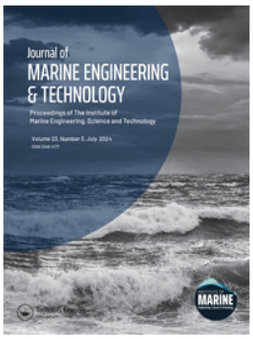


Journal of Marine Engineering & Technology

ISSN: 2046-4177 (Print) 2056-8487 (Online) Journal homepage: www.tandfonline.com/journals/tmar20

Adversarial waypoint injection attacks on Maritime Autonomous Surface Ships (MASS) collision avoidance systems

G. Longo, M. Martelli, E. Russo, A. Merlo & R. Zaccone

To cite this article: G. Longo, M. Martelli, E. Russo, A. Merlo & R. Zaccone (2024) Adversarial waypoint injection attacks on Maritime Autonomous Surface Ships (MASS) collision avoidance systems, Journal of Marine Engineering & Technology, 23:3, 184-195, DOI: 10.1080/20464177.2023.2298521

To link to this article: https://doi.org/10.1080/20464177.2023.2298521

Published online: 26 Dec 2023.

Submit your article to this journal

Article views: 758

View related articles

View Crossmark data

Citing articles: 24 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tmar20

JOURNAL OF MARINE ENGINEERING & TECHNOLOGY
2024, VOL. 23, NO. 3, 184–195
https://doi.org/10.1080/20464177.2023.2298521

Taylor & Francis
Taylor & Francis Group

Check for updates

Adversarial waypoint injection attacks on Maritime Autonomous Surface Ships (MASS) collision avoidance systems

G. Longo [a], M. Martelli[b], E. Russo[a], A. Merlo[c] and R. Zaccone[b]

[a]Dept. of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), Polytechnic School of Genoa University, Genova, Italy; [b]Dept. of Electrical, Electronic, Telecommunications, Naval Architecture and Marine Engineering (DITEN), Polytechnic School of Genoa University, Genova, Italy; [c]Centre for Higher Defense Studies, Rome, Italy

ABSTRACT

Autonomous navigation is currently subject to particular interest for naval and commercial vessels. In addition to the regulatory effort, several pieces of research tackled the development of new guidance laws, stable and robust control algorithms, methodologies to increase situational awareness, and collision avoidance algorithms. However, most of these systems blindly trust information from navigation sensors, which can malfunction or, even worse, have their data altered by cyber-attacks. This last scenario will be increasingly common shortly and represents a dangerous threat that future highly computerised ships must face. The proposed work shows an approach capable of hijacking a state-of-the-art collision avoidance algorithm route by feeding it with rogue data. The effects of such a cyber-attack are shown via an extensive simulation campaign, testing the methodology on three different COLREG scenarios: head-on, crossing, and overtaking. The study additionally highlights the susceptibility of forthcoming navigation systems to security breaches, which still needs to be addressed in the existing literature. According to the experimental findings, attackers can attract the autonomous ship guidance toward their objective at a mean distance of 133 m. Finally, suggestions on how to harden these systems against this family of attacks are provided in the conclusion.

ARTICLE HISTORY
Received 8 May 2023
Accepted 18 December 2023

KEYWORDS
Autonomous ship; collision avoidance; cyber security; data injection attack

1. Introduction

Autonomous navigation is one of the most challenging new research topics in the maritime sector. Such a topic caters to the interests of multiple communities rooted in academia, industry, and even policymakers. While control engineers concern themselves with the development of optimal route planners and robust control systems, insurers and lawyers are analysing which implications are associated with the existence of autonomous ships. An autonomous vessel is a highly sophisticated cyber-physical system containing many digitalised onboard systems cooperating during navigation. In particular, several taxonomies exist for human-automation interaction (Parasuraman et al. 2000). The automotive sector mainly refers to the one proposed by SAE (Shadrin and Ivanova 2019). Conversely, the maritime moved from the six-level scale proposed by the Lloyd's Register in the early days (Voicu 2016) to the consolidated four-level scale set by the International Maritime Organization (IMO) (International Maritime Organization 2018), which classifies autonomous ships into four increasingly challenging and automated 'degrees':

- Degree One identifies ships with automated processes and decision support, in which seafarers are actively operating onboard, except for some operations that may be automated and, at times, unsupervised.
- Degree Two includes a remotely controlled ship with seafarers on board to take control if needed.
- Degree Three includes a remotely controlled ship without seafarers on board.
- Degree Four identifies the fully autonomous ships controlled by a digital operating system capable of taking decisions and actions without human intervention.

As a result of this large amount of safety-critical digitised systems (Kavallieratos et al. 2018), and reports of multiple accidents whose numbers are increasing over time (see Figure 1), the problem of cyber security has gained significant importance in the scientific debate (Katsikas 2017). In a future outlook in which an increasing number of vessels is going to be handled by computer systems (Kim et al. 2020), the analysis of cyber risk factors and the defence from cyber attacks will become a cornerstone element of shipping operations (Misas et al. 2022). With attackers capable of influencing the ship systems through cyber attacks, an autonomous vessel could be lured near a fixed threat such as a minefield or in an area exposing her to other risks like collision and grounding. These newfound capabilities would also enable pirates to hijack the ship's trajectory in an area where she can be seized for ransom. Economic warfare could be done by unnecessarily lengthening a trip to damage goods or disrupt a supply chain. Lastly, a ship could be forced to navigate in a restricted area or enter international waters against her will. These possible scenarios come with severe consequences that near-future automation designers should consider.

All the above-listed attacks can be performed at different levels, depending on the Attacker's available resources and technical proficiency. For instance, the Maritime Cyber Risk Assessment framework (MaCRA) (Tam and Jones 2019) uses a five-tier system based

CONTACT G. Longo giacomo.longo@dibris.unige.it Dept. of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), Polytechnic School of Genoa University, Via Dodecaneso, 35, Genova 16146, Italy

© 2023 Institute of Marine Engineering, Science & Technology

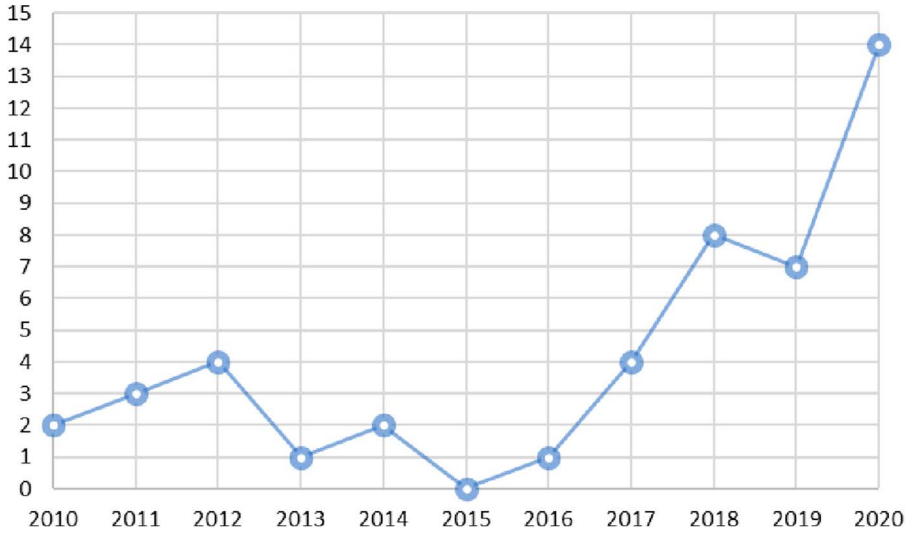


Figure 1. Number of reported cyber incidents per year (Meland et al. 2021).

on equivalences in conventional computing systems. The tiers are ranked in descending order based on the required resources and 'hacking ability' necessary for the exploit. These tiers range from tier 5, a script kiddie possessing limited resources and capabilities, to tier 1, involving acts of cyber war between nation-states with unlimited access to resources and workforce. As the level of sophistication of these attacks increases, a scientific gap has arisen for research tackling these next-generation attacks and their countermeasures to be deployed during both the operation and design phases.

This paper is an extension of the authors' previous work (Longo et al. 2022), and it explores even more scenarios in which attackers could inject a rogue waypoint inside an autonomous ship collision avoidance system. Based on the same underlying goals, this research introduces a new and improved regression method that showcases comparatively better results in achieving the attackers' objectives. Furthermore, the scope of the study has been extended to cover COLREGs rules 13 and 15, thereby achieving full coverage concerning encounters between power-driven vessels. It is worth noting that the COLREGs-complaint collision avoidance systems and the attack presented in this paper are applicable across the four IMO degrees of autonomy. It is reasonable to expect that the effectiveness of the attack is greater the higher the level of autonomy, i.e. the less human surveillance and intuition is involved. In summary, this paper provides the following contributions:

- A novel attacker model is presented and discussed: it can lead to unauthorised manipulation of a ship's trajectory by exploiting maritime-specific protocols, international regulations, and state-of-the-art Collision Avoidance Algorithm (CAA) implementations.
- A practical implementation is provided. Moreover, an extensive evaluation involving the execution of nearly half a million simulated scenarios reveals the capability of attackers to accomplish their goals.
- Possible remediation strategies to address the vulnerabilities the attack exploits or enable early detection are proposed and examined.

