This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

A survey of machine learning approaches for surface maritime navigation

Azimi , Sepinoud; Salokannel, Johanna; Lafond, Sébastien; Lilius, Johan; Salokorpi, Mirva; Porres, Ivan

Published in:
Maritime Transport VIII: proceedings of the 8th International Conference on Maritime Transport: Technology, Innovation and Research: Maritime Transport '20

Published: 01/01/2020

Document Version
Final published version

Document License
CC BY-NC-ND

Link to publication

Please cite the original version:
Azimi , S., Salokannel, J., Lafond, S., Lilius, J., Salokorpi, M., & Porres, I. (2020). A survey of machine learning approaches for surface maritime navigation. In *Maritime Transport VIII: proceedings of the 8th International Conference on Maritime Transport: Technology, Innovation and Research: Maritime Transport '20* (pp. 103). Iniciativa Digital Politècnica. http://hdl.handle.net/2117/329714

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

This document is downloaded from the Research Information Portal of ÅAU: 23. Sep. 2023

A SURVEY OF MACHINE LEARNING APPROACHES FOR SURFACE MARITIME NAVIGATION

(I) Sepinoud Azimi, (II) Johanna Salokannel, (I) Sébastien Lafond, (I) Johan Lilius, (II) Mirva Salokorpi, (I)* Ivan Porres.

(I) Åbo Akademi University, Turku, Finland. name.surname@abo.fi

(II) Novia University of Applied Sciences, Turku, Finland. name.surname@novia.fi

* Correspondence author: Ivan Porres ivan.porres@abo.fi +358 50 40 96 303

Abstract:

In this article we present the state of the art in the field of autonomous surface ship navigation using machine learning. We discuss the main challenges towards the development of fully autonomous navigation systems with the International Regulations for Preventing Collisions at Sea (COLREGs). Finally, we propose two alternative approaches that are based on machine learning. Existing COLREGs-based navigation and collision avoidance algorithms are based on traditional search-based planning and optimization algorithms. We consider that these approaches are suitable when the problem space is defined completely and rigorously. However, experts believe that is not the case for COLREGs since it leaves many aspects open to the interpretation of the captain. For example, COLREGs expects that any collision avoidance action shall be taken with due regard to the observance of good seamanship, a concept not defined in the convention. Furthermore, many rules are defined using undefined concepts like safe distance, or keywords like early, or substantial, without giving any definition. COLREGs even allow for the rules to be broken to avoid an accident. Due to this, traditional planning approaches may not be able to handle complex scenarios that are underspecified according to COLREGs.

An alternative is the use of machine learning (ML), reinforcement learning (RL) and imitation learning (IL) at the core of autonomous navigation systems. Machine learning is known to succeed and outperform traditional approaches specially in vaguely defined problem domains, where it is difficult, if not impossible, to create a full formal specification of the phenomenon under study. We consider this to be the case for COLREGs-based navigation and we conjecture that a ML-based navigation approach can outperform existing search-based and optimization algorithms.

Keywords:

Autonomous navigation; surface ship navigation; machine learning.

INTRODUCTION

Autonomy as a way to increase safety, improve the logistics, decrease the operational costs, maximize profitability and reduce the carbon-footprint of traffic, has gained a lot of interest worldwide, and is mainly focusing, in Europe and the US, in the automotive area. Maritime Autonomous Surface Ships (MASS) of the future are expected to contain functionality that handle an increasing range of self-sufficiency. Autonomous capabilities include relieving the vessel operator from constant supervision by taking over certain responsibilities of the vessels using partial or complete remote operation of vessels, or partial or complete unsupervised navigation.

Technological advances in the field of big data, data analytics and processing, cloud computing, sensor technology and improved communication infrastructures open the doors to a future where ships will sail the seas autonomously. According to the Norwegian Forum for Autonomous Ships (NFAS), an autonomous vessel, whether manned or unmanned, is a ship with a certain level of automation and self-governance. These vessels will be equipped with systems having the ability to self-initiate actions. These actions rely on sensor-based detection and analysis using algorithms guiding the automated decision-making. The level of autonomy (LoA) defines the level of engagement that is expected of a human operator and the level of performance conducted by automation. LoA will depend on the vessel and its operational purpose and objectives. Some vessels may contain maintenance crew, others operated from the shore by the remote operation center crew and some might even be fully manned with intelligent automated support systems for decision-making. The level of autonomy can vary from ship to ship raising the need for successful interaction with humans.

Parasuraman, et al. (2000), and referred by Veritas (2017), divide maritime vessel functionalities into four levels of automation: Information acquisition; Information analysis; Decision and action selection and Action implementation.

The maritime industry is already leveraging various vessel-dependent levels of automation. Shipping has always been a fore runner in acquiring modern technology. For example, first electric navigation solutions emerged in the 1930s and vessels were the first civilian user of satellite navigation technology. Further, the anti-collision radar became mandatory onboard vessels from 1974 and automatic identification transponders from 2002, [33]. The vessels of today can be equipped with an autopilot following a pre-planned route in a track-following mode executing independently turns while the officer on watch (OOW) mainly monitors the proceedings and ensures safe navigation.

