Communication and Cybersecurity Testbed for Autonomous Passenger Ship

Ahmed Amro and Vasileios Gkioulos

Norwegian University of Science and Technology, Gjøvik, Norway
ahmed.amro@ntnu.no; vasileios.gkioulos@ntnu.no

Abstract. Many industrial sectors are undergoing a digital transformation, including maritime. New technological advancements and modes of operations are being introduced to maritime infrastructure, which includes ships, ports, and other facilities. Digital transformation in maritime has among its goals reducing human involvement and improving remote connectivity. The achievement of these goals hinges on several components, including communication technologies and cybersecurity. Consequently, maritime-related communication and cybersecurity solutions are in high demand. This paper targets the development of a maritime-themed testbed utilized to evaluate and analyze several maritime use cases, including autonomous passenger ships (APS) with a prime focus on the communication and cybersecurity aspects. We have proposed abstraction of processes guiding the utilization of the testbed capabilities. Also, we proposed an approach for replicating the target system of analysis which facilitates the analysis and evaluation activities. The proposed testbed and its processes have been evaluated by discussing some of the projects that utilized it, including evaluating communication and cybersecurity architectures for an APS use case. Additionally, after comparison with the state-of-the-art in cybersecurity testbeds, the testbed was found to be supporting the majority of the concepts and properties observed in the literature while the missing elements were highlighted and designated as suggestions for future work. Moreover. we provide a discussion of the challenges in cybersecurity evaluation in maritime in general and autonomous ships in particular.

Keywords: cybersecurity · communication · testbed · autonomous passenger ship · ICS

1 Introduction

In the modern era, technological advancements are enriching several aspects of our lives. Innovations in the maritime domain have found their application in passenger transportation in inland waterways. Several projects are undergoing aiming to develop autonomous passenger ships or ferries in three regions in Norway [6] including a project named Autoferry which aims to develop an Autonomous all-electric Passenger Ship (APS) for inland water transport in the city of Trondheim [2]. The new APS operates within a new operational mode called

autoremote, this entails that the APS will be mainly autonomous, with human supervision from a remote control center (RCC) [9]. Although this unconventional mode of operation is expected to improve the provisioning of navigational services, relies on a group of interconnected Industrial Control Systems (ICS). Such ICS requires the support of various communication technologies as well as it introduces a wide range of cyber threats with possible safety impacts.

Communication and cybersecurity are considered among the biggest challenges for the advancement of the autonomous shipping concept [9]. This is based on the fact that improper communication is the main factor for maritime casualties [1] and cybersecurity has been considered among the most significant challenges in the usage of unmanned ships according to seafarers [23]. Therefore, there is a growing interest in the development of communication and cybersecurity-related solutions for autonomous ships. Cyber ranges and testbeds are commonly utilized for the evaluation of the developed solution as well as for training and awareness [27, 26]. However, during this study, we have observed a lack in the literature regarding the utility of cyber ranges or testbeds for the evaluation of cybersecurity solutions in the maritime domain in general and in autonomous shipping in particular. In the remainder of this paper, we use the terms cyber range and testbed interchangeably.

This paper proposes a testbed suitable for the analysis and evaluation of several maritime use cases focusing on cybersecurity and communication aspects. Initially, a literature review is conducted to identify relevant artifacts and approaches utilized in similar testbeds. Then the testbed is developed following the ISO 15288 standard [17]. Finally, the identified state-of-the-art is utilized to evaluate the testbed focusing on the comprehensiveness and utility of the included capabilities. Our contributions in this work can be summarised as follow:

- We propose a communication and cybersecurity testbed for several maritime use cases. The testbed capabilities are comprehensive compared to the state-of-the-art and provide a novel introduction for such testbed in the maritime domain.
- We propose an abstraction of three processes that can be followed during the utilization of cybersecurity testbeds namely, system replication, system analysis, and technical management.
- We propose an approach for the system replication process based on standardized system elements. The system elements can be utilized as guidelines for replicating the target system for analysis.

2 Background and Related Work

In this section, we provide a brief background regarding the motivation for this study as well as several relevant works regarding cybersecurity testbeds in general and in maritime in particular. Regarding the motivation, the testbed proposed in this paper is mainly developed to evaluate artifacts that were designed based on a group of established communication and cybersecurity requirements for an autonomous passenger ship or ferry (APS). The requirements were collected from

several APS stakeholders, analyzed, and adopted in our earlier work [9]. The communication requirements were utilized to define and design a communication architecture for the APS that allows it to communicate with its operational context and support several navigational services such as autonomous navigation and autonomous engine monitoring and control [10]. On the other hand, the cybersecurity requirements in addition to a group of risk analysis processes for the APS as a cyber physical system [8, 11] were utilized to define and design a cybersecurity architecture for the APS [7]. Additionally, the testbed capabilities enable the exploration of additional use cases allowing the advancement of cybersecurity research in maritime. Moreover, the testbed is evaluated using qualitative functional evaluation and through comparison with the state-of-the-art. The captured state-of-the-art of cybersecurity testbeds relies on the works summarized in the remainder of this section since a comprehensive literature survey is outside the scope of this paper.

