A Fused Data Based Real-Time Collision Warning System for Ferries in the Yangtze River

Xiaojia Liu
Ke Zhang
Han Xue
Yingtang Li

Follow this and additional works at: https://jmstt.ntou.edu.tw/journal

Part of the Fresh Water Studies Commons, Marine Biology Commons, Ocean Engineering Commons, Oceanography Commons, and the Other Oceanography and Atmospheric Sciences and Meteorology Commons
A Fused Data Based Real-Time Collision Warning System for Ferries in the Yangtze River

Xiaojia Liu a, Ke Zhang b,*, Han Xue a, Yingtang Li a

a School of Navigation, Jimei University, 1 Jiageng Road, Xiamen, Fujian 361021, China
b School of Navigation, Wuhan University of Technology, 1178 Youyi Road, Wuhan, Hubei 430070, China

Abstract

The risks for ferries in the Yangtze River are relatively high, as they frequently cross the main traffic flows, leading to more intersections with other upwards and downwards ships. Although some studies have developed many models to assess collision risks in the Yangtze River, collision warning studies on ferries are scant. Meanwhile, most of the current collision studies evaluate risk based on AIS data, which are incapable of providing real-time ship information as they are discrete-time series data. In this work, fused data combining radar and AIS data are applied in a real-time ship collision warning model to assess the dynamic risk for ferries in the Yangtze River. Firstly, data fusion technology is proposed to acquire refined ship trajectories from AIS and radar data. Then, a widely used geometric collision model is enhanced to assess the real-time collision risk for ferries. And lastly, to illustrate the model, a real case of a ferry crossing through the Yangtze River is studied. The real-time risk values of the ferries are calculated based on fused data inputs, and the output results indicate that the use of fused data provides more accurate and continuous real-time ship risks. Thus, the proposed approach is evidenced to support the development of smart maritime surveillance.

Keywords: Collision risk, Kalman filter, Data fusion, Yangtze river

1. Introduction

With the increasing demand for transportation in China, the past decades have witnessed a great development of inland transportation. The Yangtze River has become the world’s busiest waterway that crosses China from the east mainland to the west shore [21]. As water traffic becomes busier, some potential safety risks, such as collision accidents, are becoming serious. To ensure safety, maritime managers have endeavoured to prevent collision accidents by identifying the collision risks in the Yangtze River, while scholars have also proposed several risk models. In general, the quantification of collision risk models is used to aid mariners in assessing collision situations and to aid collision avoidance decision-making [11,28]; qualitative risk models are applied to evaluate the overall water risks by considering risk impact factors such as weather, traffic conditions and human impacts [23].

Many studies (e.g., Bukhari et al.,[4], Goerlandt and Montewka, [10], Montewka et al., [17]) have focused on the assessments of ship collision risk, which are useful for collision accident evaluation and prevention at sea. For individual ship encounter collision situations (i.e., crossover, head-on and overtaking situations), two crucial parameters are frequently used in collision risk studies: the distance at the closest point of approach (DCPA) and the time to the closest point (TCPA) [19]. The DCPA represents the closest distance at which two encountering ships pass each other, and the TCPA represents the time from the current position to the closest point. Moreover, the DCPA and TCPA are calculated from ship dynamic variables, including ship speed, position, course, etc. The
collision risk of two ships in an encounter situation is aggregated from the parameters [1]. For instance, Kao et al. [13] used fuzzy logic approaches to develop a risk model. A ship’s safety zone was delineated to detect any invading ships, and any potential collision candidates were identified if their DCPA and TCPA were smaller than the threshold values. Zhang et al. [29] applied collision parameters with a ship domain approach to establish a probabilistic ship collision risk model. The model was tested by calculating the commercial ship’s collision risk in the Singapore Strait. Moreover, Zhang et al. [30] proposed a fuzzy-based collision risk model to assess the ship collision risk in Tianjin Port. In those studies, mathematical approaches are used to integrate the risk factors, which are collected from real situations [16].

