The Journal of Navigation (2022), **75**:1 76–94
doi:10.1017/S0373463321000849

CrossMark

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

A novel regional collision risk assessment method considering aggregation density under multi-ship encounter situations

Rong Zhen,[1,2,3] Ziqiang Shi,[1] Zheping Shao,[1] and Jialun Liu[3,4]*

[1] Navigation College, Jimei University, Xiamen, China.
[2] Hubei Key Laboratory of Inland Shipping Technology, Wuhan, China.
[3] Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, China.
[4] National Engineering Research Center for Water Transportation Safety, Wuhan, China.
*Corresponding author. E-mail: jialunliu@whut.edu.cn

Received: 23 June 2021; **Accepted:** 16 October 2021; **First published online:** 19 November 2021

Keywords: ship collision risk; ship aggregation density; clusters of encounter ships; maritime surveillance

Abstract
The regional ship collision risk assessment for multiple ships in restricted waters is of great significance to the early warning of ship collision risk and the intelligent supervision of maritime traffic. Given the existed method of regional ship collision risk assessment without considering the impact of ship aggregation density, this paper proposes a novel regional ship collision risk assessment method that considers the aggregation density (AD) of the clusters of encounter ships (CES) for intelligent surveillance and navigation. The effectiveness of the proposed method has been examined by the experimental case study in the waters of Xiamen, China, and analysis has been compared with other existed studies to show the advantages of the new proposed algorithm. The results show that the study method can more intuitively and effectively quantify the temporal and spatial distribution of regional collision risks in the restricted sea area. The proposed method can improve the efficiency of traffic management when monitoring the ship collision risks in macroscopic view, and assist the safety of manned and unmanned ship navigation.

1. Introduction

More than 90% of the world's international trade has been contributed by maritime transportation. Especially under the influence of the COVID-19 pandemic, maritime transport provides a stable and sufficient source of production and life for the world. With the significant increase in the volume of maritime traffic and the complex navigational environment, ship collisions in the complex coastal waters occur from time to time (Li and Pang, 2013), causing loss of life and property and pollution of the marine environment. To improve the navigation safety of coastal and port waters, many countries and regions have established vessel traffic services (VTS) to monitor the safe navigation status of ships. The intelligent supervision of VTS is an important auxiliary facility to reduce maritime traffic accidents, and the system is equipped with advanced sensing, communication and computer equipment to obtain maritime traffic data and conduct situation assessment, which guides the safe navigation of ship officers. The ship collision risk assessment is an important function of VTS, which is of great significance to ensure maritime traffic safety. At the same time, the ship collision risk assessment is a key technique for maritime autonomous surface ships in the future (Fan et al., 2020; Huang et al., 2020a).

To accurately model the collision risk of ships at sea, there are extensive studies conducted by relevant scholars. A relevant literature review can be found in Ozturk and Cicek (2019) and Huang et al. (2020b). According to the research methods and objectives, the research can be divided into three categories:

(1) Methods based on microscopic perspective.
(2) Methods based on macroscopic perspective.
(3) Methods based on regional perspective.

The microscopic ship collsion risk is modeled between own ship (OS) and target ships (TS) for the purpose of automatic collision avoidance of give-away ship. Most studies of ship collision risk have been carried from a micro perspective. Firstly, Kearon (1997) used different weights assigned to DCPA and TCPA to consider collision risk. Lisowski (2002) introduced safety distance and safety time when using DCPA and TCPA, and used the Euler norm to calculate collision risk. Szlapczynski (2006) used the two-norm form to calculate the collision risk and considered the shipdomain. The ship domain was proposed by Fujii and Tanaka (1971), and it is developed by Wang (2010) into fuzzy quaternion ship domain. The dynamic ship collision risk is developed by the basic multipliers of TCPA and DCPA, which are calculated from the AIS data by Mou et al. (2010). Based on the ship domain model, Zhang and Meng (2019) proposed a probabilistic ship domain, which depicts the ship domain boundary for ship collision risk assessment. Many researchers added a lot of detail, considering relative bearing (Ren et al., 2011) and velocity ratio (Zhang et al., 2015; Ozturk and Cicek, 2019). Researchers tended to use fuzzy theory to infer collision risk (Ahn et al., 2012; Bukhari et al., 2013; Zhang et al., 2017). With the development of machine learning and deep learning, Gang et al. (2016) and Zheng et al. (2020) applied a support vector machine to classify the risk of ships. Zhang et al. (2020) presented a new approach based on convolutional neural networks (CNNs) and image recognition to interpret and classify ship–ship collision risks in encounter scenarios. Yuan et al. (2021) proposed a risk-informed collision risk awareness approach based on uncertainty nonlinear velocity obstacles (UNLVO) for real-time operating conditions, and dynamic and uncertainty characteristics of ship motion. In the microscopic perspective of ship collision assessment studies, the collision risk is usually calculated between the OS and TS for collision avoidance by OS. However, there are always many ships in the restricted waters, and multi-ship encounter situation occurs frequently, so the collision risk assessment methods for OS and TS need to be explored further.

The macroscopic ship collision risk is usually for the probility-based safety assessment for the defined waters from historical traffic flow data and accident data. Fujii and Shiobara (1971) first proposed the evaluation method of collision probability, which used statistical accident data to obtain probability. Yip (2008) investigated port traffic risk issues by discussing historic accidents in Hong Kong Port from 2001 to 2005. In Kujala et al. (2009), detailed accident statistics during the last 10 years were described and the risk of ship collisions was estimated by theoretical modeling in two locations within the Gulf of Finland. Goerlandt and Montewka (2014) offered a Bayesian network (BN) model for reasoning under uncertainty for the assessment of accidental cargo oil outflow in a ship–ship collision where a product tanker is struck. Kum and Sahin (2015) applied a root cause analysis to investigate and reveal the cause of marine accidents/incidents in Arctic waters from 1993 to 2011. The BN was applied to the analysis of collision risk and accident factors in Jiang et al. (2020) and Fan et al. (2020). Zhang et al. (2021) explored the spatial patterns and characteristics of maritime accidents on a global scale, the characteristics of global maritime accidents are found to be diverse in different regions. Considering the practical considerations of implementing this method to characterise navigational risk between waterways and scenarios, the likelihood of ship collision in the UK waters was implemented by the ship domain model (Rawson and Brito, 2021). The collision risk assessment methods of macroscopic perspective are usually for the historical accident analysis and give a reference for maritime traffic management in the studied sea area; it cannot be applied in real-time maritime surveillance.

In recent years, to facilitate visual and interactive perception of multi-ship collision risk, the study of collision risk assessment from a regional perspective has been a rising trend. The regional ship collision is formulated from the microscopic ship collsion risk using specific functions, and it shows the real-time regional ship collision risk for defined waters. Wen et al. (2015) introduced a marine traffic complexity model to evaluate the status of the traffic situation and the degree of crowding and risk of collision, which support mariners and traffic controllers to get the traffic situation awareness. Zhen et al. (2017) used DBSCAN clustering and factors of DCPA \ TCPA to identify high-risk vessels in multi-ship encounter situation areas from a macro-perspective. Liu et al. (2019, 2020) introduced the game theory and molecular dynamics to obtain macro-regional collision risk after identifying risk from a micro-perspective based on clustering. Chen et al. (2021) measured the collision risk between multiple ships from the perspective of velocity obstacles on the basis of clustering to obtain the macro-regional collision risk. The regional ship collision risk indicator is usually composed of the micro-based ship collision risk indicators, but the impact factors to be fully considered and the merged methods are still needed to be further studied.

Previous studies have been able to fully identify regional ship collision risk to some extent, and the existing regional ship collision risk identification method has some limitations. In particular, the impact of other ships to the OS is not taken into account when calculating the collision risk. In the encounter cluster, ships within a certain range from the OS will have a certain impact on the collision avoidance action taken by the OS. This is the challenge that we tackle in our work, so this paper proposes a new regional collision risk assessment method considering characteristics of the aggregation density (AD) of surrounding ships in the clusters of encounter ships (CES). The AD-based ship collision risk proposed in this paper can fully reflect the AD influences by other ships around the OS and can reflect the closeness and complexity of maritime traffic. Firstly, the dynamic and static data of ships are extracted from the automatic identification system (AIS) data of ships, and then the DBSCAN clustering algorithm is used to obtain the CES occurring at the specified encounter distance. Secondly, the influencing factors of DCPA, TCPA and AD between ships in the CES are calculated. Then, the collision risk function of each factor is constructed in each encounter ship cluster, and the collision risk matrix between ships in the CES is obtained by weighted calculation. Finally, the feasibility of the proposed method is verified by an example in the waters of Xiamen Bay, and a comparative analysis has been conducted.

