Liu, Lei; Zhang, Mingyang; Hu, Yue; Zhu, Wei; Xu, Sheng; Yu, Qing

A probabilistic analytics method to identify striking ship of ship-buoy contact at coastal waters

Published in:
Ocean Engineering

DOI:
10.1016/j.oceaneng.2022.113102

Published: 15/12/2022

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Liu, L., Zhang, M., Hu, Y., Zhu, W., Xu, S., & Yu, Q. (2022). A probabilistic analytics method to identify striking ship of ship-buoy contact at coastal waters. *Ocean Engineering*, *266*, Article 113102.
https://doi.org/10.1016/j.oceaneng.2022.113102

Ocean Engineering 266 (2022) 113102

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

A probabilistic analytics method to identify striking ship of ship-buoy contact at coastal waters

Lei Liu [a,b], Mingyang Zhang [c,*], Yue Hu [d], Wei Zhu [e], Sheng Xu [f], Qing Yu [g]

[a] School of Transportation, Southeast University, Nanning, China
[b] School of Engineering, The University of Tokyo, Tokyo, Japan
[c] Department of Mechanical Engineering, Marine Technology Group, Aalto University, Espoo, Finland
[d] College of Transportation Engineering, Tongji University, Shanghai, China
[e] Guangzhou AtoN Division of NGCS, Guangzhou, China
[f] Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway
[g] School of Navigation, Jimei University, Xiamen, China

ARTICLE INFO

Keywords:
Maritime safety
Collision
Striking ships identification
Safety domain
AIS data

ABSTRACT

The identification of the ship that contact with the buoy can provide evidence for accident accountability. To this aim, the paper develops a probabilistic analytics method to evaluate the ship-buoy contact risk for the striking ship identification at the coastal areas by combining buoy domain and bounding box models. The method makes use of Automatic Identification System (AIS) data and navigational buoy data. Firstly, an AIS-based probabilistic buoy domain model is adopted for the determination of the safety boundary of the buoy to detect potential striking ships with a higher contact probability. Then, the bounding boxes of the navigational buoy and the detected potential striking ships are developed to detect the real striking ship by analyzing the interaction between the ship bounding box and the buoy bounding box. Finally, the probabilistic analytics method is demonstrated in the South China Sea and validated using historical ship-buoy contact records. Results indicated that, from a probabilistic perspective, the safety buoy domain (critical boundary) existed with diverse distances dynamically. The proposed method could assist the identification of striking ships while aiding the definition of the safety buoy domain for preventing ship-buoy contacts. As a result, it has the potential to support the development of ship-buoy contact risk management and assist surveillance operators and master on board by improving their cognitive abilities in dangerous traffic scenarios.

1. Introduction

Marine aids to Navigation are essential to enhance ship safety during shipping in the coastal area. However, ship-buoy contact accidents often result in navigational buoy damage leading to navigation risks to ships at the coastal area (e.g., Mullai and Paulsson, 2011; Antão and Soares, 2019). Even though ship-buoy contact events are not occurred very often (Moorits and Usk, 2012), these accidents still bring great economic losses except for the potential navigation risks. Compared with ship-ship contact detection, the difficulties of ship-buoy contact detection lie in: a) in addition to the large and inconsistent time interval of buoy data, there is no fixed pattern movement and the moving is easily influenced by multiple objects, making it impossible to accurately to locate the position of buoys at any given time; b) the buoys (about 2 m * 0.4 m) is much smaller than ships and the GPS position errors may exceed the size of buoys, resulting in the inability to determine the relative state of buoys and nearby ships. These problems make it extremely difficult to identify the striking ship in the ship-buoy contact for accident accountability. Therefore, safeguarding buoy safety by using a sufficient contact detection model for striking ship identification is crucial in the maritime risk assessment (Liu et al., 2017, 2022; Otto and Petersen, 2003; Tam et al., 2009; Zhang et al., 2015b; Zhang and Meng, 2019; Zheng and Sayed, 2020).

The development and application of innovative statistical and analytics methods to provide novel insights into collision risk assessment at sea are essential and meaningful (e.g., Zhang et al., 2020; Mannering et al., 2020; Zeng et al., 2016; Antão and Soares, 2019). The research over the last two decades mainly focuses on (a) safety domain quantification (e.g., Lei et al., 2021; Szlapczynski et al., 2018a), (b) conflict scenarios detection (e.g., Zhang et al., 2015a; Zhang et al., 2021a; Zhang et al., 2021b; Du et al., 2021), and (c) collision risk evaluation (e.g.,

* Corresponding author. Otakaari 4, 02150, Koneteknikka 1, Espoo, Finland.
E-mail address: Mingyang.0.zhang@aalto.fi (M. Zhang).

https://doi.org/10.1016/j.oceaneng.2022.113102
Received 1 March 2022; Received in revised form 5 June 2022; Accepted 5 November 2022
Available online 28 November 2022

Nomenclature

Abbreviations

AIS	Automatic identification system
GNSS	Global navigation satellite system
OBB	Oriented bounding box
TTSNB	The telemetry and telecontrol system of the navigational buoy
UTM	Universal Transverse Mercator
AABB	Axis-aligned bounding box
k-DOP	Discrete orientation polytope
IQR	Interquartile range

Glossary of variablesVariables Definitions

D_{nb}	Original navigational buoy data
$f_{\Delta(\theta)}(x)$	The probability density function
$l(\theta)$	The maximum traffic density
S_α	The probability buoy domain
(x_0, y_0)	The center coordinate
L	The ship length
$(x_{i,}^{\prime} y_{i,}^{\prime})$, $i \in (1, 2, 3, 4)$	The i_{th} vertices of the circumscribed rectangle along the positive direction y-axis
r	The radius of the bounding box of navigational buoys
p_i	The i_{th} points in trajectory
n	The number of points in a trajectory
l	The grid side length
D_{ais}	Original AIS data
$\delta(\theta)$	Traffic density
$r_{\min}(\theta)$	The minimum traffic density
$(x_{0,}^{\prime} y_{0,}^{\prime})$	The positioning coordinates
θ	The ship heading
B	The ship width
(x_i, y_i), $i \in (1, 2, 3, 4)$	The i_{th} vertices of the circumscribed rectangle after rotating θ clockwise
$(x_{i,}^{j} y_{i,}^{j})$, $i \in (1, 2, 3, 4, r)$ $j \in (L, B, C)$	The i_{th} projection coordinate on j_{th} separating axis
λ	The average predicted distance error
R	The range of water area near specific navigational buoy
distance	The spatial distance based on longitude and latitude coordinates

Rawson and Brito, 2021; Zhang et al., 2020; Huang and Van Gelder, 2020a). Mostly these approaches qualify risks by evaluating the geographical positioning of ships and the difficulty of avoiding collisions to present spatial-temporal maps to demonstrate collision risk levels, which meet the requirements for the evaluation of ship-ship collision risk or ship-to-offshore structures contact risk (Goerlandt and Montewka, 2015). Notwithstanding, they fail to be applied in ship-buoy contact evaluation and striking ship identification. This is because the navigational buoy moves within dynamical a restricted circular area under the impacts of hydrometeorological conditions. And the impacts of interactions between ships and the navigational buoys in real operational conditions are often underestimated in the existing collision risk evaluation models. The parameters of the existing models are difficult to use in ship-buoy contact detection. Thus, the paper introduces a striking ship detection method by using the probabilistic analytics method and the bounding box method. Firstly, the probabilistic analytics method is used to determine the navigational buoy domain, and the potential striking ships are identified using the developed buoy domain. Then, based on the ship size of the detected the potential striking ships and buoy domain, the ship bounding box and buoy bounding box are determined. Finally, the interactions between the ship bounding box and buoy bounding box are analyzed to identify the real striking ships. Furthermore, the proposed method is applied in the South China Sea and validated using historical accident records. Consequently, the real striking ships that cause the ship to buoy contacts can be identified.

Overall, the paper proposes a method to identify striking ship after ship-buoy collision, which may provide important evidence for accident accountability. Meanwhile, the results outlined present important support for advanced mitigation of risks, they may assist surveillance operators and master on board by improving their cognitive abilities in dangerous traffic scenarios. Also, in the short term the approach presented may be useful for forensic investigations and in the medium to long term within the context of the safety case approach. The remains are organized as follows: a brief literature review in Section 2, the method for ship-buoy collision detection is elaborated in Section 3. The application and valuations are presented in Section 4. Sections 5 and 6 provide discussion and conclusion, respectively.

2. Literature review

In this section, a brief literature review on the aspect of collision/contact risk evaluation at sea is presented. The research problem is also elaborated in previous studies. Maritime risk assessments are essential to provide deep insights for improving traffic safety management standards. A collision can happen by striking another ship (regardless of whether or not she is underway) or multiple ships (EMSA, 2020). Research in collision risk mostly focuses on (1) ship-ship collisions and (2) ship collisions with offshore structures.

The existing models adopted to evaluate ship-ship collisions are known as distance/time to the Closest Point of Approach (CPA) (Lopez-Santander and Lawry, 2017; Zhao et al., 2016), Ship Safety Domain (Szlapczynski et al., 2018a, 2018b; Weng and Yang, 2015), Vessel Conflict Ranking Operator (VCRO) (Zhang et al., 2015a a), and Velocity Obstacle model (Yu et al., 2021). Zhang et al. (2021a) and (2021b) proposed Avoidance Behavior-based Collision Detection Model (ABCD-M) using CPA for identifying potential collision scenarios and collision risk index by using historical AIS data and hydrometeorological data of ships in the Gulf of Finland. Lei et al. (2017) proposed a ConflictFinder for maritime traffic based on DCPA and TCPA and constructed a behavior pattern extraction model (CapatternMiner) by analyzing the change of ship course. Zhang et al. (2015a) and Zhang et al. (2020) proposed a VCRO model to detect near miss using AIS data. Chen et al. (2020) and Huang and van Gelder (2020b) proposed a velocity obstacle model to identify the ship movements and proposed a real-time ship collision risk perception model based on risk information in encounter scenarios. Besides that, Yuan et al. (2021) and Du et al. (2020, 2021b) adopted the velocity obstacle model to detect traffic conflict using AIS data. The typical ship domains and its applications for the evaluation of ship-to-ship collision risk have been presented. Review papers by Szlapczynski and Szlapczynska (2017) present the rationale of the most influential models/methods. Du et al. (2021) summarized the typical ship domains and its applications as shown in Fig. 1. Overall, these ship domain models may be useful to determine safety domain to avoid ship-ship collisions. However, these approaches can be used to detect ship-ship conflict/potential collision scenarios using AIS data but fail to identify the striking ship in ship-buoy contact. This is because the size of the navigational buoy is very small and the parameters of the existing models are difficult to use in ship-buoy contact detection. Besides that, the sailing ships often underestimate the ship-buoy contact risk, despite the ship-buoy contact that occurred, which makes it extremely difficult to evaluate ship to buoy contact risk using the existing models. Additionally, the motion behaviors of navigational

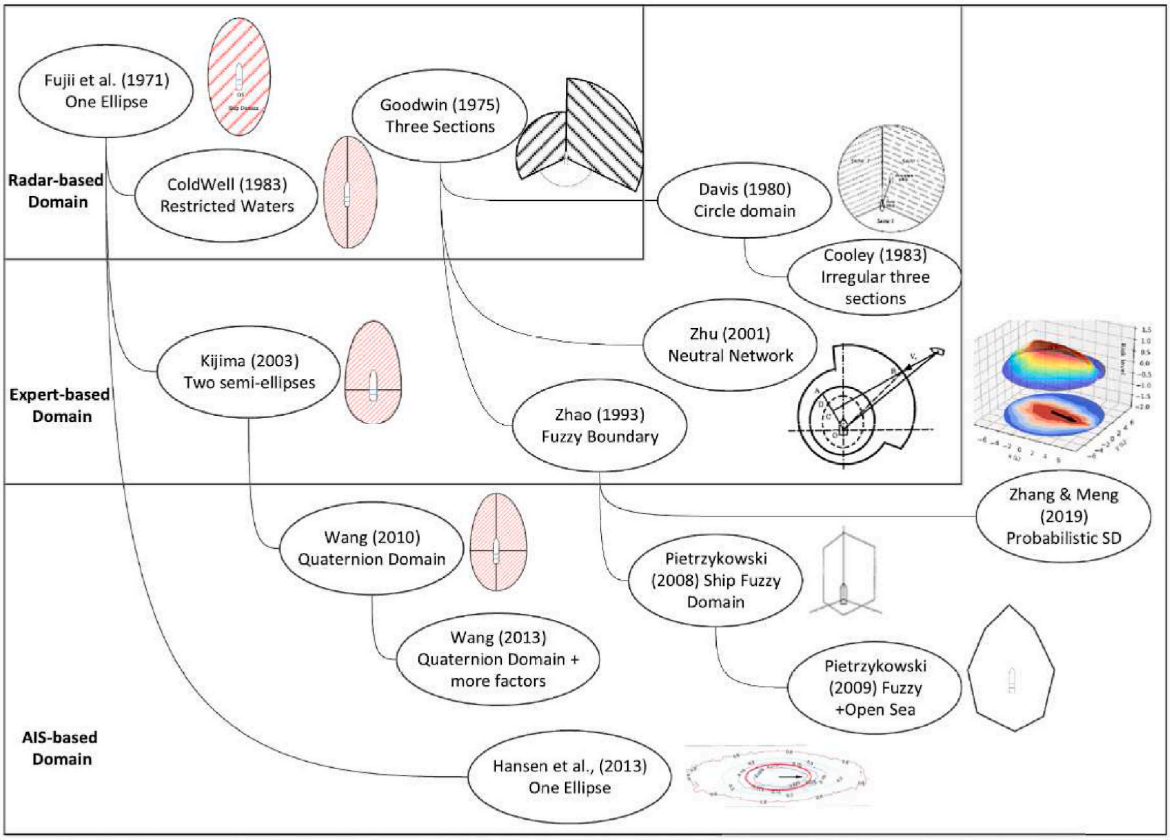

Fig. 1. Ship domains and its applications (Du et al., 2021a).

buoys are restricted in a circular area under the impacts of hydrometeorological conditions with uncertainties instead of stationary. But the sailing ships around navigational buoys often misunderstand the safety distance to the navigational buoy, resulting in a ship-buoy contact. The striking ship often escaped after ship-buoy contact. Thus, it is important to develop a navigational buoy domain to present safety indicator to the master onboard to keep a safety distance for ship-buoy contact risk. On the other hand, if the ship-buoy contact occurred, the buoy domain can also be used to detect the potential striking ships.