Today's unmanned vessels are mainly small crafts engaged in research and scientific activities, like for example marine environment exploration or deep ocean surveys. Some vessels perform underwater operations such as mine seeking and military has some autonomous surveillance vessels, e.g. US Navy's Sea Hunter Drone Ship. Although, at the moment, there is no fully functional autonomous maritime vessels in operation, there have been various attempts in that direction. Regarding the merchant vessels, the Yara Birkeland vessel is under development and is stated to be the world's first autonomous container ship sailing in the near future. Its design is mainly motivated by environmental factors. Also, some autonomous small passenger ferries have been tested such as NTNU's Autoferry and autonomous passenger ferries are expected to sail sometimes soon. Successful test runs have been done within the project SVAN by Rolls-Royce (currently Kongsberg Maritime Finland Ltd.) and Finnferries, where they successfully demonstrated the world's first fully autonomous ferry in the Finnish Archipelago. Similar tests have been done also by ABB (remote control of Suomenlinna -ferry) and Wärtsilä (successful tests with the Norwegian ferry Folgefonn).

The operational environment of a ship can vary from open sea navigation without traffic to highly complex dense traffic areas. Prevailing circumstances like weather conditions have also an impact on collision avoidance. Collision avoidance is very situationally bound, where concepts like "safe distance" or "in ample time" change according to the prevailing circumstances. Machine learning is known to succeed and outperform traditional approaches specially in vaguely defined problem domain, where it is difficult, if not impossible, to create a full formal specification of the phenomenon under study. We consider this is the case for COLREGs-based navigation and we conjecture that a ML-based navigation approach can outperform existing search-based and optimization algorithms.

1. THE DRIVING FORCES OF AUTONOMOUS SHIPPING

Maritime Autonomous Surface Ships of the future are expected to contain functionalities that handle an increasing range of self-sufficiency. Autonomous capabilities include relieving the vessel operator from constant supervision by taking over certain responsibilities of the vessels using partial or complete remote operation of vessels, or partial or complete unsupervised navigation.

An important motivation for autonomous functions in ships is to avoid human errors resulting from distraction, tiredness, lack of focus, etc. Despite the advances in technologies and the high level of reliability found in navigational equipment on a vessel, ship safety is still considered as one of the major issues in the maritime shipping industry. Several studies have shown a predominant impact of human behavior in maritime accidents and casualties, with an estimated 89% to 96% of ship collisions directly due to human error [2]. In 2017 alone, 3301 accidents were reported by the European Maritime Safety Agency and over 53% of all reported accidents were collisions, contacts or grounding occurrences, all due to navigational error [10]. The development of autonomous navigational capabilities is seen as a possible solution to dramatically reduce the number of accidents due to navigational error.

Several researchers have tackled this problem by proposing automatic COLREGs-based navigation algorithms and even complete systems [45, 13]. COLREGs prescribes a set of rules that all ships should follow to ensure safe navigations.

The origin of COLREGs goes back to the ending of Age of Sail in mid-19th century. The appearance of steam ships raised the concern about shipping safety and rules of the road at sea started to evolve. For decades the rules have been evolved according to the needs of maritime traffic and COLREGs were published in modern form in 1972. Now the maritime traffic is facing a new revolution, the autonomous vessels to which the traditional rules are challenging to apply. The COLREGs have been written to humans while in interaction with other humans in a qualitative format. Due to complex maritime domain, they also are general in nature to make them applicable to as many situations as possible [32]. The qualitative nature of COLREGs poses challenges for algorithm creation used for collision avoidance as the creation of anti-collision software requires the rules to be transformed into quantitative format. Definitions like "good seamanship" guiding the decision making of seafarers or changing concept of safe speed that is dependent on situational circumstances pose challenges for quantification. Further, the rules change according to weather conditions like restricted visibility or signals of the counter vessel such as restricted maneuverability. Therefore, COLREGs needs to be interpreted always in the context of prevailing navigational conditions and the correct application of the rules can vary from one situation to the next as the prevailing conditions change.

Porathe [31, 32], discusses the matter in his articles and raises the question for new set of rules for autonomous vessels. Although the question cannot be answered yet, he points out that if the autonomous vessels need to interact with humans, the actions of an autonomous vessel need to be transparent and understandable by human mariners. Endsley [11] states that poor understanding of the functioning of an autonomous system leads to inappropriate interaction with automation, resulting in poor situational awareness. Therefore, autonomous vessels do need to operate within the context of COLREGs to ensure that their actions are understood. Research on human behavior in various, specified circumstances and navigational areas can be the way to overcome challenges posed by quantification of COLREGs.

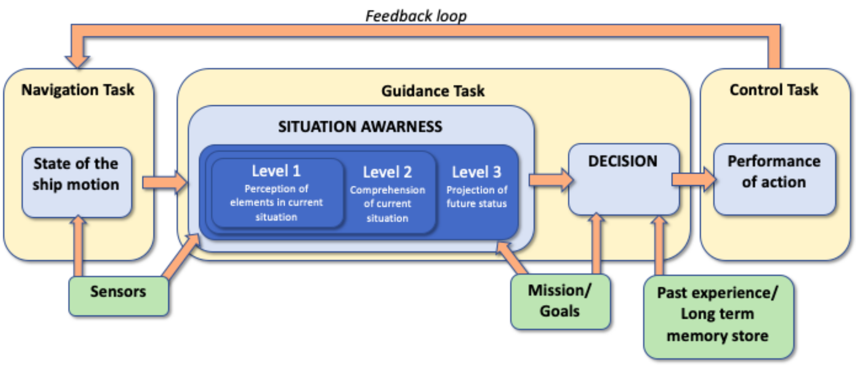
Maritime traffic has been coordinated in dense traffic areas by Traffic Separation Lanes, but these lanes do not cover all areas. Complex traffic situations with multiple vessels entering from various directions is just an ordinary day at sea for professional seafarers. Belcher [6] studied COLREGs from a societal point of view giving an example of complex traffic situations where the rules of COLREGs conflict with each other leaving the decision making and courses of action for individual interpretations of officers. Similar example was given by Porathe [32] pointing out that the actions of one ship in a complex situation will affect the decision making and actions of others. In addition, a collision avoidance situation can be solved with various solutions e.g. course change, speed change or both. These decisions also have an impact on which rule to apply as one rule can change to another dynamically in the same collision avoidance situation between vessels depending on the chosen courses of action.