Yamin et al [27] conducted a systematic literature survey (SLR) and presented the state-of-the-art in cyber ranges and cybersecurity testbeds by highlighting several aspects such as environment building, scenarios, monitoring, learning, teaming, and management. Moreover, the authors discussed the observed approaches for testbed evaluation. We mapped our testbed capabilities, processes, and evaluation based on the artifacts highlighted in this work.

Kavak et al [19] surveyed several works and presented the state-of-the-art related to the utility of simulation in the cybersecurity domain. The authors have highlighted the efforts observed in the literature during the construction of the testing environment which is referred to as "Representative environment building" and the utility of both physical equipment as well as virtual equipment in both simulating or emulating cyber exercises in security evaluation and testing.

Tam et al [26] have discussed the concept of cyber ranges in the maritime context. The authors aimed to enhance the state-of-the-art by discussing cyber ranges in a maritime context, scalability, and the coordination of cyber ranges (i.e. federation). Regarding inserting the maritime context into cyber ranges, the authors have presented a layer representation of ships and ports components in maritime to aid the development of cyber ranges. This demonstrates the utility of the concept of facilities in cyber ranges in maritime, which refers to the separation of the different arrangement of components based on their geographical location or functionality. In autonomous and remotely operated vessels this is also relevant due to the interoperability among the vessels and shore facilities. Regarding scalability, the authors have discussed the utilization of both simulation/emulation components in addition to real equipment in an attempt to maintain a balance between cost, scalability, repeatability, and realism. In this paper, we follow the same approach. Finally, the authors have highlighted the utility of cyber ranges for generating data that can be used to enhance other processes such as risk assessment and machine learning algorithms. A notion which we have adopted as well and have applied during one of the projects that utilize the testbed (Section 4.2).

3 Testbed Architecture

The testbed is aimed to include a group of capabilities that allow the analysis and evaluation of design and implementation artifacts for several maritime use cases focusing on communication and cybersecurity aspects. These use cases currently include an autonomous passenger ship and traditional integrated bridge systems. Considering the undergoing digitalization in maritime, the testbed is aimed to have a flexible design in order to accommodate several traditional and futuristic ship models and operational modes. The testbed model is a hybrid; consisting of both physical and virtual components. Moreover, the testbed provides both remote and on-site testing capabilities in addition to having a mobility feature.

3.1 Concepts and processes

Fig. 1 reflects a view of the testbed processes. It includes three main processes inspired from the ISO 15288 standard [17], namely, system replication, system analysis, and technical management.

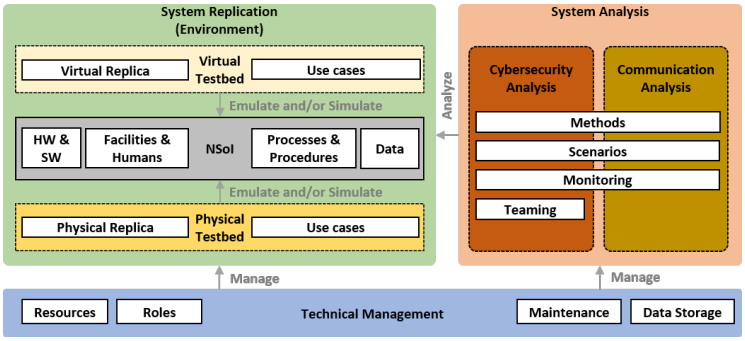

Fig. 1. Process view of the testbed

System Replication: also referred to as "Representative environment building" [19] during this process, the Narrowest System of Interest (NSoI) is constructed utilizing physical and/or virtual components emulating and or simulating the real system under investigation. The system description is intended to be comprehensive to facilitate the system analysis process. The ISO 15288 standard [17] details the different system elements that can describe the manner in which a system is configured. As a guideline for capturing each NSoI, we propose using this system element abstraction. The outcome of this process is a constructed replica of the NSoI as well as an architecture description of it. The different system elements and their replication mechanisms are depicted in Table 1.

Table 1. Replication mechanisms for the different system elements

System Element	Replication Mechanism	Example
Hardware	- Simulation/Emulation tool - Physical equipment	Automatic Identification System (AIS) replicated using physical equipment or a AIS simulator software
Software	- Tool	*OpenCPN* chart plotter software
Data	- Simulation/Emulation tool - Physical equipment - Traffic generation tools (e.g. stubs, fuzzing, replay)	Captured sensor data (e.g. lidar) transmitted through a traffic generation tool (e.g. *Tcpreplay*)
Humans	- Human - User behavior generation tool	A Remote operator role emulated using a human or a user behavior generation tool.
Processes, and Procedures	- Scenarios - Tools - Physical equipment - Human - User behavior generation tool - Facilities	Ship-to-Ship communication emulated using a group of physical equipment with relevant technology (e.g. VHF), people at another ship (i.e. facility), following a certain scenario for collision avoidance.
Facilities	- Physical location - Arrangement of physical equipment and tools	sites 1 and 2 shown in Fig.2

