Learning with Navigation Feature: Quantitative Risk Analysis for the Navigation of Autonomous Ships

Wei Chian Tan[1], Kie Hian Chua[2], and Yanling Wu[3]

[1]Nanyang Technological University
[2]Technology Centre for Offshore and Marine, Singapore
[3]University of Science and Technology of China
[3]Corresponding author: yanlwu@ustc.edu.cn

Abstract

This work presents a data-driven approach for the automated risk estimation of the voyage of a vessel or ship. While the industry is moving from a compliance-based framework with existing rules to a risk-based one, there is also a need to monitor the risk of a vessel from the perspective of the navigation. This is of even higher importance for the case of autonomous ships. Built based on the state-of-the-art mathematical representation, the navigation feature, each existing voyage is transformed into a corresponding series of points in d-dimensional space. During the stage of pre-processing, given a set of historical Automatic Identification System (AIS) data, those records that belong to the same vessel within a certain period of time are taken as a voyage and mapped to the corresponding space of the navigation feature. After the pre-processing and during the online monitoring, the current trajectory of the vessel is transformed into the corresponding representation in the same way. Based on a nearest-neighbour search scheme, the distance from the nearest neighbour is taken as the risk of the current voyage. In other words, the deviation from the closest route in the historical data is taken as the risk. The developed method has demonstrated encouraging performance on a set of challenging historical AIS data from the Australian Maritime Safety Authority, covering three regions in the Australian territory, namely the Bass Strait, the Great Australian Bight and the North West.
Keywords: Autonomous ships, navigation planning, historical AIS data, machine learning.

1 Introduction

Autonomous ships have attracted significant amount of attention from the marine and offshore industry lately. Compared to conventional manned vessels, the huge potential of these autonomous vessels towards a new level of operational efficiency makes them a promising candidate as a technological solution for the next step of the industry.

As each ship and offshore platform is a massive engineering system by itself, the need of ensuring the seaworthiness of the vessel or platform is of ultimate importance. Usually performed by an independent third party such as a classification society [10], the process of engineering a ship is required to be certified and classed. However, a trend or movement from a compliance-based framework to a risk-based one has been observed in the domain. As the complexity of the massive engineering systems nowadays is going beyond the capability of a binary (pass or fail) evaluation, there needs to be a way to ensure the reliability of these systems, beyond a compliance according to established rules. A risk-based approach can come in as an effective alternative, as it does not only take the part of compliance into a consideration, but also looking at the risk of the system, sub-systems and components quantitatively. This would give the engineers and the relevant authorities a sense on the level of confidence in terms of performance and reliability of the system.

Navigation at sea typically requires compliance with applicable rules and regulations including but not limited to Convention on the International Regulations for Preventing Collision at Sea, 1972 (COLREGs) [11] from the International Maritime Organisation [12]. These are essentially the rules of traffic at sea, globally and regionally. As an inappropriate movement of a vessel can lead to risky situations such as two vessels being too close and a potential collision, ensuring a real-time compliance by all the vessels is of significant importance for the safety and the smoothness of operations of all.

Hence, while the industry is moving from a compliance-based system towards a risk-based one, together with the recent focus on autonomous vessels, there is a need for an approach to do automated risk estimation and monitoring of the voyage of each vessel on a (near-) real-time basis. This work aims to address the issue through a historical data-driven approach. Based on a set of historical AIS (Automatic Identification System) data, the correspondence among the AIS points which belong to the same vessel is established to turn the collection of points into a voyage. After a list of historical voyages is obtained, given the current location of the vessel, the distance from the nearest point of a historical voyage in the space of navigation feature is returned as the risk estimation of the current voyage at the moment. The developed method has demonstrated encouraging performance on an existing set of historical AIS data from the Australian Maritime Safety Authority [3], covering three regions in Australian territory, namely the Bass Strait, the Great Australian Bight and the North West [4] [1]. Figure 1 shows the three regions studied, and figure 2 to 4 show the visualisation of the historical AIS data for each of the regions.