The paper is organised as follows. Section 2 addresses the multi-ship encounter situation recognition based on density clustering. Section 3 elaborates the modeling process of regional collision risk intra-CES. In Section 4, the experimental case study in Xiamen waters has been conducted. The discussions about the models are given in Section 5, followed by the conclusion and future work in Section 6.

2. Multi-ship encounter situation identification based on density clustering

When calculating the collision risk between ships in restricted waters, it is necessary to calculate the risk of the ship and all other ships around it. Then, the collision risk of the ship itself is synthesised, so that every ship must traverse all the ships in the study area. To reduce the burden of calculation, and improve the efficiency of ship collision risk identification, the ships with actual collision risk are identified from a large number of sailing ships. As shown in Figure 1, to calculate the effective collision risk, it is necessary to extract the CES that occurs from a large number of sailing ships. The calculation of the collision risk of a ship will be carried out between the ships in the cluster where the ship is located, without traversing all AIS data. The DBSCAN can get different shapes of clusters without specifying the number of clusters in advance, and only two parameters are needed, which is easy to analyse and obtain in the clustering process.

The DBSCAN algorithm can divide ships in the study area into three categories: The first kind is the core ship, ship's (x_i) neighborhood ε contains at least *MinPts* ships, namely $N_{\varepsilon}(x_i) \geq$ Minpts. The second category is boundary vessels, with fewer vessels in the neighborhood ε of vessel (x_i) than *MinPts*, namely $N_{\varepsilon}(x_i) <$ Minpts, but the ship (x_i) is in the neighborhood ε of the core ship. The third category is noise, which is neither a core ship nor a boundary ship. In the algorithm, the values

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

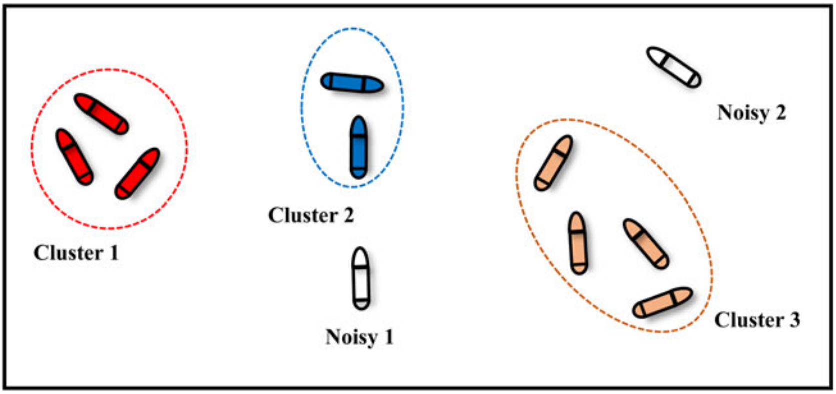

Figure 1. *Diagram of clusters of encounter ships.*

of ε and MinPts must be determined. The specific calculation process of the DBSCAN-based CES identification algorithm is as follows:

Input:
D: a dataset containing n objects
ϵ: distance parameter
MinPts: neighborhood density
Output: A set of clusters based on density
Methods:
1. Mark all objects as unvisited;
2. DO
3. Randomly select an unvisited object p;
4. Mark p for visit;
5. If p has at least MinPts objects in its ϵ neighborhood;
6. Create a new cluster C and add p to C;
7. Sets of objects in ϵ neighborhoods where N is p
8. For each point in N p
9. If p′ is unvisited;
10. Mark p′ for visit;
11. If p′ has at least MinPts objects in its ϵ neighborhood, add them to N;
12. If p′ is not a member of any cluster, add p′ to C;
13. End for;
14. OutputC;
15. Else marker p is noise;
16. Until there is no object marked unvisited;

---

3. Regional collision risk modeling in CES

3.1. Calculation of DCPA and TCPA

Firstly, the collision risk between ships in the encounter cluster is calculated. This paper uses the method of collision risk based on AIS data. Since the AIS data are received intermittently and the interval time is determined according to the ship type and ship speed, the AIS data of all ships in the study area require interpolation calculation. In this paper, the cubic spline interpolation is used to process AIS data

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

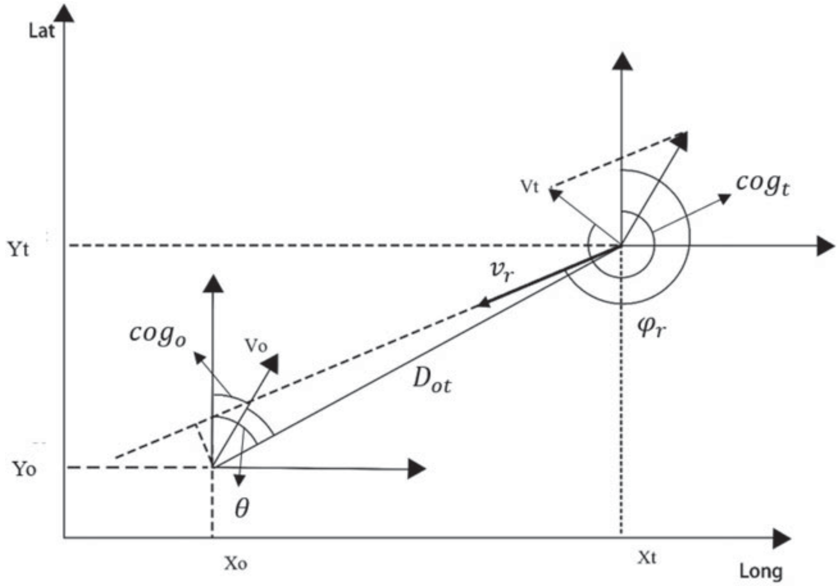

Figure 2. *The relative position of ships.*

to obtain AIS data of ships at the same time (Kouibia and Pasadas, 2012). At the same time, there are many small ships not equiped with AIS in the restricted waters; in this situation, the ship position data should be obtained by radar, and our method can also process the ship position, speed and course from the radar data.

After obtaining accurate AIS data, the DCPA and TCPA between ships are calculated. For any pair of ships, one of them is identified as OS, and the ship parameters are $(\text{long}_o, \text{lat}_o, \text{sog}_o, \text{cog}_o)$, and the TS parameter is $(\text{long}_t, \text{lat}_t, \text{sog}_t, \text{cog}_t)$. The DCPA and TCPA calculation method between ships in the encounter cluster is as follows (Figure 2):

The relative distance between the two ships is D_{ot}. Because the position information in AIS data is latitude and longitude, in order to obtain a more accurate distance between the two ships, this paper uses the Mercator method to calculate the real distance. The Mercator method is the longitude-difference calculation method, based on the characteristics of Mercator projection with equal angle and loxodrome as a straight line. The specific formulas for calculating the distance between the two ships are

$$\text{MP} = 7915{\cdot}70447\,\text{lg}\left(\tan\left(\frac{\pi}{4} + \frac{\text{lat}}{2}\right)\left(\frac{1 - e\sin(\text{lat})}{1 + e\sin(\text{lat})}\right)^{e/2}\right) \tag{3.1}$$

$$\text{DMP} = \text{MP}_t - \text{MP}_o \tag{3.2}$$

$$C = \arctan\left(\frac{D\text{long}}{\text{DMP}}\right) \tag{3.3}$$

$$S = \text{Dlat}{\cdot}\sec(C) \tag{3.4}$$

Where *MP* is meridianal part, *lat* is latitude, e is the eccentricity of Earth and e in index *e/2* is Euler number, $e = 2{\cdot}7182828459045$ and *DMP* is the difference of meridianal parts between the two ships. The *Dlong* is the longitude difference between the two ships, *Dlat* is the latitude difference between the two ships and *S* is the distance between the two ships. The unit is *n* mile.

