# Collision-Risk-Aware Ship Routing

Patrik Thomas Michalski
ptm@informatik.uni-kiel.de
Dept. of Computer Science
Kiel University
Kiel, Schleswig-Holstein, Germany

Niko Preuß
nam@informatik.uni-kiel.de
Dept. of Computer Science
Kiel University
Kiel, Schleswig-Holstein, Germany

Matthias Renz
mr@informatik.uni-kiel.de
Dept. of Computer Science
Kiel University
Kiel, Schleswig-Holstein, Germany

Andreas Tritsarolis
andrewt@unipi.gr
Dept. of Informatics
University of Piraeus
Piraeus, Greece

Yannis Theodoridis
ytheod@unipi.gr
Dept. of Informatics
University of Piraeus
Piraeus, Greece

Nikos Pelekis
npelekis@unipi.gr
Dept. of Statistics & Insurance Science
University of Piraeus
Piraeus, Greece

## ABSTRACT

This paper addresses short-term Collision-Risk-Aware ship route planning while utilizing a deep learning-based Vessel Collision Risk Assessment and Forecasting (VCRA/F) framework to quantify risks. Lacking a clear boundary between risky and viable routes, we propose a Pareto-optimal search for alternative routes, balancing collision risk and voyage time. Our main contribution is a novel framework that integrates VCRA/F for Pareto-optimal route queries in dynamic environments. We model maritime routes using a hexagon-based graph network on the sea. Our experiments on real-world AIS data validate the effectiveness of Skyline-VCRA/F while highlighting areas for further improvement.

## CCS CONCEPTS

• **Information systems → Data mining**; **Spatial-temporal systems**; • **Computing methodologies** → *Machine learning*.

## KEYWORDS

Maritime Safety, Machine Learning, Pareto Principal Optimization, Ship Routing

**ACM Reference Format:**
Patrik Thomas Michalski, Niko Preuß, Matthias Renz, Andreas Tritsarolis, Yannis Theodoridis, and Nikos Pelekis. 2024. Collision-Risk-Aware Ship Routing. In *The 32nd ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '24), October 29-November 1, 2024, Atlanta, GA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3678717.3691316

## 1 INTRODUCTION

Modern ship sensor technologies like the Automatic Identification System (AIS) offer extensive vessel positioning data for various maritime analytics and for the development of new techniques [1]. Careful planning and precise execution in ship routing are very crucial to improving marine safety, including route forecasting and collision risk assessment [3, 5, 15].

Collision-risk-aware ship routing requires particular understanding, proactive assessment, and strategic management of hazards. It involves adhering to international regulations, integrating best practices, and applying new technologies to minimize the risks of collisions and their impacts.

A contentious aspect of Vessel Collision Risk Assessment (VCRA) and Forecasting (VCRF) is determining the short-term collision risk, such as in maritime route planning. VCRA and VCRF aim to quantify the risk of collision of vessels, both currently and in the future, using the collision risk index (CRI), which indicates the likelihood of a collision. In our work, we use a deep learning-based approach [16] to assess these risks, helping to avoid unsafe routes and select the safest, shortest paths. Since there is no clear distinction between risky and safe routes, we propose an alternative approach called route-skylines. This framework optimizes a Pareto-optimal set of routes, balancing collision risk and voyage time, thus allowing for more flexible and effective route planning.

In summary, we propose a novel framework for efficient route planning on the sea that takes advantage of the latest Pareto-optimal and VCRA/F approaches in dynamic and uncertain search spaces.

## 2 RELATED WORK

Kriegel et al. [6] introduce an advanced routing algorithm that uses lower bounds to compute route skylines while utilizing a Lipschitz embedding for lower-bound approximation. They prune those routes that cannot be extended further as other routes to the same destination dominate them. Shekelyan et al. [10] introduce a two-phase method to determine all non-dominated paths while computing supported solutions and the final linear skyline.