2. STATE OF THE ART IN AUTONOMOUS NAVIGATION SYSTEMS

Maneuvering a ship is composed of three fundamental tasks: navigation, guidance and control [42]. The navigation task aims at estimating the state of the ship motion, which can be done with the help of a variety of sensors (for example inertial measurement unit, compass, global position system). The guidance of a ship is the task in charge of determining the desired ship state motion and is usually based on a set of input values related to physical environment of the ship, the ship mission or goals and associated costs (for example time, energy, and risk). The control task is implementing the required actuator commands to reach the desired ship state motion. A possible composition of tasks for an autonomous ship manoeuvring architecture is illustrated in Figure 1.

Understanding the vessel's environment, i.e., **situational awareness** is the corner stone for the guidance task. It includes components on several levels, level 1: perceive all environmental elements defining the current situation (for example, weather environment, type, location and speed of other vessels, etc.), level 2: comprehend the current situation (for example, understand and quantify a risk of collision), and level 3: project the future state of the ship (for example predict the near-future way points of the ship). The existing works for autonomous ships put a focus on the situation awareness components of the guidance task. Specifically, collision avoidance or path planning approaches have attracted attention through the past few years.

Figure 1. A decision-making architecture for an autonomous ship.



Recently, Zhou, et. al. [46] proposed a quantitative model of situational awareness based on the system safety control structure of remotely controlled vessel. This study provides a theoretical ground for further work in the area of situational awareness. The model developed in this study considers a high level of abstraction while using probabilistic approaches to provide guidelines for evaluating the navigation safety of autonomous ships. In another study, Murray and Perera [25], proposed an approach to facilitate situational awareness by predicting the other vessels' trajectories accurately. They used a data-driven approach taking advantage of the information collectable from historical AIS data. They employed a clustering based Single Point Neighbor Search Method along with a Multiple Trajectory Extraction Method. Endsley [12] proposed a three levels situational awareness model defined as Level 1: perception of element of the environment within a volume of time and space; Level 2: the comprehension of the current situation and Level 3 projection of the status in the near future.

Xu, et.al. [44] proposed an autonomous **collision avoidance** method which is designed based on the visual technique similar to the human's visual system. They used the recorded navigational manoeuvres for pre-defined scenarios performed on a simulator to train a deep convolutional neural network. The model is used to predict the collision avoidance manoeuvres based on the image scenes fed to the model as the input. During the same year, Xidias and Zissis [43] proposed an optimization approach to identify an optimum collision free path, based on the input obtained by the visibility graphs. Their optimization model is designed in such a way that the optimum path (1) should be minimal, (2) should respect the three COLREGs rules and (3) should avoid collisions with both static and moving obstacles. Target detection is another area of study that contributes in designing of collision avoidance approaches. For example, Stateczny, et. al. [37] provided an empirical analysis of the maritime surface target detection approaches, which could be utilised for the design of tracking and anti-collision systems for autonomous surface vehicles (ASV). Their research focused on identifying the field of views in which different surface targets could be detected. Their analysis included the objects that are typically detectable in the water environment, including a boat and other floating objects.

Path planning is also another area of interest. A popular choice is to develop Kalman filter-based algorithms for **path planning**. For example, Liu, et. al. [24] proposed a Kalman filter-based predictive path planning designed to predict the trajectories of moving ships, and the maritime vessel's own position in real time. Based on this information, this method, evaluates the potential collision risk of the vessels of interest. Allotta, et. al. [3, 4] conducted two studies to compare the efficiency of two different Kalman filter-based algorithms, namely, Extended Kalman filter (EKF) and Unscented Kalman filter (UKF). Kuwata, et. al. [22] took an algorithmic view towards solving the problem of identifying optimal path for the maritime vessels. As the input they consider the near-term waypoint, a reference speed, and a list of contacts representing moving and static hazards. The objective of their algorithm is to find the optimal velocity command in which the surface ship could avoid the hazards and follow the COLREGs. The algorithm computes the closest point of approach with the current position and velocity of the vessel and other traffic vessels and evaluates if any COLREGs rules need to be applied at all and change the current route based on the identified COLREGs rule. Later, Jeong, Lee, and Lee [20] developed a technique to devise a motion planning approach using real-time data. This is a risk analysis method which assesses and visualizes maritime traffic risk as the foundation for the route planning of an Autonomous Surface Vehicles (ASV). The optimal route that matches the desired objectives is identified using the risk contour map and the ASV's data processing.