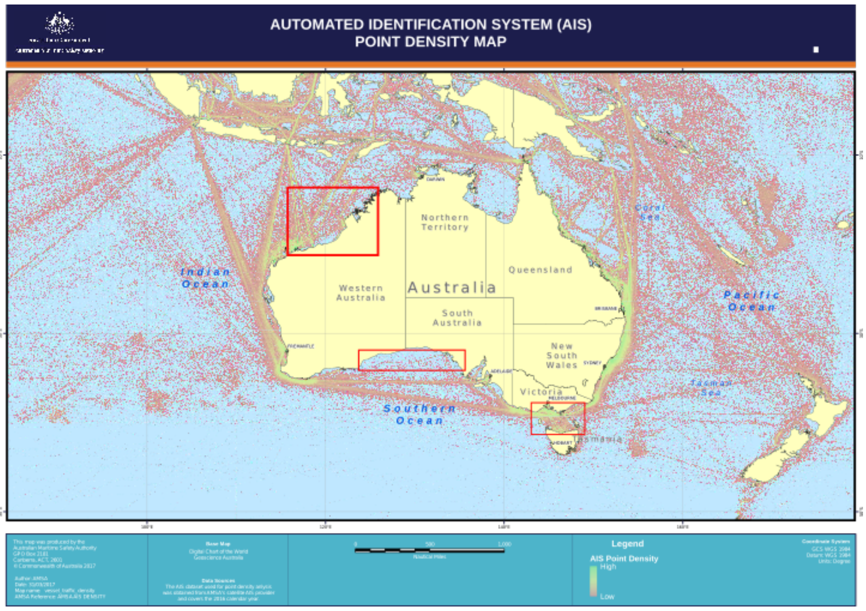

Figure 1: Maritime regions studied in this work [4].

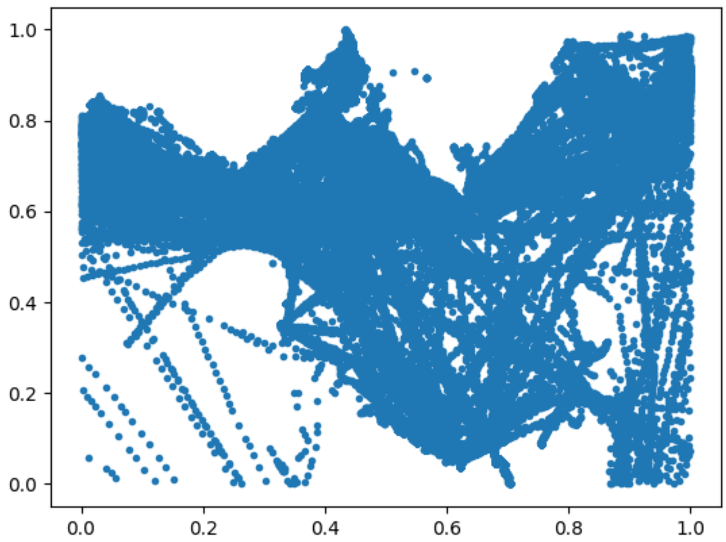

Figure 2: A visualisation of the historical AIS data for the Bass Strait [4].

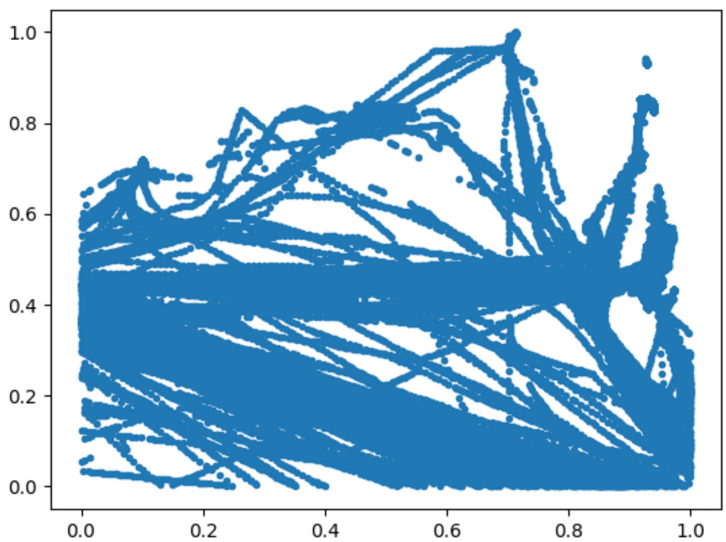

Figure 3: A visualisation of the historical AIS data for the Great Australian Bight [4].

