

Doctoral thesis

Doctoral theses at NTNU, 2023:137

Ahmed Amro

Communication and cybersecurity for autonomous passenger ferry

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical
Engineering
Dept. of Information Security and
Communication Technology



Norwegian University of
Science and Technology

Ahmed Amro

Communication and cybersecurity for autonomous passenger ferry

Thesis for the Degree of Philosophiae Doctor

Gjøvik, May 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology



Norwegian University of
Science and Technology

NTNU

Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology

© Ahmed Amro

ISBN 978-82-326-6160-2 (printed ver.)
ISBN 978-82-326-5701-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:137

Printed by NTNU Grafisk senter

Declaration of Authorship

I, Ahmed Amro, hereby declare that this thesis and the work presented in it are entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

(Ahmed Amro)

Date: April 2023

Abstract

Recent innovations in the smart city and the maritime domains have led to the proposition of a new mode of transportation utilizing Autonomous Passenger Ships (APS) or ferries in inland waterways. The novelty of the APS concept has raised a wide range of challenges regarding the interconnection of various components for the provisioning of navigational tasks. Additionally, the new mode of operation has influenced the cyber risk paradigm and led to different considerations regarding attack objectives, techniques as well as risk management approaches.

Due to the fact that the APS technology is recent, defining the technical scope is the first challenge this thesis is addressing. This is sought through the identification of the APS expected operational context, relevant stakeholders, standards, guidelines, and functions. In addition to that, this thesis addresses the technical challenges related to interconnecting the APS components with their operational context in a secure and safe manner. This is sought through the definition of a suitable communication architecture for the APS and a cyber risk management process to develop a cybersecurity architecture capable of identifying and managing the cyber risks against the APS.

To realize that, the design science research methodology (DSRM) is followed with a group of relevant system engineering standards and processes. At each phase of the research, the academic and industrial perspectives are gathered to design, develop, demonstrate and evaluate the artifacts that are needed for achieving the research objectives.

The work in this thesis has resulted in the design, implementation, and evaluation of a suitable communication architecture for the APS technology supporting the current technology posture and includes flexible, modular, and resilient principles that designate it as candidate architecture for future iterations of the technology. Additionally, a suitable cyber risk management approach has been proposed and evaluated to measure its suitability for the APS technology. The cyber risk management approach named Threat Informed Defense in Depth (TIDiD) combines two cybersecurity strategies, namely, Threat Informed Defense and Defense in Depth. TIDiD includes a cyber risk assessment approach which is another result of this thesis. The approach is named FMECA-ATT&CK as it is based on the Failure mode, effects, and criticality analysis (FMECA) that is enhanced with the knowledge and semantics in the Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework. FMECA-ATT&CK supports the efforts for com-

prehensive and continuous cyber risk assessment and management through the identification of cyber risks in the APS components and proposes suitable risk mitigation measures. Then, later steps of TIDiD process aim to integrate the proposed risk mitigation measures into a cybersecurity architecture for risk analysis, monitoring, and treatment. Some areas were further explored including navigation data anomaly analysis and detection and the utility of the Automatic Identification System (AIS) in establishing covert channels for command and control activities during the development of cyber attacks. Each produced artifact was demonstrated and evaluated through a combination of evaluation methods including simulation, checklists, adversary emulation, and engagement of experts.

Trials involving existing communication technologies have shown success for the APS as a novel maritime transportation technology. By using existing solutions and processes, including those in this thesis, the security of the system has been enhanced. There are still many areas that require additional attention in order to improve the capabilities of remote monitoring and the cybersecurity posture of the APS. Therefore, APS technology and similar maritime technologies are worthy of exploration in the future.

Acknowledgments

My doctoral study has been an incredible journey enriched with the support and companionship of those surrounding me. The past four years have been an era of personal growth I haven't witnessed before.

For that, I would like to start with my colleague, life partner, best friend, soul mate, and wife Lama, for the way you supported me throughout my life and particularly during my doctoral study. From the bottom of my heart, thank you, love.

To my daughters, Alma and Tia, the transition and journey to adaptation you had to endure for me to engage in my doctoral study will always be remembered. Your compassionate, lively, and creative personalities provided me with the needed distractions and love to fight through my journey. From your daddy thank you my monkeys.

To my father Walid, mother Dunya, sisters Maysa and Sahar, and brother Mohammad. Although you're far away, your prayers, support, and warm wishes were felt and welcomed. I hope your little son and brother made you proud.

To my supervisors Vasileios and Sokratis. I don't know where to begin. Your guidance and support have transformed my mindset, skills, and attitude. I'm a better researcher, engineer, scientist, and person because of you.

Finally, to my colleagues in the CISaR group, other PhD candidates, and postdoctoral researchers. Particularly, Mohamed Ali, Georgios Kavallieratos, Benjamin Knox, Aida Akbarzadeh, Aybars Oruc, Nabin Chowdhury, and Xhesica Ramaj. I wouldn't have done it without you. Your companionship and brilliant insights have influenced my work immensely and enriched my personal life. For that, I will always be in debt.

Ahmed Amro

Gjøvik 2023.

Abbreviations

Table 1: Abbreviations Table

Abbreviation	Description
AI	Artificial Intelligence
AIS	Automatic Identification System
APS	Autonomous Passenger Ship
ATT&CK	Adversarial Tactics, Techniques, and Common Knowledge
AUV	autonomous underwater vehicle
C&C	Command and control
C-ES	cyber-enabled ship
CIIs	Critical Information Infrastructures
CPS	Cyber-Physical System
CRWs	Cooperative Robotic Watercrafts
CTI	Cyber Threat Intelligence
CVSS	Common Vulnerability Scoring System
CYSM	collaborative security management system
D2D	Device to Device
DiD	Defense-in-Depth
DMS	Document Management Services
DREAD	Damage, Reproducibility, Exploitability, Affected Users, and Discoverability
DSRM	Design Science Research Methodology
ECDIS	Electronic Chart Display Information System
FMECA	Failure Mode, Effects & Criticality Analysis
GMDSS	Global Maritime Distress and Safety System
IACS	International Association of Classification Societies
IALA	International Association of Lighthouse Authorities
ICS	Industrial Control Systems
ICT	Information and Communication Technologies
IMO	Maritime Organization
INS	Integrated Navigation System
IoT	Internet of Things

IT	Information Technology
LTE	Long-Term Evolution
MANET	mobile ad hoc network
MASS	Maritime Autonomous Surface Ship
MBR	maritime broadband radio
ML	Machine Learning
MUNIN	Maritime Unmanned Navigation through Intelligence in Networks
NFAS	Norwegian Forum for Autonomous Ships
NIST	National Institute of Standards and Technology
NMEA	National Marine Electronics Association
OppNet	hop-by-hop network
OT	Operational Technology
PHA	Preliminary Hazar Analysis
PKI	Public Key Infrastructure
RAS	Risk Analysis Service
RCC	Remote Control Center
RF	Radio Frequency
RMS	Risk Management Services
ROS	Robot Operating System
RTSP	Real Time Streaming Protocol
SCL	Shore Control Lab
SIEM	Security Incident and Event Monitoring
SIP	Strategy Implementation Plan
SLR	Systematic Literature Review
SoS	System of Systems
SSM	Six-Step Model
STRIDE	spoofing, tampering, repudiation, information disclosure, denial of service, elevation of privilege
TIDiD	Threat-Informed Defense-in-Depth
UAV	unmanned aerial vehicle
USV	Unmanned Surface Vehicle
VDES	VHF Data Exchange System

Contents

Abstract	iii
Acknowledgments	v
Abbreviations	vii
Contents	ix
Figures	xiii
Tables	xv
I Introductory Chapters	xvii
1 Introduction	1
1.1 Research Problem and Motivation	2
1.2 Research Objectives and Questions	3
1.3 List of included publications	4
1.4 List of additional publications	5
1.5 Scope of the research	6
1.6 Thesis Outline	7
2 Background	9
2.1 Scoping the autonomous ship technology	9
2.2 Communication aspects for the APS Technology	10
2.3 Cyber Risk Management in the APS Technology	11
3 Related Work	15
3.1 Attacks, Threats, and Risks	16
3.2 Requirement Elicitation	17
3.3 Communication Architectures	18
3.4 Cyber Risk Management	19
4 Methodology	23
5 Summary of Papers and Contributions	29
5.1 Article 1: Connect and protect: requirements for maritime autonomous surface ship in urban passenger transportation	29
5.2 Article 2: Communication architecture for autonomous passenger ship	29
5.3 Article 3: Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship	30

5.4	Article 4: Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework	30
5.5	Article 5: Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth	31
5.6	Article 6: Navigation Data Anomaly Analysis and Detection	32
5.7	Article 7: From Click to Sink: Utilizing AIS for Command and Control in Maritime Cyber Attacks	32
5.8	Article 8: Communication and Cybersecurity Testbed for Autonomous Passenger Ship	33
5.9	Article 9: Evaluation of a Cyber Risk Assessment Approach for Cyber-Physical Systems	33
5.10	Article 10: Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework	34
5.11	Thesis Impact	34
5.12	Communication Architecture Implementation	35
6	Limitations and Future Work	37
6.1	Limitations	37
6.1.1	Related to Article 1: Connect and protect: requirements for maritime autonomous surface ship in urban passenger transportation	37
6.1.2	Related to Article 2: Communication architecture for autonomous passenger ship	38
6.1.3	Related to Article 3: Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship	38
6.1.4	Related to Article 4: Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework	39
6.1.5	Related to Article 5: Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth	39
6.1.6	Related to Article 6: Navigation Data Anomaly Analysis and Detection	39
6.1.7	Related to Article 7: From Click to Sink: Utilizing AIS for Command and Control in Maritime Cyber Attacks	39
6.1.8	Related to Article 8: Communication and Cybersecurity Testbed for Autonomous Passenger Ship	40
6.1.9	Related to Article 9: Evaluation of a Cyber Risk Assessment Approach for Cyber-Physical Systems	40
6.1.10	Related to Article 10: Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework	41
6.2	Future research	41
6.2.1	Communication Technologies for the operations of autonomous vessels	41
6.2.2	Maritime Cyber Risk Management	41
6.2.3	Forensics Readiness in the Maritime Sector	44
7	Conclusions	45

Bibliography 49

II Published Research Papers 59

Paper I 61

Paper II 79

Paper III 103

Paper IV 119

Paper V 157

Paper VI 195

Paper VII 225

Paper VIII 247

Paper IX 265

Paper X 291

Figures

1.1	Research flow, questions, and publications	3
2.1	Overview of the APS investigated scope	9
2.2	Overview of the APS investigated communication aspects	11
2.3	Overview of the investigated cyber risk management aspects	12
4.1	Design Science Research Method with requirement elicitation process and artefact types.	23
4.2	The activities during the DSRM stages with respect to the first research question, and included papers.	24
4.3	The activities during the DSRM stages with respect to the second research question, and included papers.	25
4.4	The activities during the DSRM stages with respect to the third research question, and included papers.	26

Tables

1	Abbreviations Table	vii
1.1	List of Included Publications and the Relevant Research Questions .	5
1.2	List of Additional Publications	6
3.1	Surveyed Communication architectures for autonomous and traditional maritime systems	19
5.1	APS Communication Architecture influence on Implementation . . .	36

Part I

Introductory Chapters

Chapter 1

Introduction

Our globally interconnected world relies on a variety of transportation methods to carry commodities, provide services, and move people throughout the world. As a result, the transportation industry is seen as a crucial component of infrastructure globally. There are five acknowledged modes of transportation in the European Union: aviation, road, rail, maritime, and inland waterways [1]. This thesis specifically focuses on the area where interior waterways and the maritime domain meet. The maritime transportation sector is connected to Europe's citizens' security, prosperity, and well-being [2]. Additionally, it accounts for 90% of all worldwide trade in goods [3]. Moreover, an era of digital transformation is unfolding in the domain leading to drastic changes in business models, processes, as well as technology [4] making it a field deserving of increased attention.