Eventually, this work aims to raise awareness among stakeholders, including researchers, industry practitioners, policymakers, and legal experts, of the evolving cyber threats in maritime operations and the need for proactive security measures. The paper is structured as follows: Section 2 discusses related work; the target collision avoidance module and the attacker model and capabilities are described in Sections 3 and 4; Sections 5 and 6 describe the attack guiding concepts and implementation. Eventually, Section 7 discusses the attack and its effectiveness, Section 8 proposes some remediations and recommendations, and Section 9 draws the necessary conclusions.

2. Related work

Several studies examine the cybersecurity risks and the threat landscape in the maritime context, including autonomous and traditional vessels. Notably, two works consider the collision avoidance system. The first (Kavallieratos and Katsikas 2020) assesses the cyber risk of cyber-physical systems on board digitalised contemporary and future ships: the authors include the collision avoidance system between critical components. Moreover, they identify an attack similar to this paper's scenario, i.e. '*an adversary may spoof the existence of another ship in the vicinity, thereby causing the vessel to change its route*', and assign a high impact and risk to it. The second paper related to collision avoidance (Bolbot et al. 2020) emphasises that installing malware onboard ships can disrupt collision avoidance and situation awareness systems, leading to potentially severe consequences. Nevertheless, the risk assessments were conducted at a high level and did not thoroughly investigate the exploitation of vulnerabilities or the execution of actual system attacks.

In this context, several studies present strategies targeting components that could compromise navigational integrity and the input data utilised by the collision avoidance module. In line with the goal of this paper, Bhatti and Humphreys (2017) investigate and showcase the Attacker's capacity to manipulate a maritime surface vessel using fake civil GPS signals. RADAR and information about surrounding ships that it conveys via ARPA represent another primary input for collision avoidance modules. Still, RADAR has also been subjected to attacks (Longo, Russo, et al. 2023) that have demonstrated the ability to introduce deceptive vessels into the system. Similarly, Meucci et al. (2023) demonstrate how APTs can leverage GANs to craft deceptive ISARs images, thus undermining navigational integrity. Another significant input for collision avoidance modules is information extracted from camera images. In particular, AI technologies and support vector machines can be used to detect and locate objects. In such a framework, adversarial attacks (Walter et al. 2023) can mislead the decisions made by collision avoidance. Other attacks exploit

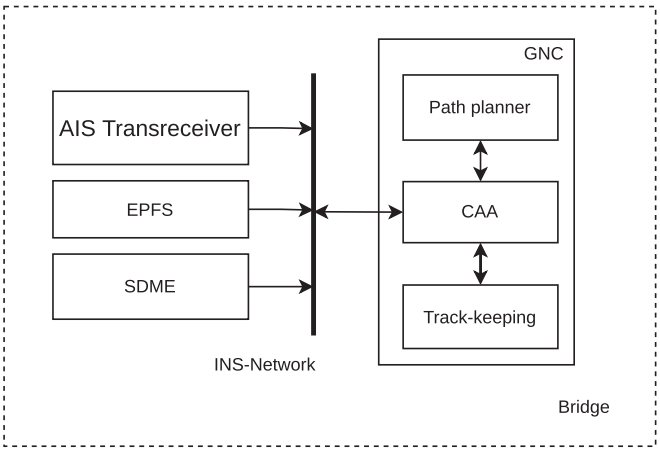


Figure 2. An overview of the autonomous navigation bridge architecture.

the weaknesses of the NMEA standard (International Electrotechnical Commission 2016) during the transmission of sensors and AIS data (Hemminghaus et al. 2021). In particular, Longo, Orlich, et al. (2023) shows an example attack featuring an adversary who injects rogue NMEA sentences to acquire control over an engaged autopilot. However, to the authors' knowledge, no existing studies focus on practical attacks against collision avoidance systems and prove the capability to impact ships' trajectories through such means.

3. Collision avoidance

The collision avoidance capabilities of an autonomous surface vessel are ensured by a GNC system. Figure 2 presents the architecture of a prototypical autonomous navigation bridge. Within the INS, data flows via a computerised network from the onboard sensors to provide situational awareness of the surrounding environment (see Section 4.1), including fixed obstacles or other ships. The GNC system includes a path-planning module, a CAA module, and a track-keeping module. During its operation, the INS shares information about surrounding obstacles and the vessel's position, speed and heading with the CAA module. Consequently, the path planning module computes a collision-free route based on the available information. The route is updated at a fixed rate according to up-to-date obstacle information: the motion planning module recomputes the optimal path each time. To ensure collision avoidance and COLREGs compliance, the path planning algorithm predicts future motions of the moving obstacles: periodic re-planning allows updating the forecasted obstacle data with any unpredictable changes in the obstacles' motion. Eventually, the route waypoints are fed to the track-keeping module, ensuring that the vessel follows the computed route.

In this article, the motion planner is based on the RRT∗ algorithm and provides COLREGs-compliant (International Maritime Organization 1972) evasive manoeuvres, considering the manoeuvring capabilities of the vessel (Zaccone et al. 2019). The algorithm has been extensively tested in simulation in realistic scenarios (Zaccone and Martelli 2020; Zaccone 2021). The presented architecture is designed for open-water navigation. Therefore, the own vessel is supposed to move at a medium to high speed and to keep a reasonable distance from the other ships in the scenario, in the order of a minimum of some ship lengths or at least twice the tactical diameter. Moreover, the evasive manoeuvres do not require a reduction of the own vessel's speed, and a course change is sufficient.

While both speed and course alterations are recognised as equal means of collision avoidance, the course change is preferable in most scenarios, as reported in the statistics presented by Li et al. (2021). In addition to the relevant statistics, the assumption is also supported by operational considerations; depending on the propulsion system's performance and the ship's dynamics, a speed reduction and a following acceleration to reach the cruise speed again can be time and energy-consuming, affecting the voyage (ETA) and fuel consumption. Moreover, engine performance can limit speed increment (power output close to the MCR, adverse weather conditions, etc.). Conversely, excessive speed reduction can result in low manoeuvrability performance. Eventually, it is worth noting that, in some cases, the speed reduction might not show the clear intention of avoiding an obstacle as prescribed by the COLREGs. Therefore, while speed alteration is more common in restricted water, as smaller manoeuvring areas may exist, course alterations are the first-choice collision avoidance action in the open sea.

In summary, the considered collision avoidance system tries to generate CAMs according to the following set of rules:

(1) If traffic ships are detected, the CAA tries to perform a COLREGs-compliant CAM, keeping a safety distance of about 1 nautical mile and limiting the route elongation;
(2) If no COLREGs-compliant CAMs are possible, the CAA tries to perform a non-COLREGs-compliant manoeuvre respecting the safety distance;
(3) If no CAMs are possible, the algorithm progressively reduces the safety distance until a feasible CAM is found.

4. Attacker model

4.1. Assumptions

The layout of Figure 2 represents the supposed architecture of the victim ship, including the network hosting an INS (International Maritime Organization 2007) and an excerpt of data sources related to connected sensors. In particular, the following sensors are considered:

- An AIS transreceiver (International Telecommunication Union 2014), which picks up radio signals and transmits over the network the position reports of other AIS-equipped vessels;
- The EPFS and SDME, which report the position and speed of the own ship, respectively.

The above data sources represent the inputs of the GNC system, which sends the computed route waypoints and cross-track error back to the network. All the transmissions are expected to leverage the standard NMEA protocol. It is assumed that a ship under the guidance of a CAA tries to meet two basic objectives:

(1) COLREGs-compliant guidance;
(2) A minimal deviation w.r.t. the original route whenever a CAM is required.

While the first assumption stems from the obligations a ship engaged in an international voyage has to follow International Maritime Organization (2020), the second avoids any unnecessary deviations from the original path. The autonomous ship is assumed to follow the routes generated by its collision avoidance module by manually operating its controls or automatically by a track-keeping autopilot or similar device. Eventually, the attackers are supposed to have already gained access to the ship systems and are connected to its INS network. For instance, infiltrating the ship systems could have

been performed during maintenance operations (Lund, Gulland, et al. 2018) or exploiting supply chain vulnerabilities (Mitre Corporation 2023, T1195). Malware can be injected in the automation workstations by connecting USB devices (BIMCO 2021, p. 32, 36) or by leveraging known weaknesses in their software (Dyryavyy 2014; Lund, Hareide, et al. 2018; Svilicic et al. 2019).

4.2. Capabilities

In the present study, adversaries are supposed to be skilled entities with the expertise to gather comprehensive information for devising and experimenting with new forms of attack. According to the Maritime Cyber Risk Assessment framework, these adversaries are classified as Tier 3 attackers. Under the assumption that the attackers are connected to the INS network (see Figure 2), they can operate like any other onboard instrument and interact by reading and writing NMEA data. This condition enables the following different attackers' capabilities:

- Observe the position (as reported by EPFS), speed (as sensed by SDME), and route waypoints (as planned by GNC) of the ship under attack.
- Gain an approximate knowledge of some properties of the CAA in use (see Section 5).
- Inject fake information into the INS regarding actual or fictitious ships navigating in the area, with a predetermined position and speed, i.e. forge and transmit AIS messages by leveraging the shared communications bus nature of the INS.