Tan, et. al. [40] presented a different approach to path planning which joins AIS locations of the same vessel at different times and locations in a region into a route. Next, it automatically computes navigation plans using nearest neighbor-based path retrieval relying on two representations, Ship Feature and Navigation Feature. Then, it utilizes the available AIS data, including ship properties, and the preprocessed corresponding route which are accessible from the form of Ship and Navigation Feature. This approach considers the available navigational constrains in the form of a vector data, then identifies the nearest neighbor to this query vector in the space under investigation and eventually points out the recommended navigational path as the output. Taking into account the importance of respecting COLREGs, the path planning approach developed by Park, Choi, and Choi [29] is based on defining all potential trajectories' uncertainty. In this approach, the authors utilize a tracking filter that estimates the motion information. They use the error covariance if this filter as the basis of their modelling. Furthermore, based on this developed model, a probabilistic approach, they identify a considered collision risk zone (CRZ) for the predicted trajectory. Finally, considering the dynamic characteristics of the vessel and in order to avoid the identified CRZ, an optimal path is proposed. In another study, Bibuli, et. al. [7] developed a two layered approach in order to devise a multiple unmanned surface vehicles navigation framework. In the top layer of the architecture, a robust path planner is adopted to generate optimal waypoints, which are later smoothed using the polyfitting operation. This smoothed trajectory is given as an input to the bottom layer of the guidance system based on virtual target approach integrated with a swarm aggregation algorithm based on attraction- repulsion strategy.

Later, Hinostroza, Xu, and Soares [16] focused on designing a motion-planning unit, which is based on the angle-guidance fast-marching square method, which is specially developed for operation in dynamic and static environments. The collision avoidance unit is based on fuzzy-logic formulation, the guidance unit uses the vector-field guidance formulation and the control unit is composed by a PID heading controller and a speed controller.

3. MACHINE LEARNING APPROACHES FOR THE GUIDANCE TASK

Several ML-based navigation algorithms have been successfully employed in other application domains, like aerospace and automotive. Reinforcement learning, imitation learning and generative adversarial imitation learning are the most relevant approaches for autonomous navigation.

3.1 REINFORCEMENT LEARNING

Reinforcement learning (RL) models identify the main elements of the problem understudy by analyzing how an agent interacts with its environment to reach a maximum reward while pursuing an ultimate goal, see Sutton and Barto [38]. In an RL model, the agent is defined to derive useful information from the environment in order to perform actions which have influence on the environment based on a predefined reward function. The goal in an RL model is to formulate different steps from collecting information to taking action and achieving goal in such a way that they are simple as well as efficient and non-trivialized. One of the main obstacles in implementing an RL model is its reliance on the reward function. As designing a reward function requires involved and complex experiments to imitate the behaviors close to the natural one, many RL models tend to settle for a local minima rather than a global one, [27].

3.2 IMITATION LEARNING

Imitation learning (IL) models are designed to learn a policy based on the data collected from humans' demonstration to imitate an expert behavior. The policies learned by an IL model typically is captured through an expressive model such as neural networks [21] IL have been widely adopted when it comes to learning car navigation. One of such adoptions is Behavioral Cloning (BC) [30]. BC is a supervised learning approach in which the driver's behaviors is learned via building a regression model based on the dataset including tuples of state and actions recorded from human experts' demonstrations. The shortcoming of BC is using a regression model on the existing data and any deviation from the data could dramatically reduce the accuracy of the model. Such a shortcoming leads to the model being unstable, for example, Ross and Bagnell [34], argue that as the dataset of experts' demonstration lack the careless driver behavior and thus lacks recovery actions, the BC model could lead to errors and undesirable driving decisions. Ross et al [35] argues that a BC model, by nature, violates the independent and identically distributed random variables (i.i.d) assumption as the future behavior heavily relies on the previous actions. Thus, the BC model could potentially showcase poor performance.

Apprenticeship learning (AL) [1] is another IL method that learns an unknown reward function which is the basis of the expert demonstration. AL could be considered superior to BC, as it focuses on learning the reward function rather than the behavior from the dataset which could potentially include non-optimal demonstrations. Syed and Schapire [39], proposed a new AL approach based on the work of Abbeel and Ng [1], where they consider the unknown reward function to be the combination of a set of known and observable features. They claim that their approach has the potential to produce a policy, in times, better than the expert's. In contrast to

Syed and Schapire [39] approach, Levine et al. [23], assumes the reward function to be a non-linear one and use Gaussian processes to learn it. They argue that such an assumption, makes it feasible to capture more complex behaviors. As such, it is worth considering that AL is computationally expensive and it could only be utilized when the resources are available [17].

3.3 GENERATIVE ADVERSARIAL IMITATION LEARNING (GAIL)

As mentioned in the previous section, the implementation of the BC and AL algorithms is not always feasible. Several different approaches have been proposed to overcome the aforementioned issues. Inverse reinforcement learning (IRL) is one of such approaches, where the goal is to find a cost function which facilitates the expert's uniquely optimal behavior. However, IRL is also a computationally expensive approach as it has a middle RL loop. To alleviate this issue, Ho and Ermon [17], proposed a new approach called Generative Adversarial Imitation Learning (GAIL). GAIL tends to learn the policy directly from the data without the need for the middle RL loop. GAIL uses Generative Adversarial Networks (GAN) [15] to fit distributions of states and actions leading to identifying the expert's behavior. GAN consists of two main components: 1) a generator (producing new data samples and 2) a discriminator (evaluating whether the new sample belongs to the actual training dataset or not). GAN the generator and discriminator are learned by a gradient. In contrast to GAIL which is a model-free approach, Baram et al. [5],used a model-based approach and reparameterization techniques to identify the gradient of the discriminator based on the state and action information. Finally, similar to the study by Baram et al. [5], Blondé and Kalousis [8], proposed a new approach to deduce the gradient of the discriminator. Rather than using intermediary approaches, they directly deduce it based on the actions using a deterministic policy.