A brief history and some of the various definitions of digital transformation are provided by Schallmo et al. [5], who summarize it as a process that aims to achieve novel value creation, process optimization, enhancement of experience, and establishment of new foundational capabilities. According to Heilig et al. [6], port management and logistics witnessed the first observed substantial application of digital transformation in the maritime industry. Soon after, new Information and Communication Technologies (ICT) were introduced with a direct focus on ships, attempting to improve how they are constructed, run, and maintained.

As a result of this process, research and innovation activities have been directed toward developing unique and environmentally friendly maritime transportation technologies. The Norwegian Forum for Autonomous Ships (NFAS) currently lists several ongoing and finished initiatives that seek to create both entire platforms and enabling technology [7]. This thesis originated from the work in one of such projects, namely, the "Autoferry" project which aims to create an all-Electric Autonomous Passenger Ship (APS) or ferries for Urban Water Transport [8].

In a previous study, Havdal et al. [9] described the difficulties involved in creating an APS, including those unique to the interaction with the environment and the navigation of the autonomous system, with the major goal of maintaining the safety and security of users, systems, and the surrounding environment.

As mentioned by Patraiko [10], several attempts have been made to advance

E-Navigation, which are also coordinated through the International Maritime Organization (IMO). E-Navigation has been defined by the International Association of Lighthouse Authorities (IALA) as "the harmonized collection, integration, exchange and presentation of maritime information aboard and ashore by electronic means to enhance berth-to-berth navigation and related services, safety and security at sea, and the protection of the marine environment" [11].

E-navigation was first proposed as a solution for open sea navigation. The analysis of the IMO's e-navigation Strategy Implementation Plan (SIP) and Korea's national e-navigation SIP, as reported by Kwang [12], shows that E-navigation services are crucial for inland navigation as well. Additionally, Kwan [12] suggested that the introduction of E-navigation services is greatly facilitated by the use of digital communication services like Long-Term Evolution (LTE) and Automatic Identification System (AIS). This also applies to APSs used for passenger inland transportation. Supporting E-navigation within the context of the APS raises various challenges in many aspects including communication and cybersecurity to ensure safe navigation.

The cyber attack against the Mærsk shipping company, which resulted in weeks of operations being disrupted and losses exceeding \$300 million US [13] in addition to the denial of service attack against COSCO shipping company [14] are only a few examples of how disruptive strikes against the maritime domain can have disastrous effects. Research in the maritime domain has also shown that maritime systems and procedures lack adequate security. To mention a few, Tran et al [15] examined the inadequate authentication, encryption, and validation in a widely used protocol in the maritime domain named NMEA (National Marine Electronics Association) while Balduzzi et al [16] showed a variety of attacks against AIS, including spoofing, jamming, and other forms of misuse.

Positively, the current situation in the maritime domain calls for the examination of cyber risks and cyber risk management. IMO has issued Resolution MSC.428(98) [17], which calls on all parties involved in the maritime sector to include cyber risk management in their safety management systems. The resolution lays out requirements and principles for managing cyber risks [18]. The recommendations call for a constant evaluation of the dangerous environment facing maritime infrastructure.

1.1 Research Problem and Motivation

The recent adoption of the autonomous and remotely controlled ferry for passenger transportation referred to in this thesis as "APS", introduces an element of cyber threats due to reliance on cyber components for the delivery of navigational functions. The operational modes for this new class depend heavily on communication, control, and monitoring capabilities. Such capabilities, if not properly secured, could result in undesired consequences possibly leading to the loss of human life. Additional challenges are relevant to the research area itself. Our research is impacted by the novelty of the autonomous shipping domain as well as its

contextual and temporal complexity. In the absence of a legal framework to govern the technology, contextual complexity arises. On the other hand, the lack of a unified industrial vision regarding the technology projects a temporal complexity. IMO recently completed a regulatory scoping exercise for the Maritime Autonomous Surface Ship (MASS); the class of ship to which the APS belongs. The next steps are yet to be determined [19]. The development of the APS technology is also underway as observed across multiple projects [20], including the Autoferry project [8] which is the prime focus of this thesis. This means that the current investigated technologies, protocols, and functions are subject to change because most of the components governing and supporting autonomous operations are yet under development.

Therefore, the development of a resilient, flexible, and modular communication architecture in addition to a comprehensive and continuous risk management approach for the identification, treatment, and monitoring of cyber risks is required and is the main premise of this thesis.

1.2 Research Objectives and Questions

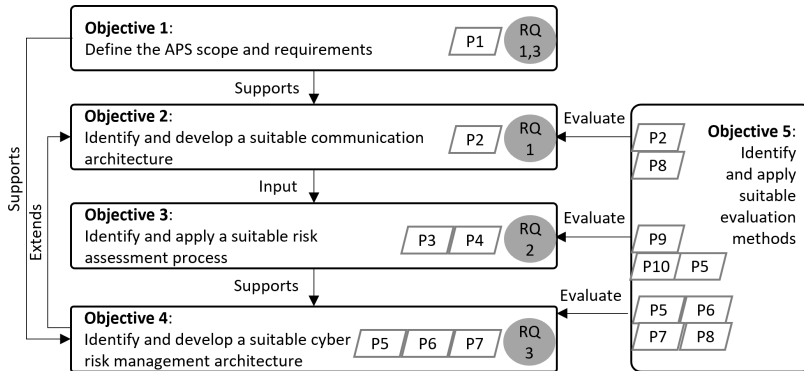


Figure 1.1: Research flow, questions, and publications

An overview of the research flow is depicted in Figure 1.1. This thesis aims to achieve five objectives. The first objective is to define the scope of the APS technology. This encompasses its operational context, relevant standards, stakeholders, and their communicated requirements related to the establishment of solutions regarding communication and cybersecurity. Then, the identified scope and requirements are utilized to identify relevant artifacts to support the second objective which is to develop a suitable communication architecture allowing the APS to interconnect with its operational context. Afterward, the third objective is to investigate suitable cyber risk assessment methods and later apply them to identify possible cyber risks in the defined scope. The outcome of the risk assessment is

then utilized to investigate suitable approaches and methods for cyber risk management toward the fourth objective which is the development of a suitable architecture that extends the previously proposed communication architecture. Finally, when all the previous solutions are proposed, the final objective is to identify and apply suitable approaches for evaluation. This entails identifying means for evaluating the soundness of the proposals introduced in this thesis and consequently conducting the evaluations to provide the reader with sufficient evidence to construct a fair judgment. Achieving these objectives is attempted by answering the following Research Questions (RQ):

- RQ 1: What communication architecture of the APS should be defined, implemented, and evaluated in order to ensure compliance with regulations, guidelines, and standards, and at the same time satisfy operational and functional requirements related to reliability and cybersecurity?

Prior to the work in this thesis, the APS technology did not exist, at least to the extent that allowed conducting research related to its communication and cybersecurity. Therefore, this question explored the relevant artifacts in the literature and the relevant methods for the development of such technology in a manner that allows further research. This also suggests the need to consider design principles that allow future changes in the relevant technologies.

- RQ 2: What risk assessment method is most suitable for assessing the cyber risks in a cyber-physical system such as the APS in order to be able to identify and model potential threats in different scenarios, for further use toward the integration and evaluation of defensive mechanisms?

After the scope of APS technology was defined, the characteristics of a suitable risk assessment process were identified. This entails the inclusion of several technology domains such as Information Technology (IT) and Operational Technology (OT) and the need for comprehensive and continuous risk assessment. This question investigated the existing risk assessment methods in order to identify a method that addresses those characteristics.

- RQ 3: What is a suitable approach for managing the cyber risks against the APS to ensure the safety of passengers and the security of the system?

Following the identification and application of a suitable risk assessment method, the risks against the APS were identified. The next step aims to integrate relevant methods and capabilities that supports the management of such risks. This question explores the relevant cyber risk management approaches in the APS application domain as well as their suitable evaluation methods.

1.3 List of included publications

The work in this thesis has produced a range of publications in national and international peer-reviewed journals and conferences. The list is depicted in Table 1.1.

it is noteworthy that the author of this thesis is the lead contributor and first author of all the listed publications except article 10. A summary of the papers can be found in Chapter 5.