Eventually, attackers can rely on highly predictable outcomes of some manoeuvres. This capability assumes that the CAA operates according to the COLREGs, an acceptable hypothesis for a system running on a vessel engaged in an international voyage.

5. Attack description

Attackers aim to lure an autonomous ship towards a predetermined malevolent location by making its CAA generate a route passing as close as possible to it. For a vessel following its CAM instructions, reaching such an objective enables attackers to force a given future position. This attack is named '*waypoint injection*': its execution consists of two distinct phases: *reconnaissance* and *exploitation*. During the *reconnaissance* phase, attackers analyse the CAA response to various scenarios. As a result, the attackers obtain the insights needed to characterise an approximate model of its operation. The use of the generated model is twofold. First, it allows attackers to maximise their probability of success by understanding which CAMs are possible. Then, during the *exploitation* phase, attackers can leverage it to calculate the trajectory of a to-be-injected fake target ship, which is subsequently fed to the INS. Such an attack does not need *a-priori* knowledge of either the system's working principle or the algorithm the system relies on, thus it can target the majority of CAA. The following subsections discuss the reconnaissance and exploitation phases in detail.

5.1. Reconnaissance phase

The reconnaissance phase is divided into two steps, the first aimed at identifying the expected behaviour, namely *profiling of baseline conditions*, and the second at characterising the CAA, namely *estimating obstacle detection range and CAA deviation*.

The *profiling of baseline conditions* is the first reconnaissance step, and it aims at establishing a baseline to correctly distinguish between standard navigation and CAMs. For that purpose, attackers monitor

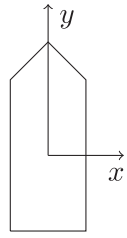


Figure 3. Reference frame.

the ship's cross-track error while underway. This activity allows them to establish the maximum deviation from the straight-line course to the next waypoint under non-collision avoidance conditions. Then, the attackers classify any future course alteration exceeding the found threshold value T_h as belonging to a CAM. Equations (1)–(3) describe a procedure to automatically detect whether an observation performed at instant n of $k+1$ ship positions belongs to the beginning of a CAM. In particular, X, P_s, and P_e are 2D vectors representing positions in cartesian coordinates (see Figure 3 for an illustration of the adopted reference frame), X_i indicates the victim ship position at instant i, $\bigwedge$ is the logical AND, and $(\ldots)_x$ suggests access to the x component of a vector.

Equation (1) defines route elongation function Δ_R in describing the position differential w.r.t. the ideal trajectory conjoining P_s and P_e, respectively the start (P_s) and end (P_e) waypoints of a route segment:

$$\Delta_R(X, P_s, P_e) \coloneqq X - \left[\frac{X \cdot (P_e - P_s)}{(P_e - P_s) \cdot (P_e - P_s)} (P_e - P_s) \right] \quad (1)$$

Leveraging the above-introduced quantity, a port CAM is individuated by the function IsPortCAM defined in Equation (2), checking the following conditions:

(1) the horizontal component of the position monotonically increases.
(2) the horizontal component of the elongation is greater than T_h.

$$\begin{aligned} &\text{IsPortCAM}(X_n, \ldots, X_{n-k}, P_s, P_e) \\ &\quad \coloneqq \left(\bigwedge_{i=n-k+1}^{n} X_{i_x} \geq X_{i-1_x} \right) \wedge [\Delta_R(X_n, P_s, P_e)_x \geq T_h] \end{aligned} \quad (2)$$

Likewise, Equation (3) defines the IsStbCAM condition to identify a starboard turn:

$$\begin{aligned} &\text{IsStbdCAM}(X_n, \ldots, X_{n-k}, P_s, P_e) \\ &\quad \coloneqq \left(\bigwedge_{i=n-k+1}^{n} X_{i_x} \leq X_{i-1_x} \right) \wedge [\Delta_R(X_n, P_s, P_e)_x \leq -T_h] \end{aligned} \quad (3)$$

Once the Attacker has profiled the baseline conditions, the *Estimating obstacle detection range and CAA deviation* phase begins. In accordance with COLREGs rule 8(d) – Action to avoid collision, '*Action taken to avoid collision with another vessel shall be such as to result in passing at a safe distance*'. Therefore, each CAM performed by the CAA must keep a safe distance between vessels. The value of such distance is not mandated to a particular value in the regulations, thus it is treated as a designer-defined function or constant value in the CAA implementation. In addition, the attackers can constrain which

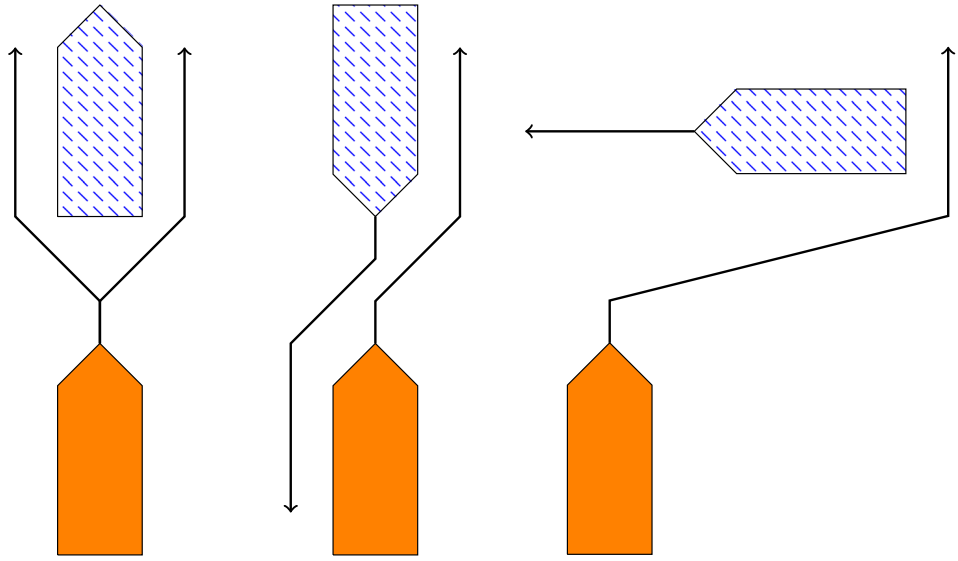


Figure 4. Encounters described by COLREGs rules 13–15. The solid coloured vessel is in a give-way situation (Enevoldsen and Galeazzi 2021, F.1). (a) 13 (Overtaking). (b) 14 (Head on) and (c) 15 (Crossing).

areas are reachable by the attack by measuring how the distance from other vessels influences the ship's route while avoiding collisions. The first problem the attackers face is finding an approximate relation between the CAA desired safe distance SD whenever faced with an obstacle O deemed in a potential collision. This relation is modelled in Equation (4) as an unknown functional λ depending on the vessel position X_v, obstacle position X_o, and their respective speed vectors s_v, s_o:

$$\mathrm{SD}(O) \coloneqq \lambda(X_v, s_v, X_o, s_o) \tag{4}$$

Then, attackers have to estimate which waypoint position $R(O)$ the CAA heuristic ξ generates in response to an obstacle O deemed potentially dangerous. Equation (5) presents a mathematical description of this second problem, describing the unknown relation ξ similarly to $\mathrm{SD}(O)$, subject to the condition that every route waypoint R_i is at least at $\mathrm{SD}(O)$ distance from O:

$$R(O) \coloneqq \xi(X_v, s_v, X_o, s_o) \quad \text{given } \|R_i - X_o\| \geq \mathrm{SD}(O)\ \forall\, i \tag{5}$$

For a generic CAA, an analytical formulation of ξ is unavailable to the attackers. As a result, the malicious actor's effectiveness is directly related to the estimated ξ accuracy. In particular, correctly predicting the codomain of ξ allows attackers to understand which waypoints can be generated during an attack.

5.2. Exploitation phase

The exploitation phase involves planning and executing the rogue ship injection. During the first step, attackers must find where to inject the fake ship to lure the victim ship towards the desired waypoint X_a. Equation (6) summarises the above step:

$$O(X_a) \coloneqq \Gamma(\lambda, \xi) \tag{6}$$

The attackers need to find a heuristic Γ for calculating the properties of a rogue vessel O, dependent on the inferred CAA behaviour λ and ξ so that the target position X_a will belong to the CAM for O.

In the second step, attackers inject the generated vessel O by sending to the INS network rogue NMEA sentences containing AIS messages (Balduzzi et al. 2014).