Yan and Zissis et al. [17] and Zissis et al. [18] introduce a novel method for extracting maritime traffic routes using AIS data to transform ship trajectories into Ship Trip Semantic Objects (STSOs). STSOs are integrated into a directed graph through graph theory, enabling the extraction and representation of shipping routes.

The current state-of-the-art in VCRA/F includes formulaic and ML-based approaches, using kinematic equations combined with ML models, e.g., Li et al. [7]. In contrast, the latter uses ML methods directly to assess the risk of collision between two vessels.


*SIGSPATIAL '24, October 29-November 1, 2024, Atlanta, GA, USA*

ACM ISBN 979-8-4007-1107-7/24/10
https://doi.org/10.1145/3678717.3691316

For example, Ma et al. [9] formulate the problem as time series classification and introduce a DL-based solution that employs bidirectional LSTMs [13] and the Attention mechanism [2] to map the behavioral features of the encountering vessels to their corresponding CRI in the future. However, Sang et al. [12] propose a novel model to predict short-term vessel trajectories using historical AIS data while improving CPA-based calculations by incorporating speed and turn rate. This method effectively identifies and warns of anomalous vessel behavior in real time.

Recently, Tritsarolis et al. [16] propose an efficient and modular ML-based solution that performs VCRA/F using an extension of the model proposed in [15], and to further assist decision-making, perform VCRF using efficient route forecasting algorithms [11]. To the best of our knowledge, this work is considered state-of-the-art in (short-term) VCRA/F, achieving an accuracy of around 96% for VCRA and 77% for VCRF.

In the field of collision avoidance, García et al. [4] recommend a thorough incorporation of COLREG 72 into Decision Support and Collision Avoidance Systems. This should improve the reliability and communication of the vessel. However, Liu et al. [8] propose a Multi-Task Deep Learning framework to predict ship trajectories and collision risks, simplifying scenarios into dynamic two-ship encounters.

Our framework differs from the aforementioned ones by not considering vessel collaboration or regulations. Instead, it reports potential routes in advance and calculates the Collision Risk Index (CRI) globally.

## 3 PROBLEM FORMULATION

We aim to find all Pareto optimal routes considering collision risk and voyage time for a given ship's start position, destination, and time. Such a set of all non-dominated objects forms a Pareto-front, representing all possible optimal trade-offs between the objectives, also called a skyline. skyline queries are essential for several multi-attribute decision-making applications [14].

For the collision risk assessment and forecasting task, we use the VCRA/F model proposed in [16], which uses the azimuth angle and relative direction of the target vessel, as well as its corresponding speed, course and haversine distance. Given these features, VCRA/F aims to compute the CRI of encountering vessels, a number that expresses their likelihood of collision.

*Definition 3.1 (Multi-attribute Network Graph (MAG)).* A Multi-attribute Network Graph (MAG) is a directed graph containing a weight vector of dimension $d$, where $d \in \mathbb{N}$.

Each of the $d$-dimensional weight vectors contains $d$ attributes that determine the costs for each dimension within the graph.

*Definition 3.2 (Time-dependent MAG (TMAG)).* An extension to MAG is TMAG, a Time-dependent Multi-attribute Network Graph (TMAG), where each cost function is time-dependent.

*Definition 3.3 (Route-skyline query).* Let $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ be a TMAG, a source $u$ and destination $v$, where $u, v \in \mathbb{V}$ and start time $t \in \mathcal{T}$.

A Route-skyline query computes the skyline from the set of all possible paths starting in $u$ and ending in $v$ at the start time $t$.

The problem we are addressing in this work is to answer Route-skyline queries considering collision risk and voyage time.

For this purpose, we propose the so-called Skyline-VCRA/F framework, the details of which are presented in the section that follows.

## 4 METHODOLOGY

Figure 1 illustrates the architecture of Skyline-VCRA/F: The *Offline Preparation* part is responsible for preprocessing the incoming AIS data, training the VCRA/F model, and creating the underlying graph and embedding. Afterward, the *Online Route Search* uses the graph mentioned above, the embedding, and the predicted CRI of the ships for efficient filtering and target-oriented routing. In contrast, the *Result* processes and visualizes all Pareto-optimal routes.