Ship collision/contact with other floating objects (e.g., oil & gas platforms, wind/tidal farms, etc) are much more safety-critical types of casualty events (EMSA, 2020; Li et al., 2012). Ship and offshore structures contact risk assessment aim to reduce ship contact risk with fixed structures. The Pedersen collision/contact model (Pedersen, 2010) is the most widely applied one that focuses on collision/contact probability. Common modeling tools are known as COLLIDE (e.g., Hassel et al., 2021), COLWT (e.g., Yu et al., 2021; Lei et al., 2021; Otto et al., 2003), and MARIN (e.g., Koukaki and Tei, 2020; Yu et al., 2021). In the Pedersen model, collision candidates are calculated with the ship traffic trajectory distributions and causation probabilities (e.g., Hassel et al., 2017; Yu et al., 2018; Mujeeb-Ahmed et al., 2018). However, the Pedersen collision model and common modeling tools still reveal a high level of uncertainties because the used causation factors are determined with generic parameters which are not commonly agreed (e.g., Szlapczynski et al., 2018a; Gil, 2021; Zhang et al., 2021a, 2021b). Besides, these models assess the collision risk using statistical approaches, which ignore ship motions and the risk dynamicity. The safe boundary approach is also widely applied to assess the contact risk between ships and offshore structures (e.g., Szlapczynski et al., 2018a; Gil, 2021; Zhang et al., 2021a, 2021b). The approach defines safety criteria in various situations, including the size of offshore structure and ship, ship maneuverability, traffic situation, and the blind area. For the safe boundary approaches, the existing models adopted are known as the empirical safe boundary model (e.g., Hansen et al., 2013; Wang, 2010), the knowledge-based safe boundary model (e.g., Zhu et al., 2001; Zhang et al., 2021a), and the analysis-based safe boundary model (e.g., Zhong et al., 2019; Gil, 2021; Gil et al., 2020; Pietrzykowski and Wielgosz, 2021). These models are used to calculate the collision risk according to the overlapping areas between safe boundaries. Notwithstanding, these models probably fail to detect real striking ship in the ship-buoy contact because it is challenging to identify accidents by quantifying overlapping areas, which is difficult to be extracted with the uncertainties of the motion behaviors under the impact of external environmental conditions based on existing models.

Currently, the Telemetry and telecontrol system of the navigational buoy is used for navigational buoy monitoring. The TTSNB (Telemetry and Telecontrol System of the Navigational Buoy) can measure, control, and monitor navigational buoys remotely. Typically, TTSNB is composed of a monitoring center, data acquisition terminal, and communication system (Delgado Román et al., 2008). However, the system cannot be used to detect the striking ship in the ship-buoy contact. In addition, AIS data include static information (call sign, IMO number, length, etc.) and dynamic information (location, speed, course, etc.)(Liu and Shi, 2020; J. Zhang et al., 2022; M. Zhang et al., 2022). Navigational buoy motion data (location records) present the motion of the buoy. Therefore, developing a probabilistic analytics method is necessary to identify both the ship-buoy contact and striking using navigational buoy motion data and AIS data. Restoring detailed time-dependent traffic situations and navigational buoy motions at the time provides insights into ship-buoy contact risk evaluation and striking ship identification.

Overall, ship-buoy contact risk estimation and striking ship identification are necessary. Yet, many challenges exist in the determination of safety buoy domain to prevent the ship-buoy contact by traffic management, and identification of striking ship in the ship-buoy contact in real operational conditions. To fill the research gaps, the paper explores a probabilistic analytics method to identify striking ship by using buoy domain and bounding box models. The proposed method quantifies the critical boundary of the navigational buoy to mitigate ship-buoy contact risk by warning ships to keep a safe distance. Based on the safety buoy domain and adopted bounding box models, the real striking ship can be detected for accident accountability.

3. Methodology

The logic framework of the probabilistic analytics method for ship-buoy contact detection and striking ship identification is shown in Fig. 2.

- Step (i) using AIS data to model probabilistic buoy domain. By restoring ship traffic and buoy motions at the time using AIS data and buoy motion data, the process is used to propose a probabilistic buoy domain that depicts the buoy domain boundary as the probability value rather than the crisp value with a forbidden boundary. The probabilistic buoy domain is further investigated to provide the safety domain of navigational buoys and extract potential striking ships with relatively higher contact probability of the buoy.
- Step (ii) determining bounding boxes of ships and navigational buoys. The bounding box method (Tu and Yu, 2009; Kim et al., 2021; Adhikari and Huttunen, 2021) is adopted to define the bounding boxes of the ship and navigational buoy under Step (i) according to the ship size and buoy motion behaviors. Then, the mathematical model is developed to present the geometrical relations between the

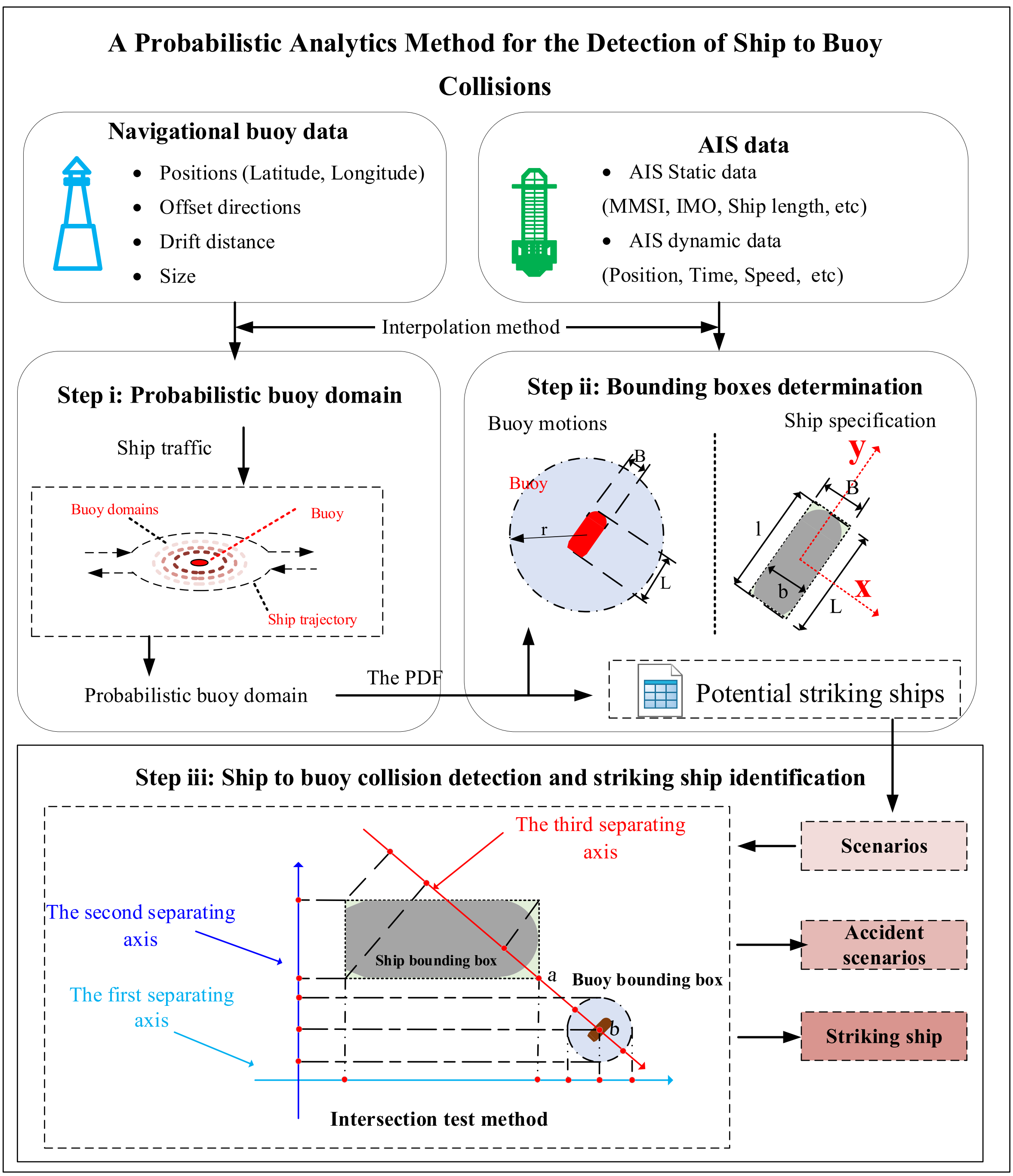

Fig. 2. The logic framework of the ship-buoy contact detection.

bounding box of the ship and the bounding box of the buoy for the intersection test.

- Step (iii) identifying a potential striking ship in a ship-buoy contact accident. The bounding boxes of potential striking ships and buoys are used for ship extraction under Step (ii). The ship-buoy contact scenario is recovered using bounding boxes by the intersection test, and the striking ship and spatial-temporal information about ship-buoy contact accidents are identified for accident accountability.

The algorithm of ship-buoy contact detection is shown in Table 1.

3.1. Step i: using AIS data to model probabilistic buoy domain

3.1.1. The definition of probabilistic buoy domain

The sailing ships require a safe distance to keep clear of buoys so that the masters of ships have time and space to take evasive actions and avoid ship-buoy contacts. However, the boundary of the buoy is vague and uncertain due to different ship motion behaviors around the navigational buoy. Fig. 3 shows the probability distribution of the minimal passing distance between the ship and buoy according to the probabilistic buoy domain based on AIS traffic data. And the buoy domain is obtained by calculating the probability density function using traffic data based on statistical analysis approaches (Zhong et al., 2019; Lei et al., 2021; Gucma and Marcjan, 2012; Xin et al., 2023; Zhang and Meng, 2019).

The probability density function $f_{\Delta(\theta)}(x)$ is used to quantify the probabilistic buoy domain. Therefore, the probabilistic buoy domain S_α is defined as the probability of the interval $[0, 1]$, as shown in Formula (1).

$$S_\alpha = \left\{ (r_{\min}(\theta) + \delta(\theta), \theta) \middle| \int_0^{\delta(\theta)} f_{\Delta(\theta)}(x)dx \le \alpha, 0 \le \delta(\theta) \le l(\theta) - r_{\min}(\theta), 0^0 \le \theta \le 360^0 \right\} \tag{1}$$

Where, $f_{\Delta(\theta)}(x)$ denotes the probability density function of traffic density around navigational buoy, x is the distance from buoy, and $\delta(\theta) = l(\theta) - r_{\min}(\theta)$ defines the same space $[r_{\min}(\theta), l(\theta)]$. $l(\theta)$ represents the variable of the maximum traffic density from the buoy to the boundary of buoy domain. $r_{\min}(\theta)$ denotes the variable of the minimum traffic density from the buoy to the boundary of the forbidden. $r(\theta)$ denotes traffic density from the own ship to the boundary of the forbidden. θ denotes the bearing angle from buoy to ships around.

Accordingly, for a given value α, the buoy domain can be depicted by Formula (2).

$$R_\alpha = \left\{ (r_{\min}(\theta) + \delta(\theta), \theta) \middle| \int_0^{\delta(\theta)} f_{\Delta(\theta)}(x)dx = \alpha, 0 \le \delta(\theta) \le l(\theta) - r_{\min}(\theta), 0^0 \le \theta \le 360^0 \right\} \tag{2}$$

Where, α denotes the probability of the α-boundary of ship domain, and α is 0, and the α-boundary of ship domain denotes the forbidden boundary R_α. Within the forbidden boundary, the ship-buoy contact occurs with highest probability.

By designing the buoy domain above, the buoy domain can be used to warn ships to keep a safe distance. The forbidden boundary of the navigational buoy could be applied to detect potential ship-buoy contacts. The algorithm of probabilistic buoy domain modelling based on AIS data is shown in Table 2.