With the wide use of Automatic Identification System (AIS) in the field of maritime studies, e.g., AIS-based vessel emission monitoring [22], traffic flow characteristics [25] and collision studies [24], AIS data are becoming popular for evaluating ship collision risk, as they provide discrete time series data to describe ship dynamic states during voyages. However, a drawback of AIS data is that a collision warning system requires continuous data (e.g., ship position, speed, and course) to ensure real-time ship position tracking and to calculate the potential conflict probabilities, while raw AIS data contain errors and need to be filtered before use in risk calculations. For instance, the transit rates of the AIS system on ships in the Yangtze River waters are different, ranging from 2 s to over 30 min. As some encounter situations remain for a few minutes, the lack of continuous ship tracking will lead to unpredictable danger for the ships involved. To overcome these difficulties, this paper proposes a fused data technology to aid real-time collision assessment, which not only improves the reliability of the risk model but also realizes continuous risk target tracking. The two types of radar and AIS data are fused to ensure real-time ship monitoring, while the collision risks for surrounding ships are calculated. To test the proposed approach, the risk for the voyage of a ferry crossing over the Yangtze River is calculated and monitored. The results show that the fused data-based model is capable of identifying the collision candidates for an individual ship and sends warning information to users (e.g., Vessel Traffic Service (VTS) and crewmembers on board) to provide help for the mariners on the ship and maritime administration officers to take action to avoid collisions.

However, different aspects can be more relevant when characterizing the risk in maritime transportation. The risk of a given ship can be composed of static and dynamic components. Static risk quantifies the risk related to ship characteristics, such as the ship type, flag, and size. It is derived from historical accident data and inspection records, i.e., ship deficiencies and detentions. Dynamic risk describes the risk of a ship accident when navigating in a specific geographical area and maritime traffic conditions. For instance [33], applied the safety assessment model for shipping and offshore in the North Sea to utilize real-time risk calculations and suggested that the dynamic risk can be the combination of some geometric collision parameters (e.g., encounter angle, the closest point of approach and time to closest point approach), which are all used to assess the collision risk for ship encounter scenarios. On such a basis, in this paper, “dynamic risk” only considers the collision risk of the ship in the study area.

The remainder of the paper is organized as follows. In Section 2, risk definitions and state-of-the-art collision models are introduced. In Section 3, an AIS & Radar data fusion approach is introduced, while a dynamic collision risk model is established by using a popular geometric collision risk analysis method (i.e., fuzzy logic-based collision models). In Section 4, a case is studied by using the proposed data fusion approach and the collision risk model. In Section 5, the improvements and limitations are discussed, and in Section 6, the conclusion of this study is drawn.

2. Related works

2.1. Risk definitions

The overview of some categories of definitions of risk was concluded by Aven [3]. The conceptual classes are based on the parameters considered in each definition, which provides insight into how risk is defined in the different application areas. A total of nine categories are summarized. The simplest definition is that risk is the probability of an undesirable event (i.e., incident) or probability of loss [10]. Meanwhile, the understanding of risk is sometimes subjective; a risk definition suggests that the risk is considered to exist independent of an assessor and thus can be physically described with an accident consequence [9]. Based on these definitions, the parameters are extended to consider the consequence of the event so that risk is defined as the combination of the probability of occurrence of an event and consequence. Despite the above-mentioned definitions, risk definitions are numerous, including the expected value of the event
occurrence and utility of consequence, objective uncertainty, understanding or statistical variations, etc. [2].

In the maritime risk assessment domain, probability and consequence are two main parameters used to assess the collision risk [15]. However, the objects of the studies are significantly different. Probability studies aim to calculate the collision probability for ships in encounter situations using the obtained results to identify the collision candidates and potential incidents [8]. The studies in terms of the consequence viewpoint are more focused on the damage/loss of collision accidents [14]. They use technologies such as finite element analysis to simulate the collision consequence under different situations and then calculate the consequence to assess the risk [5]. As our study aims to develop a real-time warning model, the risk perspective of collision probability is applied.

### 2.2. State-of-the-art collision warning studies

Collision warning systems (CWSs) enhance the situational awareness of VTS officers and crew-members on board and aid in collision-avoidance decision making. The most widely used CWS in the maritime industry is the Automatic Radar Plotting Aids (ARPA), which is fixed in current radars and VTS [4]. The ARPA tracks nearby ships and indicates the collision candidates if the ship distance is less than a certain distance (e.g., 6 nautical miles) and provides DCPA and TCPA to users for further risk evaluations. Although ARPA is the pioneer of CWSs, several drawbacks remain. First and most importantly, CWSs based on ARPA ignore static states such as the ship type, tonnage and size, so they cannot provide collision avoidance decisions that are consistent with the requirements in the International Regulations for Preventing Collisions at Sea [18]. Second, there are no common agreements of warning distance value setting. Third, ARPA selects targets based on simple criteria and can lead to false alarms in high-density traffic waters.