To efficiently compute non-dominated routes between nodes in a given MAG, we use the approach of Kriegel et al. [6] for multi-attribute Route-skyline queries and combine it with the VCRA/F model by Tritsarolis et al. [16].

Kriegel et al.'s Advanced Route skyline Computation (ARSC) dynamically computes skylines in a road network graph, reducing search space and efficiently considering multiple preferences like travel cost and voyage time. Their algorithm prunes non-extendable routes dominated by others at intermediate nodes, using lower-bound approximations to exclude paths worse than any known skyline route.

Mapping ocean routing into a graph is not straightforward, unlike road networks, because, unlike road networks, there are no predefined streets or intersections. The H3 geospatial indexing system addresses this, creating a sphere's global hexagonal grid. Hexagons are ideal because they maintain equal distance to all neighbors and minimize error during projection. This hexagonal grid is applied to the sphere's surface using a specialized projection, enabling accurate ocean network analysis.

The process is based on the A-star algorithm. Thus, it begins with the source node being added to a priority queue, which is dynamically updated to prioritize the most promising paths. To minimize the number of exploration steps and avoid searching the entire graph, a lower-bound approximation is used to prune sub-paths early in the process that others will dominate.

One-hop path extensions must be generated in every iteration as the algorithm progresses. At this point, VCRA is invoked to assess the risk of each edge. Specifically, we identify all neighboring vessels at this time and calculate the pairwise CRIs between them and our vessel. The resulting extended paths either result in a final skyline representing an optimal route between the source and destination or another sub-path that is added back into the priority queue for further exploration. Each step ensures that dominated sub-paths are pruned to maintain efficiency. The process continues until the priority queue is empty. Taking into account collision risk and voyage time, all Pareto-optimal skylines are identified from the source to the destination node.

## 5 EXPERIMENTAL STUDY

This section evaluates our Skyline-VCRA/F framework using a historical real-world AIS dataset[1] for our experimental study. The Skyline-VCRA/F framework is implemented in Python. The source code and historical AIS data used in our experiments are available at https://github.com/stu207680/risk-prediction.

[1]The Piraeus AIS dataset; Source: https://doi.org/10.1016/j.dib.2021.107782.