3.1.2. Restoration of ship traffic and buoy motion at the time

This section presents data extraction of AIS data and buoy data, and coordinate conversion. Further, the ship traffic and buoy motions are restored at the time to establish a probabilistic buoy model.

- Data extraction

Data extraction includes the selection of AIS data and navigational buoy data. The navigational buoys that have experienced contacts are selected as target buoys to analyze the closest distance under the ship-buoy contact accident. Additionally, setting the corresponding spatial scope to filter AIS data is complicated and unnecessary because navigational buoys only move within a restricted area under the action of anchorages and the anchor chains, as shown in Fig. 4. In the paper, the navigational buoys movement center is set as the center of AIS data filtering, and the navigational buoy data within a specified distance from the center are selected. Then, the coordinate system can be converted from earth fixed coordinate system (longitude and latitude in AIS ship trajectory) to a navigational buoy-centered, which can better present the distance between the navigational buoy and ship trajectory (e.g., Zhang et al., 2021a; Xin et al., 2021; Zhang et al., 2022).

- Coordinate conversion to restore ship traffic and buoy motion at the time

To restore ship traffic and buoy motion at the time using AIS data and navigational buoy data from the navigational buoy perspective, the coordinate system should be converted from the earth fixed (AIS ship trajectory) to a navigational buoy. The entire coordinate conversion includes two steps: (1) the longitude and latitude coordinates need to be converted into UTM (Universal Transverse Mercator) for calculating the distance between ships and navigational buoys, and the boundary is extracted using the grid statistic method; (2) the positions of four vertices (given that circumscribed rectangle surrounds the ship) are obtained to determine the grid occupied by the entire ship. The GNSS (Global Navigation Satellite System) terminal of AIS is not installed in the center of the ship. The location information returned by AIS needed to be converted into the center position. Therefore, the coordinates of the four vertices of each ship can be obtained according to the size and heading information of the ships.

The installation position of the GNSS terminal is related to the installation position of the bridge. Most ships adopt the design of the bridge in the stern. It is assumed that the GNSS terminal is installed at the central axis position, 1/3 times the length of the ship from the stern (given that the equipment installation location data in the AIS are not obtained). Therefore, the coordinate conversion of the center position of a ship is shown in Fig. 5. It is assumed that the positioning coordinates of

Table 1
The pseudocode of the probabilistic analytics method to detect the ship-buoy contact.

Algorithm 1: Algorithm of collision detection and striking ship identification
Input: motion data D_{nb} of crashed navigation buoy, AIS data D_{ais} near the navigational buoy
Output: the striking ship
1: **for** each E^i in D_{ais} **do**
2: Transform the longitude and latitude coordinates of E^i into UTM coordinates
3: **end for** 4: **for** each E^i in D_{ais} **do**
5: Transform the longitude and latitude coordinates of into UTM coordinates U_{cor}
6: Calculate the central position c_p of each ship according to U_{cor} and ship size in E^i
7: Calculate the four vertices v_s of the circumscribed rectangle of each ship according to c_p, COG, and ship size in E^i
8: **end for** 9: Cluster AIS points in D_{ais} according to MMSI and obtain trajectory cell T_{mmsi} of N ships
10: Initialize ship trajectory cell T_{ships}
11: **for** each T^i_{mmsi} in T_{mmsi} **do**
12: Obtain segmented track T_s of the ship after separating T^i_{mmsi} according to timestamp
13: Add T_s into T_{ships}
14: **end for** 15: Initialize suspicious ship trajectory cell T_{sus_ships}
16: **for** each T^i_{ship} in T_{ships} **do**
17: **if** ship trajectory T^i_{ship} intersects with buoy domain **then**
18: Add T^i_{ship} into T_{sus_ships}
19: **end if**
20: **end for** 21: Initialize suspicious ship matrix M_{sus_ships}
22: **for** each $T^i_{sus\_ship}$ in T_{sus_ships} **do**
23: **for** each p^j in $T^i_{sus\_ship}$ **do**
24: **if** the ship OBB of p^j intersect with the box of navigational buoy **then**
25: Add p^j into M_{sus_ships}
26: **end if**
27: **end for** 28: **end for** 29: **if** there are more than one ship in M_{sus_ships} **then**
30: **for** each p^i_{ship} in M_{sus_ships} **do**
31: **if** navigational buoy moves abnormally or returns no data at p^i_{ship}'s time **then**
32: Return the ship information in p^i_{ship}
33: Break
34: **end if**
35: **end for** 36: **else**
37: Return the ship information in M_{sus_ships}
38: **end if**

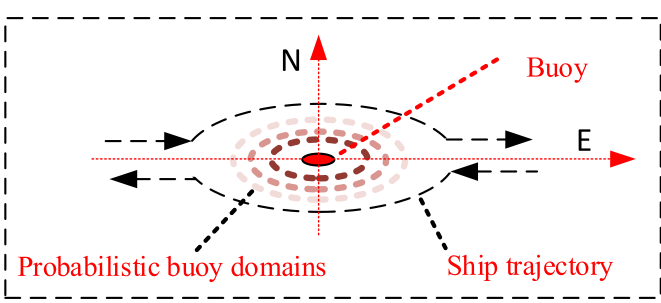

Fig. 3. The probabilistic buoy domain modeling using AIS data, inspired by (Zhang and Meng, 2019).

the ship in the AIS after UTM coordinate conversion is $(x_0{}', y_0{}')$, the center coordinate is (x_0, y_0), the heading is θ, and the distance from the center of the ship to the center of GNSS terminal is 1/6 times of the ship length. Then the calculation formulas for the center point of the ship are as follows:

$$x_0 = x_0{}' + L\sin\theta/6 \tag{3}$$

$$y_0 = y_0{}' + L\cos\theta/6 \tag{4}$$

The four vertices of the circumscribed rectangle along the positive direction y-axis are assumed as (x_1', y_1'), (x_2', y_2'), (x_3', y_3'), and (x_4', y_4'),

Table 2
The pseudocode of probabilistic buoy domain modeling.

Algorithm 2: Algorithm of buoy domain construction
Input: motion data D^i_{nb} of N navigation buoy, AIS data D^i_{ais} corresponding to each the navigational buoy, $i \in N$, the range R of water ear the navigational buoy and the grid side length l
Output: the domain model of navigational buoy
1: Separate the water area near each navigational buoy into grids according to R and l
2: **for** each E^j in D^i_{nb} **do**
3: Transform the longitude and latitude coordinates of E^j into UTM coordinates
4: **end for** 5: **for** each E^j in D^i_{ais} **do**
6: Transform the longitude and latitude coordinates of into UTM coordinates U_{cor}
7: Calculate the central position c_p of each ship according to U_{cor} and ship size in E^j
8: Calculate the four vertices v_s of the circumscribed rectangle of each ship according to c_p, COG, and ship size in E^j
9: **end for** 10: Initialize the grid superimposed zero-element matrix S_0
11: **for** $i = 1 \rightarrow N$ **do**
12: Initialize the grid zero-element matrix S_i
13: **for** each E^j in D^i_{ais} **do**
14: Obtain the circumscribed rectangle of E^j according to v_s and c_p in E^j
15: Obtain the grid IDs G_{IDs} in S_i which intersect the circumscribed rectangle of E^j
16: Add 1 to the element in G_{IDs} in S_i
17: **end for** 18: $S_0 \leftarrow S_0 + S_i$
14: **end for** 15: Initialize the boundary distance vector D
16: **for** $k = 0 \rightarrow 7$ **do**
17: Obtain the boundary grid G^k_{ID} in direction $0 + k * 45$ based on S_0
18: Obtain the boundary distance D^k in direction $0 + k * 45$ based on G^k_{ID} and l
19: Save D^k into D
20: **end for** 21: Construct the circled domain model of navigational buoy according to D

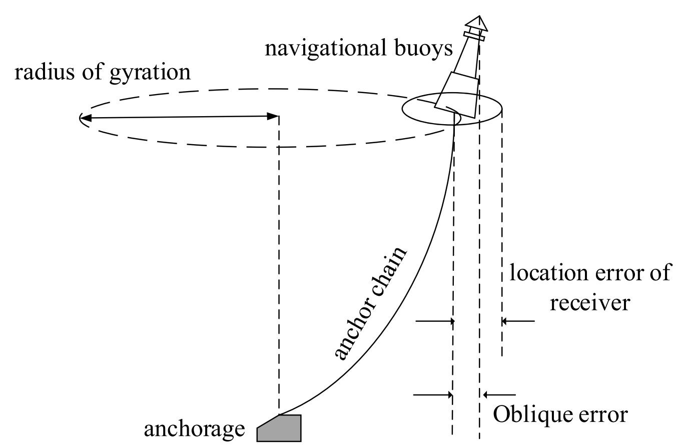

Fig. 4. Schematic diagram of navigational buoy drifting.

respectively. After rotating θ (heading) clockwise, four vertices become (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4), as shown in Fig. 6.

Taking the calculation of (x_1', y_1') and (x_1, y_1) as examples, the coordinate of (x_1', y_1') is obtained based on the center point (x_0, y_0), ship length L, and ship width B, as shown in Formulas (5) and (6):

$$x_1' = x_0 - B/2 \tag{5}$$

$$y_1' = y_0 + L/2 \tag{6}$$

Since the point (x_1, y_1) is the point (x_1', y_1') rotated θ clockwise around the point (x_0, y_0), the coordinate calculation formula of the point (x_1, y_1) is shown in Formulas (7) and (8).

$$x_1 = (x_1' - x_0) * \cos(\theta) + (y_1' - y_0) * \sin(\theta) + x_0 \tag{7}$$

$$y_1 = -(x_1' - x_0)\sin(\theta) + (y_1' - y_0) * \cos(\theta) + y_0 \tag{8}$$

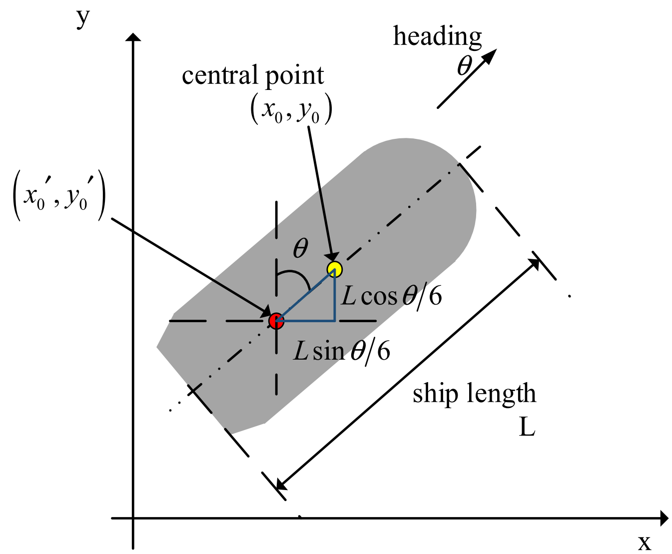

Fig. 5. Schematic diagram of coordinate conversion of the ship center point.

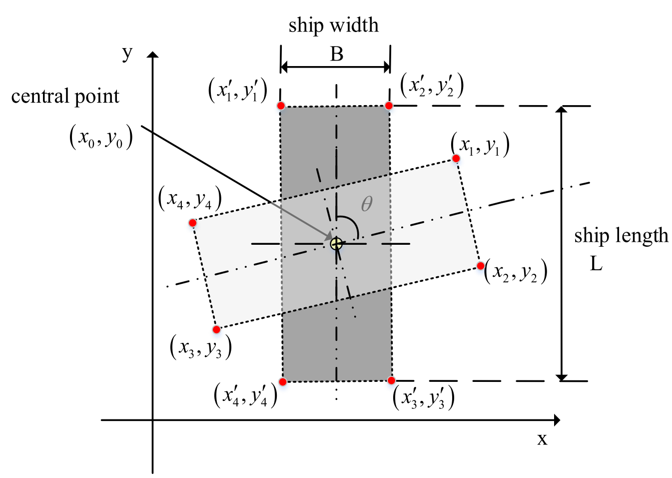

Fig. 6. Schematic diagram of coordinate rotation of ship circumscribed rectangle.

Similarly, the coordinates of (x_2, y_2), (x_3, y_3), and (x_4, y_4) can be obtained. Therefore, the ship traffic and buoy motion are restored at the time after coordinate conversion using AIS data and navigational buoy data with the consideration of the navigational motion and ship size.

3.1.3. Probabilistic buoy domain determination by griding traffic distribution

The traffic map is gridded, and the relationships between ships and navigational buoys are analyzed to develop a probabilistic buoy domain. The critical safety distance between the ship trajectory and navigational buoy can be calculated, and the buoy domain model can be established.

- The map grid generation and ship position distribution statistics

After coordinate conversion, the water area around the navigational buoys is gridded to obtain the accurate distribution of ships, thereby precisely obtaining the boundary of the buoy domain. A certain grid length ensures the validity of navigational buoy data within the specified distance from the center point. In the paper, the grid pattern is 100 m * 100 m. Then, a ship will occupy some grids, as shown in Fig. 7. When the ship occupies different grids, the frequency of corresponding grids increases by 1. Therefore, the distribution of the ship in grids should be determined to test the critical safety distance between ships and navigational buoys.