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

True azimuth of the target ship θ:

$$\begin{cases} \Delta x = \text{long}_t - \text{long}_o \\ \Delta y = \text{lat}_t - \text{lat}_o \end{cases} \tag{3.5}$$

$$\theta = \begin{cases} \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right) & \Delta x \geq 0, \Delta y > 0 \\ \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right) + \pi & \Delta x \geq 0, \Delta y < 0 \\ \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right) + \pi & \Delta x < 0, \Delta y < 0 \\ \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right) + 2\pi & \Delta x < 0, \Delta y > 0 \\ \frac{\pi}{2} & \Delta x > 0, \Delta y = 0 \\ \frac{3\pi}{2} & \Delta x < 0, \Delta y = 0 \end{cases} \tag{3.6}$$

The speed component of the OS on the coordinate axis:

$$\begin{cases} v_{ox} = v_o \cdot \sin(\text{cog}_o) \\ v_{oy} = v_o \cdot \cos(\text{cog}_o) \end{cases} \tag{3.7}$$

The speed component of the TS on the coordinate axis:

$$\begin{cases} v_{tx} = v_t \cdot \sin(\text{cog}_t) \\ v_{ty} = v_t \cdot \cos(\text{cog}_t) \end{cases} \tag{3.8}$$

The relative velocity (v_r) of the two ships and their components on the two axes:

$$\begin{cases} v_{rx} = v_{tx} - v_{ox} \\ v_{ry} = v_{ty} - v_{oy} \end{cases} \tag{3.9}$$

$$v_r = \sqrt{(v_{rx})^2 + (v_{ry})^2} \tag{3.10}$$

Relative course of two ships (φ_r):

$$\varphi_r = \begin{cases} \tan^{-1}\left(\frac{v_{rx}}{v_{ry}}\right) & v_{rx} \geq 0,\ v_{ry} > 0 \\ \tan^{-1}\left(\frac{v_{rx}}{v_{ry}}\right) + \pi & v_{rx} \geq 0,\ v_{ry} < 0 \\ \tan^{-1}\left(\frac{v_{rx}}{v_{ry}}\right) + \pi & v_{rx} < 0,\ v_{ry} < 0 \\ \tan^{-1}\left(\frac{v_{rx}}{v_{ry}}\right) + 2\pi & v_{rx} < 0,\ v_{ry} > 0 \\ \frac{\pi}{2} & v_{rx} > 0,\ v_{ry} = 0 \\ \frac{3\pi}{2} & v_{rx} < 0,\ v_{ry} = 0 \end{cases} \tag{3.11}$$

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

The DCPA can be calculated as

$$\mathrm{DCPA} = D_{ot}\sin(\varphi_r - \theta - \pi) \tag{3.12}$$

If the TS passes from the starboard side of the OS when it reaches the distance to the closest point of approach if the DCPA is negative, the TS is in the head direction of the ship, but if the DCPA is positive, the TS is astern of the OS. If the TS passes from the port side of the OS when it reaches the distance to the closest point of approach if the DCPA is negative, the TS is in the stern direction of the ship, and if the DCPA is positive, the TS is in the bow direction of the ship.

The TCPA is calculated as

$$\mathrm{TCPA} = \frac{D_{ot}\cos(\varphi_r - \theta - \pi)}{v_r} \tag{3.13}$$

If TCPA > 0, the two ships have not yet entered the closest point of approach point. If TCPA < 0, the two ships have passed the closest point of approach.

3.2. *Modeling of spatial and temporal ship collision risk*

There are many methods to study ship collision risk. For example, Mou et al. (2010) incorporated basic risks into the risk measurement of the studied waters and expressed the relationship between DCPA and TCPA and risks in a negative exponential form. Zhen et al. (2017) considered the impact of DCPA and TCPA on risk and used the negative index method to calculate the collision risk. In addition to collision risk related to DCPA and TCPA. In this paper, we incorporate the safe distance (D) and time (T), to the traditional negative exponential function from the aspect of collision avoidance, to quantify the relationship between collision risk and DCPA and TCPA.

The spatial collision risk caused by DCPA can be expressed as

$$\mathrm{CRI_{DCPA}} = k_1 e^{\alpha_1 \cdot \frac{\mathrm{DCPA}}{D}} \tag{3.14}$$

where D is the safe distance between two ships, and α_1 and k_1 are the regulating coefficients.

The temporal collision risk caused by TCPA can be expressed as

$$\mathrm{CRI_{TCPA}} = k_2 e^{\alpha_2 \cdot \frac{\mathrm{TCPA}}{T}} \tag{3.15}$$

where T is the safest time of the two ships, the shortest time for the two ships to avoid collision by manipulating the ship, and α_2 and k_2 are the adjustment coefficients.

3.3. *Collision risk modeling of ship aggregate density in clusters of encounter ships*

When a ship encounters another ship at sea, it is often necessary to adjust the navigation state of the ship in order to avoid the urgent situation of a collision. Due to the large inertia of a ship, taking speed change for collision avoidance will cause more waste of resources, so most collision avoidance operations are to change the ship course. The situation of surrounding ships should be considered when changing course. The smaller the distance between the surrounding ship and the OS is, the more factors must be considered when the OS changes its course. According to the *International Regulation for the Prevention of Collision at Sea* (COLREGS), allowing give-away ships to turn starboard when encountering to avoid collisions in a crossing situation, it can be concluded that under the same circumstances, the ships have a greater risk of collision on the right side than on the left side of the ship. The collision risk of the ship's aggregation coefficient (AC) is the aggregation risk of OS by a single ship in the CES. The AD-based collision risk is the result of the collision risk synthesis of the AC of the ship and the surrounding ships, which reflects the density of the ships around the ship. In this paper, the smooth circular ship domain by Davis et al. (1980) is used to calculate the ship domain overlap index, which is used as the influencing

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

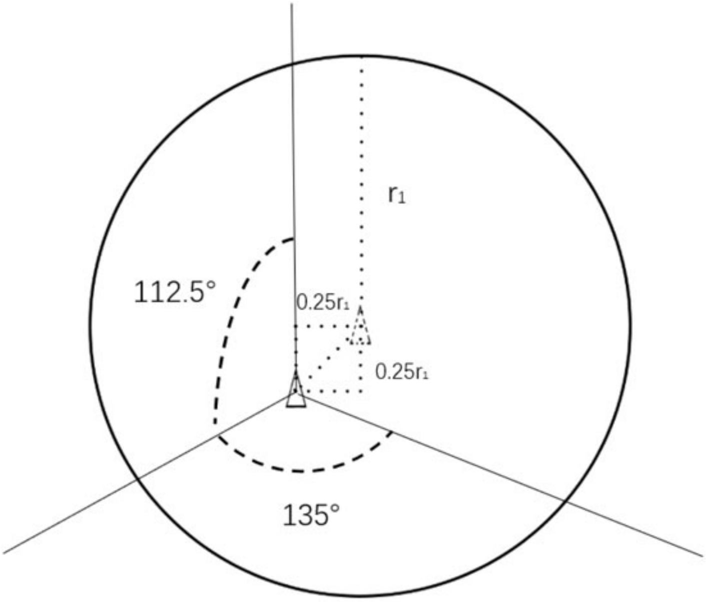

Figure 3. *The ship domain used in this paper.*

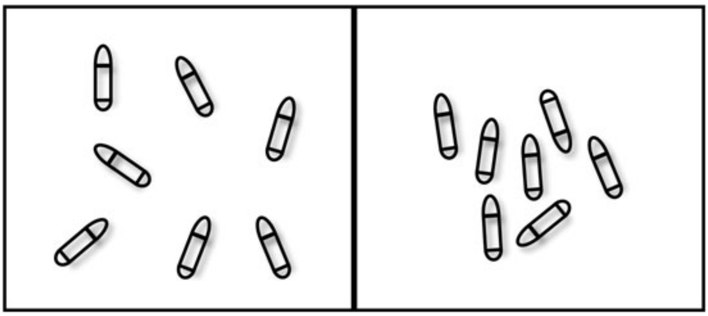

Figure 4. *Diagram of ship density.*

factor of the aggregation coefficient. Using the ship domain overlap index as an influencing factor not only represents the influence of the distance between ships but also reflects that the risk of the right side of own ship is higher than that of the ship on the left side of the ship under the same conditions. The center of the ship domain is located on the upper-right of the ship, and the specific parameters are shown in Figure 3.