Previous studies have proposed several methods to improve CWSs. In early studies, most works on CWSs established widely accepted collision criteria. For instance, Hilgert and Baldauf [12] proposed a set of heuristic criteria and suggested that the criteria are useful to categorize collision risk. A few years later, new methods and technologies are used in CWSs, in which fuzzy-based approaches (e.g., fuzzy system and fuzzy logic approaches) are highlighted. For example, Kao et al. [13] proposed a fuzzy-based method to calculate the ship domain, which is more accurate than the experience-based domain (e.g., Fuji ship domain). Chin and Debnath [7] developed a CWS framework to regress ship collision risk in port waters. Goerlandt et al. [10] proposed a risk-informed ship collision alert system and applied the system in an open seawater case. A comparison between the new system and earlier proposed CWSs was provided in the study to show the improvement of their frameworks.

In the collision assessment field, there has recently been a focus on foundational collision algorithms and theories. For instance, a velocity obstacle approach was used in a study to calculate the ship collision frequency in ship encounter situations (Huang et al., 2019). Bayesian learning approaches have been used to analyse the collision risk of ship traffic [22,26]. Although many approaches and models have been proposed, accurate data are crucial in the assessment.

### 2.3. Collision studies in the Yangtze River

The Yangtze River is the most important inland transport system in China. The accident records in the past 10 years show a decreasing tendency, but the number of accidents remains at a high level [29]. This is because of the high density of vessel traffic and complex water environments. The field of the water transportation system in the Yangtze River contains various study topics, and water transportation safety is the key topic. In terms of the ship collision domain, previous studies applied different collision approaches to analyse ship safety in the Yangtze River. For instance, Zhang et al. [32] used an analytic hierarchy process to establish a navigational risk model and identify the risk factors in the Yangtze River. Meanwhile, Zhang et al. [31] applied formal safety assessment (FSA) and Bayesian network to evaluate navigational risk in the Yangtze River. The study used historical accident data and followed the FSA risk framework to develop a Bayesian network. Compared to the field of navigational risk assessment, studies related to individual ship collision risks are scant. Chai et al. [6] considered the factors of ship domain, ship type, time and others and proposed a linear function to aggregate the collision risk. Wu et al. [20] studied the collision risk between vessels and bridges using a geometric collision model.

Although the above-mentioned studies provide useful insights to analyse the collision risk in the Yangtze River, gaps between research works and reality remain. The current topics mainly focus on upstream and downstream ships, and collision studies related to ferries are scant. The sinking accident of the Dongfangzhixing ship shows the
catastrophic consequence of a passenger ship accident. Collision assessments for ferries are urgently needed, but the limitation of AIS data leads to a lack of reliability in real-time collision assessment, which requires further improvements.

3. Data fusion technology

This section aims to propose a technology to establish a hybrid database to overcome the difficulties of using single databases in the collision warning model. In the technology, real-time AIS data and radar monitoring data are the two main sources. In the fused data, AIS data provide the ship static data, while the dynamic data from AIS are supplemented and modified by radar data. As shown in Fig. 1, there are mainly two real-time data sources, AIS and radar, which are collected and processed, involving a unifying coordinate system, spatiotemporal registration and data fusion. With a Kalman filter-based fusion framework, the AIS data are used to provide ship static data, while the dynamic data from AIS are supplemented and modified by real-time radar monitoring. The details are introduced in the following.

3.1. AIS & radar data

The variables used in the fusion framework involve position, speed and heading, in which the AIS and radar data can be collected from a shore-based observation station or onboard ships. The data receivers (i.e., AIS and radar) need to be calibrated to obtain the GPS coordinate and direction angle.

The AIS receivers automatically receive the orientation message from surrounding ships. In the AIS system, the orientation is denoted as longitude and latitude coordinates, and navigation aid facilities such as gyrocompasses and odometers are linked to the AIS system to provide ship position information. Heading and speed information is collected from GPS and logs and broadcast through AIS messages.

The orientation data of the radar are derived from its echo. The echo signal is converted to orientation data through signal processing algorithms, e.g., analogue-to-digital conversions and feature extractions. The orientation radar data provide ship positions with relative bearings and distances, which are not longitude and latitude. Therefore, it is necessary to unify the orientation outputs from AIS and radar into one coordinate system. In our work, we converted the orientation outputs of the radar to the GPS coordinate. Suppose that the radar in the longitude and latitude system with coordinates \((x_0, y_0)\) (unit: degree) detects an object and outputs relative orientation information \((d, \theta)\) (unit: km, degree); the longitude and latitude \((x, y)\) of the object can be calculated by

\[
\begin{align*}
    x &= x_0 + \frac{d}{c} \cos \theta \\
    y &= y_0 + \frac{d}{c} \sin \theta
\end{align*}
\]

where \(c\) is the conversion parameter to convert the latitude and longitude difference to distance. As one latitude degree is equal to approximately 111 km, in this paper, we assign \(c = 111\).