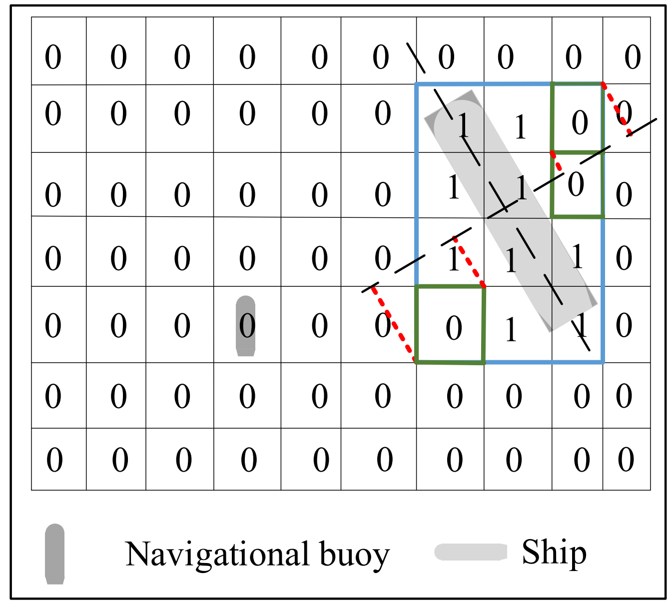

Fig. 7. The statistical analysis of the critical distance between ships and navigational buoys with the consideration of ship size and navigational buoy motions.

The determination process is separated into three steps: (1) The rectangular area where the ship is located and the corresponding network is determined according to the coordinates of the vertices of circumscribed rectangular, as shown in the blue rectangle in Fig. 7; (2) The grids that intersect with the circumscribed rectangle of the ship are separated from the non-intersecting grids; (3) The intersecting grids are extracted for statistical analysis. To identify the intersecting grids, the definition of the separating axis should be clear first: when a polyhedral *A* and a polyhedron *B* are projected vertically on the line, and the projection areas do not overlap, the line *L* is called the separating axis of polyhedron *A* and polyhedron *B*. If there is a separating axis between polyhedral *A* and polyhedron *B*, the closed regions formed will not intersect. Therefore, determining whether there is a separating axis between each grid within the blue rectangle and the circumscribed rectangle can identify the intersecting grids, as shown in Fig. 7. In a two-dimension graph of the ship, two separating axes need to be examined: the straight line perpendicular to the four sides of the circumscribed rectangle of the ship and the straight line passing through the center of the rectangle, as shown in Fig. 7. Besides, the green rectangle does not intersect with the projection of the circumscribed rectangle of ships on the transverse axis, indicating that ships are not distributed over these grids. Therefore, using the above method, the relationship between ships and navigational buoys in real operational conditions can be described by reviewing all ship trajectories around the navigational buoy, providing available data for ship-buoy contact detection.

- Probabilistic buoy domain development using critical boundary analysis

To extract the boundaries of the buoy domain from a probabilistic perspective, the ship distribution curves of the grid distribution map are analyzed from the inside to the outside, respectively, as shown in Figs. 3 and 13. The inflection points of the ship distribution curve from keeping stable to rising, are taken as the boundary point of each direction. In different directions, the boundary distances from the center point are probably different. In addition, the critical safety boundary (the forbidden boundary) (α is 0) with the lowest probability is obtained from

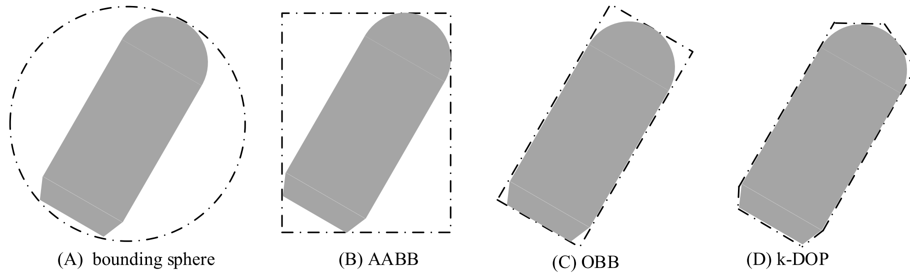

Fig. 8. Four types of common bounding boxes.

non-accident ship data and evaluated by specific contact data, as shown in Formulas (1) and (2). Thus, a probabilistic buoy domain is developed, which can be used to identify potential striking ships.

The probabilistic buoy domain is irregular. Navigational buoy motions are circle areas, as shown in Fig. 4. Therefore, the probabilistic buoy domain should be modeled as a circle according to the forbidden boundaries in different directions. Moreover, the circle model is more convenient for calculating the intersection point of the buoy domain and the ship trajectories and designing the bounding sphere of navigational buoys.

3.2. Step ii: determining bounding boxes for ships and navigational buoys

The detection method based on the buoy domain is a fuzzy detection method. Due to the wide scope of the probabilistic buoy domain, the detection method based on the buoy domain can detect potential striking ships, including non-accident ships with higher contact risk. Therefore, the detection method of the bounding box should make up for the deficiency of the detection method based on the buoy domain to accurately identify the striking ship and obtain the spatial and temporal information of contact accidents.

As shown in Fig. 8, there are four common types of bounding boxes, including the bounding sphere, axis-aligned bounding box (AABB), oriented bounding box (OBB), and discrete orientation polytope (k-DOP) (Tu and Yu, 2009; Kim et al., 2021; Adhikari and Huttunen, 2021). The direct detection method of contact using a large number of ship trajectories will take a long time (Adhikari and Huttunen, 2021). However, the detection method based on the bounding box improves detection efficiency by reducing the number of effective contact objects. For different bounding boxes, construction methods, intersection test methods, and application scenarios are also different.

The current research about the bounding box method mainly applies to virtual reality due to the differences between reality and virtual reality. For example, due to the discontinuity and error of object position

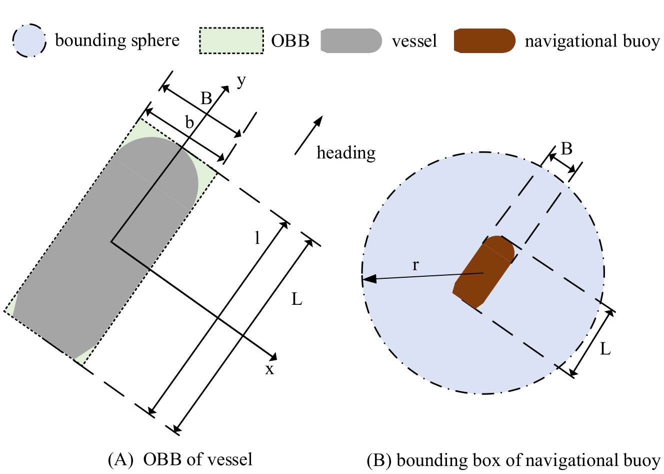

Fig. 9. The bounding box design of ship and navigational buoy.

information in the real environment, the true contact sometimes cannot be detected. However, data in the virtual environment are continuous and accurate to avoid missed detections. There are some effective methods to deal with these problems in the real environment. The bounding boxes of navigational buoys and ships are designed (see in Fig. 9), and the intersection test method between the two bounding boxes is analyzed in Section 3.3 for striking ship identification. It can be seen from Fig. 8 that the bounding sphere and AABB are not suitable for ships and increase the rate of false detection due to the existence of some blank areas. Additionally, k-DOP is more complicated than the OBB, and the parameters of k-DOP are difficult to obtain from AIS data. Therefore, the bounding box of the ship is designed based on the OBB. When using OBB, determining the size and direction of the OBB for the ship is the most important and relatively simple. The size of the OBB can be set according to the length and width of ships, and direction can be determined based on the heading of ships, as shown in Fig. 9(A). The ship length is L, and the ship width is B. The headings of ships are length l, width b, and the positive direction of the long axis of the OBB, respectively.

Unlike ships, the size of navigational buoys is much smaller. As mentioned above, when performing contact detection, the difference in various directions is not obvious. Therefore, the bounding sphere is appropriate for the navigational buoy, as shown in Fig. 9(B). Furthermore, location errors have a great negative impact on contact detection due to the small size of navigational buoys. It is noteworthy that the time interval of the data returned from navigational buoys is much larger than that of AIS data. Even if some effective interpolation algorithms are used to predict specific timestamp data, errors still occur. To reduce the negative influence of positioning errors, enlarging the radius of bounding spheres is an excellent choice.

3.3. Step iii: identifying striking ship of ship-buoy contact

Under the process of Steps i and ii, historical contact data are used to evaluate the forbidden boundary of the buoy domain, and the radius of the bounding box is determined based on the forbidden boundary. The basic principle of selecting the radius is to ensure that the bounding sphere of the navigational buoy can be touched by the OBBs of ships. The radius is not too large to avoid long-term contact or even fusion between the two bounding boxes, making it feasible to determine the time and location of the ship-buoy contact. Therefore, the radius of the bounding sphere should be smaller than that of the forbidden boundary of the buoy domain, which needs to be set according to the analysis results of specific ship-buoy contact records.

After designing the bounding sphere of the navigational buoy and the OBB of the ship, the intersection test method of the bounding sphere and the OBB is further determined. For the OBB, the separating axis theory is still valid. Moreover, there are two separating axes parallel to the long side and the short side of the OBB, respectively, as shown in Fig. 10. Besides, a third separating axis passes through the center of the bounding sphere and the vertex of the OBB closest to the bounding sphere.

Specifically, the process of the intersection test could be divided into

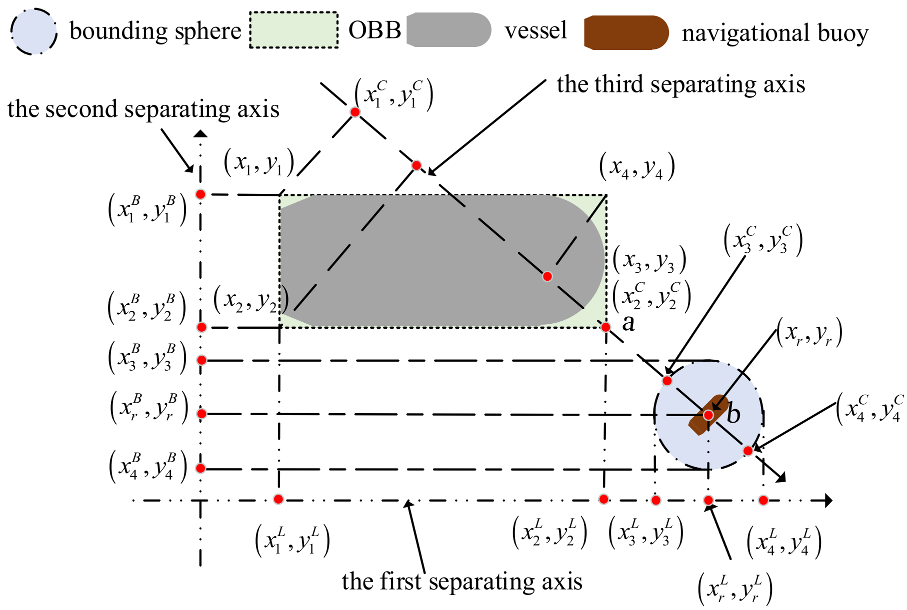

Fig. 10. Intersection test method of the bounding sphere of the navigational buoy and the OBB of the ship.

three steps:

a) Calculation of the vertex coordinates of the OBB of the ship

This calculation process is the same as the coordinate conversion in the previous section, as shown in Fig. 6. Based on location data from AIS, the coordinates of the center point of the ship are obtained by transforming. Then the coordinates of the four vertices are obtained according to the size and heading of the ship.

b) Calculation of projection points

For different separating axes, the projected areas of the OBB of the ship and the bounding sphere of the navigational buoy are calculated. For example, the first separating axis is parallel to the vector $((x_2,y_2),(x_3,y_3))$, the projection range of the OBB is $[x_1^L,x_2^L]$, and the projected range of the bounding sphere is $[x_3^L,x_4^L]$. The projection of the bounding box on the second separating axis is similar to that on the first separating axis. To determine the third separating axis, the distance between the bounding sphere and the four vertices of the OBB of the ship is calculated firstly, and the closest vertex is selected. Projection coordinates can be obtained from the projection formula of the point on the vector. Taking point (x_r,y_r) as an example, the projection formulas are as follows:

$$\left(x_r^L,y_r^L\right)=\left(x_1^L,y_1^L\right)+\frac{\left(x_2^L-x_1^L,y_2^L-y_1^L\right)\left(\left(x_2^L-x_1^L\right)*\left(x_r^L-x_r^L\right)+\left(y_2^L-y_1^L\right)*\left(y_r^L-y_r^L\right)\right)}{\left(x_2^L-x_1^L\right)^2+\left(y_2^L-y_1^L\right)^2} \tag{9}$$