Taking into account the specific considerations related to the maritime domain, mainly the need of safety requirements and lack of training data, specific approaches need to be developed for the autonomous maritime domain. As such, we propose the development and evaluation of the following approaches: Safe Reinforcement Learning and Simulation-based Multi-agent Imitation Learning.

3.4 SAFE REINFORCEMENT LEARNING

The overall goal of Reinforcement Learning is to learn a decision-making policy to maximize a reward-based objective function. The reward function can be defined in terms of the mission objectives, for example arriving to the next way point of a previously planned route. However, Reinforcement Learning does not take into account possible risks or actions to avoid even if they would contribute to achieve the goal.

Constrained Reinforcement Learning [9] can be considered an extension of Reinforcement Learning where only policies that abide a given set of constraints are considered. We can introduce the concept of Safe Reinforcement Learning as an application of Constrained Reinforcement Learning where the constraints represent the safety requirements of our problem domain.

The safety constraints are formulated in terms of cost functions that represent states to avoid. In the case of autonomous navigation, one cost function can be used to estimate the risk of collision. This, in turn, can be quantified using the Velocity Obstacles approach [22] or the risk assessment function presented by Hu et al. [19]. Once we are able to quantify the risk of collision as a real function, we can define a constrain limiting what is the maximum risk that we are able to allow during the training of the agent. In this way, the policy implemented by the autonomous agent must find a balance over maximizing the reward function, i.e. achieving the mission objective, while abiding the constrains imposed by the cost function, i.e. bounding the risks of unsafe operation to an acceptable risk.

Safe Reinforcement Learning has the potential to combine the benefits of Reinforcement Learning while taking into account safety requirements. However, there are some challenges to this approach. The first one is the actual definition of the reward and cost functions. As discussed, different authors have proposed methods to calculate the estimated time and distance to a possible collision between two vessels, but often these models are rather simplistic and often assume that the ships are following a straight line. Finally, these models require a precise estimation of parameters such as the position and speed vector of the own and other vessels, and this task can also be challenging in case of equipment malfunction or failure.

3.5 SIMULATION-BASED MULTI-AGENT IMITATION LEARNING

The use of simulation platforms to create and test traffic patterns is a common approach used in development of autonomous land vehicles. Augmenting the simulator with real world inputs promises to provide richer and more authentic environment to test the reliability and safety of automatic navigational technology. However, unlike autonomous land vehicles, maritime traffic is sparser with rarer occurrence of navigationally complex events. In addition, unlike navigational guidelines for land vehicles, maritime navigational guidelines are intentionally vague due to a large permutation of possible scenarios, making it difficult to rely only on sets of expert-written rules and interpretations of maritime navigation regulations. The situation calls for a reliable dataset of various vessel, environment and traffic scenarios with data quality ideally from real-world measurements or mimics the real-world as close as possible.

The main challenge in multi-agent imitation learning lays in preparing the simulation environment, which is highly dependent on the task to be performed. The model has to figure out how to avoid a collision in a safe environment. The first step to build such a model, a dataset is to be constructed by directly recording the state as human experts navigates a ship, to facilitate making multiple renders of both the environment and data with varied environmental conditions, other vessels specifications and positions, and different routes. Two main components need to be introduced to build the required dataset: a set of sequential decision-making environments in the simulator and a corresponding public large-scale dataset of human demonstrations. In order to facilitate the training of multi-agent imitation learning models, simulations run by multiple users on the same scenario on different simulator bridges must be recorded.

The aim of the approach is to learn multiple parametrized policies that imitate the behavior of all experts. As the data are collected from multiple players in the same scenario, the model would consider interactions among all the vessels and reflects individual differences. This approach has an adversarial component, called a critic, to compute the difference between the off-course trajectories and true trajectories. During training, the algorithm iteratively updates the policy to minimize the difference and updates the critic to maximize the difference. When the algorithm converges, the critic cannot distinguish between the true trajectory and the roll-out trajectory, which implies point-wise convergence of the policy.

4. CONCLUSION

The fundamental challenge for the maritime industry is to achieve safety for the IT systems that are being deployed. A substantial amount of research has been done for the automotive area, but these results do not easily transfer to the maritime area. This is due to differences in the physics of the vessel's vs that of a car, differences in weather conditions leading to different sensing needs, and differences in traffic conditions making the datasets collected for the automotive area pretty useless in the maritime area.

There is a need to find an approach to address the general problem of autonomy for maritime vessels, by establishing a model of "good seamanship" using modern machine learning techniques like safe reinforcement learning, and multi-agent imitation learning.

In this paper we surveyed the field of algorithms for autonomous navigation for maritime vessels. It is clear that although there have been several attempts to find a solution to design surface ship autonomous navigation system, there is a need to explore new ways to learn the behaviors of actual humans. In this paper we proposed the development and evaluation of the following approaches: Safe Reinforcement Learning and Simulation-based Multi-agent Imitation Learning.

