy reached 91.59%, which proves that the BN-based method here is efficient in the identification of moving vessels. In fact, recognition mistakes are also easily made by experienced operators.

Particularly, the BN-based identification is implemented by the software program described in Section 4.1. As shown in Figure 13, the probabilities of blips being from real moving vessels are represented as orange numbers (0 - 1) on each corresponding blip's right-top side. It is worth mentioning that the BN-based recognition approach can be used for different locations, as the conditional probability tables and the DAG are learned from the verified samples of the corresponding locations.

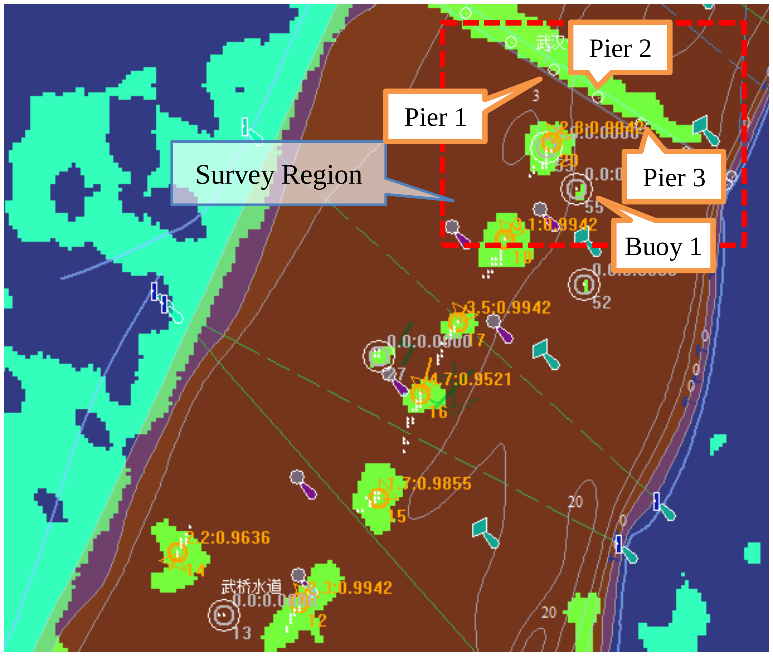

Figure 13 Blip authenticity inference

4.3 Step 2: The modelling of collision potential fields

The following issue is to estimate the collision potentials of adjacent positions, which might be estimated in accordance with the behaviours of passing vessels as described in Sections 3.2 and 3.3.

It is worth emphasising that many factors will affect the behaviours of passing vessels, including local regulations, fuel saving, weathers, and berths. However, it is too complicated to take all the factors into consideration. The behaviours of vessels will be determined by the corresponding collision potentials where the avoidance of collision becomes a major concern for ship handing as described in Section 3.3. Particularly, the depth of this waterway is only 4.5 meters. Hence, the vessels sailing in this waterway are smaller than 4000t, and their breadths are most likely smaller than 15 meters. Therefore, every single vessel is considered as a point in the APF model for simplicity. In the future research, the dimensions and the dynamic characteristics of a vessel may be taken into consideration.

In this light, a survey region in Figure 13 is chosen and marked as a red dotted rectangle, which contains three piers, a buoy, and two major channels. In Figure 14, the survey region is also represented with the S57 e-chart format. In this figure, the small blue circles and lines represent the passing vessels that crowded in the two channels; the piers are indicated with black circles; Buoy 1 is represented as a green circle at the bottom; the yellow dotted line between the centres of Pier 2 and Pier 3 is selected to be the examined cross profile that has been described in Section 3.3, namely profile K_1.

As described in Section 3.2, all the piers and buoys can be modelled as the sources of collision potential fields with the APF model, and the corresponding collision potential distribution can be obtained with the Yokawa potential function. The bold red rectangles in Figure

14 indicate the *potential edge rectangles* of Pier 2 and Pier 3, defined in Eqs. (2) and (4). The corresponding collision potential fields are represented as two highlighted red regions. In addition, the collision potential field of Buoy 1 defined in Eqs. (2) and (3), is represented as a highlighted green eclipse.