The use of simulation and emulation in cybersecurity testbeds and exercises is widely common as indicated in the literature [27, 19, 26]. Such tools can be utilized to replicate several system elements such as hardware or data streams. Yamin et al. [27] highlighted the utilization of traffic generation and behavior generation tools. The traffic generation tools are utilized for generating realistic data streams for creating different attack and normal operational scenarios while the user behavior generation tools are utilized to emulate human behavior. Additionally, Tam et al [26] have highlighted the different types of data generated in cyber ranges, particularly, data needed to meet minimum requirements and allow services to function (i.e. stubs), data simulating all types of input to systems without applying logic (i.e. fuzzing), more realistic data based on simulation, and date that is replayed after being captured. Our testbed aims to provide data replication capabilities based on the data generation mechanisms discussed in [27, 26] and focus on data streams that are relevant to the maritime domain. Therefore, several tools and physical equipment from the maritime domain are included (Tables 2 and 3 respectively).

Additionally, several maritime processes and procedures are addressed including the different communication functions specified in the APS communication architecture [10], namely, Ship-to-Shore, Ship-to-Ship, and Internal Communication. Ship-to-Shore communication targets the communication links between the ship and the shore for remote monitoring, control, and maintenance. Ship-to-Ship communication focuses on the communication channels between the ship and other ships for safe navigation. Internal communication focuses on the communication between internal ship systems. The ship systems include Information Technology (IT) as well as Operational Technology (OT). Examples of such systems are control servers (e.g Dynamic Positioning System), and Programmable Logic Controllers (PLC) for controlling several safety systems. More details can be found in our earlier work [10]. Moreover, the representation of system's facilities in maritime has been observed to provide improved system analysis capabilities.

Materials and naturally occurring entities are other physical system elements discussed in the ISO 15288 standard [17]. Nevertheless, they have been found to be irrelevant to the current objectives of our testbed as the later focuses on cybersecurity and communication aspects of maritime use cases.

System Analysis this process consists of a group of activities to analyze the constructed replica of the NSoI. In our testbed, the system analysis can follow two main directions, communication analysis, and cybersecurity analysis. For each direction, different aspects are specified, namely, methods, scenarios, monitoring, and teaming. Brief discussion for each aspect is provided below:

- **Methods**: Several methods for communication analysis are observed in the literature such as wireless coverage analysis [18] and performance analysis [22]. On the other hand, cybersecurity analysis methods include; among others, risk assessment, adversary emulation, and evaluation of security solutions [7]. Additionally, the cybersecurity analysis approaches; depending on the use case under analysis, can be conducted using black box, grey box, or white box analysis techniques [20].
- **Scenarios**: a scenario describes the storyline which specifies the steps for conducting a test or training exercise [27]. Scenario definitions should include a purpose, environment, storyline, type, domain, and tools. For the cybersecurity analysis, scenario types should include both normal operation scenarios (e.g. navigational scenario) as well as attack scenarios.
- **Monitoring**: this includes the methods, tools, and focus of the real-time monitoring of the exercise. In our testbed, this is mostly related to documentation and data collection. Network traffic capture, screen capture, and manual documentation are among the supported monitoring methods.
- **Teaming**: Cybersecurity analysis can be conducted through the utilizing of the concept of teaming. Several teaming formations have been observed in the literature including red teams conducting offensive security testing, blue teams conducting defensive security, white teams responsible for scenario creation, green teams involved in monitoring the scenarios, and autonomous teams utilized for automating the roles of other teams [27]. Additionally, a recent teaming concept, namely purple teaming [24], integrates the activities of red and blue teams extending the exercises toward further evaluation and improvement of the security posture of the target system. In our testbed, we aim to include several formations of such teams within different cybersecurity operations, namely, offensive security, defensive security, and offensive defense. Moreover, these cybersecurity operations are supported by white teams and autonomous teams for creating and automating the analysis process.
  - **Offensive Security**: This includes the identification and implementation of attack scenarios within the testbed components by conducting various penetration testing activities (i.e red team activities). The $ATT\&CK$ framework [25] is utilized to structure and formalize the description of these activities. $ATT\&CK$ was chosen based on our earlier

works [8, 7] due to its comprehensive threat model and updated common knowledge. Additionally, the utility of the ICS matrix in $ATT\&CK$ has been demonstrated in our earlier work [8] and resulted in several ICS specific attack scenarios which are target for analysis in our testbed. For instance, the manipulation of view [5] and denial of view [3] are two identified attack techniques with considerable risk against the APS system. Their risk is being evaluated in one of the project utilizing the testbed (refer to Section 4.2). The testbed provides capabilities to conduct attack techniques across the different cyber kill chain phases, including; among others, reconnaissance, initial access, discovery, impair process control, and inhibit response function. Performing these activities within the maritime context is expected to identify and evaluate novel and relevant attack techniques.

- **Defensive Security**: This includes the identification and implementation of defensive capabilities within the testbed (i.e. blue team activities). The NIST framework as well as the defense-in-depth strategies are both considered for mapping and updating the defensive capabilities to facilitate defensive operations. For instance, the testbed includes defensive capabilities allowing for threat identification, protection, and detection as well as capabilities for incident response and recovery from cyber-attacks. The choice for NIST and defense-in-depth is based on our previous work [7] which identified both among the most referenced risk management strategies. Performing these activities within the maritime context is expected to identify and evaluate novel and relevant defensive capabilities.
- **Offensive Defense**: This includes the implementation and analysis of the purple teaming concept in which red team and blue team activities are intertwined toward improving the security posture of a target system [24]. To the best of our knowledge, the introduction of this concept in the maritime domain is novel.