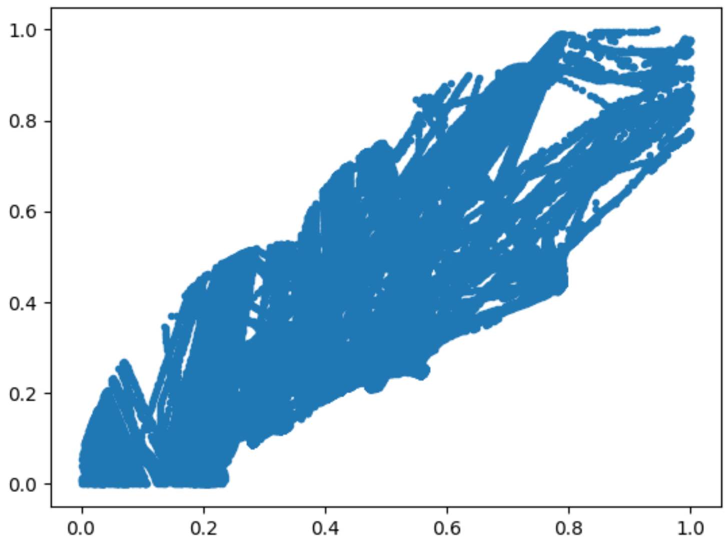

Figure 4: A visualisation of the historical AIS data for the North West [4].

2 Related work

There is a significant volume of existing work on the domain of autonomous ships, covering things like situations at high seas [9], path generation [22] to collision detection and avoidance [13] [14]. In addition, the aspect of connectivity of autonomous ships has been looked into [6]. An interesting idea to augment the sensing capability of an autonomous ship via the support from an unmanned aerial system has also been explored [8]. Along the topic of landing a quadrotor on the deck of an autonomous ship, an invariant ellipsoid method has been developed [16].

Closely relevant to the aim of this work, risk based methods have been seen in maritime for different purposes such as autonomous systems [21] and unmanned merchant ships [15]. Similarly, an approach of learning using corrosion feature with non-linear (Support Vector Machines) SVM [5] has been introduced to determine the potential corrosion mechanisms which may happen based on a set of design and operating conditions in an automated manner [19].

In a recent approach to augment the existing navigation planning with the use of historical AIS data, a weighted nearest-neighbour search in the space of ship and navigation feature has been introduced for the retrieval of a suitable route for an upcoming voyage from a database of historical routes [20].

As the industry typically deals with massive engineering systems such as ships or offshore platforms, the importance of compliance according to established applicable rules has been put at the first place. A Histogram of Connectivity and linear SVM based approach has been introduced for the evaluation of the piping design of a ship [17]. On the same topic, by fine-tuning pre-trained deep convolutional neural networks, the hypothesis that common visual features learnt can be reused for ship design has been validated to a significant extent [18].

Based on the discussion above, there appears to be a gap between the state-of-the-art and the aim of this work, which is to do an automated estimation of the risk of a voyage based on historical AIS data.

3 Methodology

Details of the main method developed in this work are presented in this section.

3.1 Representation: Navigation Feature

Before the idea of doing a risk-based estimation for a voyage can be achieved, there is a need to establish a representation or a feature space. This starts with linking those AIS points that belong to the same vessel as a voyage. Next, for each voyage identified in the database, a representation is established accordingly.

Similar to previous work, the idea of navigation feature [20] is adopted as the representation here. Assuming that there are d attributes that need to be taken into consideration, each of these attributes constitutes a dimension and this leads to a navigation feature for that set of attributes. In other words, a set of d navigation attributes is transformed into a point $\mathbf{x} \in R^d$. For each point in a voyage, a navigation feature is established and this process is repeated for all the points in the voyage. The process of forming navigation features for a voyage is illustrated in algorithm 1 and 2.

The current form of navigation feature has the advantage of being flexible in terms of attributes to be taken into consideration. Depending on the actual situation, the variables such as the speed or the operating status of the engine on-board can be included for a better analysis. This can be very helpful when certain signals become unreliable and these sources can just be excluded from the analysis meanwhile. Four variables are taken into consideration in this work, including the longitude, the latitude, the speed and the course, hence the analysis is done in four-dimensional space.