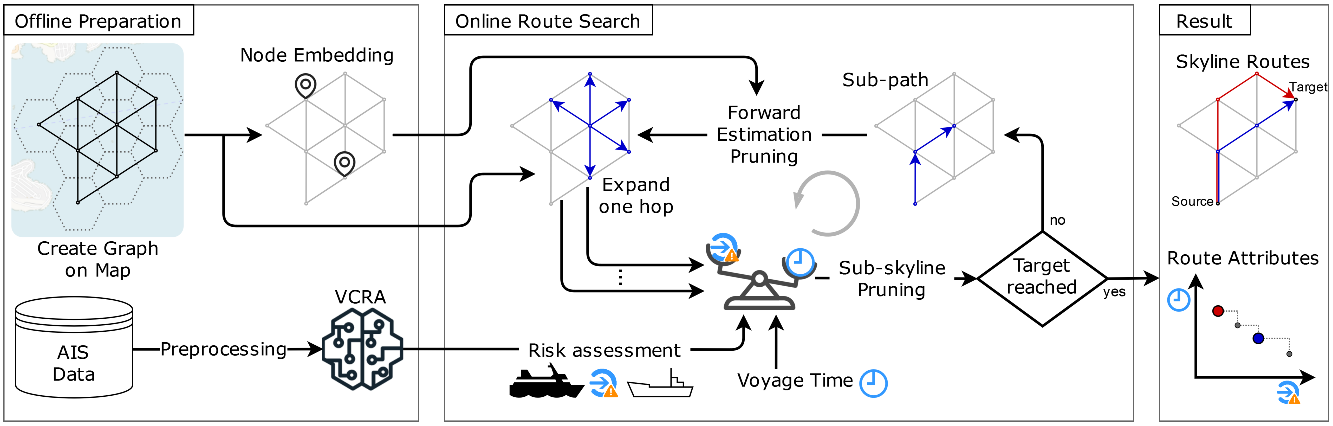


**Figure 1: Architecture overview of the proposed Skyline-VCRA/F framework.**

Unless otherwise specified, all experimental results include 100 data points per graph resolution. These data points were obtained by querying routes between 25 random locations in the Saronic Gulf and the entry to the port of Piraeus, each at four different times of the day: 9 a.m., 12 noon, 3 p.m., and 6 p.m.

To obtain a baseline for runtime comparison, we used a naive approach that runs 100 one-dimensional shortest-path queries, each with a different risk threshold to ensure complete coverage of all risks. Then, a simple Pareto filter is applied to obtain the results of Skyline-VCRA/F. Figure 2 shows the run-time comparison of Skyline-VCRA/F. It is clear to see that our approach outperforms this baseline in terms of runtime. We use a logarithmic scale to visualize the differences between the compared methods at different resolutions. The "1D" variant refers to a single one-dimensional shortest-path query, while "Pareto" represents the method to match Skyline-VCRA/F results. Different resolutions, which determine the size of the graph, are shown in Table 1.

| H3 resolution | Number of nodes | Number of edges |
|---|---|---|
| 8 | 283 | 1548 |
| 9 | 1973 | 11462 |
| 10 | 13794 | 81706 |
| 11 | 96532 | 576550 |

**Table 1: Saronic Gulf graph sizes for different H3 resolutions.**

Further, we studied the effect of allowing different speeds during exploration, i.e., we changed the speed over ground (SOG). These speeds are mapped within the graph as virtual multi-edges, i.e., $k$. More precisely, $k$ affects the granularity at which changes in SOG are possible by evenly distributing our SOG interval, i.e., eight to 14 knots, to $k$ virtual multi-edges. Figure 3 shows the effect of varying the number of speeds of travel allowed in the model. Due to the significant effect on runtime, not all H3 resolutions are feasible.

To understand this effect, let us assume that slowing down on a sub-path of the route reduces the risk of collision. In that case, another sub-path can be created at the following edge with each possible reduction in speed so that these paths do not dominate each other. However, as these sub-paths arrive at the next node at different times, the search is continued independently for each route, with a cascading effect that impacts the runtime.

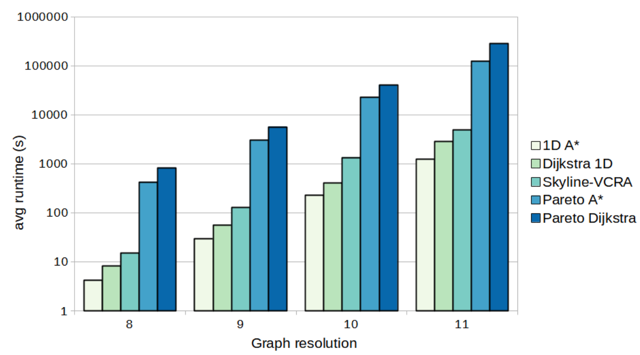


**Figure 2: Average runtime comparison for varying graph size.**

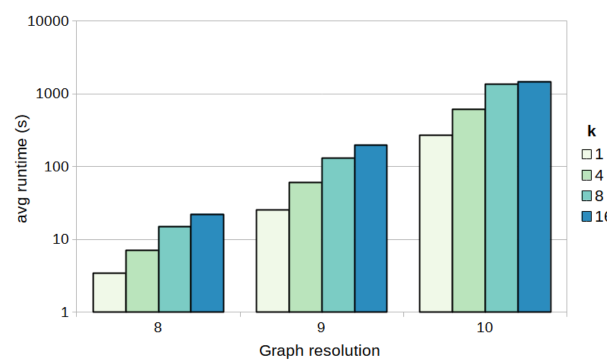


**Figure 3: Runtime comparison with varying *k*.**

Although the Saronic Gulf is relatively densely packed with AIS signals, close encounters are rare along a given route. Our experiments have shown that 68% of the queries returned a single result, often because the shortest path had no risk or because of the number of other vessels near the harbor entry, which naturally increased the risk of collision, all other routes are dominated by the shortest path due to the shorter voyage time.