References:

[1] Abbeel, P.; Ng, A. Apprenticeship learning via inverse reinforcement learning. In: *Proceedings of the 21st International Conference on Machine Learning, Banff, Canada*. New York: Association for Computing Machinery, 2004. pp. 1-8, ISBN 9781581138382. Date of access: July 2020. Available from: <https://ai.stanford.edu/~ang/papers/icml04-apprentice.pdf>

[2] Al-Shammari, N.; OH, J. Effects Of Human Error On Marine Safety: Case Study. *Journal of Engineering Research and Applications*, Ghaziabad: International Organization of Scientific Research, 2018. Vol 11(8), pp. 11-19. Date of access: July 2020. Available from: <http://www.ijera.com/papers/vol8no11/p2/B0811021119.pdf >

[3] Allotta, B. et al. Development of a navigation algorithm for autonomous underwater vehicles. *IFAC-PapersOnLine*, Elsevier, 2015. 48(2), pp. 64-69. Date of access: July 2020. Available from:< https://doi.org/10.1016/j.ifacol.2015.06.011 >

[4] Allotta, B. et al. An unscented Kalman filter based navigation algorithm for autonomous underwater vehicles. *Mechatronics*, Elsevier, 2016. vol. 39, pp. 185-195. Date of access: July 2020. Available from:< https://doi.org/10.1016/j.mechatronics.2016.05.007 >

[5] Baram, N.; Anschel, O.; Caspi, I.; Mannor, S. End-to-end differentiable adversarial imitation learning. In: *Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017*., JMLR, 2017 pp. 390-399. Date of access: July 2020. Available from:< http://proceedings.mlr.press/v70/baram17a/baram17a.pdf >

[6] Belcher, P. A sociological interpretation of the COLREGS. *The journal of navigation.* The Royal Institute of Navigation. Cambridge University Press, 2002. 55(2), pp. 213-224. Date of access: July 2020. Available from: < https://doi.org/10.1017/S0373463302001686 >

[7] Bibuli, M. et al. A two layered optimal approach towards cooperative motion planning of unmanned surface vehicles in a constrained maritime environment. *IFAC-PapersOnLine*, Elsevier, 2018. 51(29), pp. 378-383. Date of access: July 2020. Available from: <https://doi.org/10.1016/j.ifacol.2018.09.458 >

[8] Blondé, L.; Kalousis, A. Sample-Efficient Imitation Learning via Generative Adversarial Nets. In: *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS) 2019, Naha, Okinawa, Japan. The Proceedings of Machine Learning Research, 89*. 11 p. Date of access: July 2020. Available from: <http://proceedings.mlr.press/v89/blonde19a/blonde19a.pdf >

[9] Chow, Y.; Ghavamzadeh, M.; Janson, L.; Pavone, M. (2017). Risk-constrained reinforcement learning with percentile risk criteria.. *The Journal of Machine Learning Research*. 2018, 18(1), pp. 6070-6120. Date of access: July 2020. Available from: <http://www.jmlr.org/papers/volume18/15-636/15-636.pdf >

[10] European Maritime Safety Agency. *Annual Overview of Marine Casualties and Incidents 2018*, s.l.: EMSA, 2018. Date of access: July 2020. Available from: <http://www.emsa.europa.eu/news-a-press-centre/external-news/item/3406-annual-overview-of-marine-casualties-and-incidents-2018.html >

[11] Endsley, M. The limits of highly autonomous vehicles: an uncertain future: Commentary on Hancock. *Ergonomics*, [Londres]: Taylor & Francis, 2019. 62(4), pp. 496-499. Date of access: July 2020. Available from: < https://doi.org/10.1080/00140139.2019.1563330 >

[12] Ensley, M. Toward a theory of situation awareness in dynamic systems. *Human factors.* Santa Monica, CA: Human Factors and Ergonomics Society, 1995. Volume 37, pp. 85-104. Date of access: July 2020. Available from: <https://doi.org/10.1518/0018720957790495439781581138382 >

[13] Eriksen, B.; Bitar, G.; Breivik, M.; Lekkas, A. Hybrid Collision Avoidance for ASVs Compliant with COLREGs Rules 8 and 13-17. *Frontiers in Robotics and AI*, [Lausanne, Switzerland]: Frontiers in Robotics and AI, 2020. Volume 7, p. 11. Date of access: July 2020. Available from: < https://doi.org/10.3389/frobt.2020.00011 >

[14] Goodfellow, I. et al. Generative adversarial nets. *Advances in neural information processing systems.* 2014. Volume 27. pp. 2672-2680. Date of access: July 2020. Available from: < https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf >

[15] Goodfellow, I. Tutorial: Generative Adversarial Networks. *Advances in neural information processing systems.* 2016. Date of access: July 2020. Available from: <https://arxiv.org/pdf/1701.00160.pdf >

[16] Hinostroza, M.; Xu, H.; Soares, C. Cooperative operation of autonomous surface vehicles for maintaining formation in complex marine environment. *Ocean Engineering*. New York, NY: Elsevier Science Pub. Co., 2019. Volume 183, pp. 132-154. Date of access: July 2020. Available from: < https://doi.org/10.1016/j.oceaneng.2019.04.098 >