The outcome of this process is data and information for understanding the technical aspects of the NSoI. This allows for informed decision-making regarding the system development throughout its life cycle as well as support research activities in maritime communication and cybersecurity.

Technical Management This process includes several management activities related to both the system replication and the system analysis processes for each project (i.e. test), such as; among others, resource management, maintenance, role management, and data storage. Brief discussion for each activity is provided below:

– **Resource Management**: this entails the identification and allocation of computational resources (e.g. memory), disk storage, and required components for conducting tests [27].

– **Role Management**: this entails the specification and distribution of roles during the different tests. For instance, during an attack scenario targeting a certain navigational operation, an attacker role is expected as well as a navigational role (e.g. Officer on Watch OOW).
– **Maintenance**: management of the testbed equipment such as inventory, licensing, and support.
– **Data Storage**: the management of any data related to the testbed. This includes the generated data during the analysis process, the different software binaries as well as backups of the different devices.

3.2 Tools and Equipment

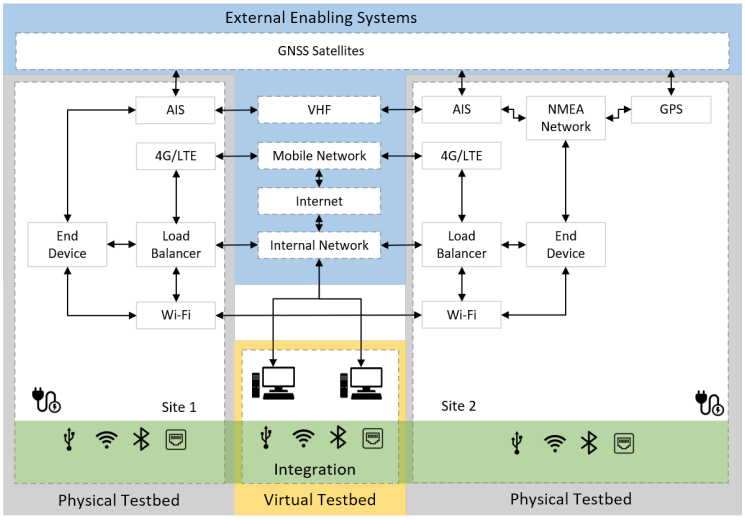

Fig. 2. Layout View of the testbed

Fig. 2 depicts a layout view of the testbed reflecting the different physical and logical components that are utilized during the different processes discussed in Section 3.1. The components can be organized in different configurations in order to emulate several use cases. Overall, the testbed is organized into three main sections, a physical testbed, a virtual testbed, and an integration of both. The virtual testbed consists of a group of workstations with several tools providing different capabilities. A summary of the included tools is depicted in Table 2 highlighting their categories and the process during which they are mainly utilized. On the other hand, the physical testbed consists of a group of hardware equipment providing different capabilities. A summary of the included equipment is depicted in Table 3. Finally, both the physical and virtual testbeds have

advantages and disadvantages which are depicted in table 4. Therefore, an integration between the two sections is proposed to enrich the system replication and analysis processes. The virtual and physical testbeds are integrated through a group of interfaces utilizing different technologies such as USB, Wi-Fi, Bluetooth, and Ethernet.

Table 2. Tools utilized in the virtual testbed

Process	Category	Tools	Description
System Replication	Emulation/ Simulation	*Bridgecommand*	Customizing and building cooperative navigational scenarios.
		NMEASimulator	Customization of navigational scenarios.
		GNS3	Generation of complex networks and functional components through virtualization technology. It can be used to emulate the network and configuration of the NSoI.
		VMWare	Utilized alone or along with the GNS3 simulator to create virtual machines.
		Virtualbox	
	Navigation	*OpenCPN*	A chart plotter software.
	Traffic Generation	*Tcpreplay*	Replay recorded packet capture containing sensor data or other types of traffic.
		Python Scripts	
		IMU + GPS	Generate and transmit Inertia measurements and GPS information from a mobile app.
		PacketSender	Transmit data or recorded packet capture over the network.
	Cybersecurity Controls	*Snort*	Open-source Intrusion Detection System(IDS).
		Wazuh	Open-source Security Information and Event Management (SIEM).
		Duo	Two Factor Authentication (2FA) software from Cisco.
		OpenLDAP	Role-Based Access Control (RBAC) software for access management.
		ClamAV	Antivirus software.
		BorgBackup	Backup software supporting encryption and compression as well as remote storage.
System Analysis	Monitoring	*Wireshark*	Packet capture and analysis.
		Screen Recorder	Record video and snapshots during experiments.
	Cybersecurity Testing	Ettercap	Man-in-the-middle tool.
		Kali Linux	Utilized as an attacker node.
		Nmap	Network scanner tools.
		Caldera	Breach and attack simulation platform for automating and emulating adversarial behavior (i.e. autonomous team).
		Scikit-learn	Machine learning library for python programming. Utilized for model building, training, and evaluation toward anomaly detection solutions.
	Communication Testing	*Iperf*	Network performance measurements.
		NetAnalyzer	App for analyzing Wi-Fi signals and LAN networks.
		WiFiAnalyzer	App for analyzing Wi-Fi signals.