6. Attack implementation

In this paper, the focus is on three COLREGs-defined conditions: *overtaking*, *head-on*, and *crossing*. All of them are characterised by highly predictable outcomes which an attacker can exploit:

- Rule 13 – Overtaking (Figure 4(a)) applies when one vessel is overtaking another. The overtaking vessel must keep out of the way of the overtaken vessel and avoid crossing ahead of the vessel until it is clear. The overtaken vessel should maintain its course and speed. In this encounter scenario, the Attacker's outcome is to induce a turn to port or starboard.
- Rule 14 – Head on (Figure 4(b)) applies when two vessels are approaching each other in a head-on situation, meaning that the vessels are on courses that would cause them to collide if they continue their present courses. In this situation, both ships must alter their courses to starboard so that they can see each other's red light. In this case, the Attacker can induce a starboard turn.
- Rule 15 – Crossing (Figure 4(c)) applies when two vessels cross each other's paths, meaning they are approaching at an angle that could lead to a collision. The vessel with the other vessel on its starboard side must give way and avoid crossing ahead of the other vessel. The other vessel must maintain its course and speed. The Attacker can leverage this rule to induce a starboard turn.

The implementation of the two phases of the attack, reconnaissance and exploitation, are discussed in detail in the following subsections.

6.1. Reconnaissance implementation

This phase consists of observing the victim ship's behaviour and traffic trajectories between pairs of waypoints provided by the CAA, namely segments.

P_s and P_e identify each segment's starting and ending points. In particular, the attackers overhear the INS network to identify segments and, for each one, determine the speeds and positions of the ship under attack and other obstacles. Then, they classify segments into the following categories.

(1) No traffic has been encountered along the way.
(2) A possible overtaking situation has unfolded during the segment.
(3) A possible head-on situation has unfolded during the segment.
(4) A possible crossing situation has unfolded during the segment.

The segments that do not fall into these categories will be discharged.

During this procedure, the CAA is continuously stimulated by the varying conditions encountered during navigation, resulting in a quasi-Monte Carlo sampling of its responses. Once the number of gathered segments of every category is deemed sufficient (see below for a reasonable heuristic), the attackers can build their adversarial CAA model. In the following, the calculation of T_h (Equations (1)–(3)), automatic classification of segments, end of the reconnaissance phase, the estimation of λ (Equation (4)) and the codomain of ξ (Equation (5)) are described.

To estimate T_h, the attackers calculate for each acquired sample belonging to the first category the point-line distance between the victim ship position X and the $P_e - P_s$ line (as described in Equation (1)). Then, they characterise a Student's t-distribution from the obtained samples. Concerning a Normal distribution, the t-distribution considers the uncertainty in the ship's position X, the error of which is not known to the Attacker. For a set P value, T_h will be the higher bound of the found confidence interval.

Then, the segment classification phase begins. For each segment S, attackers can obtain a vector of ship and obstacle positions and speeds by listening on the INS network. While the classification of no-traffic segments is straightforward, the other three categories require checking whenever some conditions happen during the segment. A segment belongs to a class whenever many of its contained points might individuate a COLREGs scenario. Attackers can repurpose a legitimate collision avoidance COLREGs classification module for this task. For instance, Zaccone et al. (2019) describes a mathematical formulation compatible with this methodology.

Understanding when to stop sampling new segments and start with the exploitation phase is a crucial choice the attackers must make. To that purpose, a simple heuristic is proposed:

(1) A given minimum number N of segments has been gathered for each class.
(2) A to-be added segment would not add significant information to the Attacker's knowledge.

While N will necessarily be an almost arbitrary choice, the information content of a segment can be intuitively derived by observing that:

(1) the more a vessel deviates from its usual route, the more uncommon this route becomes;
(2) segments containing obstacles in odd positions probably exercise unobserved areas of the input space of ξ.

Equation (7) summarises the information content I of a n samples long segment S_n containing n_o obstacles. The information content comprises two components: the *route innovation* and the *traffic innovation*.

$$I(S_n) \coloneqq \underbrace{\frac{\alpha}{n}\sum_{i}^{n}\|X_{v_i} - \tilde{X}_v(X_{v_i})\|}_{\text{Route innovation}} + \underbrace{\frac{\beta}{n}\sum_{i}^{n}\sum_{j}^{n_o}\|X_{o_{i,j}} - \tilde{X}_{o_j}(X_{o_{i,j}})\|}_{\text{Traffic innovation}} \tag{7}$$

Those two elements implement the observations above. Route innovation, weighted by some arbitrary coefficient α, is a sum of vessel position divergences over every element belonging to S. Such divergence is calculated between each vessel position X_{v_i} and the closest past position $\tilde{X}_v$ found in previous segments. Similarly multiplied by a factor β, traffic innovation is a sum over every obstacle in S of obstacle position divergences. Like before, the divergence is calculated between obstacle positions $X_{o_{i,j}}$ and the closest past position $\tilde{X}_{o_j}$ found in previous segments. It is worth noting that this formulation assumes a constant value of n_o during a segment, i.e. that the number of obstacles does not change, a limitation that can be lifted by replacing the second sum upper bound n_o with its time-varying counterpart n_{o_i}. This equation can be calculated efficiently leveraging geospatial indexes (Beckmann et al. 1990) for the $\tilde{X}_{\dots}$ terms. Alternatively, reconnaissance can be stopped once an accuracy metric obtained during the construction of Γ exceeds a given value. For instance, with a neural network regressor (see Section 6.2), a valid stopping condition may be imposed on the value of its loss function.

Eventually, the attackers must estimate λ and the codomain of ξ. λ is estimated to be a function of the victim ship speed. This assumption stems from the COLREGs definition of safe speed. For that purpose, attackers calculate the minimum distance to traffic ships and the speed s at which the victim ship was going at that moment of the closest encounter for every segment where traffic is present. Then, they use the above samples to calculate via least squares regression a slope m and intercept q to define $\lambda = m \cdot |s| + q$ as the safe distance function. Attackers can estimate the codomain of ξ by constructing a vector field of ship positions relative to a space relativised around the route segment. Points with high values indicate positions that can be easily reached during the execution of a CAM. Accordingly, points not reached during any sampled segment will not belong to the codomain. Attackers will choose the possible injected waypoint position X_a from this codomain.

6.2. Exploitation implementation

In this phase, the attackers estimate the ship's properties to be injected, i.e. the shape of Γ, and inject the rogue ship into the INS. For estimating Γ, attackers can leverage a neural network, acting as a universal function approximator (Schölkopf et al. 2001). In particular, the neural network is trained on the data gathered during the reconnaissance phase to output, as a function of the victim ship position and a time delay:

- The heading h_o of the obstacle;
- The velocity vector s_o of the obstacle;
- The lateral distance Δ_x of the obstacle;
- The horizontal distance Δ_y of the obstacle;

Figure 5 represents how the injection of a rogue ship evolves during an attack. In the initial state (see Figure 5(a)), the victim ship navigates towards P_e at a constant speed $\|s\|$. X_a identifies the target position within the codomain of ξ where the attackers want to lure the victim ship. To inject the rogue ship (see Figure 5(b)), the attackers must calculate its starting position X_o. In particular, they can obtain Δ_x, Δ_y by feeding Γ with X_a and predicting the time Δ_t the victim needs to reach X_a. In this article, Δ_t is estimated as $\frac{\|X_a - X\|}{\|s\|}$. The rogue ship starting position X_o can be calculated by adding the obtained deltas to X_a. Attackers start injection of the rogue ship with the initial data obtained from Γ. Then, Γ gets periodically interrogated with updated estimates of Δ_t, yielding different headings, speeds, and target positions for the injected ship. Directly moving

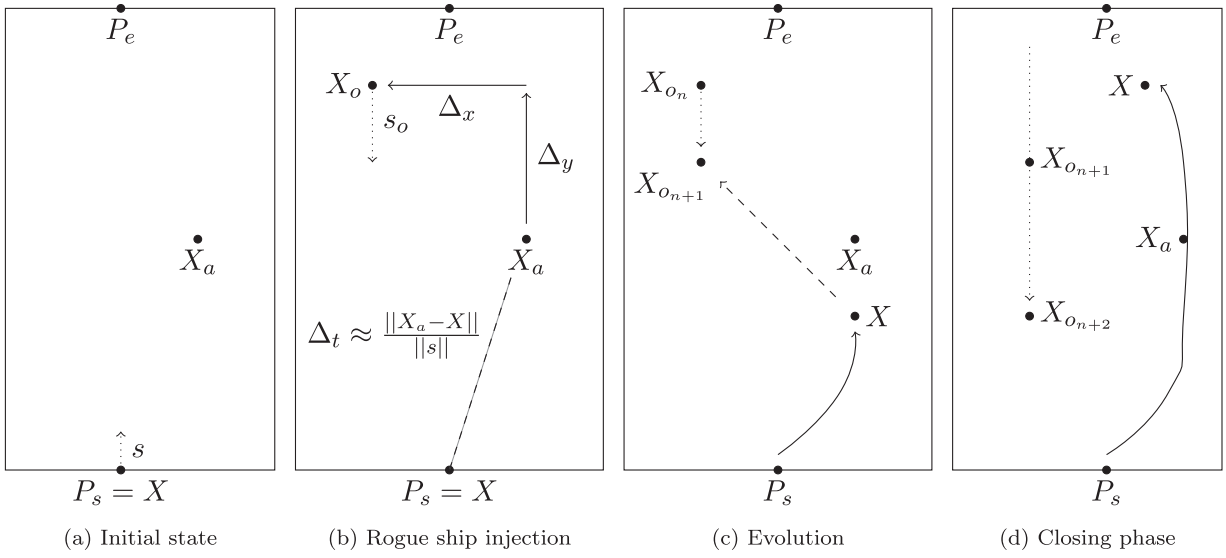


Figure 5. Evolution of the injected ship. (a) Initial state. (b) Rogue ship injection. (c) Evolution and (d) Closing phase.

the rogue ship to these newly acquired setpoints might make suspicious movements, so instead of applying them directly, attackers will steer the injected ship according to a simplified dynamic model as in Figure 5(c). In particular, attackers will evolve the injected vessel by smoothly changing its heading and speed to reach the newly calculated pose. This evolution is continued until the victim ship has reached X_a. Then, the injected ship will continue following its present attitude (see Figure 5(d)) to not raise suspicion by suddenly disappearing.