[17] Ho, J.; Ermon, SGenerative adversarial imitation learning. *Advances in neural information processing systems*. s.l: s.n., 2016. Volume 29. pp. 4565-4573. Date of access: July 2020. Available from: <https://papers.nips.cc/paper/6391-generative-adversarial-imitation-learning.pdf>

[18] Ho, J.; Gupta, J.; Ermon, S. Model-free imitation learning with policy optimization. In: *ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning*. s.l: Journal of Machine Learning Research., 2016. Volume 48, pp. 2760-2769. ISSN: 2640-3498. Date of access: July 2020. Available from: <http://proceedings.mlr.press/v48/ho16.pdf >

[19] Hu, L. et al. A multiobjective optimization approach for COLREGs-compliant path planning of autonomous surface vehicles verified on networked bridge simulators. *IEEE Transactions on Intelligent Transportation Systems.* Los Alamitos, CA : IEEE Computer Society, 2019. 21(3), pp. 1167 – 1179. Date of access: July 2020. Available from: <http://doi.org/10.1109/TITS.2019.2902927 >

[20] Jeong, M.; Lee, E.; Lee, M. An Adaptive Route Plan Technique with Risk Contour for Autonomous Navigation of Surface Vehicles. In: *OCEANS 2018 MTS/IEEE Charleston MTS/IEEE Charleston, OCEANS 2018. :1-6 Oct, 2018.* IEEE, 2018. pp. 1-4. Date of access: July 2020. Available from: < http://doi.org/10.1109/OCEANS.2018.8604638 >

[21] Kuefler, A.; Morton, J.; Wheeler, T.; Kochenderfer, M. Imitating driver behavior with generative adversarial networks. In: *2017 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2017.pp. 204-211. Date of access: July 2020. A Available from: <http://doi.org/10.1109/IVS.2017.7995721 >

[22] Kuwata, Y.; Wolf, M.; Zarzhitsky, D.; Huntsberger, T. Safe maritime autonomous navigation with COLREGS, using velocity obstacles. *IEEE Journal of Oceanic Engineering*, Los Alamitos, CA : IEEE Computer Society, 2013. 39(1), pp. 110-119. Date of access: July 2020. Available from: < http://doi.org/10.1109/JOE.2013.2254214 >

[23] Levine, S.; Popovic, Z; Koltun, V. Nonlinear inverse reinforcement learning with gaussian processes. *Advances in Neural Information Processing Systems.* s.l.:s.n., 2011, Volume 24, pp. 19-27. Date of access: July 2020. Available from: < https://papers.nips.cc/paper/4420-nonlinear-inverse-reinforcement-learning-with-gaussian-processes>

[24] Liu, Y.; Liu, W.; Song, R.; Bucknall, R. Predictive navigation of unmanned surface vehicles in a dynamic maritime environment when using the fast marching method. *International Journal of Adaptive Control and Signal Processing*, Hoboken, NJ : John Wiley & Sons, 2017. 31(4), pp. 464-488. Date of access: July 2020. Available from: <https://doi.org/10.1002/acs.2561 >

[25] Murray, B.; Perera, L. A Data-Driven Approach to Vessel Trajectory Prediction for Safe Autonomous Ship Operations. In: *2018 Thirteenth International Conference on Digital Information Management (ICDIM).* Los Alamitos, CA : IEEE Computer Society, 2018. pp.

240-247. Date of access: July 2020. Available from: <https://doi.org/10.1109/ICDIM.2018.8847003 >

[26] Ng, A.; Russell, S. Algorithms for inverse reinforcement learning. In: *ICML '00: Proceedings of the Seventeenth International Conference on Machine Learning*. s.l., ICML, 2000. pp. 663-670. Date of access: July 2020. Available from: <https://ai.stanford.edu/~ang/papers/icml00-irl.pdf >

[27] Ng, A.; Harada, D.; Russell, S. Policy invariance under reward transformations: Theory and application to reward shaping. In: *ICML '99: Proceedings of the Sixteenth International Conference on Machine Learning*. s.l., ICML, 1999. pp. 278-287. Date of access: July 2020. Available from: <https://people.eecs.berkeley.edu/~pabbeel/cs287-fa09/readings/NgHaradaRussell-shaping-ICML1999.pdf >

[28] Parasuraman, R.; Sheridan, T. ; Wickens, C. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, Los Alamitos, CA : IEEE Computer Society, 2000. 30(3), pp. 286-297. Date of access: July 2020. Available from: < https://doi.org/10.1109/3468.844354 >

[29] Park, J.; Choi, J.; Choi, H. COLREGS-compliant path planning considering timevarying trajectory uncertainty of autonomous surface vehicle. *Electronics Letters*, New York, NY : Antennas and Propagation Society of the Institute of Electrical and Electronics Engineers, 2019. 55(4), pp. 222-224. Date of access: July 2020. Available from: <https://doi.org/10.1049/el.2018.6680 >

[30] Pomerleau, D. Alvinn: An autonomous land vehicle in a neural network. *Advances in neural information processing systems*. Cambridge [etc.] : The MIT Press, cop. 1989 , volume 2. pp. 429-435. Date of access: July 2020. Available from: < https://papers.nips.cc/paper/95-alvinn-an-autonomous-land-vehicle-in-a-neural-network.pdf >