Algorithm 1 Formation of a Navigation Feature.

```
Given d attributes
Initialise x∈R^d
for 0≤i<d do
    x_i = i-th attribute
end for
return x
```

Algorithm 2 Pre-processing of a voyage.

```
Given a voyage with N_p points
for 0≤i<N_p do
    Form i-th navigation feature for the i-th point, f_i∈R^d
    Store f_i
end for
```

3.2 Risk-based estimation via nearest-neighbour search

The idea of doing risk estimation of the current voyage based on a list of historical voyages in the space of navigation feature is introduced in this section. Based on a database of historical voyages or AIS data, the deviation of the current voyage from the closest historical one is determined as the risk. Theoretically, the risk is determined by the distance between the query and the closest point in the space of navigation feature.

The process is mainly divided into two stages, including the pre-processing and the online estimation. The stage of pre-processing involves the process of establishing a correspondence among all the AIS points which come from the same vessel, registering them as a voyage and repeating the process for all the points in the database. These points are mapped to the space of navigation feature subsequently. This is summarised in Algorithm 3. During the online retrieval, given the current AIS location of the voyage, the information is transformed into the same representation as a navigation feature, which is used as the query to search for the nearest neighbour from the database of registered voyages subsequently. The distance to the nearest neighbour is returned as the risk estimation. This is summarised in Algorithm 4.

Algorithm 3 Pre-processing.

```
for each voyage in the historical data do
    Establish the series of navigation features and store
end for
```

Algorithm 4 Risk estimation with nearest-neighbour based search.

```
Given a navigation feature as the query
Search for the closest points in the historical voyages
Return the distance between the query and the nearest neighbour
```

4 Experimental study

The details of the experimental study carried out are presented in this section.

4.1 Setup

The implementation of algorithm 1 to 4 was done in Python. Each online retrieval took less than five seconds. The three existing sets of historical data were separated randomly into two equal sets, one for training (or establishing the database during pre-processing) and another for testing.

4.2 Results

Table 1 shows the performance of the developed algorithms. The results are reported in the form of the mean and the standard deviation, for the distance of the query point to the closest voyage.

In addition, as a visualisation towards the method for monitoring the real-time risk of a voyage, figure 5 to 7 show an example of the risk estimation for the case of the Bass Strait, the Great Australian Bight and the North West. The red point in each of the figures refers to the query point, green points refer to the closest voyage found and blue points are the AIS data in the training set.

Table 1: Performance of the risk estimation of a voyage.

Dataset	Distance ($\sigma \pm \nu$)
Bass Strait	0.0067 ± 0.012
Great Australian Bight	0.0095 ± 0.014
North West	0.0046 ± 0.0064

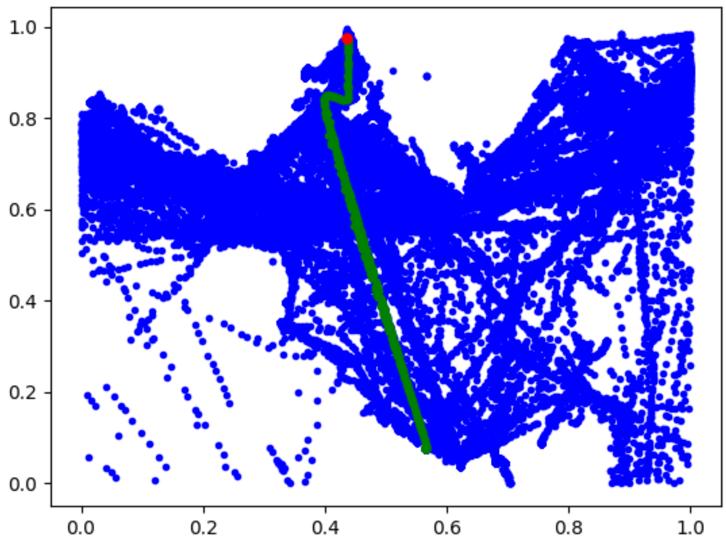

Figure 5: An example of the risk estimation for the Bass Strait.