The ship density refers to the number of ships in a certain instantaneous unit area of water. This can also reflect the ship aggregation in the study area, but the effect of other ships on the ship can influence the evasive actions. As shown in Figure 4, although the density of ships in the area is the same, the aggregation of each ship in the area is not the same, so the density of ships cannot effectively represent the risk of the location of the ship in the unit water area.

It is assumed that the radius of the ship domain is r_o, the ship domain of the target ship is r_t, and the distance between the virtual positions of the two ships is D_{ot}, and the ship domain overlap index can be expressed as (Liu et al., 2019)

$$\mathrm{SDOI}_{ot} = \frac{D'_{ot}}{r_o + r_t} \tag{3.16}$$

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

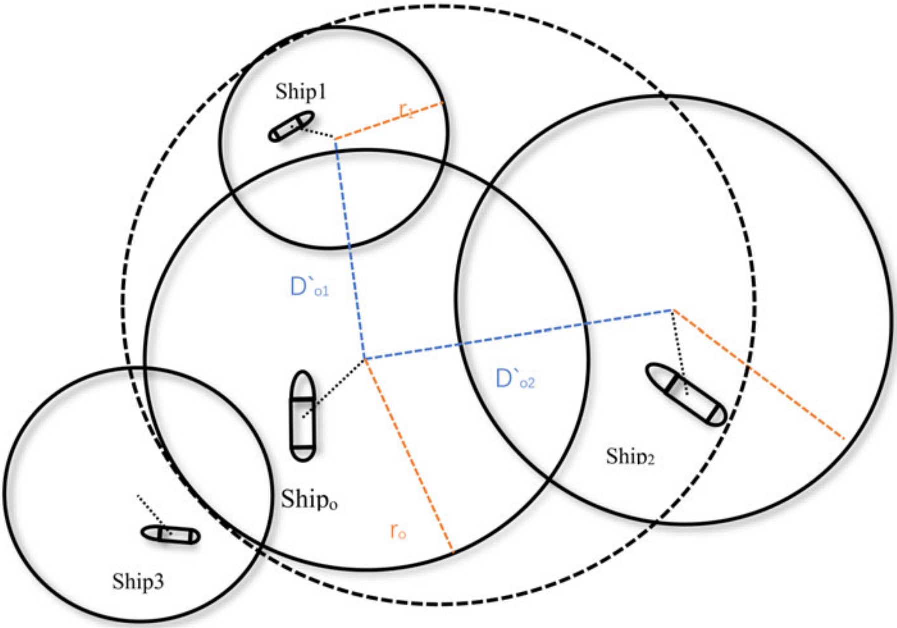

Figure 5. *The ship domain and Arena.*

The AC essentially reflects the influence of other ship's state on the ship's manipulation. The smaller the SDOI, the greater the impact of the other ships on the own ship, so the SDOI of the own ship and the other ships are negatively correlated with the collision risk of the ship's AC. Combined with spatial collision risk and temporal collision risk, this paper also uses the negative exponential to express the relationship between SDOI and ship AC collision risk. The details are

$$\mathrm{CRI}_{\mathrm{AC}_{ot}} = k_3 e^{\alpha_3 \mathrm{SDOI}_{ot}} \tag{3.17}$$

According to Davis's research, the scope of the threat to the officers is further than the ship domain boundary, and a super-domain is determined. The scope of the super-domain is that the officers start to take action to avoid the distance from other ships when an urgent situation occurs, also known as the 'Arena'. Since the AC is the influence on the ship maneuvering, it is necessary to calculate the ship aggregation within the dynamic range. Figure 5 shows the Arena model adopted in this paper and the ships involved in calculating the AC around the ship.

The dashed line circle represents the Arena and the solid line circle represents the ship domain. The Arena is tangent to the ship domain and the tangent points are collinear with two centers. In Figure 5, only ship$_1$ and ship$_2$ are required to calculate the AC of the ship. Ship$_3$ is not in range of there Arena, and here we assume there is no effect on the OS's manipulation.

In Figure 6, the OS is indicated in center, and the relative positions of the other ships are shown around in four directions. Assuming that the length of the own ship and other ship are both 100 m, and Figures 7–10 show the AC-based collision risk around OS based on Equations (3.16) and (3.17). These figures show that at different encounter angles, even if the distance from the OS is the same, the collision risk of the AC is different.

When there is only one ship within the OS Arena, the aggregation coefficient is equal to the AD. When there are many ships within the OS Arena, it is necessary to synthesise the aggregation coefficient of all ship pairs to obtain the comprehensive AD. To reduce the interference of human factors on the AD calculation, and fully reflect that the closer the ship is, the greater the risk of AD is. The calculation

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

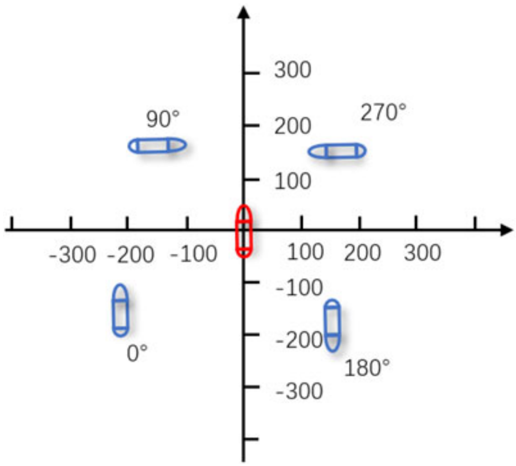

Figure 6. *Diagram of relative course between OS and other ships.*

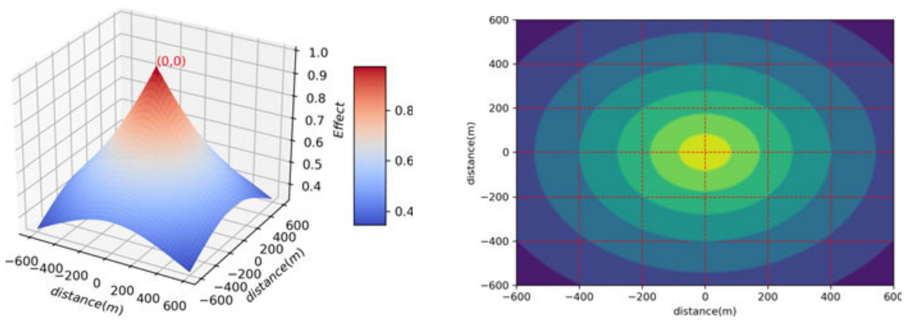

Figure 7. *The relative course of 0°.*

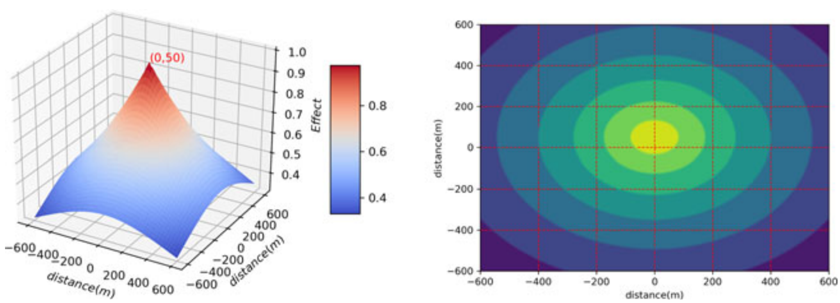

Figure 8. *The relative course of 90°.*

formula of AD-based collision risk from all the influences in the CES is

$$\mathrm{ADRI}_i = 1 - \prod_{j \neq i} (1 - \mathrm{CRI}_{\mathrm{AC}_{ij}}) \tag{3.18}$$

For example, when the AC of three ships in the dynamic range of the ship is 0·6, 0·4 and 0·2, respectively, the calculated AD is 0·808. In principle, there should be no other ships in the Arena range