Figure 4 shows 14 different routes to the port of Piraeus. The shortest, darkest route took 41 minutes with a 30% CRI, while the longest, brightest route took 58 minutes with only a 6% CRI, highlighting the trade-offs between speed and safety.

The mapping of real-world water bodies on a graph often results in angular stairs, as seen in Figure 4, due to the limitations of the grid.

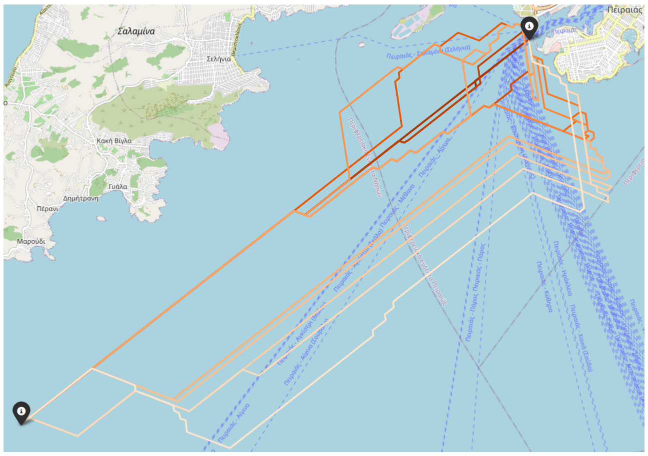


**Figure 4: Example skyline; A brighter color corresponds to lower risk.**

## 6 CONCLUSIONS AND FUTURE WORK

In this paper, we proposed Skyline-VCRA/F, a collision-risk-aware ship routing framework that takes advantage of Pareto-optimal routing and collision risk assessment. We showed through our experiments on real-world AIS data its efficiency, transparency, and applicability to scenarios with autonomous vessels, contributing to the advancement of intelligent Maritime Transportation Systems (MTS).

We plan to further optimize Skyline-VCRA/F by focusing on feasible solutions, faster pruning the search space, and addressing issues like spatial restriction by the grid-like graph, which can affect path optimization. Potential solutions include post-processing steps or exploring new methods to enhance flexibility. Moreover, we aim to extend Skyline-VCRA/F to a multi-agent approach, optimizing routes for our own and targeting vessels within an area of interest.

## ACKNOWLEDGMENTS

This paper was supported by Marispace-X and the Horizon Framework Program of the European Union under the grant agreement No. 101070279 (MobiSpaces, https://mobispaces.eu).

Further, this research was supported in part through the high performance computing resources available at the Kiel University Computing Center. The caucluster system was sponsored by the German Research Foundation (DFG) - project number 440395346.

## REFERENCES

[1] Alexander Artikis and Dimitris Zissis. 2021. *Guide to Maritime Informatics*. Springer Cham. https://doi.org/10.1007/978-3-030-61852-0

[2] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural Machine Translation by Jointly Learning to Align and Translate. In *Proceedings of the 3rd International Conference for Learning Representations (ICLR).* https://doi.org/10.48550/arXiv.1409.0473

[3] Eva Chondrodima, Petros Mandalis, Nikos Pelekis, and Yannis Theodoridis. 2022. Machine Learning Models for Vessel Route Forecasting: An Experimental Comparison. In *Proceedings of the 23rd IEEE International Conference on Mobile Data Management (MDM).* https://doi.org/10.1109/MDM55031.2022.00056

[4] Jesús A. García Maza and Reyes Poo Argüelles. 2022. COLREGs and their application in collision avoidance algorithms: A critical analysis. *Ocean Engineering* (2022). https://doi.org/10.1016/j.oceaneng.2022.112029