Table 1.1: List of Included Publications and the Relevant Research Questions

Article	Type	Publisher/Conference	RQ(s)	Reference #	Chapter #
1	Conference post-proceedings	Amro, A., Gkioulos, V., and Katsikas, S. (2019). Connect and protect: requirements for maritime autonomous surface ship in urban passenger transportation. In <i>Computer Security</i> (pp. 69-85). Springer, Cham.	RQ1 , RQ3	[21]	I
2	Journal (Level 1)	Amro, A., Gkioulos, V., and Katsikas, S. (2021). Communication architecture for autonomous passenger ship. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 1748006X2111002546.	RQ1	[22]	II
3	Conference proceedings	Amro, A., Kavallieratos, G., Louzis, K., and Thieme, C. A. (2020, November). Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship. In <i>IOP Conference Series: Materials Science and Engineering</i> (Vol. 929, No. 1, p. 012018). IOP Publishing.	RQ2	[23]	III
4	Journal (Level 2)	Ahmed Amro, Vasileios Gkioulos, and Sokratis Katsikas. 2022. Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework. <i>ACM Trans. Priv. Secur. Just Accepted</i> (November 2022). https://doi.org/10.1145/3571733	RQ2	[24]	IV
5	Journal (Level 2)	A. Amro and V. Gkioulos, 'Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth,' <i>International Journal of Information Security</i> , Nov. 2022, ISSN: 1615-5270. DOI: 10. 1007/s10207-022-00638-y. [Online]. Available: https://doi.org/10.1007/s10207-022-00638-y 151	RQ3	[25]	V
6	Journal (Level 1)	Amro, A., Oruc, A., Gkioulos, V., and Katsikas, S. (2022). Navigation Data Anomaly Analysis and Detection. <i>Information</i> , 13(3), 104.	RQ3	[26]	VI
7	Conference proceedings	Amro, A., and Gkioulos, V. (2022). From Click to Sink: Utilizing AIS for Command and Control in Maritime Cyber Attacks. In <i>European Symposium on Research in Computer Security</i> (pp. 535-553). Springer, Cham.	RQ3	[27]	VII
8	Conference post-proceedings	Amro, A., and Gkioulos, V. (2021, October). Communication and Cybersecurity Testbed for Autonomous Passenger Ship. In <i>European Symposium on Research in Computer Security</i> (pp. 5-22). Springer, Cham.	RQ1, RQ3	[28]	VIII
9	Journal (Level 1)	Amro, A., and Gkioulos, V. (2023, mARCH). Evaluation of a Cyber Risk Assessment Approach for Cyber-Physical Systems: Maritime- and Energy-Use Cases. <i>Journal of Marine Science and Engineering</i> . 2023; 11(4):744.	RQ2	[29]	IX
10	Journal (Level 1)	Oruc, A., Amro, A., and Gkioulos, V. (2022). Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework. <i>Sensors</i> , 22(22), 8745.	RQ2	[30]	X

1.4 List of additional publications

During the period of conducting the work in this thesis, additional publications were produced but are not included in this thesis due to a different scope. The list of these publications is depicted in Table 1.2.

Table 1.2: List of Additional Publications

Article	Type	Publisher/Conference	Contribution
1	Conference proceedings	Amro, A. (2020). Iot vulnerability scanning: a state of the art. <i>Computer Security</i> , 84-99.	Sole Author
2	Conference proceedings	Amro, A. (2021). Cyber-Physical Tracking of IoT devices: A maritime use case. In <i>Norsk IKT-konferanse for forskning og utdanning</i> (No. 3).	Sole Author
3	Conference proceedings	Amro, A., Yamin, M. M., & Knox, B. J. (2020, July). Applications of an Online Audience Response System in Different Academic Settings: An Empirical Study. In <i>International Conference on Human-Computer Interaction</i> (pp. 165-175). Springer, Cham.	Lead author

1.5 Scope of the research

APS technology and its associated contextual and temporal complexity led to a change in the scope of this research as it progressed. The initial scope was rather generic focusing on addressing the communication technology aspect and then the cybersecurity aspects from a technical and engineering perspective to forecast the shape of the required technologies. Basically, what is the APS context or ecosystem, how to connect the APS to it, and then how to protect it? Then, as the project advanced and with it the technology itself, the scope was adjusted to focus on the relevant aspects.

During the first years of the project that led to this thesis, the shape of the technology was unclear, the author struggled to identify the scope of the technology due to limited relevant literature in the field and the lack of clarity surrounding the technology. That is why the first objective of this thesis was to determine the scope which was accomplished in [21]. The defined scope was consequently utilized to drive the later objectives. The next iteration in defining the scope was related to the definition of the communication architecture of the APS. This defined the scope of the study at a technical and component-level allowing the architecture to be shaped as a clearly defined use case for concrete subsequent progression. After that, the scope was categorically changed to focus on cyber risk management rather than the communication aspects. This included focusing on risk identification, analysis, treatment, and monitoring. Which led to the proposition of new interconnected risk assessment and risk management approaches. Finally, the produced solutions were targeted for evaluation. The scope of the evaluation was determined based on the surveyed literature but mainly focused on system-level evaluation.

Regarding the application domain of this thesis, the APS is mainly intended for inland passenger transportation. Additionally, several identified stakeholders and system components are related to the maritime domain. Therefore, the transportation and maritime domains are the main identified application domains of the APS.

Categorically, the operational context of the autonomous ferry has been approached as a System of Systems (SoS). The autonomous ferry which constitutes a major element in the scope of this thesis has been conceptualized as a Cyber-Physical System (CPS). Additionally, the explored technology domains include

traditional enterprise networks consisting mostly of IT and ICT. In addition to Industrial Control Systems (ICS) hosting a variety of OT, a group of Radio Frequency (RF) technologies, and a limited observation of Machine Learning (ML) and Artificial Intelligence (AI) technologies. In one work, [31], maritime cyber components, some of which are hosted on the ferry have been categorized as Internet of Things (IoT) devices. Other Autoferry project members have utilized concepts and technologies from the robotics domain, however, this categorization has not been considered in this thesis.

1.6 Thesis Outline

This thesis consists of two parts. Part I contains the overview of the research project, and Part II consists of the research papers. In Part I, background information is provided in Section 2 in which a basic foundation is presented for the understanding of this thesis. Chapter 3 presents several related works. Chapter 4 discuss the applied research methods and how they were applied throughout the project duration. Then, a summary of the included research papers is presented in Chapter 5 including highlights of the key contributions of this thesis to research, education, and technology development. Later, acknowledged limitations and directions for future work are discussed in Chapter 6. Finally, concluding remarks are presented in Section 7. In Part II, the papers are presented in the sequence observed in Table 1.1.

Chapter 2

Background

2.1 Scoping the autonomous ship technology

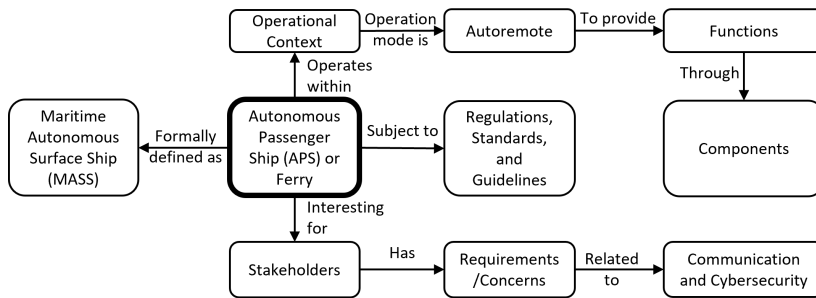


Figure 2.1: Overview of the APS investigated scope

Autonomous shipping is a relatively recent concept. This was reflected in the small amount of relevant literature at the beginning of the work in this thesis. Still, some regulations, guidelines, characteristics, and systems components in traditional shipping might be relevant to the technology. Therefore, the first portion of this thesis is dedicated to shaping the scope of the technology utilized for autonomous passenger transportation. An overview of the viewpoints investigated for defining the APS scope is depicted in Figure 2.1. Regarding definition, a document published by NFAS [32] provided definitions for autonomous ships, their operational context, functions, and other related content. Based on the operational areas (underwater or surface), control modes (remote control or autonomous), and manning levels (continuously manned to continuously unmanned), a classification of autonomous maritime systems is proposed by NFAS. MASS is the focus of this thesis, with an application for passenger transportation in urban waterways called the Autonomous Passenger Ship (APS) or ferry. The terms ship and ferry are used interchangeably throughout this thesis. The autonomous ferry fits the definition of a ship according to NFAS: "a ship is a vessel with its own propulsion

and steering system, which executes commercially useful transport of passengers or cargo and which is subject to a civilian regulatory framework". Additionally, an autonomous ship is defined as "a ship that has some level of automation and self-governance". The APS could as well be considered a cyber-enabled ship (CES) which is a term coined by Lloyd register in 2016 referring to ships equipped with ICT for enhancing monitoring, communication, and connection capabilities [33]. The operational mode of autonomous ships that is similar to an APS is called "autoremode". In this mode, a ship operates autonomously but can be taken over completely by humans in emergency situations [34]. Also, some works have been produced as deliverables of the Maritime Unmanned Navigation through Intelligence in Networks (MUNIN) project. Particularly, D10.1 [35] and D4.3 [36] were of substantial utility to this thesis with regard to defining the scope and the identification of the initial list of relevant components and aspects. Deliverable 10.1 discussed several expected issues arising from the utilization of autonomous ships on the current state of affairs in the shipping domain and discuss some solutions for the different identified constraints. Deliverable 4.3 presented the results of experts' engagement for drafting the technical pathway toward ship-to-shore communication links through the development of general requirements and evaluating different technological options in regard to satisfying those requirements. Additionally, the Danish maritime authorities have published a report analyzing the regulatory barriers to the use of autonomous ships [37]. This report aided in understanding the expected challenges in adopting the new technologies of autonomous ships as well as the expected stakeholders. The US National Institute of Standards and Technology (NIST) framework for improving critical infrastructure cybersecurity [38] was the most referenced framework regarding cybersecurity in the maritime context. Several implementation profiles of the NIST framework were published by the United States Coast Guard including a profile for passenger vessels [39]. This document was utilized for shaping the initial understanding of expected mission objectives regarding cybersecurity in the context of a passenger vessel such as the APS. Moreover, several documents related to autonomous ships were published by several members of the classification society in the maritime domain including DNV [34], Bureau Veritas [40], and The International Association of Classification Societies (IACS) [41]. These documents were thoroughly studied and analyzed in order to produce the initial list of communication and cybersecurity requirements.

2.2 Communication aspects for the APS Technology

An overview of the investigated aspects related to communication within the context of the APS is shown in Figure 2.2. After the initial scope of the APS was defined with a clear list of requirements, exploring the relevant communication architectures for supporting those requirements came next. In this regard, several works in the literature have proposed communication architectures for several classes of ships including remotely controlled and autonomous ships. A literature

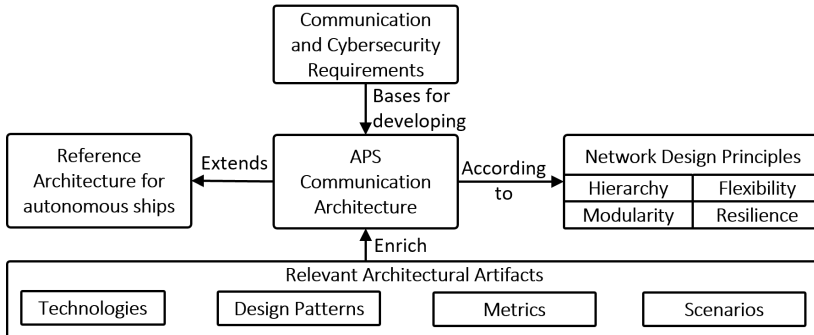


Figure 2.2: Overview of the APS investigated communication aspects

survey of these works was conducted in order to identify relevant architectural artifacts such as technologies, design patterns metrics and scenarios. The artifacts later supported the proposition of a suitable communication architecture for the APS. A reference communication architecture was proposed by Rødseth et al. [42] as part of the MUNIN project and was adopted and improved with the relevant artifacts from the literature. The architecture was developed in order to address the temporal and contextual complexity of the technology by applying hierarchical, flexible, modular, and resilient design principles inferred from Cisco Inc. [43]. Network hierarchy definition is a critical step. An access network; where end devices reside, a distribution network, and a core network are generally the three tiers of a computer network's hierarchical structure. Having a modular design is also beneficial due to the isolation it provides and the ability to seamlessly update or upgrade technologies. By avoiding single points of failure, redundancy helps realize the resilience principle, which is about maintaining operability under normal and abnormal conditions. Additionally, the network design should allow for flexibility in the use of technologies due to continuous changes in technology.