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

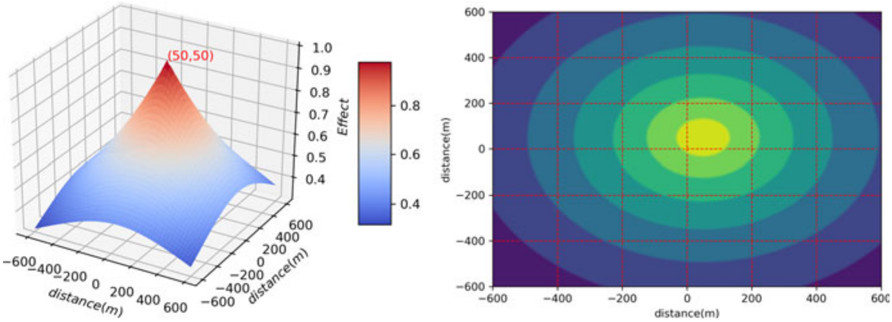

Figure 9. *The relative course of 180°.*

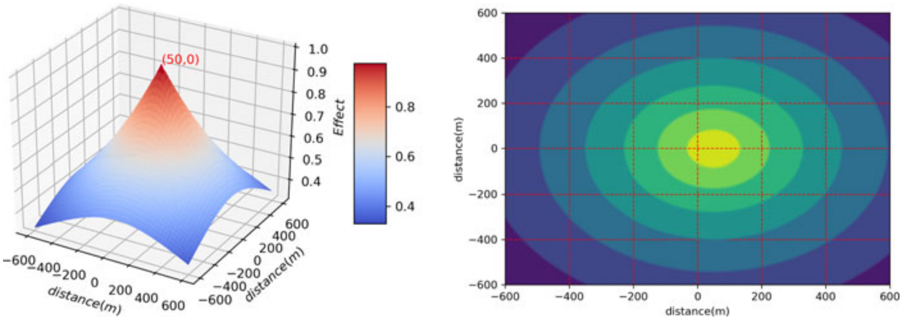

Figure 10. *The relative course of 270°.*

of the OS, so the corresponding aggregation density of the three ships will be large. This formula can fully reflect the influences of AD in the CES.

3.4. Regional collision risk modeling of encountering ship clusters

The AD of CES can reflect the influence of the surrounding ship on the ship's motion. When there is no ship within the OS Arena, the ship's AD is 0, and the ship's freedom is high. The traditional density calculation method cannot fully reflect the freedom of navigation of the ship. The method of combining ship AD, DCPA and TCPA to identify regional ship collision risk has advantages over the existing method.

All ships sailing at sea are a dynamic system, and the ships in the system interact with each other. Taking ship A as the research object, the collision risk of ship A and ship B cannot only consider the mutual influence between ship A and ship B but also consider the influence of other ships in the system on ship A. The farther the distance is, the smaller the influence is. This paper is based on the Arena of ship A, and the influence of ships outside Arena on ship A is negligible. This effect is expressed by the ship's AD. The $ADRI_i$ is the combination of AD-based risk. The ship collision risk is determined by spatial collision risk, temporal collision risk, and AD-based collision risk. Space and temporal collision risk belong to the influence between ship pairs, and AD-based collision risk belongs to the influences of all the ships in CES on OS. Collision risk can be identified by the linear combination of the three aspects of collision risk, then the specific formula of collision risk of ship_j and ship_i in CES can be expressed as

$$\mathrm{RI}_{ij} = \omega_1 \mathrm{CRI}_{\mathrm{DCPA}} + \omega_2 \mathrm{CRI}_{\mathrm{TCPA}} + \omega_3 \mathrm{ADRI}_{\mathrm{i}} \tag{3.19}$$

where the ω_1, ω_2 and ω_3 are the weights of spatial collision risk, temporal collision risk and AD-based collision risk, respectively.

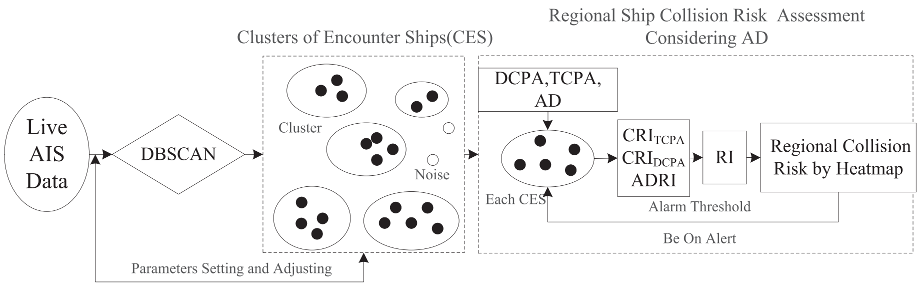

Figure 11. *The regional collision risk identification process.*

After calculating the collision risk, the researchers studied it from a deeper perspective. Zhang et al. (2017) used expert judgement to sort ship collision risks, and experts further reviewed these risks. There are few studies on the interaction between ships in the system to determine the collision risk of a single ship. Liu et al. (2019) applied the Shapley value method in game theory to determine the weight of each collision risk, and finally accumulated the collision risk of a single ship. Feng et al. (2018) constructed the risk matrix and calculated the mean value of the risk matrix as the collision risk of a single ship. These methods have some limitations to obtain the ship collision risk. For example, the risk of collision between OS and several other ships is 0·5, so whatever method is used to set the weight, the final risk is 0·5. According to common sense, if many ships influence each other at this time, the risk of the ship must be greater than these risks. To overcome this limitation, this paper proposes a method to calculate the total risk of ship collision without weight setting. According to the Equation (3.19), the value range of RI_{ij} is $0 \sim 1$. So the specific formula is

$$\mathrm{RI}_i = 1 - \prod_{j \neq i} (1 - \mathrm{RI}_{ij}) \tag{3.20}$$

Finally, the collision risk of each ship is identified from the CES, and the obtained collision risk is displayed on the map in the form of a heatmap. This not only reminds the crew of the dangerous degree of their environment, but also makes judgements on areas with high collision risk VTS operators, and reminds the ship to pay attention to navigation safety. Figure 11 shows the whole steps for identifying regional collision risks considering the AD.

4. Case study

To verify the effectiveness of the proposed model, the AIS data of Xiamen Bay was selected for the case study. The Xiamen Port is the main coastal port of China, and the ship traffic volume is large in the daytime. According to the data, there are 151 ships at this time; the specific location is shown in Figure 12.

4.1. Parameter selection of DBSCAN algorithm

Firstly, the DBSCAN algorithm is used to cluster the ship position points. The *MinPts* of DBSCAN algorithm must be equal or more than two, because the ship encounter are formed by at least two ships. Secondly, clustering results cannot produce too much noise, which will affect data integrity and reduce risk assessment. Finally, the ships in each cluster cannot be too numerous or they will lose the meaning of clustering. To get better clustering results, we have carried out many experiments on setting two parameters of *MinPts* and *Eps*, which are shown in Figure 13(a). From the clustering experiments, we know that the larger the *Eps* parameters are and the smaller the *Minpts* parameters, the smaller the amount of noise data will be.

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

Figure 12. *The real-time AIS data near Xiamen Port.*

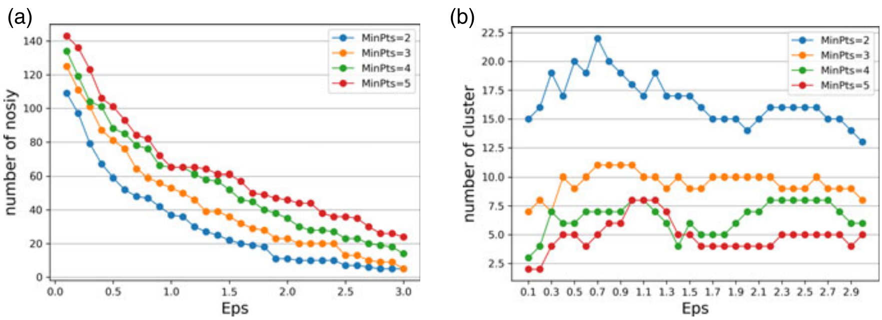

Figure 13. *(a) Relationships between Eps, minPts and noise quantitative, (b) relationship between Eps, minPts and cluster number.*

According to Figure 13(b), by setting the same *Eps*, the smaller the *MinPt*s are, the more clusters are obtained. Too many clusters will seriously segment the study area, resulting in large collision risk error. Considering the amount of generated noises and the number of clusters, *MinPts* is set to 3, and Eps is set to 1·2 nm. There are nine CES within the Xiamen Bay waters, as shown in Figure 14.