[5] Georgios Grigoropoulos, Giannis Spiliopoulos, Ilias Chamatidis, et al. 2024. A Scalable System for Maritime Route and Event Forecasting. In *Proceedings of the 27th International Conference on Extending Database Technology (EDBT).*

[6] Kriegel Hans-Peter, Renz Matthias, and Schubert Matthias. 2010. Route skyline queries: A multi-preference path planning approach. In *2010 IEEE 26th International Conference on Data Engineering (ICDE 2010).* https://doi.org/10.1109/ICDE.2010.5447845

[7] Yishan Li, Zhiqiang Guo, Jie Yang, Hui Fang, and Yongwu Hu. 2018. Prediction of ship collision risk based on CART. *IET Intelligent Transport Systems* (2018). https://doi.org/10.1049/iet-its.2018.5281

[8] Tao Liu, Xiang Xu, Zhengling Lei, Xiaocai Zhang, Mei Sha, and Fang Wang. 2023. A multi-task deep learning model integrating ship trajectory and collision risk prediction. *Ocean Engineering* (2023). https://doi.org/10.1016/j.oceaneng.2023.115870

[9] Jie Ma, Chengfeng Jia, Xin Yang, Xiaochun Cheng, Wenkai Li, and Chunwei Zhang. 2020. A Data-Driven Approach for Collision Risk Early Warning in Vessel Encounter Situations Using Attention-BiLSTM. *IEEE Access* (2020). https://doi.org/10.1109/ACCESS.2020.3031722

[10] Michael, Shekelyan, Gregor, Jossé, Matthias, Schubert, and Hans-Peter, Kriegel. 2014. Linear Path Skyline Computation in Bicriteria Networks. In *Database Systems for Advanced Applications.* https://doi.org/10.1007/978-3-319-05810-8_12

[11] Brian Murray and Lokukaluge Prasad Perera. 2021. An AIS-based deep learning framework for regional ship behavior prediction. *Reliability Engineering & System Safety* (2021). https://doi.org/10.1016/j.ress.2021.107819

[12] Ling-zhi Sang, X. Yan, Alan Wall, Jin Wang, and Zhe Mao. 2016. CPA Calculation Method based on AIS Position Prediction. *Journal of Navigation* (2016). https://doi.org/10.1017/S0373463316000229

[13] Mike Schuster and Kuldip K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* (1997). https://doi.org/10.1109/78.650093

[14] Shekelyan, Michael, Jossé, Gregor, and Schubert, Matthias. 2015. Linear path skylines in multicriteria networks. In *2015 IEEE 31st International Conference on Data Engineering.* https://doi.org/10.1109/ICDE.2015.7113306

[15] Andreas Tritsarolis, Eva Chondrodima, Nikos Pelekis, and Yannis Theodoridis. 2022. Vessel Collision Risk Assessment using AIS Data: A Machine Learning Approach. In *Proceedings of the 23rd International Conference on Mobile Data Management (MDM).* https://doi.org/10.1109/MDM55031.2022.00093

[16] Andreas Tritsarolis, Brian Murray, Nikos Pelekis, and Yannis Theodoridis. 2023. Collision Risk Assessment and Forecasting on Maritime Data. In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems.* https://doi.org/10.1145/3589132.3625573

[17] Zhaojin Yan, Yijia Xiao, Liang Cheng, Rong He, Xiaoguang Ruan, Xiao Zhou, Manchun Li, and Ran Bin. 2020. Exploring AIS data for intelligent maritime routes extraction. *Applied Ocean Research* (2020). https://doi.org/10.1016/j.apor.2020.102271

[18] Dimitris Zissis, Konstantinos Chatzikokolakis, Giannis Spiliopoulos, and Marios Vodas. 2020. A Distributed Spatial Method for Modeling Maritime Routes. *IEEE Access* (2020). https://doi.org/10.1109/ACCESS.2020.2979612

Received June, 07th 2024; accepted August, 23th 2024