4 Evaluation

In this section, we present a qualitative functional evaluation for our testbed through the discussion of some of the past and ongoing use cases utilizing it, namely, the analysis of communication and a cybersecurity architecture for an APS as well as an analysis of the security of sensor data in NMEA message format. Additionally, we provide a comparison of our testbed with the several aspects observed in the state of the art in cybersecurity testbeds. We demonstrate the utility of the testbed capabilities utilized during the system replication, system analysis, and technical management processes (refer to Section 3.1)

4.1 APS Communication and Cybersecurity Architecture

As discussed in Section 2, the main motivation for this testbed is the evaluation of a communication architecture [10] and a cybersecurity architecture [7] proposed in our earlier works based on a group of predefined communication

Table 3. Equipment utilized in the physical testbed

Process	Category	Equipment (Quantity)	Description
System Replication	Maritime Equipment	AIS A200 (1)	class A Automatic Identification System with external GNSS and VHF antenna
		AIS B921 (1)	class B Automatic Identification System with internal GNSS and VHF antenna
		Furunu GP 170 (1)	Marine GPS with external GPS antenna
		Garmin NMEA 2000 network starter kit (1)	NMEA 2000 network
		Garmin NMEA 2000 Network Updater (1)	
		Maretron IPG100 (2)	NMEA Internet Protocol Gateway
	Network Equipment	Cisco Aironet 1532E (3)	Wi-Fi outdoor lightweight access points with external directional and Omni antennas
		Cisco Wireless Controller 3504 (1)	For the management of the Wi-Fi network
		Netgear Nighthawk Mobile Hotspot Router (3)	LTE/4G router
		Cisco RV042G (2)	Load balancer, VPN router, and firewall
	Portable Power Sources	Omnicharge Ultimate (7)	Portable power source with 38400 mAh. Providing DC , AC, and USB output.
		9V power bank (3)	Additional power sources
System Analysis	Software Defined Radio (SDR)	SDRplay RSPdx (1)	Wideband SDR
		ADALM-PLUTO (4)	Active SDR learning module
Technical Management	Data Backup	LaCie 2TB (1)	2TB External Hard drive

Table 4. Advantages and disadvantages of our physical and virtual testbeds

	Advantages	Disadvantages
Physical	Wireless communication testing is possible using several technologies	Security attacks emulation is restricted due to limited possible configurations
	Built as mobile units to capture real measurements in different environments. (e.g. marine traffic).	Wired communication testing is limited due to the lack of ethernet switches.
		cost of testing autonomous navigation and control components is high due to expensive physical components (e.g. radar, lidar, cameras, etc.).
Virtual	Security attack emulation is flexible due to virtualization.	No capabilities for wireless communication testing.
	Wired communication testing is possible with advanced capabilities	Real measurements (e.g.marine traffic) cannot be effectively captured during experiments.
	Autonomous navigation and control components can be simulated,	

and cybersecurity requirements in [9] for an autonomous passenger ship (APS). The testbed in both works was utilized for the evaluation of the proposed architectures to demonstrate their fulfillment of the stakeholders' requirements and concerns. Table 5 summarizes the processes and the different aspects regarding the evaluation of both proposed architectures. A prototype of the communication architecture was implemented using the GNS3 simulator consisting of several emulated network devices with network protocols to support ship-to-ship and internal communication functions. The implementation included two networks representing both a remote control center and an APS. The role of the human operator was emulated to evaluate the provisioning of the required capabilities. Then, the implementation was subject to a test scenario to evaluate the implementation performance considering aspects such as redundancy, fault tolerance, and remote access. More details can be found in [10]. On the other hand, a prototype of cybersecurity architecture was implemented extending the implemented communication architecture. Additional equipment included two workstations emulating the two facilities for improved resource management in addition to two physical gateways (RV042G). Moreover, a group of required cy-

bersecurity controls was implemented (see Table 2) to evaluate their integration feasibility. Also, some sensor data was emulated using traffic generation tools. Then, the implemented architecture was evaluated using adversary emulation following 3 attack scenarios including red and blue team activities. The attack included several techniques including network sniffing, service scanning, ARP cache poisoning, gather victim information, and internet accessible devices using valid accounts. Although the attacks are not unique to the APS network, they were intended to evaluate the concept of layered defences within the context of the autoremote operational mode.

The testbed was found to be sufficient in evaluating the feasibility of integrating several architectural components and adequate in providing offensive security and defensive security analysis capabilities. However, the GNS3 simulator was found to be unsuitable for comprehensive performance analysis due to high latency related to virtualization.