4.2. Determine the model parameters

The parameters are set from the empirical cases and the expert's experience. Li (2019) suggested that the influences from other ships away from 1 n mile can be ignored for 20 m ships, and other ships away from 3 n miles can be ignored for 400 m ships. We use linear interpolation to obtain the negligible DCPA for ships of different lengths. In Table 1, the term '4L/1850' indicates that the DCPA is the size of the ship's domain. When DCPA is greater than the ignored distance, the risk calculated by DCPA can be ignored. When the DCPA is equal to the size of the ship doamin, the collision risk value is large. The TCPA and DCPA based collision risk calculation are are designed by Zhen et al. (2017), when TCPA is greater than 30 min, the calculated temporal collision risk is less than 0·1. When TCPA is less than 2 min, the calculated temporal collision risk is greater than 0·8. When SDOI is 0·1, the AD is about 0·9. When SDOI is 1, the AD is about 0·3. Figure 15 shows the effects of DCPA, TCPA and density on CRI.

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

Figure 14. *The DBSCAN clustering results in Xiamen waters.*

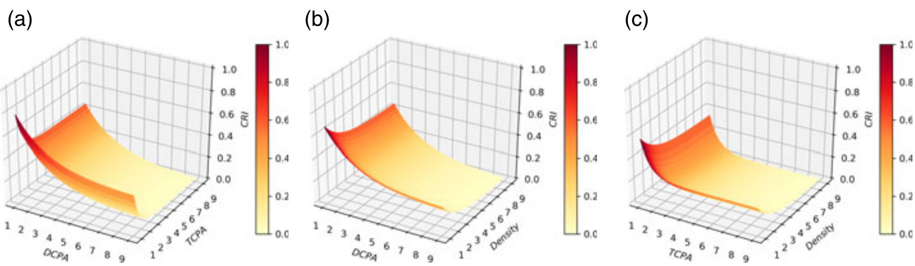

Figure 15. *(a) The effect of DCPA and TCPA on CRI, (b) The effect of DCPA and density on CRI, (c) The effect of TCPA and density on CRI.*

Table 1. *Parameters calculation of* CRI_{DCPA}.

DCPA(nm)	CRI_{DCPA}	α_1	k_1
>(170+L)/190 <4L/1852	≈0·1 ≈0·9	−0·39	1·334

4.3. Calculation of ship collision risk

The safety distance in Equation (3.14) is four times the ship length, and the safety time in Equation (3.5) is 10 min. The weight of spatial collision risk, time collision risk and AD collision risk are 1/3. To verify the effectiveness of this method, the 9th clusters of encounter ships are selected for analysis. There are five ships in this CES, and the specific parameters of these five ships are shown in Table 2.

The DCPA and TCPA matrix are obtained as follows:

$$\begin{bmatrix} 0 & -0{\cdot}025 & -0{\cdot}407 & -0{\cdot}927 & 0{\cdot}994 \\ -0{\cdot}025 & 0 & 0{\cdot}130 & -0{\cdot}685 & 0{\cdot}656 \\ -0{\cdot}407 & 0{\cdot}130 & 0 & -0{\cdot}496 & 0{\cdot}642 \\ -0{\cdot}927 & -0{\cdot}685 & -0{\cdot}496 & 0 & -0{\cdot}057 \\ 0{\cdot}994 & 0{\cdot}656 & 0{\cdot}642 & -0{\cdot}057 & 0 \end{bmatrix} \tag{4.1}$$

Table 2. *Specific parameters of ships.*

Sog	Cog	Long	Lat	Length
10·4	334·2	118·1	24·38571	129
8	358	118·1068	24·384	80
7·4	310	118·108	24·38728	70
19·4	140	118·1168	24·39089	100
18·9	0	118·1187	24·383	75

$$\begin{bmatrix} 0 & -50{\cdot}2 & 2{\cdot}4 & -0{\cdot}61 & -1{\cdot}7 \\ -5{\cdot}2 & 0 & 1{\cdot}5 & 0{\cdot}15 & 0{\cdot}21 \\ 2{\cdot}4 & 1{\cdot}5 & 0 & -4{\cdot}2 & -0{\cdot}002 \\ -6{\cdot}1 & 0{\cdot}15 & -0{\cdot}42 & 0 & 0{\cdot}80 \\ -1{\cdot}7 & -0{\cdot}21 & -0{\cdot}002 & 0{\cdot}80 & 0 \end{bmatrix} \tag{4.2}$$

According to Equations (3.16), (3.17), and (3.18), the collision risk of vessel AD can be obtained as

$$\begin{bmatrix} 0{\cdot}381 & 0{\cdot}522 & 0 & 0 & 0 \end{bmatrix} \tag{4.3}$$

According to Equation (3.19), the collision risk matrix is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{\cdot}79 & 0 & 0 \\ 0 & 0{\cdot}602 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0{\cdot}709 \\ 0 & 0 & 0 & 0{\cdot}695 & 0 \end{bmatrix} \tag{4.4}$$

According to Equation (3.10), the collision risk of a single ship in CES_5 is

$$\begin{bmatrix} 0 & 0{\cdot}79 & 0{\cdot}602 & 0{\cdot}709 & 0{\cdot}695 \end{bmatrix} \tag{4.5}$$

There are five ships displayed on the map as shown in Figure 16, the arrow length represents the relative speed of the ship. The collision risk of each ship has been influenced by other 4 ships in the 9th CES, the OS denotes each ship when calculate the collision risk from other ships in the same CES.

Due to the different lengths and relative positions of ships, the obtained risk matrix is asymmetric. $Ship_1$ and $ship_2$ have passed the CPA, so the collision risk of the two ships is 0. The $ship_1$ and $ship_3$ will encounter a distance of 0·407 n miles in the later time, which is not within the Arena range of $ship_1$ and $ship_3$, so the collision risk of $ship_3$ and $ship_1$ is 0. $ship_1$ and $ship_2$ have other ships within their Arena range. These ships will have impacts on the collision avoidance of the ships, and the AD of these ships must be considered. Based on the model proposed, the risks of $ship_1$ and $ship_2$ due to the aggregation of surrounding ships are calculated to be 0·381 and 0·522, respectively. The situation of $ship_4$ and $ship_5$ belong to the crossings situation, and the nearest encounter distance is within the dynamic range of the ship, but there is no ship within the dynamic range, so only spatial collision risk and temporal collision risk are considered. The collision risks of five ships in CES 9th are calculated as follows: 0, 0·79, 0·602, 0·709 and 0·695, respectively, as shown in Figure 17.

After calculating the collision risk of all single ships in the CES, we depict the whole spatial distribution of regional collision risk by heatmap on the nautical chart, as shown in Figure 18, the colours blue to red represents the risk from low to high. The depth of the colour depends on the magnification of the map: the higher the magnification, the more blurred the risk display, and the smaller the magnification, the more accurate the risk display. This is done so that maritime traffic management officers can find areas of higher risk quickly based on our proposed method in this paper.

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

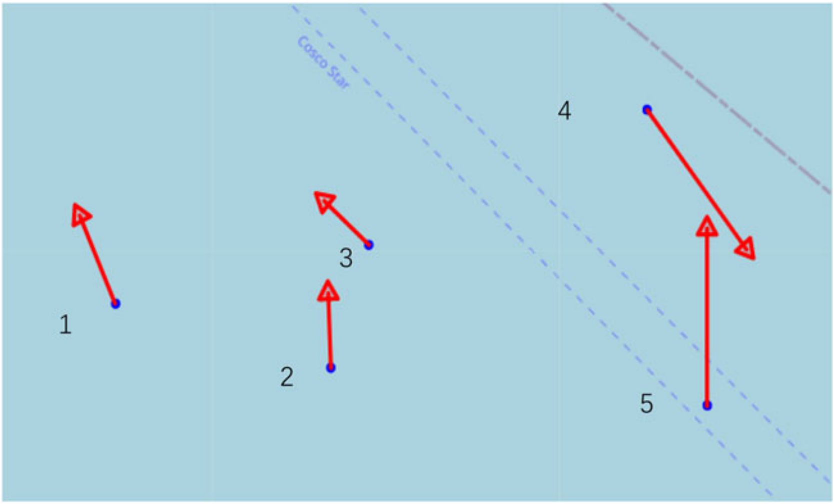

Figure 16. *The spatial distribution of ships in 9th cluster of encounter ships..*