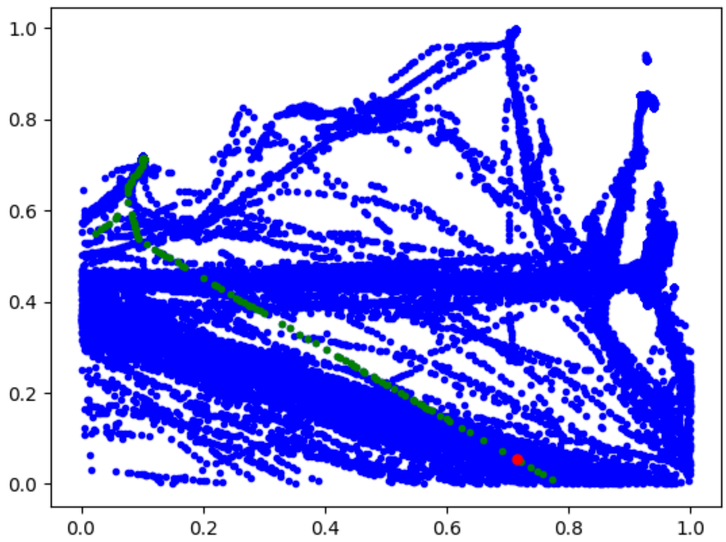

Figure 6: An example of the risk estimation for the Great Australian Bight.

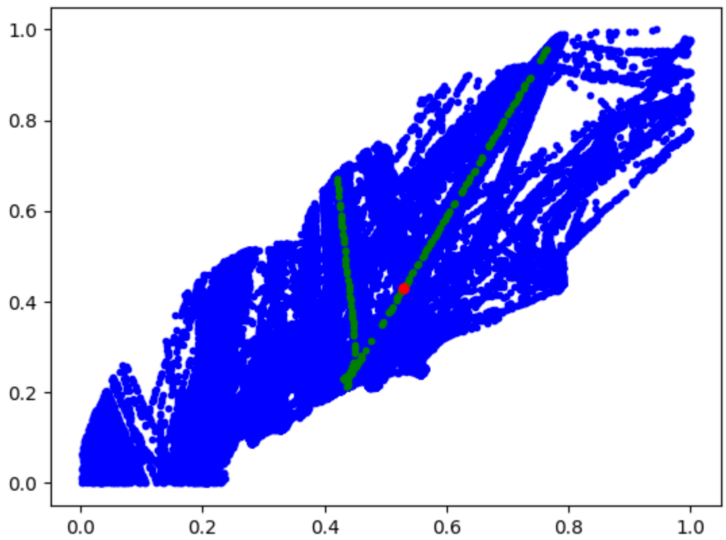

Figure 7: An example of the risk estimation for the North West.

5 Discussion

As demonstrated in the previous section, the developed method has demonstrated encouraging performance on the three existing sets of historical AIS data covering three Australian regions.

As one may notice, the distance of the query to the nearest neighbour found is returned as the estimated risk. There has not been a clear lower and upper bound determined for the navigational risk. A necessary step to take next is to identify a suitable lower and upper bound for the variable. This can be done by a mapping of the risk to a range between 0 and 1, a limit could be determined from the historical AIS points.

While the developed method has demonstrated encouraging performance, there is a need to test further. In particular, it would be necessary to compare the risk estimation with those from experienced mariners before the further development and deployment to the industry. This work serves as the first step towards the possibility of a risk-based navigation for autonomous ships, a further understanding of the requirements from the relevant industries and the authorities would also be an essential step.

Besides, there is a need to consider more parameters when the data is available. These include and are not limited to weather conditions and the time. Mathematically, this translates into a representation that moves to a higher dimensional space, covering more parameters. Achieving a balance among the various parameters or dimensions would also be necessary. While the current version of the work maps the historical data in the range of 0 to 1 individually before further processing, a better or a more suitable form could benefit the developed algorithms significantly.

The interaction between the vessels is not taken into consideration in the current version of the work. In fact, for a complete assessment, the risk coming from the other vessels especially in terms of being too close and hence a collision would need to be considered. For example, having two vessels which are relatively close especially and high estimated risks could be a good indication.

Next, while the computations so far were relatively fast, the computational complexity is still of linear basis. For a real-time monitoring or operation, an improvement to a (near-) constant complexity in terms of the time for an online search is preferred. Recent breakthroughs in achieving a reasonably close search with some additional assumptions such as approximate near-neighbour search [2] and Locality Sensitive Hashing [7] may help.