2.3 Cyber Risk Management in the APS Technology

Then, with the APS as a technology having a technically sound description of it in the form of a communication architecture, a cyber risk management process was initiated. An overview of the investigated aspects related to cyber risk management are depicted in Figure 2.3. Cyber risk management was sought with domain-specific consideration laid out in the resolution issued by IMO. IMO had urged the different maritime industry stakeholders to include cyber risk management in their safety management systems. The resolution by IMO was communicated suggesting some guidelines and requirements for cyber risk assessment and management [18]. This includes the consideration of different technology domains such as IT and OT, the consideration of operational, safety, and security

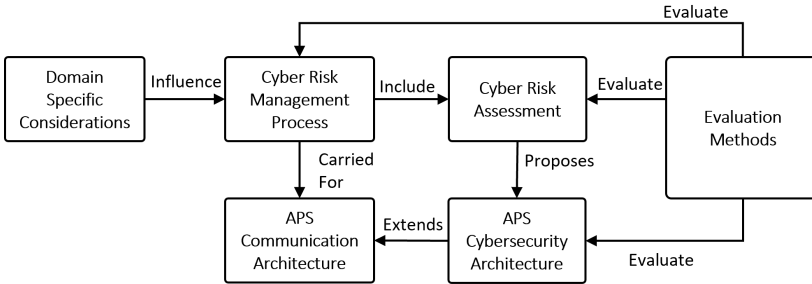


Figure 2.3: Overview of the investigated cyber risk management aspects

impacts, and the need for continuous risk assessment and management, starting with a risk assessment method followed by risk treatment and monitoring.

Firstly, a comprehensive study of risk assessment in CPS was conducted, this covered the maritime and automotive domains. A common approach for threat modeling emerged which is the STRIDE method (spoofing, tampering, repudiation, information disclosure, denial of service, elevation of privilege) [44]. The high-level abstraction of STRIDE and the limited linkage with relevant mitigation measures as well as the increased reliance on expert input was rendered limiting. This is among the reasons for considering the utility of the Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework by MITRE [45, 46] for threat modeling and beyond. ATT&CK has the power to reflect the granularity of actions that adversaries can take, how they relate to one another, their consequences related to adversarial objectives, their correlation with mitigation methods and data sources, and their targeted platforms and systems [45]. Additionally, the increased reliance on the expert judgment was observed in subsequent steps in the risk assessment including likelihood and impact estimation and risk controls proposition. ATT&CK has been found to include curated knowledge in these regards that would reduce the need for expert judgment, thus reducing the impact of bias and required efforts which would support continuous risk assessment and management activities. Additionally, concepts from graph theory [47] have been employed for providing information relevant to risk calculation drawn from systems modeled as graphs [48–50]. In a graph, nodes represent components that are linked with each other by edges. The graph can be analyzed using a number of formal measures, including measures of centrality. These efforts have led to the proposition of a semi-automated cyber risk assessment approach for CPS and were demonstrated against the APS communication architecture.

Afterward, further activities in risk management were sought including risk treatment and monitoring. After reviewing the literature in the maritime domain and the communicated guidelines of several influential entities, a security design pattern appeared to be an agreed-upon approach for cyber risk management, namely, Defense-in-Depth (DiD). DiD is defined by NIST as an “Information se-

curity strategy integrating people, technology, and operations capabilities to establish variable barriers across multiple layers and dimensions of the organization" [51]. Several cyber risk management strategies from relevant entities were investigated to identify DiD elements that are suitable for addressing the APS risks and requirements. This includes BIMCO [52], DNV [34], and other guidelines focused on DiD for Industrial Control Systems (ICS) [53]. Still, critical discussions have been raised regarding the limitations of DiD against targeted sophisticated attacks [54]. This could possibly be linked to the lack of a Cyber Threat Intelligence (CTI), one of the missing elements of DiD [55]. With CTI, defenders can constantly adjust their defenses to manage the risks targeting their assets based on the current threat landscape [56]. As part of the Threat Informed Defense strategy from MITRE [57] (i.e. Threat-Based Defense [58]), CTI is one of three pillars along with defensive engagement and focused sharing and collaboration. [58]. The three pillars interact together to provide the *ATT&CK* framework [45] which can be used as an up-to-date resource for encoded common knowledge regarding adversarial behavior. Aligning the *ATT&CK* framework and DiD layers was envisioned to allow more evolved cyber risk management capabilities. So, a cyber risk management approach was proposed named Threat-Informed Defense-in-Depth (TIDiD) which integrates elements from the two strategies, namely, the Threat Informed Defense, and DiD.

Lastly, efforts to evaluate the proposed propositions related to risk management included a systematic literature review of cybersecurity evaluation approaches in the maritime domain. Several aspects of evaluation, objectives, and methods were observed with varying degrees of fidelity. This includes the engagement of experts and stakeholders, unit testing, simulation, adversary emulation (i.e. pentesting), auditing, and others. The cyber risk assessment approach is evaluated through demonstration and the engagement of experts. The cyber risk management approach and its produced cybersecurity architecture are evaluated through a combination of evaluation, checklist, simulation, and adversary emulation.

Chapter 3

Related Work

The Autonomous Passenger Ship (APS) stands distinct from other autonomous systems, primarily due to the unique characteristics and requirements of the maritime domain. In order to ensure compliance with specific maritime regulations and guidelines, the APS must accommodate a range of devices, protocols, standards, technologies, and human tasks. These include the utilization of the AIS for marine traffic management, adherence to the NMEA protocol for communication, and the execution of specific tasks by human operators, and others.

Some existing solutions employed for autonomous systems such as autonomous cars and unmanned aerial vehicles would not fit the specific purpose of the APS due to different design characteristics and operational variables. Moreover, the maritime domain encompasses a diverse array of autonomous systems, such as Unmanned Surface Vessels (USV), which find applications in civil, military, and research contexts, including oceanography, remote sensing, weapons delivery, environmental monitoring, surveying, anti-submarine warfare, and electronic warfare [59]. Nevertheless, the APS distinguishes itself through its unique application in urban passenger transportation. Consequently, the development and operation of the APS must adhere to stringent conditions to accommodate passengers and their cargo. The present thesis, therefore, emphasizes the consideration of these unique characteristics during the development of solutions for the APS.

Due to the novelty of the APS technology, no work has addressed this topic in the context of communication technologies and cybersecurity as the main focus areas. In a study of public perception of autonomous urban ferries, experienced operators highlighted the safety and security challenges that need to be addressed [60]. Thieme et al [61] address the identification and control of hazards in the APS technology by conducting a Preliminary Hazard Analysis (PHA) while considering a few cyber attack scenarios. Guo et al [62] propose a risk assessment approach based on the Bayesian belief network for assessing the risks of collisions from a safety perspective while considering cyber attacks as a generic failure mode. Another work addresses the remote control facility for providing remote monitoring and controls for urban ferries while briefly discussing communication and cybersecurity aspects as expected challenges [63]. In the remainder of this chapter,

several aspects of the state of the art that is relevant to this thesis are highlighted, namely, attacks, threats and risks, requirement elicitation, communication architectures, and cyber risk management.

3.1 Attacks, Threats, and Risks

Attacks against maritime systems are diverse in their objectives, techniques, and consequences. In recent times, a number of incidents have occurred which have highlighted an immediate threat. These include a ransomware attack on DNV's ShipManager software, which affected 1000 vessels ¹, and a similar attack on the Jawaharlal Nehru Port Container Terminal in India, which led to the port turning away ships to other terminals in the complex near Mumbai ². Additionally, Sembcorp Marine, a Singapore-based specialist in the ship and offshore rig building, suffered a cyber breach that resulted in the exposure of employee data and Sembcorp's operations ³. Research involving 200 industry professionals has been conducted to assess the state of cyber risk management in the maritime industry, revealing that shipping companies pay an average of \$3.1 million in ransom for cyber-attacks and that 44% of the professionals surveyed reported that their organizations had been the subject of a cyber-attack ⁴.

Among the goals of this thesis is to improve cyber risk management by identifying adversarial techniques employed at different stages of cyber attacks on maritime infrastructure. A comprehensive literature review was undertaken to cover all stages of cyber attacks, from reconnaissance to impact, with a particular focus on the maritime context.

In the reconnaissance stage, the use of tools such as OpenVAS and NMAP for gathering information on vessel systems is discussed [64]. Standard et al [65] also mentioned the teaching of network reconnaissance for naval officers during a cybersecurity course for capacity development.

Attack delivery could be achieved through different methods, including the use of USB flash drives [66], compromising the supply chain [67], VSAT TCP session hijacking [68], or tricking users into downloading and executing malicious software [69].

Once access is gained, attackers aim to achieve objectives such as discovery, credential access, and collection. Techniques such as sniffing, vulnerability scanning, and eavesdropping are used for these objectives. Hemminghaus et al [70]

¹DNV, Cyber-attack on ShipManager servers – update, <https://www.dnv.com/news/cyber-attack-on-shipmanager-servers-update-237931> (accessed on 04.04.2023)

²BSM, IFC: Maritime Security Situation Mid-Year Report 2022, <https://www.bs-shipmanagement.com/media-centre/bsm-insight/ifc-maritime-security-situation-mid-year-report-2022/> (accessed on 04.04.2023)

³Rivieram, Sembcorp Marine addresses cyber-security incident, <https://www.rivieram.com/news-content-hub/sembcorp-marine-addresses-cyber-security-incident-72723> (accessed on 04.04.2023)

⁴CyberOwl, Global industry report: The great disconnect, <https://cyberowl.io/resources/global-maritime-industry-report-the-great-disconnect/> (accessed on 04.04.2023)

targeted the network for discovery through sniffing and collection of network traffic, including navigation data. Jo et al [71] categorized vulnerability scanning of ship systems, eavesdropping on Voice over Internet Protocol (VoIP), and Wi-Fi communication in the discovery stage of cyber attacks.

Privilege escalation can be achieved through different techniques, including hijacking execution flow by targeting the operating system library loading mechanism [66]. Also, the infection of stations running all the time with administrative privilege would eliminate the need for escalation. This is particularly relevant to operator stations if they are utilized as the pivot point of attacks [66].