Table 5. Use case 1: Architecture Evaluation

Process	Aspect	Communication Architecture	Cybersecurity Architecture
System Replication	Hardware	Workstation, GNS3, VMWare	Workstation, GNS3, VMWare, Virtualbox, Cisco RV042G
	Software		Cyber security Controls
	Data		Python scripts, IMU+GPS, Packet Sender
	Humans	Human (e.g. operator)	Human
	Processes, and Procedures	Ship-to-Shore, internal communication	Ship-to-Shore, internal communication, cybersecurity functions and protocols, sensor data collection.
	Facilities	Remote Control Center, APS	Remote Control Center, APS
System Analysis	Tools		Kali Linux, Nmap, Iperf
	Methods	Performance Analysis	Feasibility of security solutions, Adversary Emulation, Performance Analysis
	Scenarios	1 Scenario	3 Scenarios
	Teaming		Red team, Blue team
Technical Management	Resource Managemen		Each facility at a dedicated workstation
	Role Managemen	Human	Human, attacker
	Maintenance	✓	✓
	Data Storage	Local, Cloud	Local, Cloud and External HDD

4.2 NMEA Security

Several maritime-related protocols operate within the testbed components such as the National Marine Electronics Association (NMEA) protocol which is a standard for the communication among marine equipment including sensor data. A study is being conducted to analyze the security of NMEA messages in two use cases, the APS as well as Integrated Navigation Systems (INS) in traditional vessels [12]. Initially, a system emulating the INS and its equivalent in the APS is constructed using several tools that emit NMEA messages including the *bridgecommand* [1] simulator, *NMEA simulator* [2], and a physical GPS or AIS device. Additionally, the *OpenCPN* chart plotter software [3] is used and configured to receive the transmitted NMEA messages. Additional scripts are

[1] https://www.bridgecommand.co.uk (accessed July 2021)
[2] https://cutt.ly/NMEASimulator (accessed July 2021)
[3] https://opencpn.org (accessed July 2021)

utilized to transmit NMEA messages in certain scenarios. Several navigational procedures are emulated such as collision avoidance. Then the developed system is used to study the NMEA messages, their structure, behavior, and security. Several attack scenarios are carried as well as normal operational scenarios. This allowed for the generation of both normal and attack traffic for the application of machine learning techniques utilizing several modules in the Scikit-learn including some pre-processing modules and classifiers (e.g. decision trees) [21]. The analysis included offensive security, defensive security as well as a offensive defense by interchanging the red team and blue team activities toward an improved anomaly detection solution. The offensive security activities included several attacks among them are attacks against maritime sensor data including variations of Manipulation of View [5] and Denial of View [3] attack techniques. Table 6 depicts a summary of the processes and the different aspects related the activities in this project.

Table 6. Use case 2: NMEA Security

Process	Aspect	APS, INS
System Replication	Hardware	Workstation, Virtualbox, Bridgecommand Simulator, NMEA Simulator, Furunu GP 170
	Software	OpenCPN chart plotter
	Data	Simulated GPS, Python scripts
	Humans	Officer on Watch (OOW)
	Processes, and Procedures	Navigation status, route planning, collision avoidance, internal communication
	Facilities	Vessel
System Analysis	Tools	Kali Linux, ettercap, Scikit-learn
	Methods	Adversary emulation, anomaly detection, risk analysis
	Scenarios	Many navigational scenarios, many attack scenarios
	Monitoring	Wireshark, Screen recorder
	Teaming	Red, blue, and purple teaming
Technical Management	Resource Management	
	Role Management	Attacker, OOW
	Maintenance	✓
	Data Storage	Local, cloud, external HDD

4.3 Relevance to the state-of-the-art

Table 7 depicts a summary of the comparison between our testbed and the concepts and properties observed in the state-of-the-art cybersecurity testbeds captured by the literature discussed in Section 2. The comparison highlights the comprehensive nature of our testbeds capabilities as it supports most of the common concepts and properties. However, this comparison points to the areas of limitations. First of all, our testbed does not include components dedicated to cybersecurity learning; which is adopted by 25% of the surveyed works, this is because no requirements for such component have been communicated by the stakeholders. This also justifies the lack of education-related scenarios, scoring tools, and a green team. Additionally, no user behavior generation tools or dedicated or special management tools are utilized in our testbed. The management process is supported by several general-purpose tools such as Microsoft office word, excel, as well as commercial data backup software.

Table 7. Comparison between our proposed testbed and the concepts and properties observed in the state-of-the-art

Concepts and properties			Our testbed	Concepts and properties		Our testbed
Scenario	Purpose	Testing	✓	Environment	Emulation	✓
		Education	✗		Simulation	✓
		Experiment	✓		Real Equipment	✓
	type	Dynamic	✓		Hybrid	✓
		Static	✗	Tools	Emulation tools	✓
	Domain	Hybrid network applications	✓		Simulation tools	✓
		Networking	✓		Management tools	✗
		SCADA systems	✗		Monitoring tools	✓
		Social engineering	✗		Traffic generation	✓
		IoT systems	✗		User behavior generation	✗
		Critical infrastructure	✗		Scoring tools	✗
		Cloud based systems	✗		Security testing tools	✓
		Autonomous systems	✓	Teaming	Red team	✓
Management			✓		Blue team	✓
Learning			✗		White team	✓
Monitoring			✓		Green team	✗
Remote Access			✓		Autonomous Team	✓
Mobility			✓		Purple teaming	Yes
Scalability			Restricted			