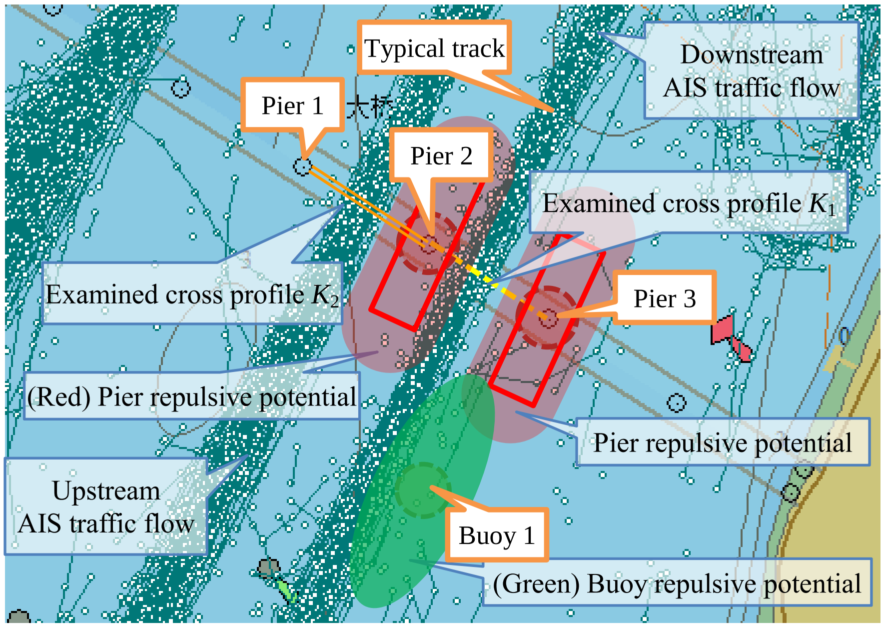

Figure 14 The survey region

Intuitively, the distribution of passing vessels on profile K_1 can be inferred based on the collision potential fields of Pier 2, Pier 3, and Buoy 1. Apparently, the vessel distribution should be symmetrical on profile K_1 if Pier 2 and Pier 3 are the only obstacles. However, Buoy 1 produces an extra collision potential field on the right side; in other words, Buoy 1 "repulses" passing vessels from the right side. Therefore, a conjecture can be made that the peak value of the vessel distribution on profile K_1 should be slightly shifted to the left hand side duo to the corresponding collision potential.

With the help of the software program described in Section 4.1, profile K_1 is analysed with 35 statistical individual points or sections in Figure 14. In this figure, each point or section denotes 3.55 meters which is the maximum resolution of the electronic-chart. In other words, the space discretization is of 3.55 meters. Based on the AIS records described in Section 4.1, the distribution of passing vessels on profile K_1 can be normalised with Eq. (8) and presented in Figure 15, where the X-axis represents the distance to Pier 2, and the Y-axis represents the normalised densities. Apparently, the densities follow a normal distribution, and the peak value is situated in the left side of profile K_1 between pier 2 and pier 3 as expected.

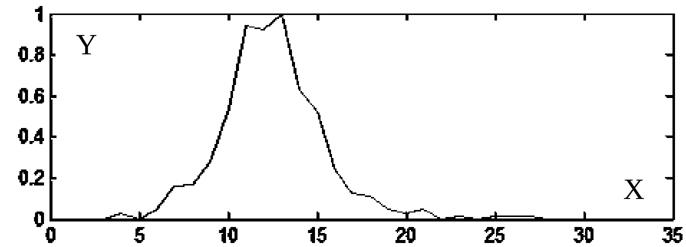

Figure 15 The normalised distribution of passing vessels on profile K_1

4.4 Step 3: The training of the coefficients using AIS records

- *coefficient training*

With Step 2, the collision potential fields of piers and buoys are established. The next task is to obtain the coefficients of these potential fields. Take profile K_1 as an example, the coefficients should make the collision potentials consistent with the distribution of passing vessels. Therefore, the coefficients can be obtained in a nonlinear optimisation model, as Eq. (9). In this occasion, i = 35, the coefficients are obtained as $\overrightarrow{para} = \{\alpha_b, \xi, \sigma_b, \alpha_b, w, l, \sigma_p\} = \{0,18.8142,0.5208, 0,12.3944,63.3068,0.3710\}$ using the '*fmincon*' function of MATLAB 2013b.

Subsequently, the collision potential distribution of profile K_1 can be calculated, and the normalised "safety distribution" is presented in Figure 16, which is defined in Eq. (7). The X-axis represents the profile positions, and Y-axis represents the normalised "safety degree". By comparing Figures 15 and 16, a good agreement can be found. In other words, the distribution of collision potentials is consistency with the distribution of passing vessels on profile K_1.

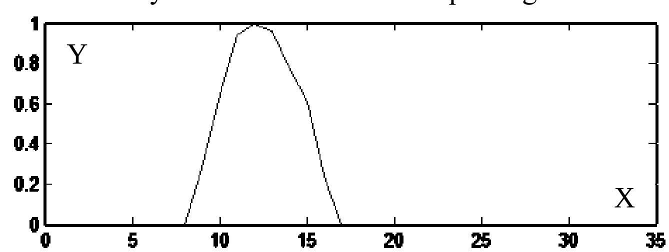

Figure 16 The normalised distribution of safety degree on profile K_1

In addition, the Bhattacharyya distance is introduced to measure the similarity between Figures 15 and 16, which is widely used to quantify the difference between discrete distributions (Kailath, 1967). For discrete distributions $p(x)$ and $q(x)$, where x is the discrete variable, the Bhattacharyya distance is defined as follows:

$$D_B(p,q) = -\ln(BC(p,q)) \quad (17)$$

where $BC(p,q) = \sum_{x \in X} \sqrt{p(x)q(x)}$. $D_B(p,q) \in [0,1]$, 0 denotes that there is no distance between p and q, or p is exactly the same as q; 1 denotes that q is completely different from q. Obviously, the formulations of $p(x)$ and $q(x)$ are probably unknown in practice. Figures 15 and 16 approximately follow a normal distribution. Then, the Bhattacharyya distance can be calculated by extracting the mean and variances of p and q distributions (Coleman and Andrews, 1979), presented as,