Finally, attacks on maritime operations can involve manipulating sensor messages [66, 70], denial of view based on navigational data [26], alarm suppression for inhibiting response functions, and spoof reporting messages to impair process control [70].

3.2 Requirement Elicitation

Requirement elicitation is an integral process supporting several other processes during the system development. It is relevant to this thesis being conducted at an early stage in the followed research methodology (more details in Chapter 4). Including the security requirement engineering process in early system development phases can increase the trustworthiness of the autonomous vehicles [72] toward defining suitable countermeasures that satisfy security goals. Several works exist regarding requirement elicitation for software and CPS such as autonomous vehicles and vessels. Oladimeji et al [73] developed a threat and risk analysis process for secure software development. The authors proposed a goal-oriented threat modeling approach that can be utilized during the requirement analysis phase to support the subsequent design and implementation processes. In the automotive domain, Islam et al [74] proposed a risk assessment framework be performed in the requirement elicitation phase during the development life-cycle in order to guide countermeasure integration in the design and development phase. The framework is aligned with known automotive processes related to functional safety and usability. Moreover, in a maritime context, Kavallieratos et al [75] proposed a systematic requirement elicitation process for eliciting security requirements for C-ES. Then the author applied the process for eliciting security requirements for the three most vulnerable CPS systems, namely, the AIS, the Electronic Chart Display Information System (ECDIS), and the Global Maritime Distress and Safety System (GMDSS).

Still, during the requirement elicitation stage in this thesis, very limited information existed regarding the APS technology, its scope, expected systems, and relevant regulations and standards. This hindered the ability to conduct a systematic and thorough risk-based requirement elicitation process similar to the aforementioned works. Additionally, eliciting requirements regarding communication technology was needed. Therefore, we referred to eliciting requirements based on stakeholders' communicated guidelines and regulations regarding the autonom-

ous shipping technology [21]. The work is included in this thesis in Chapter I. The APS stakeholders were identified and their requirements were elicited by analyzing relevant documents discussing the stakeholders' viewpoints regarding the technology. For instance, the regulatory perspective emphasizes addressing the safety hazards in the technology while the classification society perspective addresses technical challenges related to system components, services, and design considerations.

3.3 Communication Architectures

The communication technologies utilized in shipping and particularly autonomous shipping have gained considerable attention in recent years. Several works have proposed solutions for ship-to-ship, ship-to-shore as well as internal ship networks. The main challenges these works aim to achieve for sea-going navigation are the high cost associated with satellite communication, the limited data rate, and communication coverage. To mention a few, Lopes et al [76] targeted the provision of low-cost license-free communication for maritime applications by utilizing the 5.8 GHz band using Wi-Fi technology. The authors reported the ability to maintain a 1 Mbps throughput up to 7 km. Ludvigsen et al [77] conducted experiments performed using a group of different autonomous vehicles including autonomous underwater vehicle (AUV), unmanned aerial vehicle (UAV), and Unmanned Surface Vehicle (USV) each with their suitable network component to formulate a network of heterogeneous vehicles operating together to provide seabed mapping of an area. The experiment utilized various sensors (e.g. bathymetry), network devices using the proprietary maritime broadband radio (MBR) technology, and the LSTS software toolchain for situational awareness [78]. Emam et al [79] provide an analysis of Device to Device (D2D) communication on water based on WiFi technology aiming to extend network geographical coverage. The authors studied the behavior of Wi-Fi on the water in a D2D scenario. They suggested two wireless communication methods for D2D, end-to-end and mobile ad hoc network (MANET) and hop-by-hop (OppNet). They conducted field experiments using Cooperative Robotic Watercrafts (CRWs) to measure the coverage, end-to-end throughput, and delay. They concluded that OppNet extends the coverage as double as MANET but at the same time it has 60 times the delay.

Many works related to communication technologies have been surveyed during this thesis. A summary of the surveyed works, the observed communication types, network design, communication technologies, coverage, and data rates are summarized in Table 3.1. Several artifacts from these works have been found useful and were integrated during the development of the APS communication architecture presented in [22] and are included in this thesis in Chapter II. The artifacts include reference architectures, design patterns, candidate technologies, and relevant standards and guidelines. For instance, cellular communication such as LTE has been demonstrated to provide good coverage as shown in Table 3.1, yet, the data rates suggest that it would not be enough for a large amount of sens-

ory data which was expected to send from the APS to the Remote Control Center (RCC). This limitation strengthen the direction toward 5G technology which was later implemented.

Table 3.1: Surveyed Communication architectures for autonomous and traditional maritime systems

Vessels type		Work	Year	Comm. Type	Network Design	Communication Technology	Coverage	Data rate
Autonomous Vessels	1	[42]	2013	S2Sh	SH	Satellite, AIS, VHF, WiFi	Global	4 Mbps (requirement)
	2	[80]	2017	S2Sh S2S	MH, MC	Satellite, 5G, LTE, Wi-Fi, HF Radio	Global	TD*
	3	[81]	2017	S2Sh S2S	MH, MC	Satellite, 5G, LTE, Wi-Fi	Global	TD*
	4	[82]	2017	S2Sh S2S	MH	LTE, Wi-Fi	150 km (2-hop)	3 Mbps
	5	[83]	2017	S2Sh S2S	SH, MH	MBR, Radionor	22.57 km	DL: 2.98 Mbps UL: 1.81 Mbps
	6	[77]	2016	S2Sh S2S	SH, MH	MBR, Acoustic	>22 km	at 22 km DL: 2.98 Mbps UL: 1.81 Mbps
	7	[84]	2009	S2Sh	SH	Wi-Fi, GPRS/UMTS, Satellite	3 km Global	19 Mbps
Traditional Vessels	8	[85]	2019	S2Sh	SH	LTE	100 km	10 Mbps
	9	[86]	2018	S2Sh S2S	SH, MH	LTE, Fiber-In-Motion, MBR	20-50 km **	5-70 Mbps **
	10	[87]	2018	S2Sh S2S	mesh	LTE, Long-Range Wi-Fi	45-66 km **	at 45 km 740 kbps
	11	[88]	2016	S2Sh	SH, MH	LTE	N/A	N/A
	12	[89]	2015		MH	LTE, Wi-Fi	100 km	1 Mbps
	13	[76]	2014	S2Sh	SH	Wi-Fi 5.8Ghz	7 km	1 Mbps
	14	[90]	2012	Ports	Mesh	WiMAX	15 km	5 Mbps
	15	[91]	2010	S2Sh S2S	Mesh, SH, MH	Satellite, Wi-Fi, 3G UMTS	Global	TD*
S2Sh: Ship-to-Shore ; S2S: Ship-to-Ship ; SH: Single-Hop ; MH: Multi-hop ; MC: Multi-Carrier								
* Technology dependent: the paper suggested high level architecture with possible technologies and discussed the expected bandwidth for each one.								
** Depends on location from shore and employed technology								

3.4 Cyber Risk Management

Addressing cyber risks within the context of maritime transportation systems is a recent field of study. The field had attracted some attention focusing on the different maritime components such as ports, ships, and offshore units.

Focusing on risk management at ports, Karantjias et al (2014) [92] have proposed a collaborative security management system (CYSM) that aims to improve safety and security at commercial ports' Critical Information Infrastructures (CIIs). The CYSM provides several services focusing on the dual cyber and physical nature of port systems components to the port operators. The services are categorized

into three main categories. Starting with Risk Analysis Service (RAS) that analyzes both the security and safety issues in a unified way and suggests appropriate countermeasures. The Risk Management Services (RMS) is another service that interacts with the RAS services to produce security requirements and possible solutions. The Document Management Services (DMS) is the graphical and interactive interface of the system that presents and manage the information related to security and safety such as policies and legal, guidelines, and standards documentation. The authors have claimed that some components of the CYSM have been developed and implemented. However, there is no evidence of any evaluation efforts. Grigoriadis et al [93] proposed a group of security controls for improving the cybersecurity of ports including vulnerability risk assessment, communication authenticity through Public Key Infrastructure (PKI), and software hardening including weak password protection, and binary protection. The authors have implemented the risk assessment and evaluated it using stakeholder engagement. Also, they implemented the software hardening and evaluated it using unit testing. Finally, they integrated the PKI solution and engaged stakeholders for its evaluation.

Focusing on risk management in traditional ships, Svilicic et al [94] proposed and conducted a novel cyber risk assessment on board a vessel. In this paper, the authors reviewed the vessel cybersecurity management system, which consists of several security controls, including physical access, patching, and access control. The approach mainly addresses the established risk mitigation measures or the existence of vulnerabilities in the target system. The author utilized a vulnerability scanning tool to identify risks in the vessel ECDIS and engaged the vessel personnel in a questionnaire to assess the existence and awareness of controls. Rajaram et al [95] have outlined guidelines for cyber risk management for shipboard systems, emphasizing operational technology. The authors propose a checklist approach to determining vessels' cyber hygiene. Security tiers, namely low, medium, and high, were introduced as a concept in the approach to align with risk priority levels. Tiers of security reflect the necessity to implement security controls at various levels of risk.

Focusing on cyber risk management in autonomous and cyber-enabled vessels, STRIDE and DREAD (Damage, Reproducibility, Exploitability, Affected Users, and Discoverability) risk analysis techniques were used by Kavallieratos et al [96] to assess cyber risks in cyber-enabled ships, which include autonomous and remotely controlled vessels. To mitigate the identified risks, the authors followed the ISO 31000 risk management process [97]. They relied on the controls suggested by the Guide to the security of ICS [98]. Furumoto et al [99] have proposed a secure network topology for autonomous ship operations after the identification and analysis of a group of attack scenarios. The authors discussed existing typologies of ships, and the types of protocols observed onboard (e.g. CAN bus, NMEA, etc). Then, they discussed security threats such as spoofing and malware infection. Later, they proposed a zone-based network topology and the application of software-defined networking to enable automation and integration of security

services. There exist no evidence regarding the evaluation of the proposed solutions.

In summary, the observed works lacked a clear implementation of the industry trend which is the Defense-in-Depth strategy for ensuring that all layers of defenses are systematically considered. Also, existing works focus mainly on addressing the cyber risk in specific infrastructure components such as vessels and ports. Moreover, efforts for evaluations are usually very limited or nonexistent. Therefore, a systematic risk management approach has been proposed in this thesis starting with risk analysis and assessment toward risk monitoring and treatment. The approach addresses the risk spanning across shore facilities and vessels. It incorporates industrial guidelines and a comprehensive risk assessment method to systematically consider an extensive list of threats and risk mitigation measures toward the development of a cybersecurity architecture. Moreover, various evaluation methods have been implemented to provide evidence of the suitability, feasibility, and comprehensiveness of the approach. The risk assessment approach has been published in [24] and is included in this thesis in Chapter IV. Also, the risk management approach has been published in [25] and is included in this thesis in Chapter V.