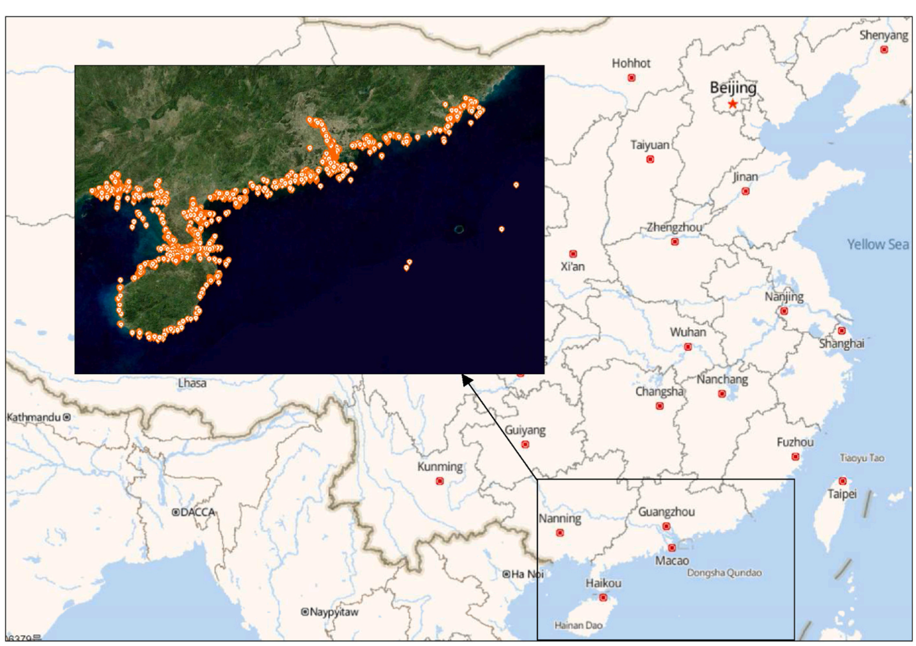

Fig. 11. Study area and distribution of navigational buoys.

$$
\begin{aligned}
(x_r^B, y_r^B) &= (x_1^B, y_1^B) \\
&+ \frac{(x_2^B - x_1^B, y_2^B - y_1^B)\left((x_2^B - x_1^B) * (x_r^B - x_r^B) + (y_2^B - y_1^B) * (y_r^B - y_r^B)\right)}{(x_2^B - x_1^B)^2 + (y_2^B - y_1^B)^2}
\end{aligned}
\tag{10}
$$

The distance between the boundary projection of the bounding box and the projection of the center is still r. The coordinates of x_3^L and x_4^L can be obtained by x_r^L minus r and x_r^L plus r, respectively. The projection coordinates of other points are calculated in the same way.

c) Intersection identification of bounding boxes

After determining the coordinates of the projection points of the OBB of the ship and the bounding sphere of the navigational buoy, judging whether a bounding box intersects with another is relatively simple. For a random separating axis, when the maximum value of its coordinate is less than the minimum value of another coordinate or its minimum value is greater than the maximum value of another coordinate, the bounding box will not intersect with another axis. It could be concluded that no accident occurred between the navigational buoy and the ship. Otherwise, the ship will be detected as a striking ship.

To validate the proposed probabilistic analytics method, a practical case study is carried out in the South China Sea, see in Section 4.

4. Case study

A real case in the South China Sea is studied. In this case, three years AIS data and buoy information is collected, which will be used for detecting potential collision between ship and buoy at coastal waters.

4.1. Study area and data preparation

In this section, the study area and the data used in this paper are introduced first. Moreover, the analysis of the time interval and drifting movement of normal data is carried out to find out the anomaly. Finally, positioning-drift data are filtered, and the interpolation algorithm is applied for AIS data and buoy data prediction.

The South China Sea is one of the most important navigational areas in China. More than 1297 public and dedicated navigational buoys in this area as shown in Fig. 11.

In the paper, a probabilistic analytics method is proposed to detect ship-buoy contacts using AIS data in real operational conditions. Navigational buoy data, AIS data, and contact data between navigational buoys and ships in the South China Sea are selected to construct the buoy domain and to verify the effectiveness of the contact detection method.

- AIS data reconstruction

AIS data consists of static and dynamic messages, see in Table A1 in Appendix. Data collection, transmission, and reception cause data errors. In addition, the time interval of AIS broadcasts varies so that AIS data can be interpolated with a predefined time interval (Chaturvedi, 2019). Therefore, AIS data reconstruction requires trajectory separation, data filtering (i.e., outlier removal), and data prediction (Zhong et al., 2019; Zhang et al., 2016). For data prediction, some methods such as Long Short-Term Memory (Zhong et al., 2019) are complex and time-consuming despite their advancement. Besides, experiments show that the data prediction performance based on piecewise cubic Hermite interpolation (Liu et al., 2019; Liu et al., 2017) meets the requirements of accuracy.

- Description of navigational buoy data

The navigational buoy data is obtained from the floating aids equipped in the AIS receiver. There are many attribute columns, such as basic information and movement information of navigational buoys, voltage, and current data of various equipment. However, some columns are null values, and some are not significant for contact detection research. Finally, the preserved data columns of the navigational buoy are illustrated in Table A2 in Appendix A. It should be noted that the lamp voltage and lamp current represent the working voltage of the navigational buoy and can be used to analyze the working status. Abnormal data can be detected based on the offset distance in navigational buoy data, combined with the interquartile range. In addition, the data needs to be predicted to get the data of ships and navigational buoys under the same timestamp. And piecewise cubic Hermite interpolation is still efficient in buoy data prediction.

- Historical ship-buoy contact records of striking ships and navigational buoys

Ship-buoy contact accidents records in 2018–2020 are collected in the mentioned study area, as shown in Fig. 12. The records contain AIS data and navigational buoy data in this period, as shown in Appendix A.

4.2. Results and analysis

As shown in Fig. 2, the first step is to construct the probabilistic buoy domain model. Then preliminary detection is carried out based on the forbidden boundary of the buoy domain to filter out ships that are unlikely to collide with navigational buoys. The second detection is conducted using the bounding box method since non-accident ships

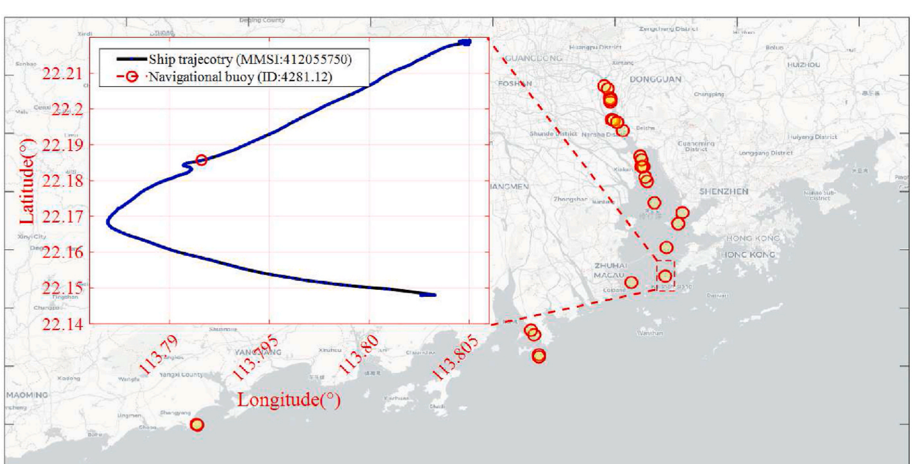

Fig. 12. The map distribution of ship-buoy contacts.

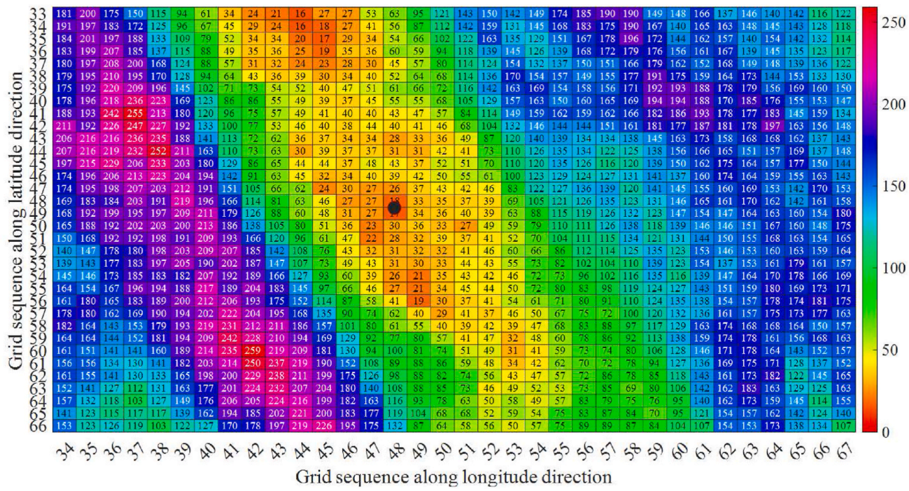

Fig. 13. Ship position distribution in the grids from ID 30 to ID 70.

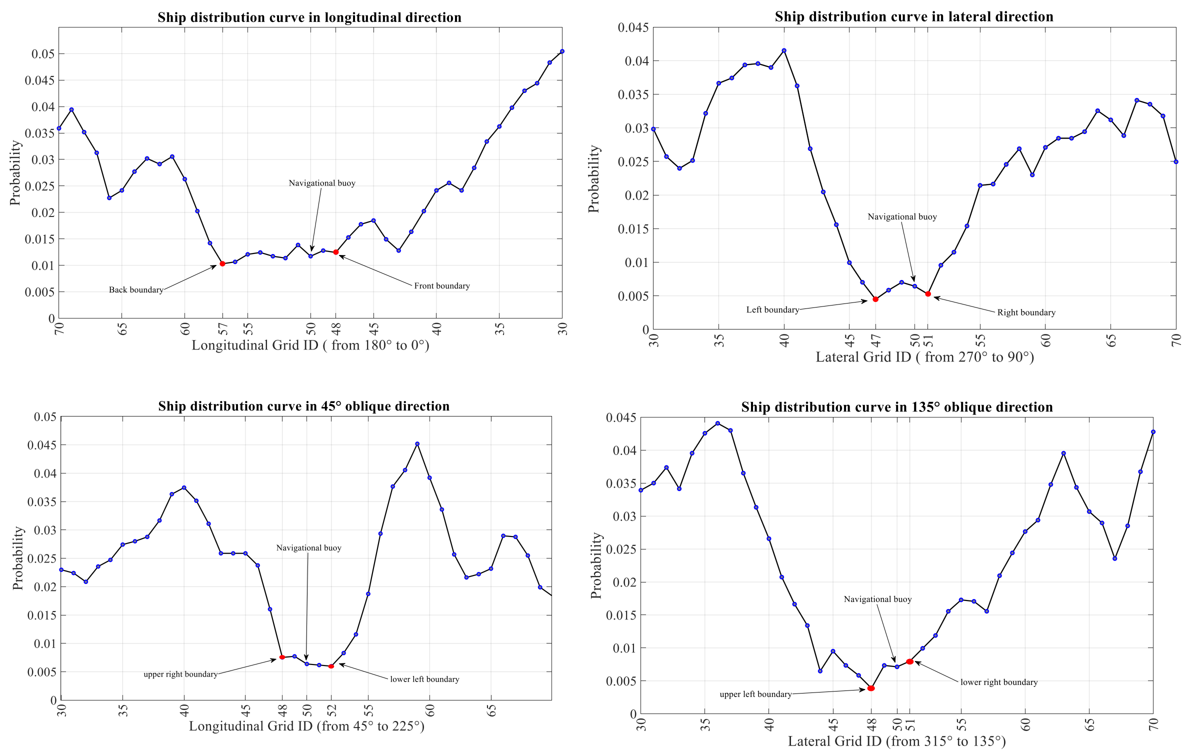

Fig. 14. The ship distribution curve in the vertical and horizontal directions.

remained in the data retained after the first detection. Moreover, the ship-buoy contact process would be entirely presented by analyzing the relative position of ships and navigational buoys in continuous time.

4.2.1. *Probabilistic buoy domain modeling*

There is no heading-related data in navigational buoy data, and the size of the navigational buoy is about 2 m * 0.4 m, which is much smaller than the size of the ship. To show the relationship between the ship and navigational buoy, the heading of all navigational buoys is assumed to be true north. To facilitate data selection, the filtering area is set to a rectangular area with a side length of 1000 m, centered on the navigational buoy. It means that the data in the circumscribed rectangle of the circle with a radius of 500 m, centered on the navigational buoy, will be used as the basic data for constructing the buoy domain model. In addition, the length of the grid side is set to 10 m, and the grid pattern is 100 * 100. The grid pattern of the ship distribution of the corresponding navigational buoys is superimposed together, and the result is shown in Fig. 13. For better visualization, only the grids from ID 30 to ID 70 in the grid pattern are displayed. There is a distinct domain of the navigational buoy, although a closed area like the ship domain is not obvious.