$$D_B(p,q) = -\frac{1}{4}\ln\left(\frac{1}{4}\left(\frac{\sigma_p^2}{\sigma_q^2} + \frac{\sigma_q^2}{\sigma_p^2} + 2\right)\right) + \frac{1}{4}\left(\frac{(\mu_p - \mu_q)^2}{\sigma_p^2 + \sigma_q^2}\right) \tag{18}$$

where σ_p and σ_q are the variance of the p and q distributions, μ_p and μ_q are the means of the p and q distributions. Hence, the Bhattacharyya distance between Figures 15 and 16 is 0.015, proving that the collision potentials are highly consistent with the real vessel distribution, and the coefficients obtained from the optimisation model are appropriate.

Then, a global distribution of collision potentials can been obtained and represented as a heat map with $\overrightarrow{para} = \{\alpha_b, \xi, \sigma_b, \alpha_b, w, l, \sigma_p\}$, which is shown in Figure 17. The red colour represents the maximum value of collision potential, the blue colour represents the minimum, and the transition colours between red and blue represent the continuous variation of collision potential. Apparently, such a heat map is consistent with the intuitive judgements of human.

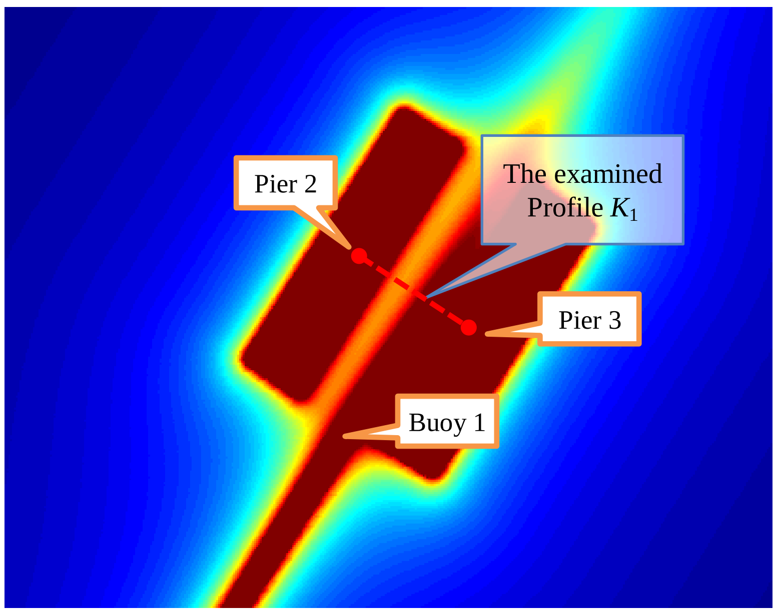

Figure 17 The heat map of collision potentials in the survey region after a nonlinear optimisation

- *Coefficient validation in another typical scenario*

Moreover, another scenario or examined profile is introduced to validate the APF model. In Figure 14, there is another examined profile K_2 between Pier 1 and Pier 2, which is also relatively close and isolated. Hence, the collision avoidance is also the major concern for ship handling on this profile. Using the coefficients $\overrightarrow{para} = \{\alpha_b, \xi, \sigma_b, \alpha_b, w, l, \sigma_p\}$ obtained previously, the heat map of collision potential fields for profile K_2 is presented in Figure 18. Based on Eq. (7), the normalised safety degree of profile K_2 is presented in Figure 19. Based on Eq. (8), the actual normalised vessel distribution of profile K_2 is presented in Figure 20. Obviously, a high agreement can also be found between Figures 19 and 20. Furthermore, the Bhattacharyya distance between collision potential distribution and the vessel distribution is 0.011 based on Eqs. (17) and (18), proving that the coefficients $\overrightarrow{para} = \{\alpha_b, \xi, \sigma_b, \alpha_b, w, l, \sigma_p\}$ and the APF model are also reasonable and applicable.