7. Results

This section elaborates on how the attack implementation has been evaluated. In particular, the simulation scenario is presented, and then the reconnaissance and exploitation phases are simulated.

7.1. Simulation scenario

Figure 6 depicts the simulating scenario used to analyse the attack implementation against the GNC described in Section 3.

The autonomous own-ship starts from P_s to reach a destination P_e located 15 nm ahead. The scenario evolves in discrete timesteps of 5 s, in which the own-ship autonomously manoeuvres (with a max rate of turn of 0.2°/s) along the waypoints generated every 30 s by the CAA. The three COLREGs scenarios have been simulated by placing a traffic ship at a random vertical distance between 2 and 13 nautical miles. For rules 13 (overtaking) and 14 (head-on), traffic was randomly displaced at most at 500 m from the $P_e - P_s$ centre line. Instead, Rule 15 traffic was randomly placed between 500 m and 4500 m on the victim vessel's starboard side. The simulation reproduced 100 no-traffic situations, 335 overtaking, 565 head-on, and 490 crossings. Those last three scenario simulation counts were obtained by simulating an attacker following the heuristic described in Section 5.1, stopping once $I(R)$ went below 100 for more than 5 consecutive samplings using $\alpha = \beta = 1$. According to their capabilities to listen for the INS traffic (see Section 4.2), the attackers recorded the victim and traffic ship trajectories at each timestep of every scenario.

7.2. Reconnaissance experimental results

The recorded simulated trajectories represent the dataset used by the attackers to perform their reconnaissance, as described in Section 6.1. Starting from the estimation of T_h, its value has been calculated on scenarios where no traffic was present, leveraging a value for P of 0.95. From the fitted distribution of distances from the centreline in Figure 7, T_h was set to 57 m.

Since attackers do not have prior knowledge about which COLREGs condition appeared in a scenario, this paper relies on the mathematical model of Zaccone et al. (2019) to identify segments. Table 1 shows the built classifier accuracy.

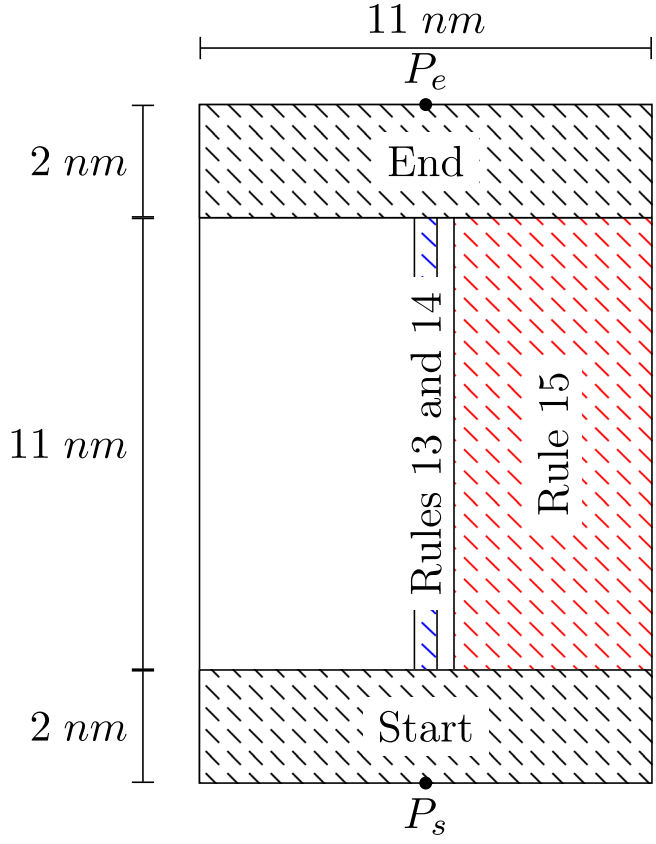


Figure 6. The environment for the experiments.

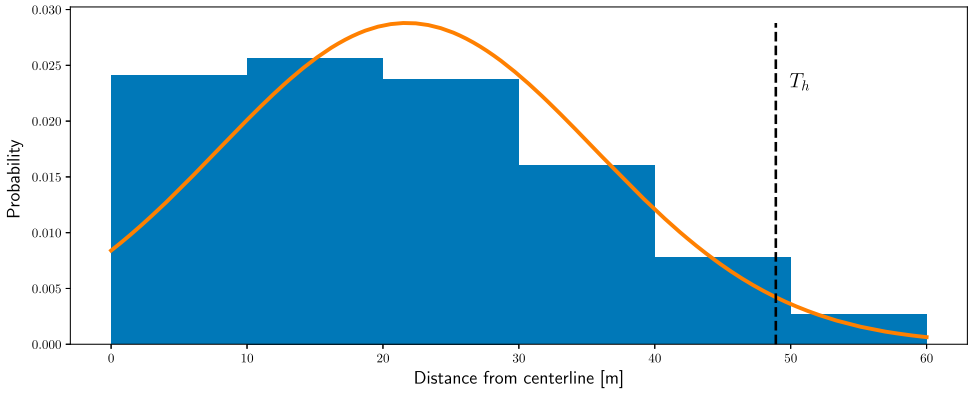


Figure 7. Distribution fitted during the estimation of T_h ($P = 0.95$).

Table 1. Attacker segment classification confusion matrix.

	Rule 13	Rule 14	Rule 15
Rule 13	**99**.7%	0.2%	0.1%
Rule 14	0.4%	**99**.6%	0.0%
Rule 15	0.1%	1.9%	**98**.0%

Overall, the accuracy figures suggest that the attackers can automatically infer the segment classification with high accuracy. However, a slight performance deterioration is seen with the crossing scenario, which could be mistaken for one of the other two for some problematic encounter geometries. Then, to calculate the safe distance function λ, attackers analyse every scenario for which either IsPortCAM or IsStbdCAM returned a favourable verdict. The closest encounter victim-traffic distance and the ships' relative speeds were extracted from these selected segments. Then, two separate linear regressions, one for each speed vector component, were performed, resulting in the formula below:

$$\lambda_x \coloneqq 0.311 s_x + 2197.3 \quad \lambda_y \coloneqq 0.298 s_x + 1993.1 \tag{8}$$

These two similar values show to the attackers that the CAA under analysis does not have significant asymmetries between the two axes. It was probably tuned to keep a minimum safe distance of approximately 2000 m w.r.t each and every obstacle. The next step in the procedure is the estimation of ξ's codomain. The trajectories undertaken by the ship have been discretised in 100×100 m squares and then added together, leveraging the previously described vector field method. Figure 8 depicts the obtained field. Within it, for every y interval, a relative distribution of the x values has been drawn. Even though most of the field has near zero probability due to the vast covered area, attackers can tell at a glance that the collision avoidance module under study strongly favours starboard turns, with a maximum achievable horizontal elongation of around 4000 m.

7.3. Exploitation experimental results

After terminating the reconnaissance phase, a neural network was trained on approximately 5 million data points from the reconnaissance dataset belonging to the scenarios classified as containing a

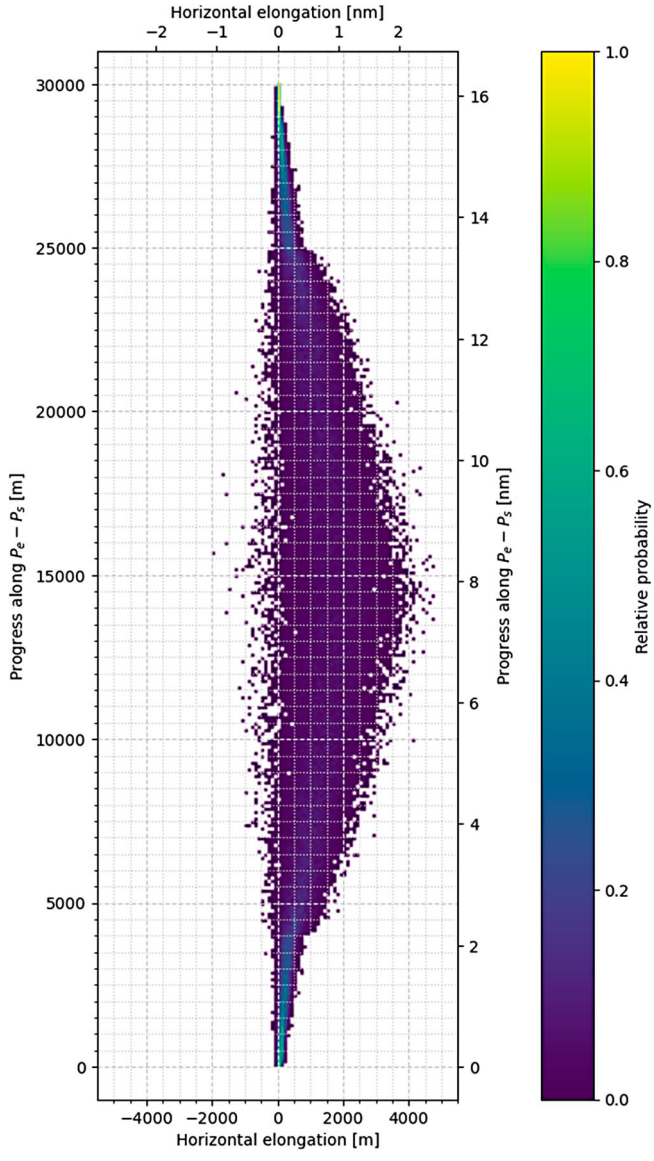


Figure 8. Vector field of observed ξ.