3.2. Spatio-temporal registration

The real-time collected AIS or radar data are expected to enhance the tracking performance by fusing the two different measurements. Before the fusion, it is required to register the data from either the AIS or the radar with tracking objects.

In general, it takes 2–4 s for the radar to scan 360°, which means that the radar data update rate is constant between 2 s and 4 s. However, the update period of AIS is not fixed and is from 2 s to 30 min,

![Fig. 1. AIS & Radar data fusion framework based on the Kalman filter.](image)
which depends on the motion state of the ship. It is obvious that the update periods of the radar and AIS are asynchronous. In this work, a temporal window filter is designed to achieve time matching between AIS and radar data. Assuming there is one radar measurement $P_1$ between two AIS measurements ($P_{11}$ and $P_{12}$), the AIS measurement at time $t$ can be calculated as follows:

$$P_t = P_{11} + \frac{(P_{12} - P_{11})}{t_2 - t_1}(t-t_1)$$

where $P_{11}$ is the AIS measurement at time $t_1$ and $P_{12}$ is the AIS measurement at time $t_2$.

On the other hand, spatial registrations are required to associate the measurements from different data. Therefore, we employ a spatial window filter with a size of 100 m and define that when the distance difference between the AIS data and the radar data at the same time is less than the window size, the two data records can be registered as one object; otherwise, the two data records are treated as two different objects. When there are several ships that are close to each other, it is difficult to associate the two different data. In such a situation, the radar system captures the objects as one 2D-point cluster and cannot distinguish how many ships there are. However, this situation is rare except for at anchorage and shoreline. We utilize a spatial filter using 100 or other metres as the window threshold, which is useful for data association in general situations.

3.3. Kalman filter-based fusion

A Kalman filter is employed as a fusion framework to fuse the radar data with the AIS data. The Kalman filter involves two parts: time prediction and measurement update. The state of one object (ship) is defined by $x = [x, y, v_x, v_y]$, which is composed of position $p = (x, y)$ and speed $v = [v_x, v_y]$, and the measurements $p'$ or $p''$ are derived from the AIS or radar. According to the Kalman filter, such a fusion can be achieved by

$$prediction: \quad \bar{x}_t = F(\bar{x}_{t-1}) + n_{4 \times 1}$$

$$update: \quad \bar{y}_t = H(\bar{x}_t) + \omega_{2 \times 1}$$

where $F(\cdot)$ is a prediction model of vessel motion, and $H(\cdot)$ is an update model of AIS or radar measurements. In this work, a constant speed is considered reasonable for any one ship in a short enough time. Therefore, the prediction matrix $F = [1, 0, \Delta t; 0, 0, 1, 0, \Delta t; 0, 0, 1, 0; 0, 0, 0, 1]$ and the measurement matrix

$$H = [1, 0, 0, 0; 0, 1, 0, 0; 0, 0, 1, 0; 0, 0, 0, 1].$$

The prediction error comes from kinematics system noise, defined by $Q = [0.25e-10; 0.25e-10; 1e-10; 1e-10]$. The measurement errors are composed of AIS and Radar noise, defined by $R_{ais} = [0.25e-10; 0.25e-10]$ and $R_{radar} = [0.25e-6; 0.25e-6]$, respectively. $p_1$ is the location expressed by longitude and latitude and is used to calculate the initial velocity $v_1$ directly in GPS coordinates.

Algorithm 1: The AIS & Radar data fusion based on the Kalman filter

| Initial | If two consecutive AIS observations of a vessel are available:
|---|---|
| | $v_{1-1} = (p'_1 - p''_1)/(t - t')$;
| | $x_{1-1} = [p'_1; v'_1]$;
| | $P_{1-1} = 0$;

| Tracking: for $t_1 = t_{1-1} + \Delta t$ |
|---|---|
| | $x_k = F(x_{k-1})$;
| | $p'_k = F(p'_{k-1})$;
| | if $p''$ is available or $p'_1$ and $p''_1$ are available:
| | $K_k = P_k \cdot H^T \cdot (H \cdot P_k \cdot H^T + R_{ais})^{-1}$;
| | $z_k = p''$;
| | $x_k = x_k + K_k \cdot (z_k - H \cdot x_k)$;
| | $P_k = (I - K_k \cdot H) \cdot P_k$;
| | if $p''$ is available:
| | $K_k = P_k \cdot H^T \cdot (H \cdot P_k \cdot H^T + R_{radar})^{-1}$;
| | $z_k = p''$;
| | $x_k = x_k + K_k \cdot (z_k - H \cdot x_k)$;
| | $P_k = (I - K_k \cdot H) \cdot P_k$;
| | if $p''$ or $p'$ is not available:
| | $x_k = x_k$;
| | $P_k = P_k$;