To obtain the forbidden boundary around the navigational buoy, the ship grid distribution curves are drawn from a probabilistic perspective, as shown in Fig. 13. For example, the lowest probability areas are 20 m, 28.2 m, 10 m, 14.1 m, 70 m, 28.7 m, 30 m, and 28.2 m from the center to the front direction ($\theta = 0°$), right front direction ($\theta = 45°$), right direction ($\theta = 90°$), right rear direction ($\theta = 135°$), rear direction ($\theta = 180°$),

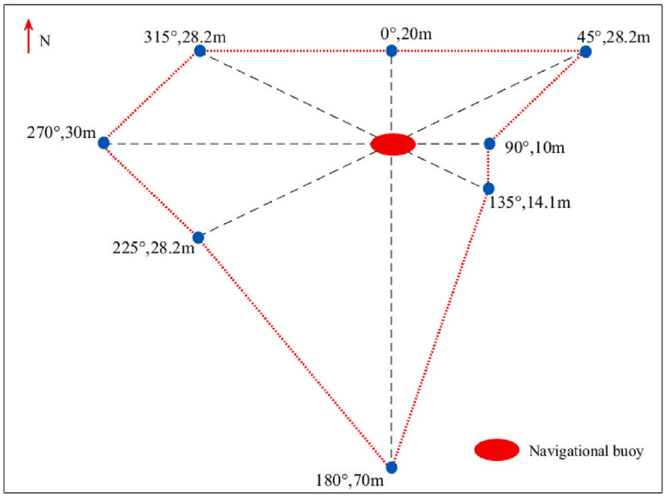

Fig. 15. The forbidden boundary around the navigational buoy (θ = 0°, 45°, 135°, 180°, 225°, 270°, 315°).

left rear direction (θ = 225°), left direction (θ = 270°), and left front direction (θ = 315°), respectively, as shown in Figs. 14 and 15. The cases of forbidden boundary around the navigational buoy are presented as Formulas (11)-(13).

$$QBD = \{(x, y)|f(x, y; \theta) \leq 1, \theta = \{0, 45, 90, 135, 180, 225, 270, 315\}\} \tag{11}$$

The boundary functions $f(.)$ can be described as follows:

$$f(x, y; \theta) = \frac{2x}{(1 + \mathrm{sgn}x)R_\theta - (1 - \mathrm{sgn}x)R_\theta} + \frac{2y}{(1 + \mathrm{sgny})R_\theta - (1 - \mathrm{sgny})R_\theta} \tag{12}$$

where the sign function sgn(.) is defined as follows:

$$\mathrm{sgn}x = \begin{cases} 1, x \geq 0 \\ -1, x < 0 \end{cases} \tag{13}$$

After determining the forbidden boundary around the navigational buoy, the forbidden boundary is combined with traffic density around the navigational buoy, indicating that 96.8% of ships evaded the region (red line in Fig. 16). The results show that the navigation buoy domain is different from to ship domain, and it has the potential to assist surveillance operators and master on board by improving their cognitive abilities in dangerous traffic scenarios.

In consideration of navigational buoy motions, the navigational buoy motion occurs in a circle area, as shown in Fig. 17. Thus, the model of the buoy domain is assumed as a circle, as described in Section 3.1. Furthermore, 99.7% of navigational buoy positions are located within 70 m. Therefore, the radius of the forbidden boundary of the buoy domain is set as 70 m (Fig. 17) to guarantee that no striking ships are missed. Finally, the buoy domain is determined and used for preliminary detection.

4.2.2. Preliminary detection based on the forbidden boundary of the probabilistic buoy domain

Preliminary detection is centered on the position of the navigation buoy using the forbidden boundary to identify potential striking ships in

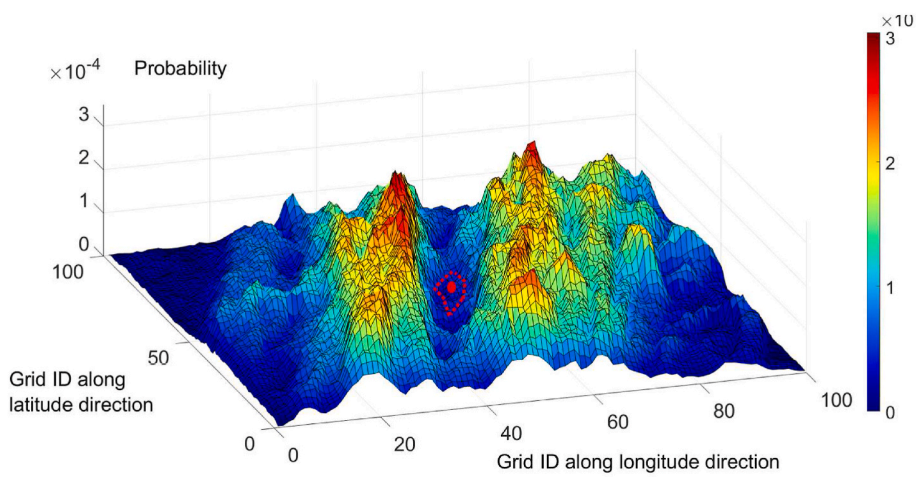

Fig. 16. The forbidden boundary and ship traffic density around the navigational buoy.

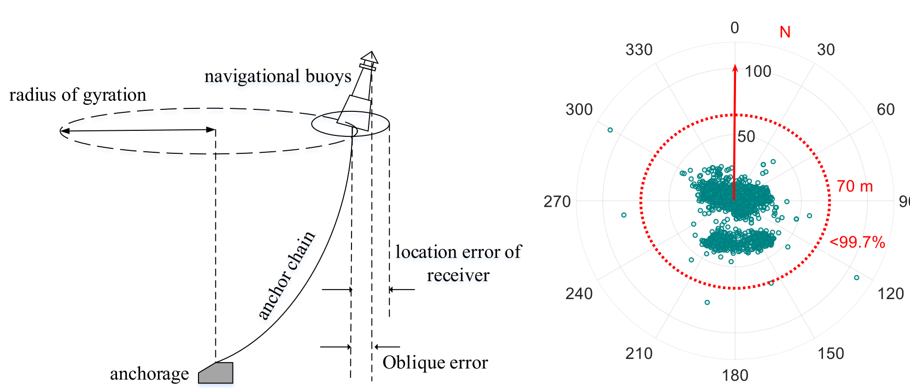

Fig. 17. Navigational buoy positions in real operational conditions.

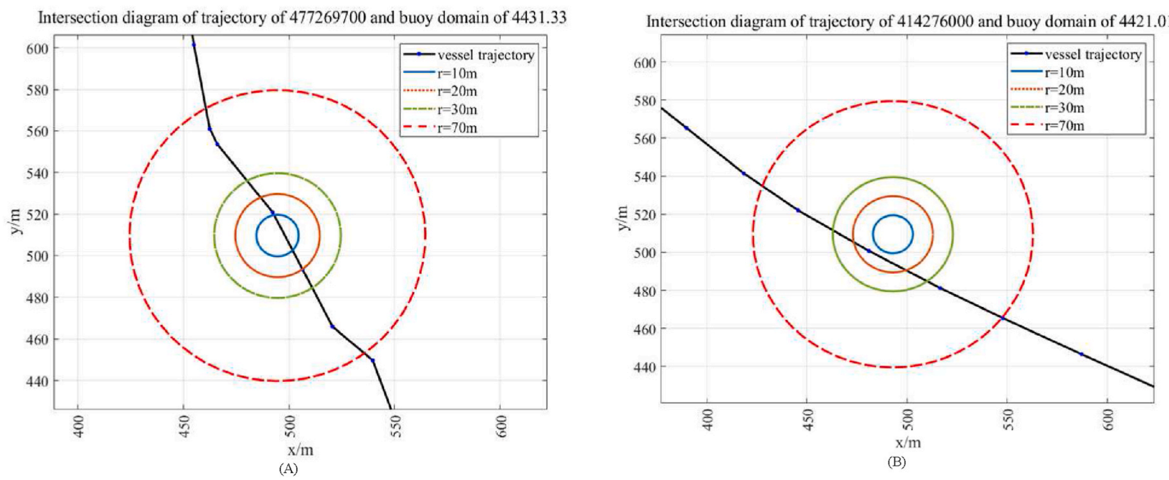

Fig. 18. Intersection diagrams of trajectories of different potential striking ships and domains of the corresponding navigational buoys.

Table 3
Results of contact detection based on the buoy domain.

Navigational buoy	Number of potential striking ships	Is striking ship included	The ratio of potential striking ships to all ships
4417.13	5	Yes	0.0781
4421.01	14	Yes	0.2090
4417.18	6	Yes	0.0343
4650.908	1	Yes	0.1667
4374.42	11	Yes	0.0564
4431.33	2	Yes	0.0606

the contact accidents. Fig. 18 shows the scene of multiple potential ships and corresponding navigation buoys with different radii to verify the effectiveness of the buoy detection method and the validity of the 70 m radius value.

According to Fig. 18, the trajectories of different potential striking ships intersect with buoy domains with different radii. Since navigational buoys can drift, it is unnecessary to ensure that navigational buoys are located at the center of the rectangle during the buoy domain modeling, and the center of buoy domains can deviate from the center of the rectangles. The trajectory in Fig. 18 (A) passes through a buoy domain with a radius of 10 m, while the trajectory in Fig. 18 (B) intersects with a buoy domain with a radius of only 20 m, showing a relatively large radius of 70 m.

To further test the detection effect of potential striking ships based on the buoy domain, all potential striking ships within 1000 m of the navigational buoys on the accident day are traced, as shown in Table 3. Although there are still some non-accident ships, the striking ship can be detected based on the buoy domain. The results present that the ratio of potential striking ships to all vessels is tiny, which indicate that the potential striking ships are identified and most non-accident ships are filtered out. And without the proposed method, we are difficult to identify the striking ships from a large number of ships around the navigational buoy. Therefore, the effectiveness of the contact detection method based on the buoy domain is verified. In addition, Table 3 also indicates that introducing a new method that can be used for ship-buoy contacts detection and striking ship identification is necessary.

4.2.3. Accurate detection based on the bounding box

To identify the striking ship among all potential striking ships and track the contact process, accurate detection is conducted based on the bounding spheres of the navigational buoy and the OBB of the ship. However, the radius of the bounding sphere is uncertain, and the size of the OBB can be determined based on the size of the ship. Section 3.2 provides more details.

- Bounding box determination for accurate detection

According to the principles described in Section 3.3, combined with the construction results of the probabilistic buoy domain model, the radius is set to 2 m (approximate length of the navigational buoy), 10 m, and 20 m. Fig. 19 shows the contact detection between the striking ship and navigational buoy 4374.42 based on the bounding box method. Only the ship positions in Fig. 19 (A) and (D) are the original AIS data, while the ship positions in Fig. 19(B) and (C) are predicted by piecewise cubic Hermite interpolation. The navigational buoy data in all subgraphs are obtained by interpolation, which means the contact will be missed if there is only Fig. 19 (A) and (D). Therefore, the prediction of missing data is necessary.

Fig. 19 (B) displays that the contact start phase is detected based on the bounding box method, showing the instantaneous contact. However, the bounding boxes of the ship and the navigational buoy overlap each other in Fig. 19 (C). The main reason is that navigational buoy 4373.42 drifted after being collided with the ship and is no longer in that position. However, the ship-buoy contact happened within 1 min around 7:14:7, whereas the next returned data of the navigational buoy is at 8:12:3, resulting in a significant prediction error.

The minimum radius r of different navigational buoys and the number of other potential striking ships detected within the radius r are illustrated in Table 4. It can be seen that the minimum radius is mainly 2 m, and the maximum radius of the striking ship is 10 m. The reasons for the result include GPS positioning errors and data prediction errors. The radius $r_0 = 10m$ is tested to determine if more potential striking ships are detected than the minimum radius r. According to the last two columns in Table 4, the results are the same. Therefore, the radius is finally set to 10 m.

- Confirmation of striking ships based on navigational buoy data

After contact detection based on the bounding box, there are still one or two non-accident ships. However, depending on the relative positions of the navigation buoy and the ship, non-accident ships cannot be separated from potential striking ships. The contact can also be solved by analyzing if any significant data change happens after the bounding sphere of the navigational buoy intersects with the OBB of the ship.