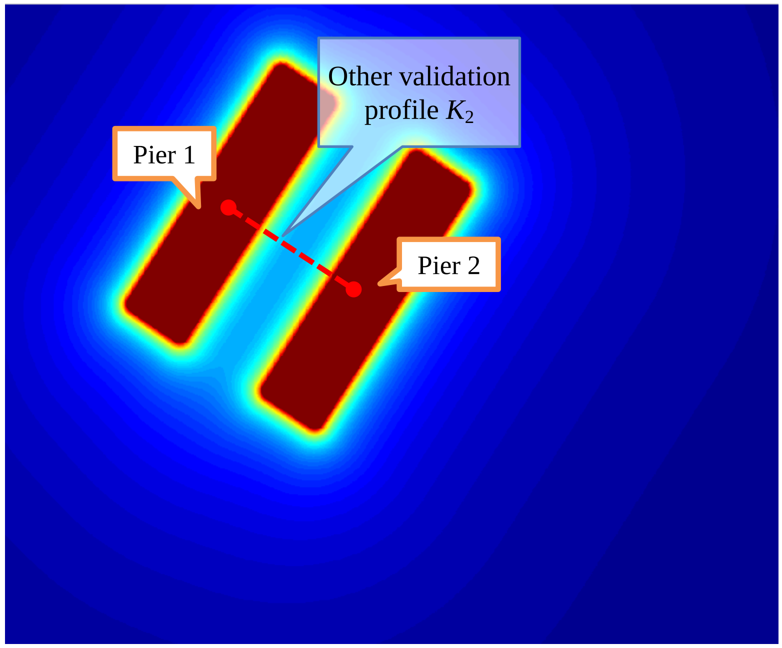

Figure 18 The heat map of collision potentials in the other survey region

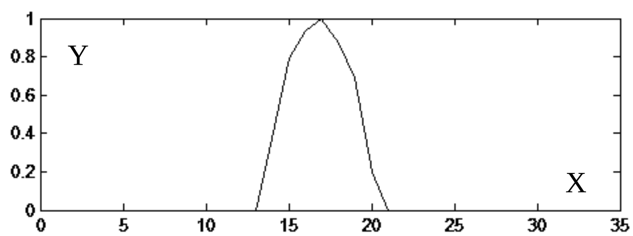

Figure 19 The normalised distribution of safety degree on profile K_2

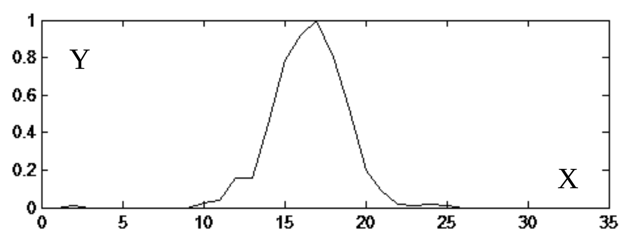

Figure 20 The normalised distribution of passing vessels on profile K_2

Furthermore, 7 other scenarios or profiles are introduced to validate the coefficients and the APF model, which have been marked in Figure 21 as red dotted lines. It is worth noting that only the profiles located at the four major archways are selected, where vessels might pass through, otherwise no vessel distributions can be obtained.

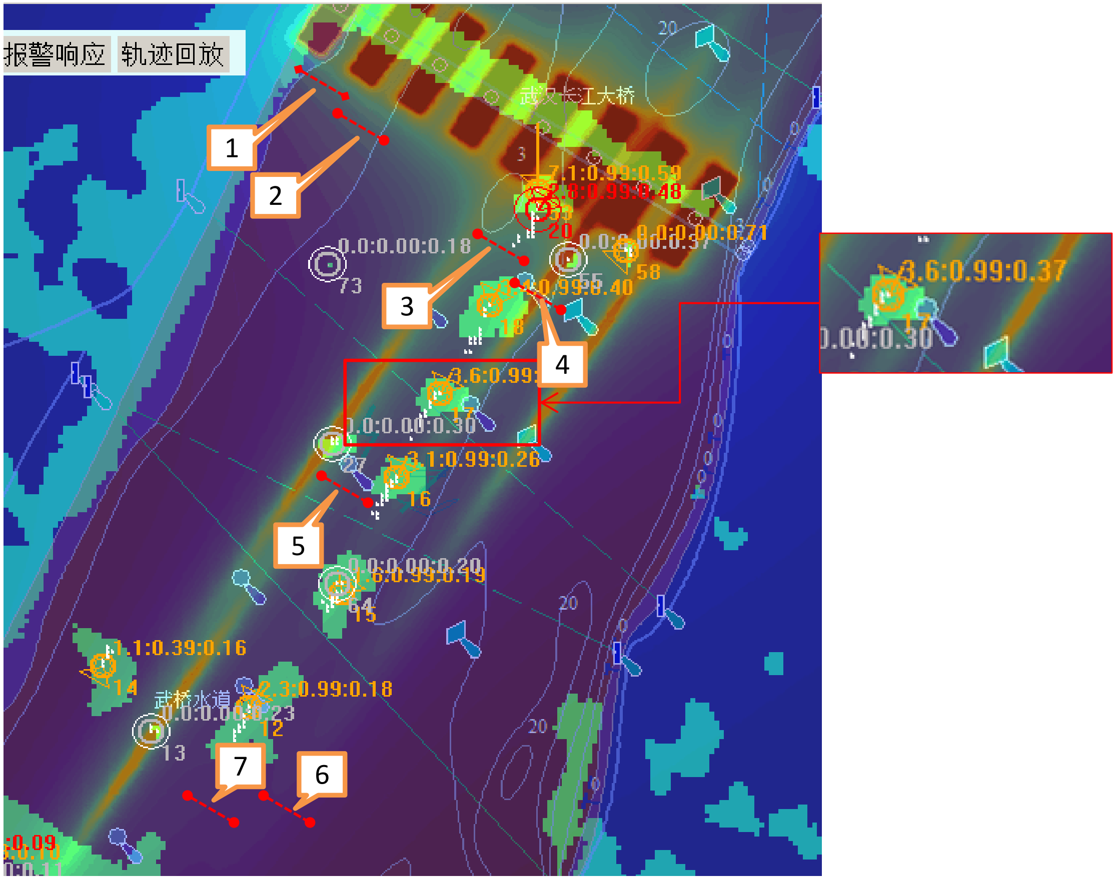

Figure 21 The presumed collision potential distribution using the APF model

The Bhattacharyya distances between the collision potential distributions from the APF model and the vessel distributions from the AIS records on these profiles are presented in Table II.