End | Current time - Last measurement time > T

The entire process of AIS & Radar fusion is shown in Fig. 1. At first, a new tracking process is initialized for a new object under the criteria that at least two consecutive AIS measurements are captured. In the Kalman filter, the object state is predicted every second, and the prediction model assumes that the motion of the object is uniform. When the AIS or radar measurement is available, the prediction states of the object are updated by the measurements. The tracking process is finished if there are no new measurements in the next 3 min.

4. Collision risk modelling

As introduced in Section II, geometric parameters are widely used in ship collision modelling. This study applies the DCPA and TCPA to calculate real-time collision risk. The DCPA and TCPA calculation process is introduced, and several risk functions are proposed to aggregate the collision risk value.

4.1. Collision variables

The calculation of the DCPA and TCPA between two ships is introduced in this section. Fig. 2 shows
an encounter situation that two ships are moving with speeds \( V_1 \) and \( V_2 \) and courses \( \theta_1 \) and \( \theta_2 \). The vector \( OA \) stands for the speed \( V_1 \) and course \( \theta_1 \) for the own ship, and the vector \( BC \) stands for the speed \( V_2 \) and course \( \theta_2 \) for the target ship. The vector \( BD \) is the relative speed and course between the own ship and the target ship. The vector \( OE \) indicates the DCPA between the own ship and the target ship, while the TCPA is the time that the target ship moves from position \( B \) to position \( E \).

The DCPA and TCPA can be indicated using the following functions:

\[
DCPA = D \times \sin(\mu - \theta - \pi) \\
TCPA = D \times \sin(\mu - \theta - \pi) / V
\]

Where \( D \) is the relative distance and \( \mu \) is the relative course between the two encountering ships, \( \theta \) denotes the true bearing and \( V' \) is the relative speed for the target ship. A more detailed function for calculating these two parameters can be found in Bukhari et al., [4].

### 4.2. Collision risk value

The DCPA and TCPA for surrounding ships are selected as the two main variables to calculate the collision risk values. Besides these two, we also consider the differences in ship type and ship tonnage. The following functions are used to calculate the collision risk value.

\[
R = \begin{cases} 
0, & \text{if } TCPA > TV_{TCPA} \text{ and } DCPA > TV_{DCPA} \\
\alpha \times \beta \times \left(1 - \frac{DCPA}{TV_{DCPA}}\right) \times \left(1 - \frac{TCPA}{TV_{TCPA}}\right), & \text{else}
\end{cases}
\]

where \( R \) is the collision risk value, \( TV_{TCPA} \) means the threshold value of the TCPA, \( TV_{DCPA} \) means the threshold value of the DCPA, \( \alpha \) is the criterion validity of ship types and \( \beta \) is the criterion validity of ship tonnage.

Then, we calculate the overall risk\(^1\) of the own ship. The overall risk for the own ship is the sum of all target ships’ collision risks. It can be calculated using the following equation:

\[
Overall\ Risk = \sum_{i=1}^{n} R_{TSi}
\]

where \( R_{TSi} \) is the collision risk value of each target ship and \( i \) is the number of target ships.

To define the collision candidates, threshold values for the DCPA and TCPA are needed. The selections of threshold values are based on the surrounding environment. In general, crew members use a DCPA of 2 nautical miles and a TCPA of 20 min in open sea waters. In the Yangtze River, the water areas are restricted, and the density of ships is higher. Thus, both the previous study results Ozturk et al., [19] and expert judgements are comprehensively referenced. As per these results, a DCPA less than 0.5 nautical miles and a TCPA less than 10 min are assigned. Target ships with a DCPA and TCPA less than the threshold values are selected as the collision candidates (i.e., \( TV_{DCPA} = 0.5 \) and \( TV_{TCPA} = 10 \)). Meanwhile, the criteria for \( \alpha \) and \( \beta \) are defined by experts.

Table 2 shows that chemical cargo ships have the largest weight, which is 2 (see Table 1). The oil tankers are less important and carry a weight of 1.5. Tugs and service ships are small and have good manoeuvring; thus, the risk values for those ships are halved. The tonnage differences of ships are ranked from high to low as follows. The weight for ships larger than 20,000 tons is 2, for ships between 3000 tons and 20,000 tons is 1.5, for ships between 500 tons and 3000 tons is 1 and for ships with tonnage less than 500 tons is 0.5.