The state-of-the-art captured by Yamin et al [27] does not capture the concept of testbed mobility. Additionally, purple teaming and remote access are discussed as concepts but the number of works that implement them were not tracked. Moreover, scalability is discussed only as a direction for future work. However, Tam et al [26] discussed testbed mobility and its utility in maritime testbeds. Also, the authors addressed scalability as a main direction for developing maritime-specific cyber ranges. Our testbed includes solutions for remote access, mobility, scalability, as well as activities implementing purple teaming. The remote access component is carried using the $TeamViewer$ software configured with the roles defined during the role management process (Section 3.1). The utility of $TeamViewer$ for remote laboratories and collaborative learning has been discussed in the literature (e.g. [15, 16]) and is found adequate in our testbed especially during the pandemic. Our testbed includes a mobility feature allowing it to be relocated to other indoor and outdoor locations. The mobility is supported through portable power sources allowing for extended experimentation periods, compact workstations in addition to specialized suite cases and mountable equipment, as well as certain waterproof equipment. Regarding scalability, our virtual testbed includes elements supporting scalabilities such as the GNS3 simulator, virtualization technology, and other simulation tools. This allows for the expansion, replication, and exportation of test scenarios. However, the scalability is restricted by the resources allowed by the testbed and identified during the resource management process (Section 3.1). The integration of a cloud-based component for the generation and execution of test scenarios is a future research direction. Lastly, the purple teaming concept has been applied in our testbed in a project targeting NMEA security (Section 4.2). This is supported by the integration of capabilities supporting red teams activities (e.g. Kali, Caldera, etc) as well as blue team activities through the different security controls.

5 Challenges and Future Work

The testbed proposed in this paper aims to support research regarding communication and cybersecurity of an autonomous passenger ship (APS) and other related maritime use cases. The novelty of the autonomous shipping domain introduces both temporal and contextual complexity that impacts our research. The contextual complexity is related to the lack of legal framework governing the technology while the temporal complexity is related to the lack of a unified industrial vision regarding the technology. The International Maritime Organization (IMO) has just recently completed a regulatory scoping exercise for the Maritime Autonomous Surface Ship (MASS); the ship class under which the APS falls. Plans for the next steps are yet undecided [4]. Moreover, several projects are undergoing regarding the development of autonomous passenger ships or ferries [6] including the Autoferry project [2] which is the prime focus of this testbed. This means that the current envisaged technology posture is subject to change because most of the components governing and supporting autonomous operations are yet under development. This leads to the possibility that certain communication and cybersecurity testing capabilities supported by the testbed might not be of relevance in the future. The contextual complexity can be addressed in the same manner when addressing the temporal complexity, particularly by using a divide and conquer approach [14]. This entails the formulation of a specific operational context (i.e. use case) containing several design alternatives to be analyzed. Then, the data generated by the analysis can lead to the generation of new possible use cases or technology adaptation of the analyzed technology. For this sake, our testbed included several components from several providers, using several technologies, and providing several capabilities. This flexible design aims to circumvent the challenges inflected by the aforementioned complexity aspects.

Additional challenges are related to the usage of licensed communication frequencies for ship-to-ship, and ship-to-shore communication. Our testbed includes two AIS devices for supporting ship-to-ship communication. AIS operates over Very High Frequency (VHF) which requires a license to operate in Norway. Until the time of writing this paper, the process for obtaining a license and permission is still undergoing. This has led to restricted testing capabilities. We have deferred to other means for getting AIS and NMEA data through utilizing simulators and previously captured data. On the other hand, the LTE routers supporting ship-to-shore communication requires monthly data subscription which adds additional management cost.

In maritime, safety and cybersecurity are inter-related aspects, recently, IMO has issued resolution MSC.428(98) dictating that ship owners and operators must address cybersecurity in their safety management system [13]. Integrating capabilities for safety management within the testbed is a future direction. This is intended to support the efforts of integrating cybersecurity capabilities in such management systems toward the development of an Integrated Ship Safety and Security Management System (IS3MS). In addition to this, several use cases are expected to be utilized in the testbed including AIS security and Breach and Attack Simulation (BAS) platforms in the maritime context. Finally, the testbed

is still under development and not available for public access at this moment. However, we can provide demonstrations of certain scenarios and capabilities.

6 Conclusion

The maritime domain is undergoing major digitization through the integration of technology and new operational aspects. Communication and cybersecurity are considered crucial aspects that could impact this major change in the industry. Therefore, in this paper, we proposed a testbed that can be utilized for the evaluation of several maritime use cases including the autonomous passenger ships (APS), and focusing on the communication and cybersecurity aspects. The testbed development is based on the observed state-of-the-art in cybersecurity testbeds and is inspired by several processes from the ISO 15288 system development standard. Our proposition includes an abstraction of three processes that can be followed for the utilization of the testbed namely, system replication, system analysis, and technical management. Moreover, we propose a system engineering approach for the system replication process that relies on standardized system elements. The three processes were followed during two projects (Sections 4.1 and 4.2) and found to help guide the progress throughout the projects. Additionally, the utilization of standardized system elements as guidelines during the system replication process led to the development of a realistic replica of the systems targeted for analysis.