Table II Distances between the predicted collision potential distribution and vessel distribution

	Profile 1	Profile 2	Profile 3	Profile 4	Profile 5	Profile 6	Profile 7
Bhattacharyya distances (0-1)	0.012	0.013	0.019	0.027	0.232	0.152	0.120

According to Table II and Figure 21, it can be inferred that the closer to obstacles profile is, the more accurate APF model and corresponding coefficients will be. It is reasonable that the closer to obstacles vessels are, the more attention on the obstacles ship operators will pay. Overall, the APF model is an efficient model in the quantification of collision potentials.

4.5 Step 4: Collision probability estimation

The collision potential of any position in the waterway can be obtained with the APF model as discussed in Sections 4.3 and 4.4. Then, a global distribution of collision potentials in the waterway can be represented as a heat map in Figure 21. It is worth mentioning that the heat map has been overlapped on the application program. In this figure, the red colour represents the maximum collision potential, the blue colour represents the minimum, and the transition colours between red and blue represent the continuous variation of collision potential.

As discussed, the probability of a blip being a real moving vessel, and the collision potential of its position are the two factors in determining whether it needs much attention in manual work. In this research, the two factors are combined with Dempster's rule as described in Section 3.4. For instance, the object 17 is enlarged on the right hand side of Figure 21; the text "3.6:0.99:0.37" on its right-top side denotes that its speed is 3.6 pixels/units per 10 frames based on Section 3.1, its probability of being a moving vessel is 0.99 (99%) given by Section 4.3, and the normalised collision potential of this position is 0.37 (37%) given by Eq. (13).

The two pieces of evidence are presented as $e_1{:}\{p(\theta_0),p(\theta_1),p(\theta_2)\}=\{0.99,0.01,0\}$, $e_2{:}\{p(\theta_0),p(\theta_1),p(\theta_2)\}=\{0.37,0.63,0\}$ based on Eqs. (14) and (15). Then, the basic probabilities about the θ_0, θ_1 and θ_2 states can be obtained as $\{p(\theta_0),p(\theta_1),p(\theta_2)\}=\{0.98,0.02,0\}$ by combining e_1 and e_2 based on Eq. (16). The collision probability of the target can be considered as $p(\theta_0)=0.98$. In fact, $p(\theta_0)$ here represents a large belief degree about the *Collision* state for reminding the supervisors that the blip needs attention.

The efficiencies of the BN-based method and the APF model have been proved individually in Sections 4.1 and 4.4. Eventually, the proposed approach was tested with the verified samples, in order to prove its validity and reliability preliminarily. 3 officers from local maritime administrations, Wuhan, China, were invited to rank blips' threats to piers and buoys manually. The validation samples are the same as those of the BN validation in Section 4.2. At last, the approach identified 35 objects that had the highest collision probability, and 32 of these objects were also inferred to be most dangerous by manual work. In other words, the accuracy can be considered as 91.43%, and a high agreement has been found. Moreover, in the testing, the ones that were close to the piers and buoys could be identified accurately; the ones that were far away from obstacles were incorrectly identified occasionally. As discussed in Section 4.4, if the vessels are close to the obstacles, and the collision avoidance becomes a major concern for ship handling, the APF model becomes more efficient.

5 Conclusion and Discussions

Coastal surveillance radar is the kernel sensor in port management. To lower the burden of supervisors, this paper proposed a BN and APF-based approach to estimate the collision probabilities to obstacles of blips preliminarily with sequential radar images and AIS records. The conclusions are given below.

1) With inter-frame differences in frames, including the velocity, course and size of blips, the BN-based method is capable of estimating the probability of a blip being a true moving vessel, whilst updating the structure and coefficients from verified samples, and high accuracy was achieved in a field test.
2) The APF model can be introduced to describe the collision potentials caused by obstacles. Moreover, the coefficients can be trained in a nonlinear optimisation model using AIS data records. According to manual work, the collision probability of a blip can be considered as the synthesis of the collision potential and the authenticity probability, and a high agreement has been found in the preliminary test. Particularly, the case study is conducted in a relatively narrow waterway. Hence, the space discretisation is based on the maximum

resolution of the corresponding electronic-chart. In other scenarios, the space discretisation can be also based on different distances in accordance with the distribution of collision potential fields generated by the obstacles and the traffic characteristics.

While the proposed approach is aimed to serve as a rigorous assessment process so that the inferred results could be used to form a sound basis for further analysis and decision making, other issues, such as the following, need to be investigated in future research for the more robust and wider application of the approach.