### 5. Case study

This case study is carried out within the scope of a ship dynamic risk assessment project. The study was undertaken in the Nantong section of the Yangtze River, supported by the local maritime administration and a ferry company. To develop the case study database, the real-time AIS and radar monitoring data for a ferry in November 2019 were

---

\(^1\) Overall risk: In this study, the overall risk is defined as the integrated risk that aggregates the risk caused from target ships under an encounter situation.
The case study was undertaken in November 2019. The own ship is a scheduled ferry sailing between two ports located upstream and downstream on the Yangtze River. The distance of a signal voyage is approximately 12 nautical miles. During the voyage, four ships are encountered, including two general cargo ships, a chemical ship and a tug. The general cargo ships and the chemical ship are up/downstream ships that follow the direction of the Yangtze River main route, and the tug crosses the river from the Northern bank to the Southern bank. The detailed ship information of the own ship and the target ships is collected and shown in Table 3.

5.2. Data details

This study establishes a fused database that records ships’ static and dynamic data during a single ship voyage. The time interval for the voyage is approximately 60 min. The own ship is fitted with Class B AIS equipment. The reporting interval for Class B shipborne mobile equipment is 30 s. The update time for the radar data is 6 s. After fusing the AIS data and radar data, a total of 3000 records are obtained, comprising 600 records for each ship. The ship trajectories are shown in Fig. 3. The red line is the trajectory of ownship, the yellow line is the trajectory of target ship 1 (i.e., general cargo ship), the green line is the trajectory of target ship 2 (i.e., chemical ship), the blue line is the trajectory of target ship 3 (i.e., general cargo ship), and the black line is the trajectory of target ship 4 (i.e., tug).

5.3. Application

The proposed data fusion technology is implemented to calculate the collision parameters of each

---

**Table 1. List of abbreviations.**

<table>
<thead>
<tr>
<th>Term</th>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance at the Closest Point of Approach</td>
<td>DCPA</td>
<td>The closest distance at which two encountering ships pass each other</td>
</tr>
<tr>
<td>Time to the Closest Point of Approach</td>
<td>TCPA</td>
<td>The time from the current position to the closest point</td>
</tr>
<tr>
<td>Automatic Identification System</td>
<td>AIS</td>
<td>A navigational aid system which is applied to maritime safety and communication between ships and between ships and the shore</td>
</tr>
<tr>
<td>Vessel Traffic Service</td>
<td>VTS</td>
<td>A system that monitors ships sailing in and out of harbours using AIS, Radar, CCTV, wireless telephone, and shipborne terminals and provides the safety information during navigation</td>
</tr>
<tr>
<td>Global Position System</td>
<td>GPS</td>
<td>A high precision radio navigation positioning system based on artificial earth satellites, which can provide accurate geographical location, vehicle speed and accurate time information anywhere in the world and in near-earth space</td>
</tr>
<tr>
<td>Collision Warning System</td>
<td>CWS</td>
<td>A system that enhances situational awareness of VTS officers and crewmembers on board, and aids collision-avoidance decision making</td>
</tr>
<tr>
<td>Automatic Radar Plotting Aid</td>
<td>ARPA</td>
<td>A radar system that can automatically track, calculate and display the echo of selected objects and predict the result of avoidance</td>
</tr>
<tr>
<td>Formal Safety Assessment</td>
<td>FSA</td>
<td>An integrated safety assessment and standardized method, which aims to improve maritime safety through risk analysis and cost-benefit assessment</td>
</tr>
</tbody>
</table>

**Table 2. Criteria for ship style and tonnage.**

<table>
<thead>
<tr>
<th>Ship style</th>
<th>$\alpha$</th>
<th>Tonnage</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical cargo ship</td>
<td>2</td>
<td>Larger than 20,000 tons</td>
<td>2</td>
</tr>
<tr>
<td>Oil tanker</td>
<td>1.5</td>
<td>Between 3000 tons and 20,000 tons</td>
<td>1.5</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td>Between 500 tons and 3000 tons</td>
<td>1</td>
</tr>
<tr>
<td>Tugs and service ships</td>
<td>0.5</td>
<td>Less than 500 tons</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**Table 3. Criteria for ship style and tonnage.**