When navigational buoys are collided by ships, there may be three abnormalities in the navigational buoys data: a) The drift distance of the navigational buoy increases because the navigational buoy is taken away by the striking ship. b) Due to the impact of the contact, the drift velocity of the navigational buoy at the time of contact is higher than

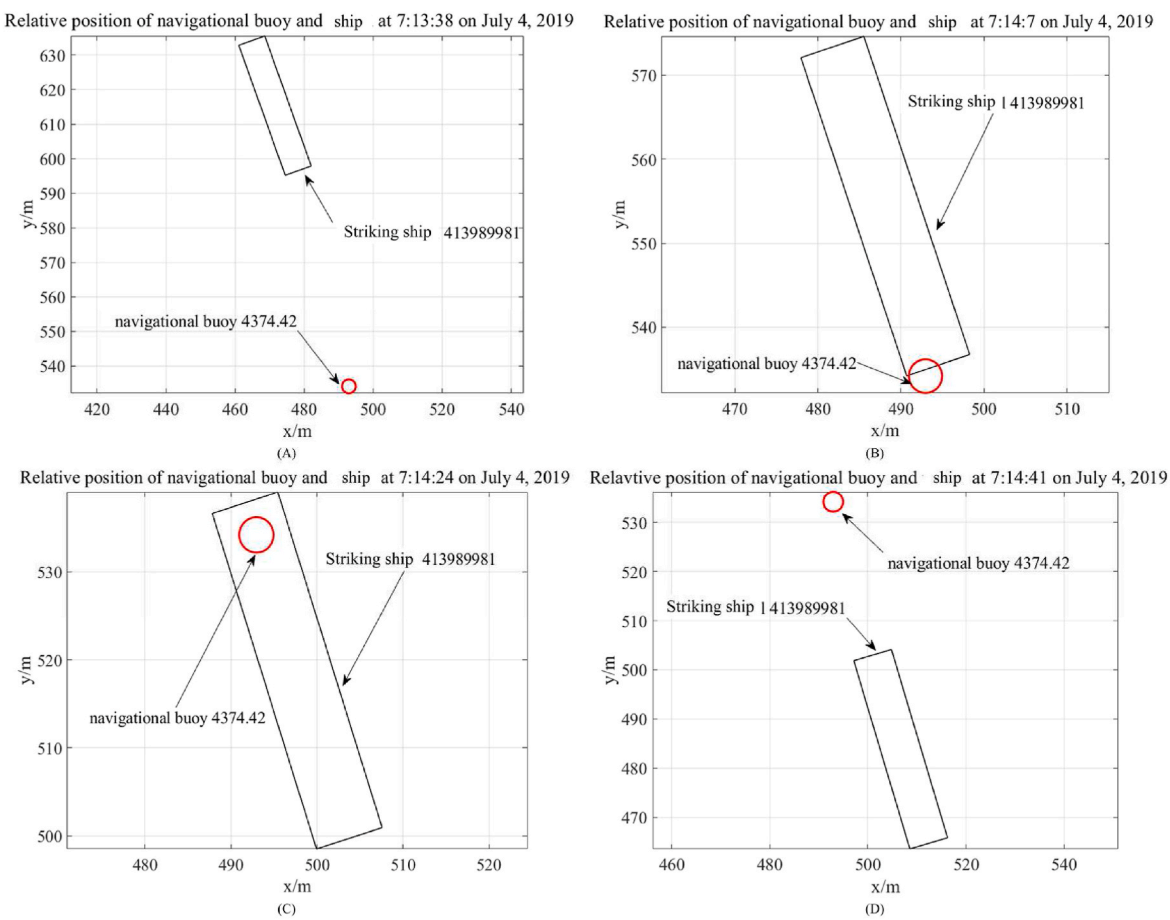

Fig. 19. Contact detection between the striking ship and navigational buoy 4374.42 with a radius of 2 m.

Table 4
The radius detection of the bounding box of navigational buoys.

Navigational buoy	The minimum radius r for the striking ship/m	Number of other ships detected within the minimum radius r for the striking ship	Number of other ships detected by $r_0 = 10m$
4417.13	2	0	0
4421.01	10	2	2
4417.18	2	0	0
4650.908	2	0	0
4374.42	2	1	1
4431.33	2	1	1

that during the normal period, although it is not taken away. c) The working voltage of the navigational buoy is normal, but the navigational buoy no longer returns data, which means it is damaged.

For navigational buoy 4374.42, there are two potential striking ships, one of which is a striking ship. Ship 413989981 collided with the navigational buoy at 7:14:03, and the other ship 413487181 collided with the navigational buoy at 11:30:31. The navigational buoy data near the two timestamps are analyzed. It indicates that the navigational buoy no longer returned data after 7:10:09, but the working voltage remained normal. The next navigational buoy data is returned at 8:12:03. Therefore, navigational buoy 4374.42 is probably damaged, and the staff reset it in time. Additionally, no abnormalities occur around 11:30:31, indicating that ship 413487181 does not collide with navigational buoy 4421.01.

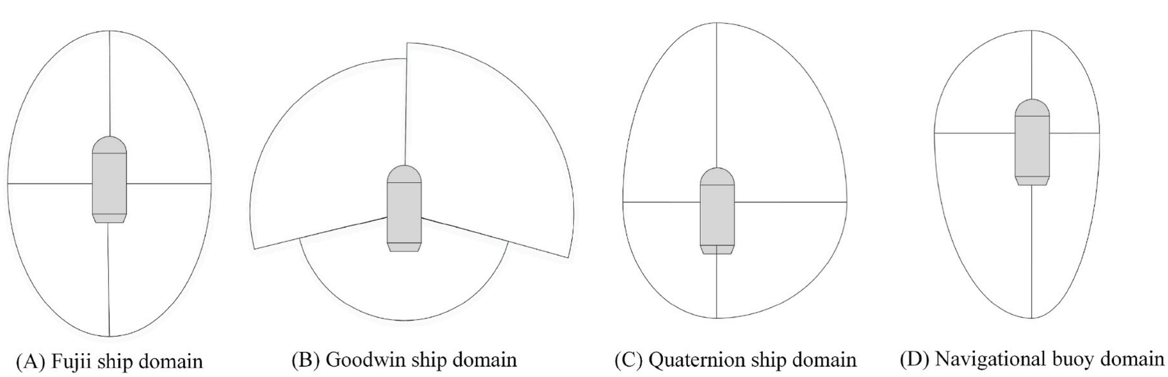

Fig. 20. Comparison between ship domain and buoy domain.

5. Discussion

In this section, the proposed buoy safety domain and the detection method of the striking ship in ship-buoy contact will be compared and discussed. In addition, the characteristics of ship-buoy contact and the striking ship will be analyzed to gain insight into contact risk mitigation.

- Features of the proposed method and its application

The paper proposes a probabilistic analytics method to detect the striking ship in ship-buoy contact. The buoy domain is constructed to exclude the potential striking ships with a higher contact probability and reduce the amount of intersection test calculation based on the bounding box, as shown in Fig. 20 (D). Different from the traditional ship domain in Fig. 20(A) and (B) (see more in Fujii and Tanaka, 1971; Goodwin, 1975), the proposed buoy domain indicates that the distances between the center and the fore, aft, port, and starboard boundaries are 20 m, 70 m, 30 m, and 10 m, respectively. And the buoy domain is similar to the quaternion ship domain in Fig. 20 (C) (Wang, 2010). However, the front-right ship domain is larger than the left-rear ship domain, while the buoy domain is the opposite. Although a circle with a radius of 70 m is set as the forbidden boundary to avoid missed detections, other values of radii can be selected according to the application requirements. In addition, the size of the navigational buoy is generally 2 m * 0.4 m, which means the fore and aft lengths of the buoy domain are 10 and 35 times the length of the navigational buoy, the widths of the port and starboard are 75 and 25 times the width of the navigational buoy, and the ratios are much greater than that of the ship domain.

Therefore, it is necessary to develop a probabilistic buoy domain for detecting a striking ship in ship-buoy contact by restoring ship trajectories and navigational buoys at times. According to the results, a safety buoy domain exists from a probabilistic perspective. The distance between the center and the forbidden boundary of the buoy domain in the fore, aft, port, and starboard are 20 m, 70 m, 30 m, and 10 m, respectively. What is outlined could provide essential support to the masters of ships as part of an intelligent decision support system to avoid the ship-buoy contact.

Under the process of preliminary detection using the probabilistic buoy domain, the bounding box method is used to determine the striking ship accurately. The OBB of ships and the bounding sphere of navigational buoy are designed, respectively. Table 4 reveals the effectiveness of the bounding box. However, since it is usually adopted in the virtual environment, information update lags, and positioning errors are encountered. The errors or incomplete information in AIS data will result in the results with uncertainties. As shown in 18 (B), the OBB of ship 413989981 coincides with the bounding sphere of navigational buoy 4374.42. The intersection is because navigational buoy 4374.42 no longer broadcasts data after a contact but stays in the original location after reset. Consequently, the position obtains by interpolation is not the actual counterpart of navigational buoy 4374.42. Few false detections still exist in Table 4. Therefore, it is necessary to analyze the navigational buoy data near the moment of each contact scenario to identify the striking ship by observing abnormalities.

In addition, the traditional determination of the striking ship is not reliable, i.e., analyzing navigational buoy data to find out the striking ships passing through the waters near the navigational buoy around the time of abnormal data. Because the abnormal states in Section 3.3 can also occur in other situations instead of only contact accidents. For example, navigational buoys no longer broadcast data due to equipment failures, and the drift velocity of navigational buoys is also affected by the hydrometeorological conditions and broadcast frequency of buoy data. However, in these cases, it is almost impossible for a ship to detect a contact by the bounding box method, ensuring the reliability of the detection framework.

- Uncertainties and limitations

The limitations or uncertainties should be clarified there. The piecewise cubic Hermite interpolation and piecewise cubic spline interpolation (Xin et al., 2021) are applied for AIS data streams spanning over an interval of 30 s. And the navigational buoy data is interpolated over an interval of 120 s. The used ship-buoy contact records contain AIS data and the navigational buoy data. We assumed that the AIS transceiver's position is located on the stern of the ship and the distance from the AIS transceiver's position to the bow is 90% of ship length (Zhang et al., 2021(a); Zhang et al., 2022). Additionally, we set AIS transceiver's position located in the centre of navigational buoy. These uncertainties have slight impacts on the results and findings. As shown in Table 3, even though the ratio of potential striking ships to all vessels is tiny, most non-accident ships are filtered out. However, more than one striking ship were detected in the study. This is because the position offset of AIS data and navigational buoy data caused by the locations of the antennas can be larger than the buoy domain and bounding box. Future work could focus on comparing the results for various ship types in this area using higher quantity data (e.g., from CCTV, radar, etc) (Zhou et al., 2022). The relation of factors such as ship type, size, and other environmental factors such as visibility and time of day could be studied. In addition, the probabilistic buoy domain will be set as a reference displayed on navigation equipment, which can provide essential support to the masters for ship-buoy contact risk mitigation. And the safe area of buoy is related to the density and movement of ships, indicating that it is necessary to construct the corresponding buoy domain in specific waters through our methodology.

6. Conclusion

The paper introduces a probabilistic analytics method for detecting ship-buoy contact using AIS data and navigational buoy data. The proposed method is developed by combining the buoy domain method and bounding box model. The buoy domain is obtained by calculating the probability density function using AIS data. And the potential striking ships are identified using the developed buoy domain as preliminary detection. Furthermore, the ship bounding box and the buoy bounding box are developed to detect the real striking ship as second detection. The proposed method is demonstrated using AIS data, navigational buoy data, and historical contact accidents covering 2018, 2019, and 2020 in the South China Sea. The data and code are provided as shown in Appendix B.

Results show that there is an obvious forbidden domain for buoys and the characteristics of the domain are different from those of the ship domains in current studies (Figs. 1 and 16). The forbidden boundaries of the buoy domain in the fore, aft, port, and starboard are 20 m, 70 m, 30 m, and 10 m, respectively. The developed buoy domain could be used to detect potential striking ships. Furthermore, based on the ship size of the detected the potential striking ships and buoy domain, the ship bounding box and buoy bounding box are determined based on the OBB. The innovative use of bounding boxes is promising and useful for the actual maritime ship buoy contact detection to find out the striking ships and the spatial-temporal contacting information. Overall, the paper proposes a striking ship detection tool to provide evidence for buoy ship contact accident accountability. Besides that, obvious forbidden domain for buoy could provide support information for ship buoy contact risk mitigation as well as assist surveillance operators and master on board to improve their cognitive abilities in dangerous traffic scenarios.

The safety buoy domain and ship-buoy contact risk are also influenced by other factors, e.g., hydrometeorological conditions, waterway complexity, ship navigation systems (specifically the autopilot and ARPA radar), operational instructions, and procedures of the shipping company. It is important to improve detection accuracy by filtering location data and reducing prediction errors in future research. Since AIS equipment can be turned off by crews on ships, a detection method based on radar or camera can be developed to monitor ships passing through the waters near the navigational buoy.

CRediT authorship contribution statement

Lei Liu: Methodology, Software, Investigation, Data curation, Funding acquisition, Writing – original draft, Writing – review & editing. **Mingyang Zhang:** Methodology, Validation, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Yue Hu:** Methodology, Software, Writing – review & editing. **Wei Zhu:** Writing – review & editing, Funding acquisition. **Sheng Xu:** Writing – review & editing, Funding acquisition. **Qing Yu:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research was supported by the National Natural Science Foundation of China (Number: 52071247) and the China Scholarship Council Scholarship (Number: 202106090198).

Appendix A. Description of data columns

Table A1
Description of AIS data.

Data type	Data label	Description
AIS data	MMSI	Maritime Mobile Service Identity
	TIMESTAMP	The timestamp of AIS data
	LONGITUDE	The longitude of the position
	LATITUDE	The latitude of the position
	SOG	The shipping speed over ground
	COG	The shipping course over ground
	HEADING	The ship heading
	LENGTH	The ship length
	WIDTH	The ship width

Table A2
Description of the data content of navigational buoys.