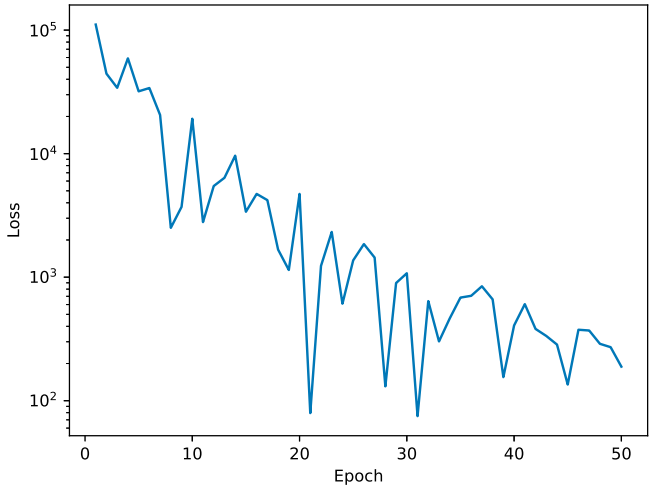


Figure 9. Neural network loss evolution during training.

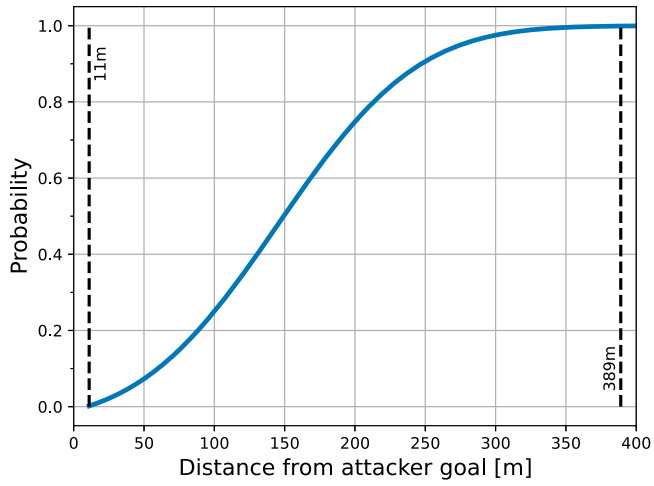


Figure 11. Cumulative distribution of minimum distances between X and X_a.

CAM. Such a neural network contains six hidden layers with 750 neurons each, leveraging a ReLU nonlinearity (Equation (9)).

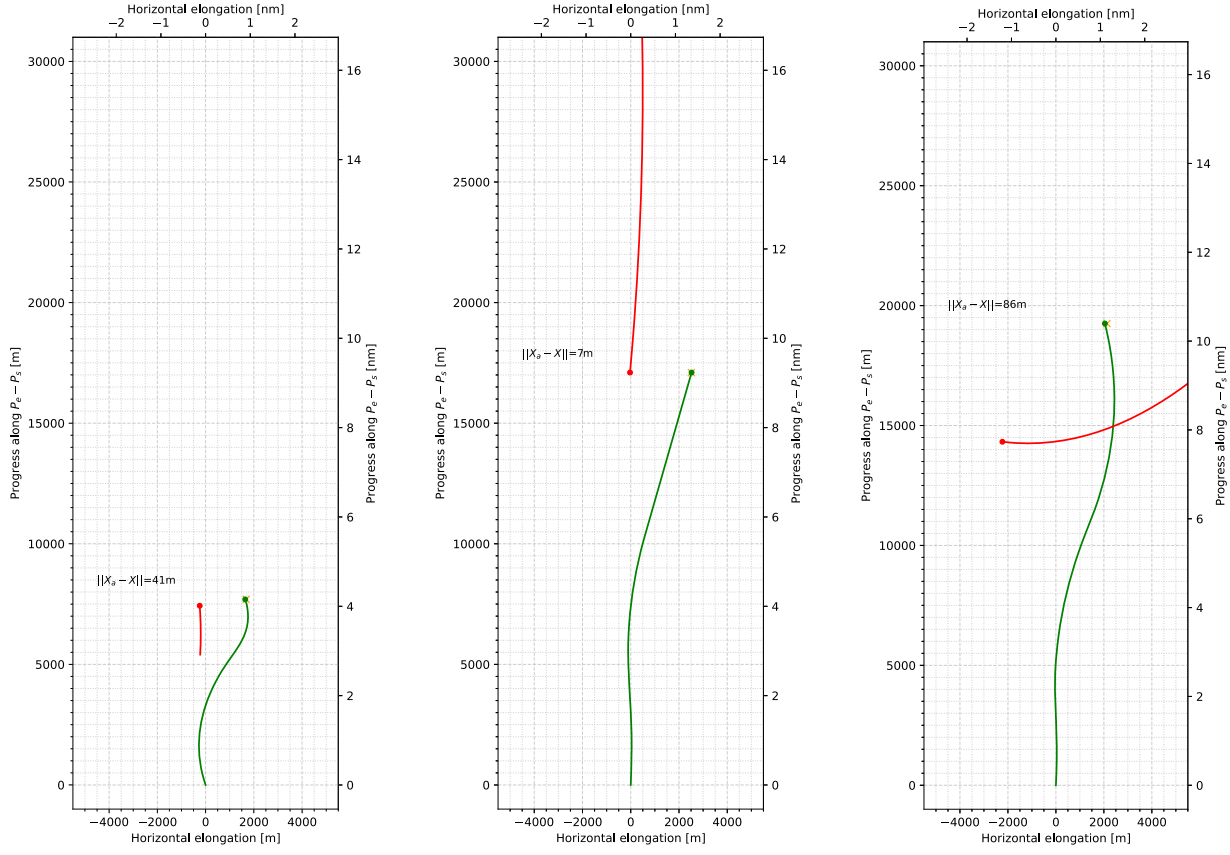
$$\mathrm{ReLU}(x) \coloneqq \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{9}$$

Equation (10) shows the loss function, with $\alpha = \beta = 10$ and $\zeta = \eta = 1$.

$$L(y, \hat{y}) \coloneqq \alpha(y_{h_o} - \hat{y}_{h_o})^2 + \beta(y_{s_o} - \hat{y}_{s_o})^2 + \zeta(y_{\Delta_x} - \hat{y}_{\Delta_x})^2 + \eta(y_{\Delta_y} - \hat{y}_{\Delta_y})^2 \tag{10}$$

Such network was trained for 50 epochs using the Adam (Kingma and Ba 2014) stochastic gradient descent method, Figure 9 summarises the training progress.

To assess the methodology effectiveness, 25 random X_a points in each 100 × 100 m square were picked from within the non-zero vector field values of the codomain of ξ (see Figure 8 for a graphical representation) to perform a waypoint injection attack as described in Section 6.2. Figure 10 depicts three example attack traces, one for each of the COLREGs analysed in the article. Figure 11 shows the resulting minimum distance distribution between X and X_a



(a) Overtaking (b) Head-on (c) Crossing

Figure 10. Three examples of waypoint injection attack traces. The victim ship always starts in (0, 0). Each trace ends with the victim close (<100 m) to the Attacker's goal. (a) Overtaking. (b) Head-on and (c) Crossing.