<table>
<thead>
<tr>
<th>Static information</th>
<th>Ship details</th>
<th>Target ship 1</th>
<th>Target ship 2</th>
<th>Target ship 3</th>
<th>Target ship 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ship type</td>
<td>Ferry</td>
<td>Cargo ship</td>
<td>Oil tanker</td>
<td>Cargo ship</td>
<td>Tug</td>
</tr>
<tr>
<td>Tonnage (tons)</td>
<td>446</td>
<td>24950</td>
<td>1</td>
<td>2450</td>
<td>348</td>
</tr>
<tr>
<td>Length (metres)</td>
<td>53</td>
<td>200</td>
<td>110</td>
<td>42.8</td>
<td>37.2</td>
</tr>
<tr>
<td>Beam (metres)</td>
<td>12</td>
<td>32</td>
<td>17</td>
<td>9</td>
<td>9.8</td>
</tr>
<tr>
<td>Draught</td>
<td>1.6</td>
<td>7.1</td>
<td>5.7</td>
<td>6.6</td>
<td>4.4</td>
</tr>
</tbody>
</table>
target ship. The distance between the own ship and the target ship, changes in the relative bearing, DCPA, TCPA, and speed are calculated.

Target ship 1: The statistical analysis results (Fig. 4) show the collision parameters of target ship 1. The closest distance between the own ship and target ship 1 is 13 min with a distance of 0.106 nautical miles. To avoid the collision risk, target ship 1 decelerates from 11 knots to 9.6 knots and then increases the speed back to 11.2 knots after the encounter situation is finished. The state of the TCPA being less than the threshold value (i.e., 10 min) remains between 8 min 30 s and 16 min 12 s. Between the start time and 4 min 6 s and between 8 min and 18 s and 14 min and 18 s, the DCPA of target ship 1 is less than 0.5 nautical miles.

Target ship 2: Fig. 5 shows the collision parameters of target ship 2. The closest distance between the own ship and target ship 2 is at 29 min and 12 s with a distance of 0.099 nautical miles. At that time, the speed of target ship 2 reaches the lowest value of 6.97 knots. The state of the TCPA being less than the threshold value (i.e., 10 min) remains between 8 min 24 s and 23 min 16 s and between 53 min and 55 min 24 s. Meanwhile, during the time interval between the start time and 3 min and between 5 min 48 s and 11 min, the DCPA of target ship 1 is less than 0.5 nautical miles.

Target ship 3: Fig. 6 reports the collision parameters of target ship 3. The closest distance between the own ship and target ship 3 is at 10 min and 24 s with a distance of 0.133 nautical miles. At that time, the speed of target ship 3 remains at a low speed of 11.95 knots. The state of the TCPA being less than the threshold value (i.e., 10 min) remains between 8 min 24 s and 23 min 16 s and between 53 min and 55 min 24 s. Meanwhile, during the time interval between the start time and 3 min and between 5 min 48 s and 11 min, the DCPA of target ship 1 is less than 0.5 nautical miles.

Target ship 4: Fig. 7 presents the collision parameters of target ship 4. The closest distance between the own ship and target ship 4 is at 20 min and 42 s with a distance of 0.027 nautical miles. The speed of target ship 4 is relatively low compared with the other target ships (i.e., 8.03 knots). The state of the TCPA being less than the threshold value (i.e., 10 min) remains between 14 min 48 s and 30 min 06 s, between 42 min 30 s and 48 min 18 s, etc. Meanwhile, during the time interval between 5 min and 6 min and between 17 min 48 s and 22 min, the DCPA of target ship 1 is less than 0.5 nautical miles.

5.4. Collision risk

This section proves the whole warning process in the case study. To calculate the dynamic risk value of the own ship, the proposed function 7 is used in real-time risk assessment. The results are shown in Fig. 8.
Fig. 8 shows that own ship has four encounter situations during its voyage. The first collision candidate in the encounter situations is target ship 3; the collision situation starts at 8 min 24 s and ends at 10 min. The highest risk value of target ship 3 is 0.0733, which is at 9 min and 24 s. The second collision candidate is target ship 1, and the encounter situation continues from 10 min 18 s to 13 min 42 s, reaching a peak value of 0.4642 at 12 min. The third collision candidate is target ship 4, which shows a collision risk between 17 min 48 s and 21 min 54 s and between 23 min and 23 min 18 s. The highest risk value of target ship 4 is 0.2089, which is at the moment of 20 min 42 s. The last collision risks caused by target ship 2 are identified in...
two time ranges. One is between 22 min 54 s and 23 min 6 s. Another one is between 28 min and 32 min 24 s. The highest collision risk between the own ship and target ship 2 is 0.505 at 30 min. The highest collision risk moments between the own ship and each target ship are shown in Fig. 9.