Data type	Data column	Description
Navigational buoy data	NAVMARKID	Identification of navigational buoy
	COLLECTTIME	Data collecting time
	LATITUDE	The latitude of the position
	LONGITUDE	The longitude of the position
	OFFSETDISTANCE	The distance from placement position
	LAMPVOLTAGE	The voltage of the lamp of navigational buoys
	LAMPCURRENT	The current of the lamp of navigational buoys
	OFFSET DIRECTION	The angle after rotating clockwise from due north relative to the placement position

Appendix B. Data availability and open-source code implementation

- The complete analysis performed, and the figures created in the paper can be found at: https://github.com/zhangmingyangliu/A-Probabilistic-Analytics-Method-to-Detect-the-Ship-Buoy-Collision-in-Real-Operational-Conditions/commit/a08de3fcc2e8f593a87b33369808ec8a2aec30d6.
- This repository also contains all datasets preprocessed and structured as they are used: https://github.com/zhangmingyangliu/A-Probabilistic-Analytics-Method-to-Detect-the-Ship-Buoy-Collision-in-Real-Operational-Conditions/commit/a08de3fcc2e8f593a87b33369808ec8a2aec30d6.

References

Adhikari, B., Huttunen, H., 2021. Iterative bounding box annotation for object detection. In: 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, pp. 4040–4046.

Antão, P., Soares, C.G., 2019. Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks. Accid. Anal. Prev. 133, 105262.

Chaturvedi, S.K., 2019. Study of synthetic aperture radar and automatic identification system for ship target detection. Journal of Ocean Engineering and Science 4 (2), 173–182.

Chen, P., Huang, Y., Papadimitriou, E., Mou, J., van Gelder, P.H.A.J.M., 2020. An improved time discretized non-linear velocity obstacle method for multi-ship encounter detection. Ocean. Eng. 196, 106718.

Delgado Román, M.D.C., Cortés, P., Onieva, L., Escudero Santana, A., 2008. Design of an intelligent waterway ambient infrastructure based on multiagent systems and wireless sensor networks, 4. In: International Journal of Systems Applications, Engineering & Development, 2, pp. 137–146.

Du, L., Banda, O.A.V., Huang, Y., Goerlandt, F., Kujala, P., Zhang, W., 2021 a. An empirical ship domain based on evasive maneuver and perceived collision risk. Reliab. Eng. Syst. Saf. 213, 107752.

Du, L., Goerlandt, F., Banda, O.A.V., Huang, Y., Wen, Y., Kujala, P., 2020. Improving stand-on ship's situational awareness by estimating the intention of the give-way ship. Ocean. Eng. 201, 107110.

Du, L., Valdez Banda, O.A., Goerlandt, F., Kujala, P., Zhang, W., 2021. Improving near miss detection in maritime traffic in the northern baltic sea from AIS data. J. Mar. Sci. Eng. 9 (2), 180.

EMSA, 2020. Annual Overview of Marine Casualties and Incidents 2020. EMSA - European Maritime Safety Agency, Lisboa.

Fujii, Y., Tanaka, K., 1971. Traffic capacity. J. Navig. 24 (4), 543–552.

Gil, M., 2021. A concept of critical safety area applicable for an obstacle-avoidance process for manned and autonomous ships. Reliab. Eng. Syst. Saf. 214, 107806.

Gil, M., Montewka, J., Krata, P., Hinz, T., Hirdaris, S., 2020. Determination of the dynamic critical maneuvering area in an encounter between two vessels: operation with negligible environmental disruption. Ocean. Eng. 213, 107709.

Goerlandt, F., Montewka, J., 2015. A framework for risk analysis of maritime transportation systems: a case study for oil spill from tankers in a ship-ship collision. Saf. Sci. 76, 42–66.
Goodwin, E.M., 1975. A statistical study of ship domains. J. Navig. 28 (3), 328–344.
Gucma, L., Marcjan, K., 2012. Probabilistic Model of Minimal Passing Distances of Vessels Navigating in Polish Coastal Waters. PSAM11 &ESREL.
Hansen, M.G., Jensen, T.K., Lehn-Schiøler, T., Melchild, K., Rasmussen, F.M., Ennemark, F., 2013. Empirical ship domain based on AIS data. J. Navig. 66 (6), 931–940.
Hassel, M., Utne, I.B., Vinnem, J.E., 2017. Allision risk analysis of offshore petroleum installations on the Norwegian Continental Shelf-an empirical study of vessel traffic patterns. WMU J. Maritime Affair. 16 (2), 175–195.
Hassel, M., Utne, I.B., Vinnem, J.E., 2021. An allision risk model for passing vessels and offshore oil and gas installations on the Norwegian Continental Shelf. Proc. Inst. Mech. Eng. O J. Risk Reliab. 235 (1), 17–32.
Huang, Y., Van Gelder, P.H.A.J.M., 2020a. Collision risk measure for triggering evasive actions of maritime autonomous surface ships. Saf. Sci. 127, 104708.
Huang, Y., van Gelder, P.H.A.J.M., 2020b. Collision risk measure for triggering evasive actions of maritime autonomous surface ships. Saf. Sci. 127, 104708.
Kim, D., Kim, S., Jeong, S., Ham, J.W., Son, S., Moon, J., Oh, K.Y., 2021. Rotational multipyramid network with bounding-box transformation for object detection. Int. J. Intell. Syst. 36 (9), 5307–5338.
Koukaki, T., Tei, A., 2020. Innovation and maritime transport: a systematic review. Case Stud. Transport Pol. 8 (3), 700–710.
Lei, J., Liu, L., Chu, X., H, W., Liu, X., Liu, C., 2021. Automatic identification system data-driven model for analysis of ship domain near bridge-waters. J. Navig. 74 (6), 1284–1304.
Lei, P.R., Tsai, T.H., Wen, Y.T., Peng, W.C., 2017. ConflictFinder: mining maritime traffic conflict from massive ship trajectories. In: IEEE International Conference on Mobile Data Management. IEEE. https://doi.org/10.1109/MDM.2017.60.
Li, S., Meng, Q., Qu, X., 2012. An overview of maritime waterway quantitative risk assessment models. Risk Anal.: Int. J. 32 (3), 496–512.
Liu, D., Shi, G., 2020. Ship collision risk assessment based on collision detection algorithm. IEEE Access 8, 161969–161980.
Liu, L., Jiang, Z., Chu, X., Zhong, C., Zhang, D, 2019. Automatic Identification System Data Restoration and Prediction. Journal of Harbin Engineering University 40 (6), 1072–1077.
Liu, L., Liu, X., Chu, X., Li, J., Jiang, Z., He, W., 2017. Coverage effectiveness analysis of AIS base station: a case study in Yangtze River. In: 2017 4th International Conference on Transportation Information and Safety (ICTIS). IEEE, pp. 178–183.
Liu, J., Zhang, J., Yan, X., Soares, C.G., 2022. Multi-ship collision avoidance decision-making and coordination mechanism in Mixed Navigation Scenarios. Ocean Engineering 257, 111666.
Lopez-Santander, A., Lawry, J., 2017. An ordinal model of risk based on mariner's judgement. J. Navig. 70 (2), 309–324.
Mannering, F., Bhat, C.R., Shankar, V., Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Anal. Method. Accident Res. 25, 100113.
Moorits, E., Usk, A., 2012. Buoy collision detection. September. In: Proceedings ELMAR-2012. IEEE, pp. 109–112.
Mujeeb-Ahmed, M.P., Seo, J.K., Paik, J.K., 2018. Probabilistic approach for collision risk analysis of powered vessel with offshore platforms. Ocean. Eng. 151, 206–221.
Mullai, A., Paulsson, U., 2011. A grounded theory model for analysis of marine accidents. Accid. Anal. Prev. 43 (4), 1590–1603.
Otto, S., Petersen, U., 2003. Offshore-Windpark Kriegers Flak. 2. Ergänzung zur Risikoanalyse. Bericht Nr 155. ERD2004.
Pedersen, P.T., 2010. Review and application of ship collision and grounding analysis procedures. Mar. Struct. 23 (3), 241–262.
Pietrzykowski, Z., Wielgosz, M., 2021. Effective ship domain-Impact of ship size and speed. Ocean. Eng. 219, 108423.
Rawson, A., Brito, M., 2021. A critique of the use of domain analysis for spatial collision risk assessment. Ocean. Eng. 219, 108259.
Szlapczynski, R., Szlapczynska, J., 2017. Review of ship safety domains: models and applications. Ocean. Eng. 145, 277–289.
Szlapczynski, R., Krata, P., Szlapczynska, J., 2018a. A ship domain-based method of determining action distances for evasive manoeuvres in stand-on situations. J. Adv. Transport. 2018.
Szlapczynski, R., Krata, P., Szlapczynska, J., 2018b. Ship domain applied to determining distances for collision avoidance manoeuvres in give-way situations. Ocean. Eng. 165, 43–54.
Tam, C., Bucknall, R., Greig, A., 2009. Review of collision avoidance and path planning methods for ships in close range encounters. J. Navig. 62 (3), 455–476.
Tu, C., Yu, L., 2009. Study on collision detection algorithm of hybrid bounding box, 2009, May. In: International Forum on Information Technology and Applications, vol. 2. IEEE, pp. 190–192.
Wang, N., 2010. An intelligent spatial collision risk based on the quaternion ship domain. J. Navig. 63 (4), 733–749.
Weng, J., Yang, D., 2015. Investigation of shipping accident injury severity and mortality. Accid. Anal. Prev. 76, 92–101.
Xin, X., Liu, K., Yang, Z., Zhang, J., Wu, X., 2021. A Probabilistic Risk Approach for the Collision Detection of Multi-Ships under Spatiotemporal Movement Uncertainty. Reliability Engineering & System Safety, 107772.
Xin, X., Yang, Z., Liu, K., Zhang, J., Wu, X., 2023. Multi-stage and multi-topology analysis of ship traffic complexity for probabilistic collision detection. Expert Syst. Appl. 213, 118890.
Yu, Q., Liu, K., Yang, Z., Wang, H., Yang, Z., 2021. Geometrical risk evaluation of the collisions between ships and offshore installations using rule-based Bayesian reasoning. Reliab. Eng. Syst. Saf. 210, 107474.
Yu, Q., Liu, K., Zhang, J., Xin, X., Guedes Soares, C., Santos, T.A., 2018. Risk analysis of ships & offshore wind turbines collision: risk evaluation and case study. Progress Maritime Eng. Technol. 511–524.
Yuan, X., Zhang, D., Zhang, J., Zhang, M., 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.
Zeng, Q., Huang, H., Pei, X., Wong, S.C., 2016. Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. Anal. Method. Accident Res. 10, 12–25.
Zhang, M., Kujala, P., Hirdaris, S., 2022. A machine learning method for the evaluation of ship grounding risk in real operational conditions. Reliab. Eng. Syst. Saf. 226, 108697.
Zhang, J., Liu, J., Hirdaris, S., Zhang, M., Tian, W., 2022. An interpretable knowledge-based decision support method for ship collision avoidance using AIS data. Reliab. Eng. Syst. Saf. 108919.
Zhang, L., Meng, Q., 2019. Probabilistic ship domain with applications to ship collision risk assessment. Ocean. Eng. 186, 106130.
Zhang, M., Conti, F., Le Sourne, H., Vassalos, D., Kujala, P., Lindroth, D., Hirdaris, S., 2021b. A method for the direct assessment of ship collision damage and flooding risk in real conditions. Ocean. Eng. 237, 109605.
Zhang, M., Montewka, J., Manderbacka, T., Kujala, P., Hirdaris, S., 2021a. A big data analytics method for the evaluation of ship-ship collision risk reflecting hydrometeorological conditions. Reliab. Eng. Syst. Saf. 213, 107674.
Zhang, M., Zhang, D., Fu, S., Kujala, P., Hirdaris, S., 2022. A predictive analytics method for maritime traffic flow complexity estimation in inland waterways. Reliab. Eng. Syst. Saf., 108317
Zhang, W., Feng, X., Goerlandt, F., Liu, Q., 2020. Towards a convolutional neural network model for classifying regional ship collision risk levels for waterway risk analysis. Reliab. Eng. Syst. Saf. 204, 107127.
Zhang, W., Goerlandt, F., Kujala, P., Wang, Y., 2016. An advanced method for detecting possible near miss ship collisions from AIS data. Ocean. Eng. 124, 141–156.
Zhang, W., Goerlandt, F., Montewka, J., Kujala, P., 2015a. A method for detecting possible near miss ship collisions from AIS data. Ocean. Eng. 107, 60–69.
Zhang, W., Montewka, J., Goerlandt, F., 2015b. Semi-qualitative method for ship collision risk assessment. In: Nowakowski (Ed.), Safety and Reliability; Methodology and Applications. Taylor and Francis Group, London, pp. 1563–1572.
Zhao, Y., Li, W., Shi, P., 2016. A real-time collision avoidance learning system for Unmanned Surface Vessels. Neurocomputing 182, 255–266.
Zheng, L., Sayed, T., 2020. A bivariate Bayesian hierarchical extreme value model for traffic conflict-based crash estimation. Anal. Method. Accident Res. 25, 100111.
Zhong, C., Jiang, Z., Chu, X., Liu, L., 2019. Inland ship trajectory restoration by recurrent neural network. J. Navig. 72 (6), 1359–1377.
Zhou, T., Wu, W., Peng, L., Zhang, M., Li, Z., Xiong, Y., Bai, Y., 2022. Evaluation of urban bus service reliability on variable time horizons using a hybrid deep learning method. Reliab. Eng. Syst. Saf. 217, 108090.
Zhu, X., Xu, H., Lin, J., 2001. Domain and its model based on neural networks. J. Navig. 54 (1), 97.