6 Conclusion

In summary, this work presents an approach for the risk estimation of a navigation. Starting by mapping the list of historical AIS points into a list of corresponding voyages, each of the points in a voyage is transformed as a point in the space of navigation feature. After the pre-processing, given a query in the form of a navigation feature, a search for the nearest neighbour from the list of historical voyages is performed and the distance to the nearest neighbour in the list of the pre-processed historical voyages is returned as the estimated risk at the moment. The developed method has demonstrated encouraging performance on a challenging dataset of historical AIS data covering three regions in Australia, namely the Bass Strait, the Great Australian Bight and the North West.

Acknowledgement

The support from ”National Natural Science Foundation of China” (Grant No. 11621202) and ”the Fundamental Research Funds for the Central Universities” (Grant No. WK2090000009) for the author Yanling Wu are gratefully acknowledged.

References

[1] https://creativecommons.org/licenses/by-nc/3.0/au/.

[2] Alexandr Andoni and Piotr Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '06, pages 459–468, Washington, DC, USA, 2006. IEEE Computer Society.

[3] Australian Maritime Safety Authority.

[4] Australian Maritime Safety Authority. https://www.operations.amsa.gov.au/spatial/.

[5] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, Sep 1995.

[6] M. Höyhtyä, J. Huusko, M. Kiviranta, K. Solberg, and J. Rokka. Connectivity for autonomous ships: Architecture, use cases, and research challenges. In *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 345–350, Oct 2017.

[7] Ke Jiang, Qichao Que, and Brian Kulis. Revisiting kernelized locality-sensitive hashing for improved large-scale image retrieval. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[8] T. A. Johansen and T. Perez. Unmanned aerial surveillance system for hazard collision avoidance in autonomous shipping. In *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 1056–1065, June 2016.

[9] O. Levander. Autonomous ships on the high seas. *IEEE Spectrum*, 54(2):26–31, February 2017.

[10] International Association of Classification Societies. http://www.iacs.org.uk/.

[11] International Maritime Organisation. Convention on the international regulations for preventing collisions at sea, 1972 (colregs).

[12] International Maritime Organisation. International maritime organisation.

[13] Lokukaluge P. Perera and C. Guedes Soares. Collision risk detection and quantification in ship navigation with integrated bridge systems. *Ocean Engineering*, 109:344 – 354, 2015.

[14] Riccardo Polvara, Sanjay Sharma, Jian Wan, Andrew Manning, and Robert Sutton. Obstacle avoidance approaches for autonomous navigation of unmanned surface vehicles. *Journal of Navigation*, 71(1):241–256, 2018.

[15] Ørnulf Jan Rødseth and Hans-Christoph Burmeister. Risk assessment for an unmanned merchant ship. *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*, 9(3):357–364, 2015.

[16] C. K. Tan, J. Wang, Y. C. Paw, and F. Liao. Autonomous ship deck landing of a quadrotor using invariant ellipsoid method. *IEEE Transactions on Aerospace and Electronic Systems*, 52(2):891–903, April 2016.

[17] W. C. Tan, I. Chen, S. J. Pan, and H. K. Tan. Analysis with histogram of connectivity: For automated evaluation of piping layout. *IEEE Transactions on Automation Science and Engineering*, 15(1):381–392, Jan 2018.

[18] W. C. Tan, I. Chen, D. Pantazis, and S. J. Pan. Transfer learning with pipnet: For automated visual analysis of piping design. In *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*, pages 1296–1301, Aug 2018.

[19] W. C. Tan, P. C. Goh, K. H. Chua, and I. Chen. Learning with corrosion feature: For automated quantitative risk analysis of corrosion mechanism. In *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*, pages 1290–1295, Aug 2018.

[20] W. C. Tan, C. Weng, Y. Zhou, K. H. Chua, and I. . Chen. Historical data is useful for navigation planning: Data driven route generation for autonomous ship. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7478–7483, May 2018.

[21] Ingrid Bouwer Utne, Asgeir J. Sørensen, and Ingrid Schjølberg. Risk management of autonomous marine systems and operations, 2017.

[22] Y. Wang, S. Wang, and M. Tan. Path generation of autonomous approach to a moving ship for unmanned vehicles. *IEEE Transactions on Industrial Electronics*, 62(9):5619–5629, Sep. 2015.