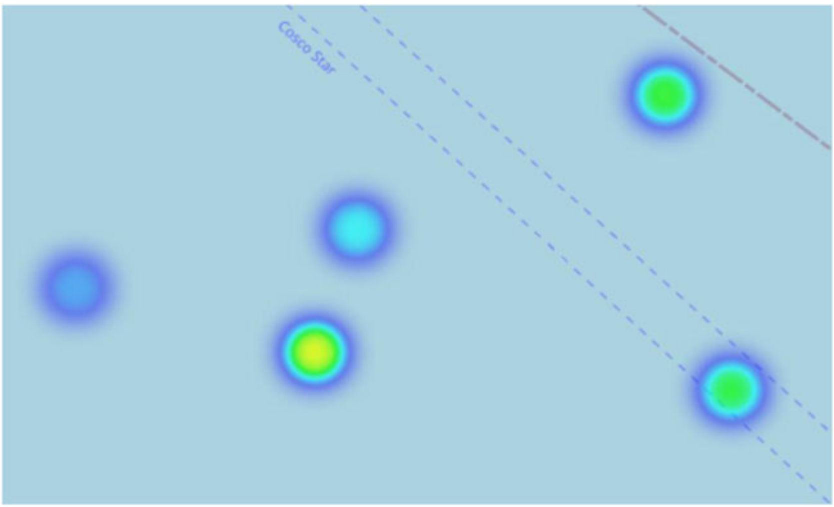

Figure 17. *The regional collision risk of 9th cluster of encounter ships.*

5. Discussion

To verify the feasibility and effectiveness of the proposed method over the existed methods, here we simulate several groups of encounter ships and calculate the regional collision risk compared with other methods. First, we simulate a scenario where two ships encounter a situation. To study conveniently and intuitively, the ship domain of the two ships is defined as 1 n mile, and the ship boundary radius is 2·828 nautical miles. As shown in Figure 19, the occurrence of a red ship and a green ship is a large-angle crossing, and collision risk occurs if the ship state is not changed. The solid line circle belongs to the field of ships, and the dashed lines belong to the dynamic boundary of ships. These illustrate the ship domain and the movement boundary position for the ship's last time location. At this time, the model is compared with the collision risk model of Zhen et al. (2017) and Liu et al. (2019).

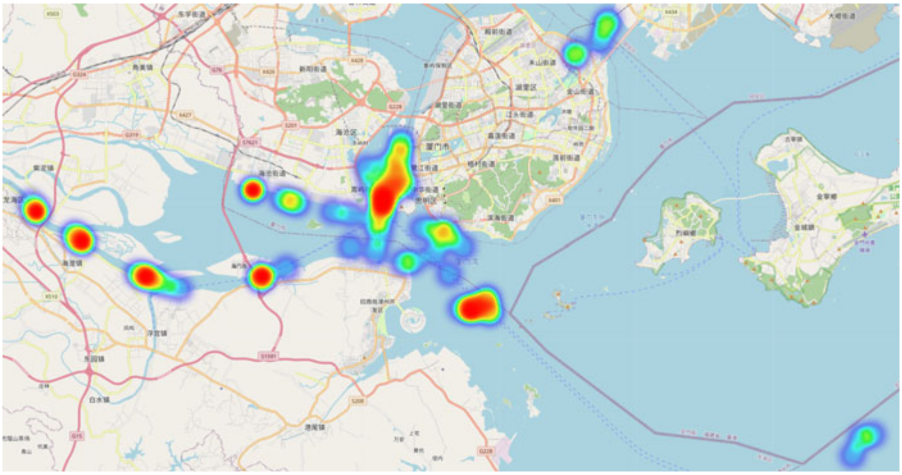

Figure 18. The distribution of regional ship collision risk within Xiamen Bay.

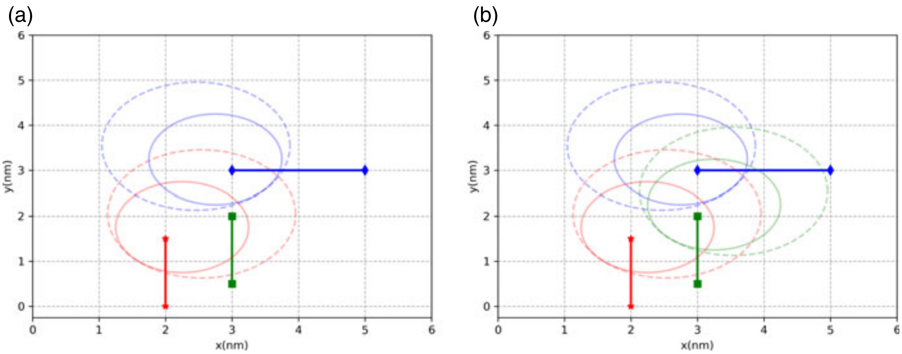

Figure 19. The simulated cases. (a) case of two ships, (b) case of three ships.

Due to the use of a smooth circular ship domain, the characteristics of the domain model are to highlight the ship's starboard, which is in accordance with the rules of ship collision avoidance. There are only two ships in the sea area of the left figure, and the nearest encounter distance of the two ships is relatively small, so measures will be taken to avoid the occurrence of collision risk. The right figure shows the encounter of three ships. The state of red and blue ships is the same as that of the left figure, but there is one more green ship. Because of the large collision risk between the red and blue ships, most of the collision avoidance measures are to change the course, so there are ships around the ship that will have an impact on the operation of the ship. The density effect considered in this paper can fully reflect the risks in the sea area. According to expert experience, it can be concluded that the risk of red and blue ships on the right map is significantly higher than that on the left map. According to the calculation results of the collision risk model in Liu et al. (2019) and Zhen et al. (2017), the risk of these two cases are the same, but the regional collision risk calculated by the proposed model are 0·573 and 0·662, respectively, it shows the advantage of proposed method in this paper. The calculated risk value is not specific, but the relative size between risks can reflect the description of the calculated risk value in the proposed method.

This paper considers the impact of DCPA, TCPA and AD. In fact, in addition to these factors, the behaviour of the officers and the speed ratio of the ship can be considered. When it is located in particularly narrow waters, the DCPA or TCPA of ships sailing, in turn, will sometimes be quite small, which has a great influence on this model. At this time, it is necessary to change the parameters or reconstruct the collision risk model for calculating particularly narrow waters. The specific collision

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

risk results are of no practical significance, and it is only meaningful when compared with a set value. Therefore, a large number of cases are needed to study the threshold setting. Setting the value size will be subjective, so we can consider a method to reduce subjectivity. It can also be calculated by the rate of change of ship collision risk, when the value is growing larger and larger, indicating that the increase of risk must change the navigation state, when the value is grows smaller and smaller, indicating that the risk is reducing need not change the navigation state. This can solve the calculation problem of collision risk in particularly confined waters.

6. Conclusions

In this paper, a novel regional ship collision risk assessment method that considers the AD of encounter ships is proposed. The collision risk between ships is composed of spatial collision risk, temporal collision risk and AD-based collision risk. The AD-based collision risk is judged by the overlap of ship areas between ships, and then we fuse the regional collision risk of each ship according to the collision risk of the OS and all TS in the CES. To prove the effectiveness of the method, the data of the sea area near Xiamen are selected for verification and the effects of ship length, speed, course, and relative position are fully considered. The results show that the proposed method can effectively display the distribution of regional ship collision risk over the existed methods, and can help traffic management officers and crews to observe the high risky sea area directly.

In the research of regional collision for the multi-ship situation, our paper considers the influences of DCPA, TCPA, AD, the ship length, heading, relative distance and density. The contribution of this study is that when calculating the ship collision risk, not only is the influence between ship pairs considered but the influence of the ship in the CES should also be considered. Finally, the spatial distribution of regional collision risk by heatmap can be applied to regional collision risk detection and management. Considering ship geometric collision probability in future work is more persuasive to synthesise single-ship collision risk. At the same time, for the crowded channel, the threshold alarm can be set and the change rate of collision risk can be considered to improve the assessment efficiency in future work.