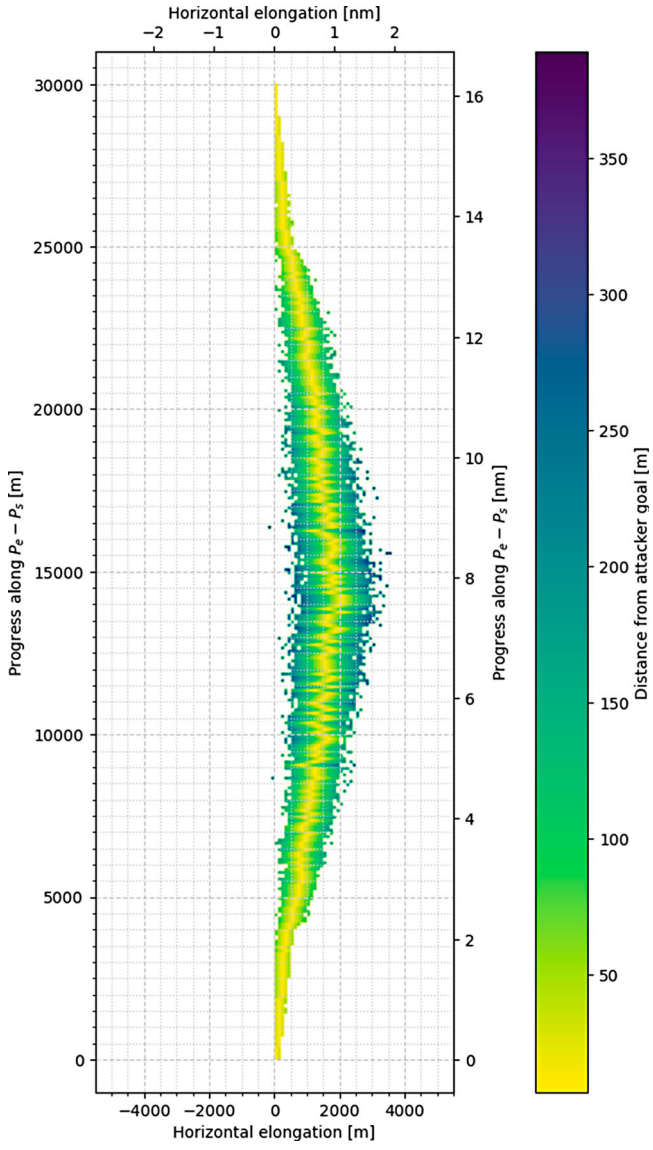


Figure 12. Minimum distances between X and X_a.

across 372,325 experiments. Attackers have been able to lure the attacked GNC to the chosen waypoint with a mean distance of 133 m with a standard deviation of 79 m, while the maximum distance from the experiments results in 389 m. It is worthy of note that the mean distance is comparable with the length of ocean-going ships. The best experiment reached a distance of 11 m w.r.t. the attacker goal. These figures show a clear improvement in the attack performance compared to the article which this work is extending (Longo et al. 2022). Carefully choosing a suitable X_a allows attackers to lure the autonomous ship to the malicious waypoint accurately. Another analysis of the same data is reported in Figure 12, where a heat map of the achieved minimum distance is depicted as a function of position. This visualisation greatly resembles the inferred codomain of ξ, confirming that attackers can better achieve their goals when X_a is chosen close to positions belonging to plausible CAMs.

8. Remediations

The emergence of vulnerabilities like the ones presented in this article highlights the need for a robust risk management strategy driven by the principle of reducing risks to ALARP. Within the constraints of the current regulatory framework, this principle obliges operators to recognise the inherent risks and develop targeted countermeasures. To that effort, mitigations in this section will be introduced in descending order of their practical applicability in real-world scenarios, starting with the most feasible and ending with the more theoretical or complex solutions.

The first proposed remediation is relying on redundant data sources. The vulnerability in question arises from the potential for malicious actors to manipulate AIS transponder data, coupled with the GNC's susceptibility to ingest this falsified data for computing its CAMs. A possible mitigation strategy involves leveraging alternative sensor arrangements, including radars, cameras, and data fusion technologies. This approach would necessitate a considerably coordinated effort for attackers to compromise all sensor systems simultaneously, thus raising the barrier to successful exploitation.

Continuously monitoring the ship's behaviour and comparing it to expected navigation patterns can help further mitigate the vulnerability. Monitoring involves implementing solutions and continuously collecting data from AIS and other ship systems. Such solutions can be provided onboard or remotely by transmitting data to ROCs, i.e. specialised onshore facilities. Collected data can be correlated or compared with expected patterns to detect unusual inconsistencies or behaviours, triggering immediate alerts. Automated responses can be configured to isolate or nullify suspicious data. In addition, ROCs can offer enhanced insight into anomalies by cross-referencing data from vessels within the same fleet. They also enable human supervision of autonomous ship operations and the capability to intervene and override automated system decisions when necessary. AI represents an established technique for implementing behavioural analysis and anomaly detection and finds extensive application within the maritime domain (Lane et al. 2010; Coleman et al. 2020; Wolsing et al. 2022).

Reducing the risk associated with cyber attacks such as the one presented in this article can also be tackled by preventing vessels from being compromised and reducing the impact of successful exploits. On the prevention side, where feasible, adherence to relevant international regulations and industry best practices can reduce the likelihood of compromise. To this end, the regulations and guidelines from the International Maritime Organization (2021), International Association of Classification Societies (2022), BIMCO (2021), Sungbaek et al. (2022), DNV (2021), and Bureau Veritas (2023) play a primary role. Still, given the criticality of the involved systems and impossibility, it is imperative for human operators to be equipped with post-exploitation measures. In particular, operators should devise robust contingency plans and integrate them within their operating procedures. Moreover, seafarers should engage in regular emergency drills oriented at tacking failures in the highly automated systems found on the vessel they are sailing or remotely managing via ROCs.

Eventually, data authentication and integrity verification could be considered. In its current form, the attack leverages the unauthenticated nature of how the AIS transponder system operates and the missing integrity verification of data flowing within the INS. The AIS system should be fitted with a means to provide authentication by leveraging countermeasures providing data authentications such as the ones proposed by Kessler (2020), Goudosis and Katsikas (2022), and Wimpenny et al. (2022). Nevertheless, adding a cryptographic layer might demand significant efforts, making it impractical for existing systems due to recertification and integration requirements.

9. Conclusion

This paper presented an extensive analysis of cyber attacks to collision avoidance modules showing enhanced attacker capabilities. The developed system successfully misled the collision avoidance

algorithm and lured it to the desired injected waypoint. As a significant result, the paper demonstrated that, by refining the logic involved in the attack, an attacker could hijack the autonomous ship's trajectory with increased precision and real-world applicability. The experimental scenarios have shown the effectiveness and impact of these waypoint injection attacks. Furthermore, the paper described some actionable topics to reduce the impact of the presented weaknesses. With open access to navigation data and the availability of commodity software-defined radio transponders, these exploits might be carried out inexpensively and leverage entirely open information sources.

The authors hope this article sheds light on these systems' potential risks, inspiring future practitioners to develop mitigations and even more robust systems. Future research should also analyse the possibility of allowing collision avoidance modules to communicate with each other to coordinate and ensure that the obstacle being avoided exists. Lastly, developing and deploying autonomous vehicle guidance systems must establish a robust international regulatory framework encompassing comprehensive certification procedures and detailed testing plans to ensure their safe and reliable operation in real-world scenarios.

Acronyms

AI Artificial Intelligence. 5, 50, 52
AIS Automatic Identification System. 5, 9, 10, 17, 50
ALARP 'as low as reasonably practicable'. 50
APT Advanced Persistent Threat. 5
ARPA Automatic Radar Plotting Aid. 5
CAA Collision Avoidance Algorithm. 4, 5, 9–11, 15–18, 29, 35
CAM Collision Avoidance Maneuver. 9–11, 14, 15, 17, 38, 47, 50
COLREGs Convention on the International Regulations for Preventing Collisions at Sea. 4–6, 9, 10, 15, 17–19, 22, 29, 32, 41
EPFS Electronic Position Fixing System. 9, 10
ETA Estimated Time of Arrival. 6
GAN Generative Adversarial Network. 5
GNC Guidance Navigation and Control. 5, 9, 10, 26, 44, 50
IMO International Maritime Organization. 3, 4
INS Integrated Navigation System. 5, 9, 10, 17, 19, 29, 50
ISAR Inverse Synthetic Aperture Radar. 5
MaCRA Maritime Cyber Risk Assessment framework. 3, 10
MCR Maximum Continuous Rating. 6
ROC Remote Operation Centers. 50
RRT Rapidly exploring Random Tree. 5
SDME Speed and Distance Measuring Equipment. 9, 10

Acknowledgments

It was carried out while Giacomo Longo was enrolled in the Italian National Doctorate on AI run by the Sapienza University of Rome in collaboration with the University of Genova.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was partially funded by the COMPASS lab of the University of Genova and the NextGenerationEU project 'Security and Rights in CyberSpace' (SERICS).

ORCID

G. Longo http://orcid.org/0000-0003-0025-7191

References

Balduzzi M, Pasta A, Wilhoit K. 2014. A security evaluation of AIS automated identification system. In: Proceedings of the 30th Annual Computer Security Applications Conference (ACSAC '14); New York, NY, USA. Association for Computing Machinery. p. 436–445.

Beckmann N, Kriegel HP, Schneider R, Seeger B. 1990. The R∗-tree: an efficient and robust access method for points and rectangles. In: Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data, Atlantic City, New Jersey, USA. p. 322–331.

Bhatti J, Humphreys TE. 2017. Hostile control of ships via false GPS signals: demonstration and detection. NAVIGATION. 64(1):51–66. doi: 10.1002/navi.183.

BIMCO. 2021. The guidelines on cyber security onboard ships. Report No.: Version 4. Available from: https://u.garr.it/RODcs.