Chapter 4

Methodology

This thesis targets a new technology; the autonomous passenger ship (APS) or ferry, which did not exist at the beginning of the project. So, among the first hurdles in this research was to determine “how such a technology that does not exist yet can be developed in a way that is reliable and secure?”. Hence, Design Science Research Methodology (DSRM) was chosen as the research methodology guiding the progress of this thesis. The detailed activities of the DSRM are shown in figure 4.1. This thesis addresses three research questions as discussed in Section 1.2. Each question is considered a problem for which we aim to find answers by applying DSRM within the context of the question. After establishing design requirements for the problem at hand, the design and development of relevant artifacts are conducted resulting in system designs, relevant methods and algorithms, guidelines, and others. Then, the developed artifacts are demonstrated and evaluated. Also, aspects and processes from the system engineering domain related to system development were found relevant and have been integrated within the different stages of the DSRM.

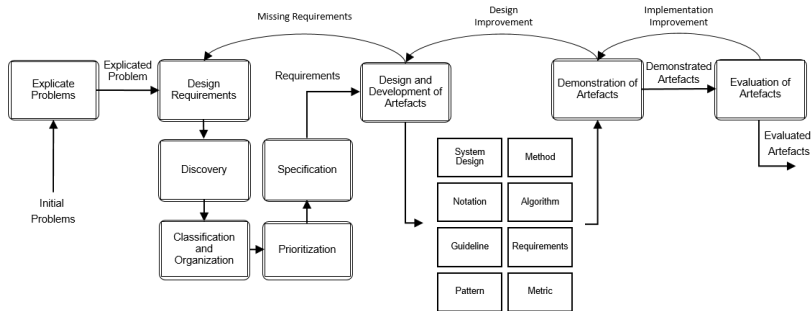


Figure 4.1: Design Science Research Method with requirement elicitation process and artefact types.

The different stages of the DSRM in the context of the first research question are depicted in Figure 4.2. The first bulk of work was related to the communication technologies in the APS context. The APS's application domain is the maritime domain. So, in addition to reviewing the relevant literature in the domain, the consideration of the view of the relevant industrial entities such as the classification societies were deemed necessary. For this, desk research was conducted aiming at collecting and analyzing the different communicated guidelines and publications. In the design requirement stage, an objective was completed by performing a requirement elicitation and analysis process. During this, the stakeholders' requirements related to communication and cybersecurity were discovered, classified, organized, prioritized, and specified. The outcome of this process has been published in [21]. With respect to the design and development stage, the development of a communication architecture has been investigated leading to defining and developing an architecture that satisfies the established requirements and supports the expected functions. The architecture definition, design, and development process followed a group of relevant standards as well as several artifacts in the literature. The outcome of this activity has been published in [22]. Finally, the proposed architecture has been demonstrated and evaluated by utilizing a communication and cybersecurity testbed. The testbed model is a hybrid between physical and virtual components enabling on-site and remote testing capabilities. The physical setting is aimed mainly to test the reliability of the communication architecture functions, namely, Ship-to-Ship, Ship-to-Shore, and limited internal communication functions. On the other hand, the virtual setting aims to test extended internal communication functions as part of the communication architecture as well as security testing capabilities. The development and evaluation of the testbed have been presented in another publication [28]. Further demonstration and evaluation of the proposed communication architecture are undergoing in the form of actual integration into the developed APS prototype named milliAmpere2 [8]. Several aspects of the proposed architecture have already been adopted, integrated, and tested. Yet, the work in this direction has not yet been documented in a research paper. Instead, the influence of the communication architecture on the implementation of the milliAmpere 2 is summarized in Section 5.12. In conclusion, the aforementioned activities were considered sufficient for addressing the first research question allowing the work to proceed to the next step.

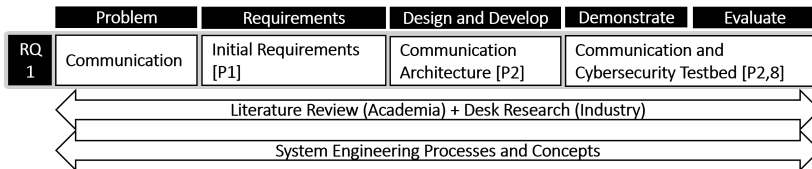


Figure 4.2: The activities during the DSRM stages with respect to the first research question, and included papers.

The different stages of the DSRM in the context of the second research question are depicted in Figure 4.3. The next group of activities targeted the identification and application of a suitable cyber risk assessment approach in order to support subsequent activities related to cyber risk management. This was accomplished in two separate risk analysis processes. First, an empirical study of a joint safety and security risk analysis was conducted utilizing the Six-Step Model (SSM) proposed by Sabaliauskaite et al. [100]. SSM was utilized to identify interrelations between safety and security risks as well as synergies between safety and security countermeasures. The outcome of this process has been published in [23]. Nevertheless, the objective to conduct a risk assessment process was not yet achieved. Therefore, a comprehensive literature review of risk assessment approaches for CPS was conducted and led to the proposition of a risk assessment method. The proposed process follows a Failure Mode, Effects & Criticality Analysis (FMECA) [101] to allow the assessment of cybersecurity threats. The threats and their analysis utilize the semantics in the MITRE ATT&CK framework [45] to reduce expert judgment and assist in automating the risk assessment process. The impact analysis utilizes graph theory centrality metrics to estimate certain impact factors in order to reduce the impact of biased estimation. The risk assessment approach has been demonstrated through application to identify risks against the proposed communication architecture. The original work proposing the risk assessment approach has been published in [24]. The approach was further demonstrated and evaluated through empirical application against a model of the milliAmpere2 APS prototype [25] and an Integrated Navigation System (INS) [30] leading to the identification of a group of cyber risks. Moreover, the risk assessment approach has been demonstrated and evaluated through expert engagement allowing the identification of limitations and areas for future improvements. The evaluation efforts are presented in [29].

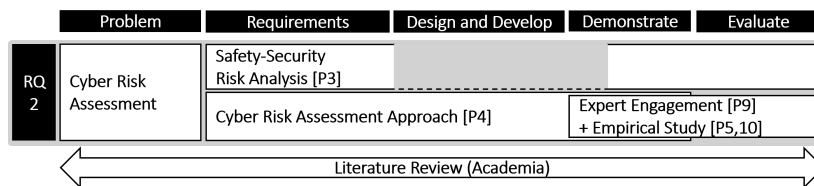


Figure 4.3: The activities during the DSRM stages with respect to the second research question, and included papers.

The different stages of the DSRM in the context of the third research question are depicted in Figure 4.4. In this direction, suitable risk management approaches are investigated through a comprehensive literature review toward the development of a suitable risk management architecture for the APS. This has led to the proposition of a new risk management approach that builds upon the proposed risk assessment and integrates elements from the DiD security design pattern. The

requirements identified in the earlier work [21] include a group of relevant cybersecurity requirements for this task. Additionally, results from the risk analysis [23] and assessment [24] processes are utilized to support the development process. Then, in the design and development stage, the architecture definition, design, and development follow a system engineering process similar to the communication architecture development but utilizing different literature artifacts. Later, a systematic literature review has been conducted to identify suitable methods for evaluating the proposed risk management approach and architecture. Then, several evaluation methods were applied. This included, adversary emulation in the simulated network within the aforementioned communication and cybersecurity testbed [28], in addition to a comprehensive adversary emulation process against the milliAmpere2 APS prototype leading to the identification of a series of vulnerabilities and cybersecurity issues in the real ferry. The outcome of these activities has been published in [25].

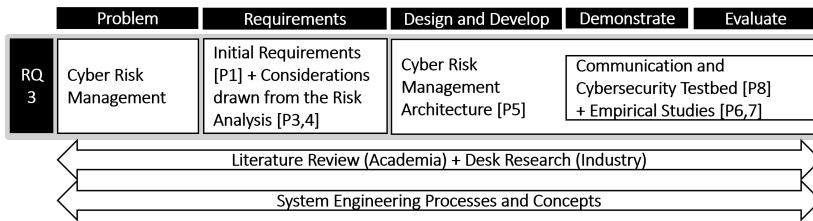


Figure 4.4: The activities during the DSRM stages with respect to the third research question, and included papers.

The proposed risk assessment and management approaches have led to the identification of new risks and areas that require increased attention related to risk management. Therefore, the work in this thesis has explored a group of domain-specific attacks that materialize and demonstrate some of the identified risks, and then discuss relevant mitigation measures to be included in the cyber risk management architecture. Two empirical studies were explored in this direction. The first one explores the field of navigation data anomaly detection and analysis [26] while the second focus on the investigation of covert channels [27]. In [26], a systematic anomaly analysis process was proposed to identify anomalies, attacks, and relevant mitigation measures in a prominent navigation protocol named “NMEA”. The identified anomalies were then demonstrated through a developed tool that can automate relevant attacks targeting the availability and integrity of the navigation data. Then, relevant anomaly detection approaches were investigated to identify suitable ones. On the other hand, in [27], the utility of the AIS for covert command and control channels was investigated. This was an identified risk through the risk assessment process and after exploring the literature for relevant works, it was identified as an area of limited coverage. The investigation has led to the development and demonstration of a proof of concept attack utilizing

real AIS devices. Then, the identified covert channel was integrated into a complete cyber kill chain against the APS technology to be utilized for cybersecurity evaluation.

The employed research methods were chosen over others because they are appropriate for addressing the research questions and provide a comprehensive approach to developing and evaluating the APS technology. The combination of DSRM, desk research, requirement elicitation and analysis, design and development, testbed model, comprehensive literature review, and empirical studies allows the author to not only develop the APS technology but also assess its communication, cybersecurity, and risk management aspects.

These research methods are considered sufficient as they address each research question, enabling the work to progress from one stage to another. By following a structured approach, the author is able to design, develop, and systematically evaluate the APS technology, while also considering the feedback and input from relevant stakeholders and experts. This ensures that the resulting technology is reliable, secure, and relevant to the needs of the industry.

Chapter 5

Summary of Papers and Contributions

This chapter summarizes the objectives, methods, findings, and contributions of each paper included in this thesis.

5.1 Article 1: Connect and protect: requirements for maritime autonomous surface ship in urban passenger transportation

In this paper [21], the aim was to determine the scope of the autonomous ferry as a technology. Due to the lack of sufficient relevant work in the literature, desk research was conducted to investigate the relevant laws, communicated guidelines, and technical reports. The targeted ship class was defined as an Autonomous Passenger Ship (APS). Additionally, the APS operational context, stakeholders, related regulations, related standards, and expected functions were specified. Finally, a list of communication and cybersecurity requirements was established towards designing a secure communication architecture suitable for APS. The list of requirements constitutes the main contribution of this paper. The full article is included in Chapter I.