Acknowledgments. The research described in this paper is supported by the National Natural Science Foundation of China (51879119, 52001134); Natural Science Foundation of Fujian Province (2020J01661); Youth Science and Technology Project of Education Department of Fujian Province (JAT190293); Doctoral Research Start-up Fund of Jimei University (ZQ2019014); Open Foundation of Hubei Provincial Key Laboratory of Inland River Navigation Technology (NHHY2020001) and Open Foundation of Jimei University (HHXY2020005). We are grateful to the anonymous reviewers whose comments and suggestions have contributed to improving the quality of research described in this paper.

References

Ahn, J. H., Rhee, K. P. and You, Y. J. (2012). A study on the collision avoidance of a ship using neural networks and fuzzy logic. *Applied Ocean Research*, **37**, 162–173.

Bukhari, A. C., Tusseyeva, I. and Kim, Y. G. (2013). An intelligent real-time multi-vessel collision risk assessment system from VTS viewpoint based on fuzzy inference system. *Expert Systems with Applications*, **40**(4), 1220–1230.

Chen, P., Li, M. and Mou, J. (2021). A velocity obstacle-based real-time regional ship collision risk analysis method. *Journal of Marine Science and Engineering*, **9**(4), 428.

Davis, P. V., Dove, M. J. and Stockel, C. T. (1980). A computer simulation of marine traffic using domains and arenas. *The Journal of Navigation*, **33**(2), 215–222.

Fan, C., Wróbel, K., Montewka, J., Gil, M., Wan, C. and Zhang, D. (2020). A framework to identify factors influencing navigational risk for maritime autonomous surface ships. *Ocean Engineering*, **202**, 107188.

Fan, S., Zhang, J., Blanco-Davis, E., Yang, Z. and Yan, X. (2020). Maritime accident prevention strategy formulation from a human factor perspective using Bayesian networks and TOPSIS. *Ocean Engineering*, **210**, 107544.

Feng Y., X., Zhang W., B., Qi Y., *et al.*Ship Collision Risk Assessment in Open Waters Based on AIS [C]. *World Transport Congress* 2018, 13, Beijing.

Fujii, Y. and Shiobara, R. (1971). The analysis of traffic accidents. *The Journal of Navigation*, **24**(4), 534–543.

Fujii, Y. and Tanaka, K. (1971). Traffic capacity. *The Journal of Navigation*, **24**(4), 543–552.

Gang, L., Wang, Y., Sun, Y., Zhou, L. and Zhang, M. (2016). Estimation of vessel collision risk index based on support vector machine. *Advances in Mechanical Engineering*, **8**(11), 1687814016671250.

https://doi.org/10.1017/S0373463321000849 Published online by Cambridge University Press

Goerlandt, F. and Montewka, J. (2014). A probabilistic model for accidental cargo oil outflow from product tankers in a ship–ship collision. *Marine Pollution Bulletin*, **79**(1–2), 130–144.
Huang, Y., Chen, L., Negenborn, R. R. and van Gelder, P. H. A. J. M. (2020a). A ship collision avoidance system for human-machine cooperation during collision avoidance. *Ocean Engineering*, **217**, 107913.
Huang, Y., Chen, L., Chen, P., Negenborn, R. R. and Van Gelder, P. H. A. J. M. (2020b). Ship collision avoidance methods: State-of-the-art. *Safety Science*, **121**, 451–473.
Jiang, M., Lu, J., Yang, Z. and Li, J. (2020). Risk analysis of maritime accidents along the main route of the maritime silk road: A Bayesian network approach. *Maritime Policy & Management*, **47**(6), 815–832.
Kearon, J. (1997). Computer programs for collision avoidance and traffic keeping conference on mathematical aspects on marine traffic.
Kouibia, A. and Pasadas, M. (2012). An approximation problem of noisy data by cubic and bicubic splines. *Applied Mathematical Modelling*, **36**(9), 4135–4145.
Kujala, P., Hänninen, M., Arola, T. and Ylitalo, J. (2009). Analysis of the marine traffic safety in the gulf of Finland. *Reliability Engineering & System Safety*, **94**(8), 1349–1357.
Kum, S. and Sahin, B. (2015). A root cause analysis for Arctic marine accidents from 1993 to 2011. *Safety Science*, **74**, 206–220.
Li, Y. P. (2019). Research on major characteristics of maritime traffic based on AIS data [D]. *Dalian Maritime University*.
Li, B. and Pang, F. W. (2013). An approach of vessel collision risk assessment based on the D–S evidence theory. *Ocean Engineering*, **74**, 16–21.
Lisowski, J. (2002). Game control of moving objects. *IFAC Proceedings Volumes*, **35**(1), 373–378.
Liu, Z., Wu, Z. and Zheng, Z. (2019). A novel framework for regional collision risk identification based on AIS data. *Applied Ocean Research*, **89**, 261–272.
Liu, Z., Wu, Z. and Zheng, Z. (2020). A molecular dynamics approach for modeling the geographical distribution of ship collision risk. *Ocean Engineering*, **217**, 107991.
Mou, J. M., Van der Tak, C. and Ligteringen, H. (2010). Study on collision avoidance in busy waterways by using AIS data. *Ocean Engineering*, **37**(5-6), 483–490.
Ozturk, U. and Cicek, K. (2019). Individual collision risk assessment in ship navigation: A systematic literature review. *Ocean Engineering*, **180**, 130–143.
Rawson, A. and Brito, M. (2021). A critique of the use of domain analysis for spatial collision risk assessment. *Ocean Engineering*, **219**, 108259.
Ren, Y., Mou, J., Yan, Q., & Zhang, F. (2011). Study on Assessing Dynamic Risk of Ship Collision. In *ICTIS 2011: Multimodal Approach to Sustained Transportation System Development: Information, Technology, Implementation* (pp. 2751–2757).
Szlapczynski, R. (2006). A unified measure of collision risk derived from the concept of a ship domain. *The Journal of Navigation*, **59**(3), 477.
Wang, N. (2010). An intelligent spatial collision risk based on the quaternion ship domain. *The Journal of Navigation*, **63**(4), 733.
Wen, Y., Huang, Y., Zhou, C., Yang, J., Xiao, C. and Wu, X. (2015). Modelling of marine traffic flow complexity. *Ocean Engineering*, **104**, 500–510.
Yip, T. L. (2008). Port traffic risks–A study of accidents in Hong Kong waters. *Transportation Research Part E: Logistics and Transportation Review*, **44**(5), 921–931.
Yuan, X., Zhang, D., Zhang, J., Zhang, M. and Soares, C. G. (2021). A novel real-time collision risk awareness method based on velocity obstacle considering uncertainties in ship dynamics. *Ocean Engineering*, **220**, 108436.
Zhang, L. and Meng, Q. (2019). Probabilistic ship domain with applications to ship collision risk assessment. *Ocean Engineering*, **186**, 106130.
Zhang, J., Zhang, D., Yan, X., Haugen, S. and Soares, C. G. (2015). A distributed anti-collision decision support formulation in multi-ship encounter situations under COLREGs. *Ocean Engineering*, **105**, 336–348.
Zhang, W., Kopca, C., Tang, J., Ma, D. and Wang, Y. (2017). A systematic approach for collision risk analysis based on AIS data. *The Journal of Navigation*, **70**(5), 1117.
Zhang, W., Feng, X., Goerlandt, F. and Liu, Q. (2020). Towards a convolutional neural network model for classifying regional ship collision risk levels for waterway risk analysis. *Reliability Engineering & System Safety*, **204**, 107127.
Zhang, Y., Sun, X., Chen, J. and Cheng, C. (2021). Spatial patterns and characteristics of global maritime accidents. *Reliability Engineering & System Safety*, **206**, 107310.
Zhen, R., Riveiro, M. and Jin, Y. (2017). A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance. *Ocean Engineering*, **145**, 492–501.
Zheng, K., Chen, Y., Jiang, Y. and Qiao, S. (2020). A SVM based ship collision risk assessment algorithm. *Ocean Engineering*, **202**, 107062.

Cite this article: Zhen R, Shi Z, Shao Z, Liu J (2022). A novel regional collision risk assessment method considering aggregation density under multi-ship encounter situations. *The Journal of Navigation* **75**: 1, 76–94. https://doi.org/10.1017/S0373463321000849