Bolbot V, Theotokatos G, Boulougouris E, Vassalos D. 2020. A novel cyber-risk assessment method for ship systems. Saf Sci. 131:104908. doi: 10.1016/j.ssci.2020.104908.

Bureau Veritas. 2023. Rules on cyber security for the classification of marine units. NR659 R02. Available from: https://iacs.org.uk/resolutions/recommendations/161-180/rec-166-new-corr2-cln.

Coleman J, Kandah F, Huber B. 2020. Behavioral model anomaly detection in automatic identification systems (AIS). In: 2020 10th Annual Computing and Communication Workshop and Conference (CCWC). IEEE. p. 0481–0487.

DNV. 2021. DNV GL - Rules For Classification - Ships - Part 6 "Additional class notations" - Chapter 5 "Equipment and design features" - section 21 "Cyber Security".

Dyryavyy Y. 2014. Preparing for cyber battleships – electronic chart display and information system security. NCC Group. Available from: https://u.garr.it/ctabM.

Enevoldsen TT, Galeazzi R. 2021. Grounding-aware RRT∗ for path planning and safe navigation of marine crafts in confined waters. IFAC-PapersOnLine. 54(16):195–201. 13th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2021. doi: 10.1016/j.ifacol.2021.10.093.

Goudosis A, Katsikas S. 2022. Secure automatic identification system (SecAIS): proof-of-concept implementation. J Mar Sci Eng. 10(6):805. doi: 10.3390/jmse10060805.

Hemminghaus C, Bauer J, Padilla E. 2021. BRAT: a BRidge attack tool for cyber security assessments of maritime systems. TransNav: Int J Mar Navig Saf Sea Transp. 15(1):35–44. doi: 10.12716/1001.15.01.02.

International Association of Classification Societies. 2022. Recommendation on cyber resilience. Rec 166 Corr.2. Available from: https://iacs.org.uk/resolutions/recommendations/161-180/rec-166-new-corr2-cln.

International Electrotechnical Commission. 2016. 61162-1 Maritime navigation and radiocommunication equipment and systems-digital interfaces-part 1: single talker and multiple listeners.

International Maritime Organization. 1972. Convention on the international regulations for preventing collisions at sea.

International Maritime Organization. 2007. Adoption of the revised performance standards for integrated navigation systems (INS). MSC.252(83).

International Maritime Organization. 2018. Framework for the regulatory scoping exercise for the use of maritime autonomous surface ships (MASS). MSC.100/20.

International Maritime Organization. 2020. SOLAS 2020 consolidated edition. IMO-publication.

International Maritime Organization. 2021. Guidelines on maritime cyber risk management. MSC-FAL.1/Circ.3/Rev.1. Available from: https://wwwcdn.imo.org/localresources/en/OurWork/Facilitation/Facilitation/MSC-FAL.1-Circ.3-Rev.1%20-%20Guidelines%20On%20Maritime%20Cyber%20Risk%20Management%20(Secretariat).pdf.

International Telecommunication Union. 2014. Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile frequency band. M.1371-5.

Katsikas SK. 2017. Cyber security of the autonomous ship. In: Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security, Abu Dhabi, United Arab Emirates. p. 55–56.

Kavallieratos G, Katsikas S. 2020. Managing cyber security risks of the cyber-enabled ship. J Mar Sci Eng. 8(10):768. doi: 10.3390/jmse8100768.

Kavallieratos G, Katsikas S, Gkioulos V. 2018. Cyber-attacks against the autonomous ship. In: Computer security. Guildford: Springer; p. 20–36.

Kessler G. 2020. Protected AIS: a demonstration of capability scheme to provide authentication and message integrity. TransNav: Int J Mar Navig Saf Sea Transp. 14(2):279–286. doi: 10.12716/1001.

Kim M, Joung TH, Jeong B, Park HS. 2020. Autonomous shipping and its impact on regulations, technologies, and industries. J Int Marit Saf Environ Aff Shipp. 4(2):17–25.

Kingma DP, Ba J. 2014. ADAM: a method for stochastic optimization. Preprint. arXiv:14126980.

Lane RO, Nevell DA, Hayward SD, Beaney TW. 2010. Maritime anomaly detection and threat assessment. In: 2010 13th International Conference on Information Fusion. IEEE. p. 1–8.

Li M, Mou J, Chen L, Huang Y, Chen P. 2021. Comparison between the collision avoidance decision-making in theoretical research and navigation practices. Ocean Eng. 228:108881. doi: 10.1016/j.oceaneng.2021.108881.

Longo G, Martelli M, Russo E, Zaccone R. 2022. Collision-avoidance capabilities reduction after a cyber-attack to the navigation sensors. In: Proceedings of the International Ship Control Systems Symposium. Available from: http://library.imarest.org/record/10729.

Longo G, Orlich A, Musante S, Merlo A, Russo E. 2023. MaCySTe: a virtual testbed for maritime cybersecurity. SoftwareX. 23:101426. doi: 10.1016/j.softx.2023.101426.

Longo G, Russo E, Armando A, Merlo A. 2023. Attacking (and defending) the maritime radar system. IEEE Trans Inf Forensics Sec. 18:3575–3589. doi: 10.1109/TIFS.2023.3282132.

Lund MS, Gulland JE, Hareide OS, Jøsok ø, Weum KOC. 2018. Integrity of integrated navigation systems. In: 2018 IEEE Conference on Communications and Network Security (CNS), Beijing, China. p. 1–5.

Lund MS, Hareide OS, Jøsok Ø. 2018. An attack on an integrated navigation system. NECESSE Royal Norw Naval Acad Monogr Ser. 3(2):149–163. doi: 10.21339/2464-353x.3.2.149.

Meland P, Bernsmed K, Wille E, Rødseth Ø, Nesheim D. 2021. A retrospective analysis of maritime cyber security incidents. TransNav: Int J Mar Navig Saf Sea Transp. 15. doi: 10.12716/1001.

Meucci G, Karahoda B, Oveis AH, Mancuso F, Jajaga E, Cantelli-Forti A. 2023. Naval cybersecurity in the age of AI: deceptive ISAR images generation with GANs. doi: 10.36227/techrxiv.23709846.v1.

Misas J, Hopcraft R, Tam K. 2022. Future of maritime autonomy: cybersecurity, trust and mariner's situational awareness. In: Proceedings of the International Ship Control Systems Symposium; 8th–10th Nov 2022.

Mitre Corporation. 2023. MITRE ATT&CK – supply chain compromise. Available from: https://attack.mitre.org/techniques/T1195/.

Parasuraman R, Sheridan TB, Wickens CD. 2000. A model for types and levels of human interaction with automation. IEEE Trans Syst Man Cybern A: Syst Hum. 30(3):286–297. doi: 10.1109/3468.844354.

Schölkopf B, Herbrich R, Smola AJ. 2001. A generalized representer theorem. In: Computational Learning Theory: 14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001; Amsterdam, The Netherlands; July 16–19, 2001; Proceedings 14. Springer. p. 416–426.

Shadrin SS, Ivanova AA. 2019. Analytical review of standard SAE j3016 «taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles» with latest updates. Avtomobil' Doroga Infrastruktura. 3(21):10.

Sungbaek C, Erwin O, Gabor V, Vasco P. 2022. Cybersecurity considerations in autonomous ships. Tallin: NATO Cooperative Cyber Defence Centre of Excellence.

Svilicic B, Brčić D, Žuškin S, Kalebić D. 2019. Raising awareness on cyber security of ECDIS. TransNav: Int J Mar Navig Saf Sea Transp. 13(1). doi: 10.12716/1001.

Tam K, Jones K. 2019. MaCRA: a model-based framework for maritime cyber-risk assessment. WMU J Marit Aff. 18(1):129–163. doi: 10.1007/s13437-019-00162-2.

Voicu M. 2016. Six levels of ship autonomy, according to lloyd's register. Rev Rom Drept Mar: 85.

Walter MJ, Barrett A, Walker DJ, Tam K. 2023. Adversarial AI testcases for maritime autonomous systems. AI Comput Sci Robot Technol. 2. doi: 10.5772/acrt.15.

Wimpenny G, Šafář J, Grant A, Bransby M. 2022. Securing the automatic identification system (AIS): using public key cryptography to prevent spoofing whilst retaining backwards compatibility. J Navig. 75(2):333–345. doi: 10.1017/S0373463321000837.

Wolsing K, Roepert L, Bauer J, Wehrle K. 2022. Anomaly detection in maritime AIS tracks: a review of recent approaches. J Mar Sci Eng. 10(1):112. doi: 10.3390/jmse10010112.

Zaccone R. 2021. COLREG-compliant optimal path planning for real-time guidance and control of autonomous ships. J Mar Sci Eng. 9(4). doi: 10.3390/jmse9040405.

Zaccone R, Martelli M. 2020. A collision avoidance algorithm for ship guidance applications. J Mar Eng Technol. 19(sup1):62–75. doi: 10.1080/20464177.2019.1685836.

Zaccone R, Martelli M, Figari M. 2019. A COLREG-compliant ship collision avoidance algorithm. In: IEEE European Control Conference – ECC2019, Naples, Italy. p. 2530–2535.