Also, after comparing our testbed to the state-of-the-art it was found to be comprehensive in the inclusion of a set of capabilities covering most of the observed concepts and properties. In addition to that, the testbed includes additional less observed features such as remote access, mobility, and purple teaming. Nevertheless, the testbed was found to be lacking some of the observed aspects such as having a learning component, user behavior generation tools, automated environment building tools, and dedicated management system tools in addition to restricted scalability. However, such limitations can induce future research directions.

References

1. Norwegian maritime authority - focus on risks 2018. =http://bit.ly/sdirRisks2018 (Sep 2017)
2. Autonomous all-electric passenger ferries for urban water transport. =https://www.ntnu.edu/autoferry (July 2021)
3. Denial of view - att&ck ics. https://cutt.ly/DoV (2021)
4. Imo completes regulatory scoping exercise for autonomous ships. http://bit.ly/IMOMASS (May 2021)
5. Manipulation of view - att&ck ics. https://cutt.ly/MoV (2021)
6. Nfas - norwegian projects. https://cutt.ly/NFAS (2021)
7. Amro, A., Gkioulos, V.: Securing autonomous passenger ship using threat informed defense-in-depth (2021), Preprint. Submitted for review to Scientific Reports

8. Amro, A., Gkioulos, V., Katsikas, S.: Assessing cyber risk in cyber-physical systems using the *att&ck* framework (2021), Preprint. Submitted for review to ACM Transactions on Privacy and Security (TOPS)
9. Amro, A., Gkioulos, V., Katsikas, S.: Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation. In: Computer Security, pp. 69–85. Springer (2019)
10. Amro, A., Gkioulos, V., Katsikas, S.: Communication architecture for autonomous passenger ship. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability p. 1748006X211002546 (2021)
11. Amro, A., Kavallieratos, G., Louzis, K., Thieme, C.A.: Impact of cyber risk on the safety of the milliampere2 autonomous passenger ship. In: IOP Conference Series: Materials Science and Engineering. vol. 929, p. 012018. IOP Publishing (2020)
12. Amro, A., Oruc, A., Yildirim Yayilgan, S.: Nmea anomaly analysis and detection (2020), to be submitted
13. Committee, T.M.S.: Maritime cyber risk management in safety management systems (2017)
14. Gaspar, H.M., Ross, A.M., Rhodes, D.H., Erikstad, S.O.: Handling complexity aspects in conceptual ship design. In: International Maritime Design Conference, Glasgow, UK (2012)
15. Gravano, D.M., Chakraborty, U., Pesce, I., Thomson, M.: Solutions for shared resource lab remote quality control and instrument troubleshooting during a pandemic. Cytometry Part A **99**(1), 51–59 (2021)
16. Hubalovsky, S.: Remote desktop access us a method of learning of programming in distance study. In: 2011 14th International Conference on Interactive Collaborative Learning. pp. 450–455. IEEE (2011)
17. ISO, I.: Iec/ieee 15288: 2015. Systems and software engineering-Content of systems and software life cycle process information products (Documentation), International Organization for Standardization/International Electrotechnical Commission: Geneva, Switzerland (2015)
18. Jo, S.W., Shim, W.S.: Lte-maritime: High-speed maritime wireless communication based on lte technology. IEEE Access **7**, 53172–53181 (2019)
19. Kavak, H., Padilla, J.J., Vernon-Bido, D., Diallo, S.Y., Gore, R., Shetty, S.: Simulation for cybersecurity: state of the art and future directions. Journal of Cybersecurity **7**(1), tyab005 (2021)
20. Khan, M.E., Khan, F., et al.: A comparative study of white box, black box and grey box testing techniques. Int. J. Adv. Comput. Sci. Appl **3**(6) (2012)
21. Komer, B., Bergstra, J., Eliasmith, C.: Hyperopt-sklearn. In: Automated Machine Learning, pp. 97–111. Springer, Cham (2019)
22. Mir, Z.H., Filali, F.: Lte and ieee 802.11 p for vehicular networking: a performance evaluation. EURASIP Journal on Wireless Communications and Networking **2014**(1), 89 (2014)
23. Norwegian Shipowners' Association: Maritime outlook 2018. Tech. rep., Norwegian Shipowners' Association (2018)
24. Oakley, J.G.: Purple teaming. In: Professional Red Teaming, pp. 105–115. Springer (2019)
25. Strom, B.E., Applebaum, A., Miller, D.P., Nickels, K.C., Pennington, A.G., Thomas, C.B.: Mitre att&ck: Design and philosophy. Technical report (2018)
26. Tam, K., Moara-Nkwe, K., Jones, K.: The use of cyber ranges in the maritime context: Assessing maritime-cyber risks, raising awareness, and providing training. Maritime Technology and Research **3**(1), Manuscript–Manuscript (2021)

27. Yamin, M.M., Katt, B., Gkioulos, V.: Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. Computers & Security **88**, 101636 (2020)