1) Stationary vessels were treated as noises in this research for simplicity. However, a new method may be needed to distinguish them from general noises. In manual judgments, for example, the continuous characteristics of a target are used as important evidence
2) Waterfronts or other encountered vessels may also need to be modelled in a similar way; this will make the collision potential more accurately estimated.
3) The concept of potential fields may need to be further investigated in order to fully realise the APF model's potential in ship collision assessment. This may be particularly useful for studying collision risks associated with berths and recommended channels.
4) The authenticity and collision potential of a blip were considered to be independent of each other and of equal weight in this research for simplification purposes. Further work may be useful to investigate how their dependency and their different weights would affect collision probability estimation.
5) In many circumstances, neighbouring vessels might take influences on the distribution of the collision potential. Therefore, neighbouring vessels are also needed to be modelled as the sources of collision potentials.
6) In this research, the APF model is only used to describe the collision potential distribution caused by static obstacles. However, it is widely acknowledged that there is coupling among static obstacles, neighbouring traffic and moving vessels in collision assessments. In other words, to model the collision potential comprehensively, the behaviours and the predicted route of the vessel are also essential. As discussed in this paper, a VTS operator might not be capable of obtaining such information of a ship when it is passing through the monitoring area directly. To address this problem, in the future research, not only the authenticity recognition investigated in this paper, but also the behaviour recognition and the route prediction may be considered.

Acknowledgements

The authors are grateful for the financial support by National Science Foundation of China (Grants no. 61503289 and 51479158), EU (Marie Curie grant REFERENCE no. 314836), and the Information Foundation of Ministry of Transport of China (Grant no. 2015-328-811-180).

Appendices

A.1

As described, there are 7,643 verified samples captured in the first two hour. These samples are saved as a text form presented in Section 4.2 and Figure 8. It is worth mentioning that there are about 10 to 100 records for a blip, as it usually stay in the observation area for about 20 to 300 seconds. 20 verified samples in the text form are presented as follows.

10 selected observations from verified vessel blips are presented as:

178,12,4,135,549.5-479.5,5,11,33,0.3
178,12,4,135,549.5-479.5,5,11,33,0.3
301,0,0,240,629.0-611.0,16,8,81,0.3
301,10,3,246,589.5-597.5,15,9,72,0.3
770,14,4,244,590.5-592.0,19,10,86,0.2
770,9,3,214,518.5-534.0,3,4,12,0.6
962,0,0,270,624.5-610.5,13,7,74,0.4
962,5,2,270,624.0-610.5,14,7,74,0.4
1206,8,8,338,564.0-296.5,6,13,59,0.4
1206,9,8,338,563.5-297.5,7,15,62,0.3

10 selected observations from verified noise blips are presented as:

0,1,0,225,533.0-181.5,23,10,94,0.2
0,1,0,225,533.0-181.5,23,10,94,0.2
1,0,0,0,521.0-189.5,10,9,56,0.4
1,1,0,90,521.0-189.5,10,9,55,0.4
1524,1,0,270,571.0-512.0,38,106,343,0.0
1524,0,0,270,571.5-512.0,39,106,343,0.0
2086,5,2,355,517.0-312.5,6,3,25,0.7
2086,6,2,355,517.0-312.5,6,3,25,0.7
2837,4,1,300,559.5-286.5,9,7,28,0.3
2837,4,1,326,561.0-287.5,6,9,25,0.3

A.2

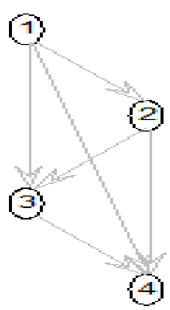

Figure 22 The output of the 'learn_struct_K2' function in the MATLAB 2013b

The learning procedure is implemented with the 'learn_struct_K2' function in the MATLAB 2013b bnt toolbox, and the output is presented in Figure 22 where the nodes named 1, 2, 3 and 4 denotes the Velocity (D), Direction (V), Size (S), and Authenticity (A). Hence, the updated DAG structure is shown in Figure 12.

A.3

Table III. The CPT of node Velocity (D)

D_1	D_2	D_3	D_4

0.257	0.132	0.593	0.018

Table IV. The CPT of node Direction (V)

	V_1	V_2	V_3	V_4	V_5
D_1	0.257	0.132	0.593	0.018	0.257
D_2	0.257	0.132	0.593	0.018	0.257
D_3	0.257	0.132	0.593	0.018	0.257
D_4	0.257	0.132	0.593	0.018	0.257

Table V. The CPT of node Slenderness (S)

	S_1					S_3					S_4					S_4					S_5				
	V_1	V_2	V_3	V_4	V_1	V_1	V_2	V_1	V_2	V_3	V_4	V_5	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5
D_1	0.024	0.057	0.102	0.049	0.315	0.048	0.018	0.043	0.052	0.047	0.008	0.080	0.062	0.115	0.060	0.315	0.320	0.289	0.418	0.213	0.605	0.526	0.503	0.366	0.539
D_2	0.133	0.007	0.109	0.096	0.089	0.100	0.031	0.065	0.062	0.079	0.011	0.007	0.065	0.144	0.087	0.089	0.182	0.109	0.356	0.198	0.667	0.773	0.652	0.343	0.508
D_3	0.040	0.011	0.058	0.067	0.103	0.008	0.005	0.041	0.089	0.037	0.087	0.008	0.031	0.076	0.011	0.103	0.219	0.157	0.337	0.177	0.762	0.756	0.714	0.430	0.701
D_4	0.000	0.070	0.111	0.006	0.000	0.000	0.070	0.111	0.076	0.000	0.000	0.116	0.056	0.045	0.050	0.000	0.302	0.222	0.191	0.000	0.000	0.442	0.500	0.682	0.950