5.2 Article 2: Communication architecture for autonomous passenger ship

In this paper [22], the aim was to establish a communication architecture that satisfies the pre-established requirements and supports autonomous and remotely controlled functions of an APS. The architecture was developed by following a system engineering approach to define the concepts and the system components. The proposed architecture was verified by showcasing the role of the different architectural components in addressing the requirements and in supporting the

expected functions in a number of operational scenarios based on the expected operations of an APS use case called "Autoferry". Furthermore, the proposed architecture has been evaluated by demonstrating its ability to achieve the expected performance according to the requirements, in simulated experiments using the network simulator GNS3. The contribution of this study is the communication architecture which will constitute a use case for the consequent risk management tasks. The full article is included in Chapter II.

5.3 Article 3: Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship

The aim of this work [23] was to study the safety and cybersecurity of the communication architecture. The SSM was utilized to facilitate joint analysis. The application of the SSM enables, among others, the capturing of relationships between cyber attacks and component failures, the assessment of safety and cybersecurity countermeasures, as well as, the synergies between them. It has been found that most countermeasures in both categories are reinforcing or are conditionally dependent on each other, while few antagonize each other. These findings will allow for improved design and implementation of the cybersecurity architecture. The main contributions of this paper are:

- A new application of the SSM model for joint safety and security analysis of autonomous ship systems.
- A metric for estimating the criticality of assets to the system functions.
- A metric for estimating the safety impact of a system failure.
- A group of cyber threats targeting the milliAmpere2 ferry elicited using the STRIDE threat modeling method.
- The results of the risk analysis are contributions to future efforts in developing cybersecurity architecture.

The full article is included in Chapter III.

5.4 Article 4: Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework

This work [24] addressed the identification and proposition of a suitable risk assessment approach for assessing the cyber risks in the APS ecosystem. A comprehensive literature review has been conducted exploring relevant risk analysis and assessment approaches for CPS. The observed approaches were found highly dependent on expert judgment and mostly not comprehensive in their coverage of risk analysis elements such as threats, mitigation methods, technologies, etc. Therefore, a new cyber risk assessment approach was proposed aiming to reduce the need for expert judgment, reduce the impact of bias estimation, and automate several stages of the risk assessment process to support its continuity. The

proposed approach is based on an FMECA and integrates the semantics and information encoded in the MITRE ATT&CK framework. Additionally, likelihood estimation follows the Common Vulnerability Scoring System (CVSS) while impact estimation relies on a combination of expert judgment and centrality metrics from graph theory. The main contribution of this paper is a new approach for assessing cyber risks in CPS, a demonstration of the approach against an interesting new use case that is the APS, and a semi-automated open source implementation. The full article is included in Chapter IV.

5.5 Article 5: Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth

This work [25] investigates the relevant cyber risk management approaches in the maritime domain, discusses the development of a cybersecurity architecture that supports cyber risk management, and investigates the existing approaches for cybersecurity evaluation in the domain. Firstly, a comprehensive literature survey was conducted in addition to analyzing communicated guidelines related to cyber risk management by industrial entities. This has led to the proposition of a new cyber risk management process that supports existing ones toward identifying implementation gaps. The approach combines two cybersecurity strategies, namely, threat-informed defense, and defense in depth. Then, a system engineering technical process is followed for the development of a cybersecurity architecture that addresses the communicated requirements, and the identified risks as well as considers the relevant elements of cyber risk management identified in the literature and communicated by the industry. Later, a Systematic Literature Review (SLR) was conducted to stand on the state of the art of cyber security evaluation in the maritime domain in order to identify suitable methods to evaluate the proposed architecture. Then, several evaluation approaches were conducted including simulation, checklists, and adversary emulation. The contributions of this paper can be summarized as follows:

- a new risk management approach named TIDiD that combines components from two cybersecurity strategies, namely, Threat Informed Defense, and Defense-in-Depth.
- a cybersecurity architecture for APS that is an outcome of the TIDiD approach.
- The results of an SLR regarding cybersecurity evaluation in the maritime domain highlight different aspects and approaches.
- The results of the conducted cybersecurity evaluation processes for an operational ferry that is a prototype implementation of an APS.
- Challenges observed in carrying cyber risk management functions in the context of the autonomous and remotely controlled operational mode of the APS.

The full article is included in Chapter V.

5.6 Article 6: Navigation Data Anomaly Analysis and Detection

In this work [26] the field of navigation data security in the maritime domain was investigated. Particularly, the NMEA message-based protocol for marine communication was studied. A comprehensive literature review captured the state of the art of navigation data security. A lack of systematic approach for analyzing the anomalies in NMEA protocols was observed. Therefore, a systematic anomaly analysis process has been proposed and applied against two use cases, a traditional INS and the APS. For each use case, the different types of navigational functions and the relevant types of NMEA messages were identified. Also, the types of expected anomalous patterns were defined within a certain anomaly categorization. Several attacks with several variations were identified and implemented to cause the identified anomalies. Finally, several detection methods were proposed. Specification-based anomaly detection has been demonstrated to provide detection capability for a wide range of anomalies. The contributions of this work can be summarized as follows:

- A novel systematic approach for anomaly detection in NMEA messages.
- An analysis of possible anomalies in NMEA messages, their cause-and-effect relationship, and a range of cyber-attacks against them.
- A method for creating synthetic datasets with both normal and maliciously tampered with NMEA messages.

The full article is included in Chapter VI.

5.7 Article 7: From Click to Sink: Utilizing AIS for Command and Control in Maritime Cyber Attacks

The idea for this work [27] was inferred after the risk assessment process identified a risk of command and control (C&C) objective related to the AIS. This has motivated us to investigate the threat landscape in the maritime domain and assess its coverage of the different adversarial objectives. For this, a comprehensive literature review has been conducted to evaluate the literature related to maritime cybersecurity regarding their coverage of adversarial behaviors. The ATT&CK framework has been utilized for mapping the related works with the discussed adversarial behaviors. This study has suggested a lack of discussion of the C&C as an adversarial objective. Therefore, an empirical study was conducted to evaluate the utility of AIS for covert channels used in C&C. A threat model was developed, and a proof of concept was implemented and evaluated. Later, two complete attack scenarios against an APS were discussed to demonstrate the utility of the discussed covert channels. The contributions of this work can be summarized as follows:

- An overview of the maritime threat landscape is presented covering the dif-

ferent adversarial behaviors.

- A Threat model, proof of concept implementation and evaluation of an AIS covert channel.
- Two realistic attack scenarios to be utilized for cybersecurity evaluation in maritime systems with similar technologies.

The full article is included in Chapter VII.

5.8 Article 8: Communication and Cybersecurity Testbed for Autonomous Passenger Ship

The aim of this paper [28] was to develop a maritime-themed testbed that can be used to evaluate and analyze several maritime use cases, including the APS, with a focus on communication and cybersecurity. The development of the testbed was conducted following a series of processes described in ISO 15288 based on a literature review of relevant aspects and artifacts. The high demand for maritime-themed testbeds and cyber ranges focusing on cybersecurity and communication has been observed. The testbed has been utilized during the evaluation of the proposed communication and cybersecurity architectures as well as the development of anomaly analysis and detection capabilities in [26] (Discussed in Section 5.6). The testbed has been evaluated based on the observed aspects in relevant testbeds and cyber ranges found in the literature. The contributions of this work can be summarized as follows

- A communication and cybersecurity testbed suitable for several maritime use cases. The testbed hosts comprehensive capabilities compared to the state-of-the-art in the domain.
- A abstraction of three processes that can be followed during the utilization of cybersecurity testbeds namely, system replication, system analysis, and technical management.
- 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.

The full article is included in Chapter VIII.

5.9 Article 9: Evaluation of a Cyber Risk Assessment Approach for Cyber-Physical Systems

In this article, [29], a comprehensive form of evaluation is proposed and conducted for the FMECA-ATT&CK risk assessment approach proposed in [24] (Discussed in Section 5.4). A group of experts was involved to conduct two risk assessment processes, namely, FMECA-ATT&CK and Bow-Tie against two use cases in different application domains, particularly, an APS and a digital substation. This allows for

the evaluation of the approach based on a group of characteristics, namely, applicability, feasibility, accuracy, comprehensiveness, adaptability, scalability, and usability. The evaluation of FMECA-ATT&CK demonstrated positive ratings with respect to the considered characteristics and identified a group of limitations as directions for future work. The contribution of this article can be summarized as follows:

- An evaluation of an open source risk assessment process that is FMECA-ATT&CK supporting its development as a semi-automated cyber risk assessment tool for CPS.
- Key characteristics for the evaluation of risk assessment methods. These characteristics can be utilized as a basis for comparison among existing and newly proposed methods for risk assessment.
- A standard-aligned methodology for the evaluation of risk assessment methods. The methodology allows for the evaluation according to a group of characteristics while reducing the impact of bias.

The full article is included in Chapter IX.

5.10 Article 10: Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework

This article [30] aims to empirically evaluate the applicability of the FMECA-ATT&CK risk assessment method [24] in assessing cyber risks in the INS; a system found on modern vessels. FMECA-ATT&CK has been extended in this work to be suitable for systems onboard ships. Then, the extended method has been applied for assessing the risks in 25 components. In total, 1850 risks were identified. Among those, 32 were classified as medium, 9 as high, and 4 as critical. The applicability of FMECA-ATT&CK has been demonstrated in this article with some needed adaptations. The adaptation also reflects the flexibility and adaptability of the method. The contributions of this article can be summarized as follows:

- Extending the impact model in FMECA-ATT&CK to assess the reputation and environmental consequences.
- First-of-a-kind complete cyber risk assessment of 25 components constituting the INS.

The full article is included in Chapter X.

5.11 Thesis Impact

The work conducted during the development of this thesis has caused a noticeable and promising impact. The impact is discussed in several directions, namely, cybersecurity education and awareness, system development, and research.

Regarding cybersecurity education and awareness, several deliverables of this

thesis have been integrated into university lectures and tutorials for master-level courses related to cybersecurity as well as webinars for industrial institutions. Additionally, a group of master projects has been executed with several ties to the work conducted in this thesis. Skarshaug [102] utilized a simulated IP network of the communication architecture in [22] in cybersecurity testing. Ruud [103] replicated a group of attacks discussed in [26] for the demonstration of vulnerabilities in INS. Solli [104] followed a similar methodology in [28] for developing a testing environment for automating cyber attacks in maritime settings. Finally, Del Riego [105] followed a similar methodology in [24] for the assessment of IT infrastructure in health care.