Meanwhile, the overall risk values of the own ship during the voyage are calculated and shown in Fig. 10. The collision risk values are dynamically calculated (given in Fig. 10 (a)), in which the areas of high collision risk are highlighted with red in Fig. 10 (b).
6. Validation

A fine-tuned model needs to be validated to ensure its reliability. A typical way to validate a dynamic risk model is to evaluate how well it performs on the collected data, i.e., check if the risk evaluation result is consistent with the ship officer’s mental sense. This is also applicable to the proposed model when real-time AIS data is available. Therefore, this section validates the developed model by using a validity method, which has been used in previous studies (Floris et al., 2015; Yu et al., [22]).

6.1. Face validity

Regarding model validation in general, previous studies from different disciplines conclude that...
proper validation should ensure the model’s credibility and relevance, especially in consistency with human experience. Face validity considers the general outputs from the model and applies expert judgements to evaluate the validity of the model behaviour.

The obtained risk results are evaluated by the ship captain on board the ferry. He agrees with the statement that the developed risk model can be considered appropriately for real-time risk evaluation as it produces rational risk evaluation and the variation of the risk is consistent with his experience.
Meanwhile, it is clear from the study that the collected data are fused in an appropriate way to ensure that the input is accurate and reliable. Therefore, the risk model is expected to provide reliable evaluations.

6.2. Model comparison

The novel approach overcomes the difficulty of the nonsequence AIS data while involving relevant geometrical parameters for collision risk assessments. The advantages of using sequential and fused data can therefore be discussed by a comparison analysis. Therefore, to validate the model and determine the superiorities of the proposed approach in real practice, the proposed risk model is then used to calculate the ship risk by using the raw AIS data alone. The obtained results are shown in Fig. 12.
Fig. 12 reports the overall risk by using two data inputs (fused data and AIS data). Generally, the two risk curves show great consistency in the risk evaluations, which proves that the data fusion method is reliable. However, compared to the AIS data-based risk curve (red line), the fused data-based risk curve (black line) is smoother, while a defect of the AIS data is that the red line does not show the collision risks between 20 min and 30 min, which is the ownship passing TS4. As a real-time warning model, this defect could lead to a serious consequence of collision if the officer relies on the model. In addition, the ferry captain explained that this is the reason why radar data are more reliable than AIS data in real-time collision warning.

The purpose of the behaviour test is to confirm that the model qualitatively corresponds to other similar studies or experiences. However, more RIFs need to be taken into consideration since these RIFs also have high impacts on ship dynamic risk, as evidenced by the results from previous studies. For instance, encounter situations, navigation rules, ship-related factors, weather-related factors and navigation-related factors are not involved and need to be expanded in future studies.

7. Conclusion

To conclude, the use of fused data in real-time collision assessment shows attractiveness and advantages. In this study, state-of-the-art technology and collision models are discussed. The drawbacks and limitations of using signal data sources (e.g., AIS data) in real-time collision warnings for ferries in the Yangtze River are highlighted. To overcome these difficulties, this study proposes a fused data-
based real-time collision risk model. In the proposed model, data fusion technology is first discussed. Then, the geometric collision risk parameters are selected, and collision aggregation algorithms are applied to develop a collision risk model. Finally, the model is tested in a real case of a ferry crossing the Yangtze River.

Based on the case study results, the attractiveness and advantages of the proposed model are threefold: 1) Compared to the AIS data-based collision model, the fused data collected from multiple data sources provide more information, thus enhancing the risk factors used in collision models, showing greater improvement than current models used in the maritime industry (e.g., APAR). 2) The applications of fusing AIS data and radar monitoring data not only increase the update rate of the ship position information but also improve the accuracy of the ship dynamic data in real-time monitoring. Any ships that show collision risk can be identified and alerted in a timely and accurate manner. 3) The calculations of the proposed model are efficient both in time and cost, which meets the real-time collision warning requirement that the collision risk among ships close to the own ship should be calculated accurately and updated rapidly. Therefore, the proposed data fusion technology and collision warning model are sufficient in real-time collision warning, which is an important topic in ship dynamic monitoring and traffic management. The model can be used to aid in onboard ship collision detection and avoidance while also providing a useful tool to support local VTS.

Meanwhile, some limitations are found. The first and the most important is that we notice that there are some false echoes in the radar observations. In a real-time scenario, false echoes must be identified and eliminated to ensure the reliability of the data. Meanwhile, more risk factors can be considered in the risk model, such as human error and machine failure rates. Finally, as there are different types of encounter situations defined by COLREGs, the model should be further studied to evaluate the collision risk of a ship under different encounter situations.

Uncited reference

[27]

Declaration of competing interest

No potential conflict of interest was reported by the authors.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

References