Table VI. The CPT of node Authenticity (A)

	A_1																								
	S_1					S_3					S_3					S_4					S_5				
	V_1	V_2	V_3	V_4	V_1	V_1	V_2	V_1	V_2	V_3	V_4	V_5	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5
D_1	0.667	0.762	0.920	1.000	0.922	0.667	0.600	0.906	0.933	0.818	0.000	0.843	0.882	1.000	0.971	0.615	0.577	0.876	0.867	0.968	0.347	0.278	0.606	0.571	0.924
D_2	0.000	0.000	0.867	0.500	0.313	0.000	0.000	0.111	0.778	0.200	0.000	0.100	0.667	0.429	0.182	0.250	0.124	0.067	0.423	0.680	0.183	0.015	0.078	0.180	0.656
D_3	0.000	0.226	0.529	0.000	0.950	1.000	0.000	0.083	0.110	0.900	0.000	0.000	0.000	0.180	1.000	0.000	0.048	0.000	0.110	0.563	0.010	0.006	0.000	0.036	0.263
D_4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

	A_2																								
	S_1					S_3					S_3					S_4					S_5				
	V_1	V_2	V_3	V_4	V_1	V_1	V_2	V_1	V_2	V_3	V_4	V_5	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5	V_1	V_2	V_3	V_4	V_5
D_1	0.333	0.238	0.080	0.000	0.078	0.333	0.400	0.094	0.067	0.182	1.000	0.157	0.118	0.000	0.029	0.385	0.423	0.124	0.133	0.032	0.653	0.722	0.395	0.429	0.076
D_2	1.000	1.000	0.133	0.500	0.688	1.000	1.000	0.889	0.222	0.800	1.000	0.900	0.333	0.571	0.818	0.750	0.876	0.933	0.577	0.320	0.817	0.986	0.922	0.820	0.344
D_3	1.000	0.774	0.471	1.000	0.050	0.000	1.000	0.917	0.890	0.100	1.000	1.000	1.000	0.821	0.000	1.000	0.952	1.000	0.890	0.438	0.990	0.994	1.000	0.964	0.737
D_4	0.000	1.000	1.000	1.000	0.000	0.000	1.000	1.000	1.000	0.000	0.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	0.000	1.000	1.000	1.000	1.000

References

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.

Bing, H., Gang, L., Jiang, G., Hong, W., Nan, N. & Yan, L. (2011). A route planning method based on improved artificial potential field algorithm. In *Proceedings of Communication Software and Networks (ICCSN) International Conference*, 550-554.

Coleman, G. B. & Andrews, H. C. (1979). Image segmentation by clustering. *Proceedings of the IEEE*, 67(5), 773-785.

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*, 86-94.

De Campos, L. M. & Castellano, J. G. (2007). Bayesian network learning algorithms using structural restrictions. *International Journal of Approximate Reasoning*, 45(2), 233-254.

Dellacherie, C. & Meyer, P. A. (2011). *Probabilities and Potential, C: Potential Theory for Discrete and Continuous Semigroups*. Elsevier.

Dong, Y. & Frangopol, D. M. (2015). Probabilistic ship collision risk and sustainability assessment considering risk attitudes. *Structural Safety*, 53, 75-84.

Eleye-Datubo, A. G., Wall, A. & Wang, J. (2008). Marine and offshore safety assessment by incorporative risk modelling in a fuzzy-Bayesian network of an induced mass assignment paradigm. *Risk Analysis*, 28(1), 95-112.

Fujii, Y., Oshima, R., Yamanouchi, H. & Mizuki, N. (1974). Some factors affecting the frequency of accidents in marine traffic: I-the diameter of evasion for crossing encounters, II-the probability of stranding, III-the effect of darkness of the probability of collision and stranding. *Journal of Navigation*, 27(2), 239-247.

Hänninen, M. & Kujala, P. (2012). Influences of variables on ship collision probability in a Bayesian belief network model. *Reliability Engineering & System Safety*, 102, 27-40.

IEC, I. 62288 (2014): Maritime navigation and radio communication equipment and systems - Presentation of navigation-related information on shipborne navigational displays - General requirements, methods of testing and required test results.

IEC, I. 62388 (2013): Maritime navigation and radio communication equipment and systems– Shipborne radar–Performance requirements, methods of testing and required results.

Islam, T., Rico-Ramirez, M. A., Han, D. & Srivastava, P. K. (2012). Artificial intelligence techniques for clutter identification with polarimetric radar signatures. *Atmospheric Research*, 109, 95-113.

Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology*, 15(1), 52-60.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 5(1), 90-98.

Kim, J. O. & Khosla, P. K. (1992). Real-time obstacle avoidance using harmonic potential functions. *IEEE Transactions on Robotics and Automation*, 8(3), 338-349.

Kujala, P., Hänninen, M., Arola, T. & Ylitalo, J. (2009). Analysis of the marine traffic safety in the Gulf of Finland. *Reliability Engineering & System Safety*, 94(8), 1349-1357.

Li, H. B., Shen, Y. Y. & Liu, Y. T. (2007). Estimation of detection threshold in multiple ship target situations with HF ground wave radar. *Journal of Systems Engineering and Electronics*, 18(4), 739-744.

Li, B. & Pang, F. W. (2013). An approach of vessel collision risk assessment based on the D–S evidence theory. *Ocean Engineering*, 74, pp. 16-21.

Lin, B. & Huang, C. H. (2006). Comparison between ARPA radar and AIS characteristics for vessel traffic services. *Journal of Marine Science and Technology*, 14(3), 182-189.

Liu, G. P., Yang, J. B. & Whidborne, J. F. (2003). *Multiobjective optimisation and control. Baldock:* Research Studies Press.

Ma, F., Wu, Q., Yan, X., Chu, X. & Zhang, D. (2015a). Classification of Automatic Radar Plotting Aid targets based on improved Fuzzy C-Means. *Transportation Research Part C: Emerging Technologies*, 51, 180-195.

Ma, F., Chen, Y. W., Yan, X. P., Chu, X. M. & Wang, J. (2015b). A novel marine radar targets extraction approach based on sequential images and Bayesian network. *Ocean Engineering*, under review.

Monti, S. & Cooper, G. F. (1998). A multivariate discretization method for learning Bayesian networks from mixed data. In *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, 404-413.

Montiel, O., Orozco-Rosas, U. & Sepúlveda, R. (2015). Path planning for mobile robots using Bacterial Potential Field for avoiding static and dynamic obstacles. *Expert Systems with Applications*, 42(12), 5177-5191.

Montewka, J., Goerlandt, F. & Kujala, P. (2012). Determination of collision criteria and causation factors appropriate to a model for estimating the probability of maritime accidents. *Ocean Engineering*, 40, 50-61.

Montewka, J., Hinz, T., Kujala, P. & Matusiak, J. (2010). Probability modelling of vessel collisions. *Reliability Engineering & System Safety*, 95(5), 573-589.

Park, M. G., Jeon, J. H. & Lee, M. C. (2001). Obstacle avoidance for mobile robots using artificial potential field approach with simulated annealing. In *Proceedings of IEEE International Symposium on Industrial Electronics*, 3, 1530-1535.

Pêtrès, C., Romero-Ramirez, M. A. & Plumet, F. (2012). A potential field approach for reactive navigation of autonomous sailboats. Robotics and Autonomous Systems, 60(12), 1520-1527.

Pedersen, P. T. (2010). Review and application of ship collision and grounding analysis procedures. *Marine Structures*, 23(3), 241-262.

Qu, X., Meng, Q. & Suyi, L. (2011). Ship collision risk assessment for the Singapore Strait. *Accident Analysis & Prevention*, 43(6), 2030-2036.

Ranganathan, A., Al-Muhtadi, J. & Campbell, R. H. (2004). Reasoning about uncertain contexts in pervasive computing environments. *IEEE Pervasive Computing*, (2), 62-70.

Statheros, T., Howells, G. & Maier, K. M. (2008). Autonomous ship collision avoidance navigation concepts, technologies and techniques. *Journal of Navigation*, 61(1), 129-142.

Volpe, R. & Khosla, P. (1990). Manipulator control with superquadric artificial potential functions: Theory and experiments. *IEEE Transactions on Systems, Man and Cybernetics*, 20(6), 1423-1436.

Wang, J. (2001). The current status and future aspects in formal ship safety assessment. *Safety Science*, 38(1), 19-30.

Williams, T. M. (1996). The two-dimensionality of project risk. *International Journal of Project Management*, 14(3), 185-186.

Yoo, J. C. & Kim, Y. S. (2003). Alpha–beta-tracking index (α–β–Λ) tracking filter. *Signal Processing*, 83(1), 169-180.

Zhang, D., Yan, X. P., Yang, Z. L., Wall, A. & Wang, J. (2013). Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. *Reliability Engineering & System Safety*, 118, 93-105.

Zhang, Q., Chen, D. & Chen, T. (2012). An obstacle avoidance method of soccer robot based on evolutionary artificial potential field. *Energy Procedia*, 16, 1792-1798.

Zhao, X., Wang, S., Zhang, J., Fan, Z. & Min, H. (2013). Real-time fault detection method based on belief rule base for aircraft navigation system. *Chinese Journal of Aeronautics*, 26(3), 717-729.

Zhou, D., Shen, X. & Yang, W. (2013). Radar target recognition based on fuzzy optimal transformation using high-resolution range profile. *Pattern Recognition Letters*, 34(3), 256-264.