[31] Porathe, T. Interaction Between Manned and Autonomous Ships: Automation Transparency. In:. *ICMASS 2018. Proceedings of the 1st International Conference on Maritime Autonomous Surface Ships, Busan, Korea*. s.l., SINTEF Academic Press, 2019. Date of access: July 2020. Available from: < http://hdl.handle.net/11250/2599020 >

[32] Porathe, T. Maritime Autonomous Surface Ships (MASS) and the COLREGS: Do we need quantified rules or is "the ordinary practice of seamen" specific enough?. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 2019. 13(3). Date of access: July 2020. Available from: < https://doi.org/10.12716/1001.13.03.04 >

[33] Rødseth, Ø.; Burmeister, H. Developments toward the unmanned ship. In*: International Symposium Information on Ships–ISIS*. S.l: Deutsche Gesellschaft für Ortung und Navigation 2012. pp. 30-31. Date of access: July 2020. Available from: <https://pdfs.semanticscholar.org/6a8a/771d52e21 0f36d48cd345a13aab294d83780.pdf >

[34] Ross, S.; Bagnell, D. Efficient reductions for imitation learning. In: *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy*. s.l.: Proceedings of Machine Learning Research, 2010. Volume 9. pp. 661-668. Date of access: July 2020. Available from: <http://proceedings.mlr.press/v9/ross10a/ross10a.pdf >

[35] Ross, S.; Gordon, G.; Bagnell, D. A reduction of imitation learning and structured prediction to no-regret online learning. In*: Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS) 2011, Fort Lauderdale, FL, USA.* s.l.: Journal of Machine Learning Research, 2011. Volume 15, pp.627-635. Date of access: July 2020. Available from: < http://proceedings.mlr.press/v15/ross11a/ross11a.pdf >

[36] Trevithick, J. Navy's Sea Hunter Drone Ship Has Sailed Autonomously To Hawaii And Back Amid Talk Of New. [Online] *The drive: The war zone.* February 4, 2019. Date of access: July 2020. Available from: https://www.thedrive.com/the-war-zone/26319/usns-sea-hunter-drone-ship-has-sailed-autonomously-to-hawaii-and-back-amid-talk-of-new-roles

[37] Stateczny, A.; Kazimierski, W.; Gronska-Sledz, D.; Motyl, W. The Empirical Application of Automotive 3D Radar Sensor for Target Detection for an Autonomous Surface Vehicle's Navigation. *Remote Sensing,* S.l: MDPI, 2019. 11(10), p. 1156. Date of access: July 2020. Available from: < https://doi.org/10.3390/rs11101156 >

[38] Sutton, R.; Barto, A. R*einforcement learning: an introduction.* s.l.: Cambridge. MA: MIT Press, 1998. Date of access: July 2020. Available from: <https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf >

[39] Syed, U.; Schapire, R. A game-theoretic approach to apprenticeship learning. In: *Advances in neural information processing systems.* s.l.: Neural Information Processing Systems Foundation, 2008. pp. 1449-1456. Date of access: July 2020. Available from: <https://papers.nips.cc/paper/3293-a-game-theoretic-approach-to-apprenticeship-learning.pdf>

[40] Tan, W. et al. Historical Data is Useful for Navigation Planning: Data Driven Route Generation for Autonomous Ship In: 2018 IEEE International Conference on Robotics and Automation (ICRA). Los Alamitos, CA : IEEE Computer Society, 2018. pp. 7478-7483. Date of access: July 2020. Available from: < https://doi.org/10.1109/ICRA.2018.8460880 >

[41] Bureau Veritas. Guidelines for autonomous shipping. *Guidance Note NI*, Paris: Bureau Veritas, 2017. p. 641. Date of access: July 2020. Available from: <http://erules.veristar.com/dy/data/bv/pdf/641-NI_2019-10.pdf >

[42] Wolek, A.; Woolsey, C. Model-based path planning. In: *Sensing and Control for Autonomous Vehicles.* s.l.: Springer, Cham., 2017. pp. 183-206. Date of access: July 2020. Available from: < https://link.springer.com/chapter/10.1007/978-3-319-55372-6_9 >

[43] Xidias, E.; Zissis, D. Real Time Autonomous Maritime Navigation using Dynamic Visibility Graphs. In*: 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV).* Los Alamitos, CA : IEEE Computer Society, 2018. pp. 1-6. Date of access: July 2020. Available from: < https://doi.org/10.1109/AUV.2018.8729821 >

[44] Xu, Q.; Yang, Y.; Zhang, C.; Zhang, L. Deep convolutional neural network-based autonomous marine vehicle maneuver. *International Journal of Fuzzy Systems*, [Berlin]: Springer, 2018. 20(2), pp. 687- 699. Date of access: July 2020. Available from: <https://link.springer.com/article/10.1007/s40815-017-0393-z >

[45] Zaccone, R.; Martelli, M.; Figari, M. A colreg-compliant ship collision avoidance algorithm. In: *2019 18th European Control Conference (ECC)* Los Alamitos, CA : IEEE

Computer Society, 2019, pp. 2630-2535. Date of access: July 2020. Available from: <https://doi.org/10.23919/ECC.2019.8796207 >

[46] Zhou, X.; Liu, Z.; Wu, Z.; Wang, F. (2019). Quantitative processing of situation awareness for autonomous ships navigation. *Transnav: International Journal on Marine Navigation and Safety of Sea Transportation*, Volume 13 (1). Date of access: July 2020. Available from: < https://doi.org/10.12716/1001.13.01.01 >