Regarding system development, the work in this thesis has contributed to the development of the autonomous ferry milliAmpere2 and the Shore Control Lab (SCL) in the city of Trondheim (more details in Section 5.12). The author was leading the activities related to building the communication link between the ferry and the SCL which was successfully demonstrated and utilized in remote monitoring and sensor data acquisition during a 3-weeks public trial of the ferry [106].

Regarding research, the work in this thesis has generated ten publications at national and international peer-reviewed conferences and journals. Several opportunities for collaboration have been created during and due to this thesis. Additionally, the thesis has identified a wide range of research areas that are under investigation, this is expected to pave the way for future research (more details in Section 6.2).

5.12 Communication Architecture Implementation

Several aspects of the communication architecture design proposed in [22] and briefly discussed in Section 5.2 were found relevant during the implementation of the autonomous ferry milliAmpere2. The author of this thesis was closely involved in the development process and was responsible for several implementation tasks. Table 5.1 summarises the influence of the proposed communication architecture on the implementation. This is reflected by the implementation of various architectural components some by the author himself and others with the consultation of the author. Most of the proposed components and features found their way into implementations while others are proposed for future work. This supports the author's claims regarding the suitability of the proposed communication architecture for APS technology.

Table 5.1: APS Communication Architecture influence on Implementation

Component Group	Function	Architectural Component		Author Responsible	Author Consulted	Future Plans	
		Proposed	Implementation				
APS	Internal Communication	The core/distribution tier	Primary and backup networks are implemented. Primary network allows for remote access similar to the proposed part (A) while the backup network is for emergency and only allow local access		Y		
	Ship-to-Shore-Communication	Mobile Communication Module (MC Module)	5G router	Y			
	Ship-to-Ship-Communication	APS-RCC Module	Mesh Radio	Mesh Radio, implemented in the same module as the APS-RCC Module. The same access point connects to the RCC and an close by ECT.		Y	
		Traffic Module	AIS Transceiver				
	Emergency Communication	Emergency Modules - Emergency control					
	Emergency Communication	Emergency Modules - Passenger Push Button	Not implemented		Y	Y	
	Autonomous Navigation	Time and Positioning Modules	Two GNSS systems, a main and a backup. In addition to an RTK receiver.		Y		
	Autonomous Navigation, and Autonomous Engine Monitoring and Control	Autonomous Ship Controller (ASC)	Not implemented as a functional component. But some of its subcomponents are implemented.		Y	Y	
	Autonomous Navigation	The Autonomous Navigation System (ANS)	Merged together in a single functional component. Built on top of ROS operating system.		Y		
	Autonomous Engine Monitoring and Control	he Autonomous Engine Monitoring and Control (AEMC)					
	Network and System Management	Network and System Management	Not implemented as a functional component. But some of its subcomponents are implemented.		Y	Y	
	Storage	Digital Logbook	Not implemented				
	Network and System Management	Domain Controller	Not implemented		Y		
		System and network documentation repository	Partially implemented using distributed documentation.				
	All communication functions	Connectivity Management (CM)	Not implemented as a functional component. But some of its subcomponents are implemented as distributed capabilities.		Y		
		Quality of Service Controller (CM-QoS)	Implemented in the 5G router		Y		
		Network Monitor and Troubleshooter (CM-NMT)	Manual network monitoring and troubleshooting		Y		
Network Software Updater (CM-NSU)		Manual periodic update of some of the network devices.		Y			
Network Segmentation Manager (CM-NSM)		The APS and RCC network is segmented.		Y	Y		
Network Security Coordinator (CM-NSC)		Manual and distributed configurations related to cybersecurity are implemented in the APS and RCC networks.		Y			
Network Device Backup Controller (CM-NDBC)		Manual backup of some of the network devices		Y	Y		
Autonomous Navigation	Navigation Systems	Lidars, cameras, and radars are interfaced with the network through junction boxes and switches					
Autonomous Engine Monitoring and Control	Machinery Systems	Redundant dynamic positioning systems with I/O cards interfacing the thrusters with the network					
RCC	Internal Communication, Ship-to-Shore Communication	Network	RCC is implemented on a remote facility, not a vessel. 5G router and mesh radio allow two communication links with the ferry. A single network tier is implemented on the RCC.	Y			
	Remote Navigation and Remote Engine Monitoring and Control	Remote Ship Control (RSC)	Implemented as single workstation		Y		
	Remote Navigation	Remote Navigation System (RNS)	Implemented as part of the RSC. A backup component is planned.	Y	Y		
	Remote Engine Monitoring and Control	Remote Engine Monitoring and Control (REMC)	Not implemented		Y	Y	
ECT	Emergency Remote Control and Emergency Remote Navigation	Emergency controller	Implemented as a dedicated component		Y		
	Passenger Safety	Emergency Alarm and Response system	Not implemented		Y	Y	
Cloud Component	Communication, Cybersecurity, and others	Several components	Several cloud-based components exist for communication, battery management, and cybersecurity.	Y	Y		
Mobile Network	Ship-to-Shore-Communication	-	A dedicated Access Point Name (APN) is implemented for the APS-RCC link	Y			
Shore Sensor System (SSS)	Docking and Charging	-	A group of sensors are distributed on shore for docking and charging. However, they're implemented as isolated systems				

Chapter 6

Limitations and Future Work

In this chapter, the identified limitations of this thesis are discussed. Also, directions for future work are highlighted.

6.1 Limitations

In this section, the limitations that are related to this thesis are discussed.

6.1.1 Related to Article 1: Connect and protect: requirements for maritime autonomous surface ship in urban passenger transportation

The followed approach for the requirement elicitation is based on stakeholders' perspectives elicited through analyzing the communicated laws, regulations, and guidelines. This approach might have limited the identification of system-specific requirements which can be identified through other systematic approaches such as the work of Kavallieratos et al [75]. Also, some of the established requirements communicated by the stakeholders lack qualitative or/and quantitative metrics to sufficiently verify their satisfaction with the architecture design. Examples of such limitations:

1. No verification metrics for reliability.
2. No Quality of Service (QoS) requirements have been defined to verify the satisfaction of certain communication requirements.
3. Measurable metric for redundancy is not provided.

Efforts to overcome these limitations were made by formalizing design-level and implementation-level verification metrics of the requirements as well as a suitable verification method. These verification metrics and methods were utilized during the evaluation of the architecture. Furthermore, additional system-specific considerations were identified through the comprehensive cyber risk management approach which provided a detailed description of the required risk controls for each system component based on the identified risks.

6.1.2 Related to Article 2: Communication architecture for autonomous passenger ship

Although the APS system functions and expected operations have been previously defined, the APS systems are still under development. This limited the ability to customize design decisions that would be more suitable in the future APS and confined the architect (myself) to best-judgment decisions based on previous experience, discussions with other project members, and future expectations discussed in the literature. Also, additional architecture analysis methods such as technical risk analysis, trade-off studies, cost analysis, usability, dependability, and maintainability analysis were deemed out of the scope of this task.

Regarding the evaluation of the architecture, limited simulation capabilities exist to simulate heterogeneous networks consisting of IP and Non-IP components. Because of this, we were unable to verify the proposed non-IP components using simulation. Therefore, we utilized scenarios to conceptually demonstrate the functionality of Non-IP components.

6.1.3 Related to Article 3: Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship

Due to the fact that the domain of autonomous ships is recent, no historical statistics are available for quantitative and more informed risk analysis. Additionally, since the APS systems are yet under development, the identification of realistic and comprehensive attack scenarios was challenging. In order to overcome these challenges, during the joint safety and security risk assessment we relied on judgment based on experience in the field. We attempted to improve the validity of this study and reduce the effect of bias by following guidance from the Risk Management standard NEK/IEC 31010:2019. A brief summary of the performed tasks is listed below:

- Different teams were organized during different steps of the analysis. The members' expertise was considered in the team organization to ensure the diversification of teams.
- The judgments that were based on assumptions are reported with comments that are later checked by another team member.
- Opportunities for team members to work individually for part of the time were provided to avoid convergent thinking.

Moreover, a more comprehensive, model-informed, and tool-assisted risk assessment method has been proposed in another work [24] focusing on the cyber-risks and including the cyber and safety impacts.

6.1.4 Related to Article 4: Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework

The proposed risk assessment approach FMECA-ATT&CK has been thoroughly evaluated throughout the duration of the thesis since its proposition. Some limitations were identified such as the component-level nature of the assessment which does not provide insights into the propagation of risks within the evaluated system. Other issues related to the categorization of components, difficulties in estimating the safety and financial impacts of cyber threats, the focus on adversarial threats, and exclusion of benign threats. We have discussed those limitations and provided suggestions for future work.

6.1.5 Related to Article 5: Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth

The risk management approach includes a defense strategy comparison algorithm for the initial assessment and comparison of different risk management strategies. The comparison is only based on risk reduction without considering other aspects such as the cost and feasibility of implementation. Furthermore, the calculation relies on the ATT&CK framework's controls. Some controls suggested in relevant guidelines do not map to a clear control in the ATT&CK framework. Thus, some controls do not account for reducing risk in the developed strategy comparison algorithm. Moreover, only a few tests against the ferry were allowed from a complete list of tests to constitute a comprehensive evaluation. Still, some vulnerabilities and areas for improvement were identified. Future works can investigate solutions for less intrusive yet comprehensive testing. Finally, the recovery and response functions are limited in the proposed risk management activities. This is related to the low technology readiness level of the APS technology. The different systems and involving human roles are still subject to change which restricts the expected suitability of incident response and recovery plans.

6.1.6 Related to Article 6: Navigation Data Anomaly Analysis and Detection

A detailed list of limitations can be found in the article in Section 6.4 in Chapter VI. In summary, the number of analyzed messages, types of attacks, and considered protocols are not comprehensive. Covering the entire scope would require a lot of time and effort and is a candidate topic for many future works.

6.1.7 Related to Article 7: From Click to Sink: Utilizing AIS for Command and Control in Maritime Cyber Attacks

The developed proof of concept for the AIS covert channel is based on the practical application of specific AIS transceivers from a certain manufacturer. This might not reflect the practical feasibility of the covert channel if other transceivers are

involved. Another investigation approach could have been followed by contacting several AIS manufacturers and querying their input regarding the feasibility of the covert channel.

Additionally, the assumption of pre-existing infection allowing the allocation of an agent node in order for the covert channel to be established has been found to be a very controversial assumption. The rationale is based on the premise that if malicious software already exists in the ferry, many undesired scenarios might occur that do not require the covert channel to exist, which can be true. Still, the covert channel does provide additional C&C capabilities and flexibility in the malware development for the attackers that would not exist without the covert channel. This is argued to increase the cyber risks associated with the application of AIS in maritime systems.

Moreover, the stealthiness of the covert channel might be affected by the geographical location of the testing procedures. There is very limited maritime traffic in the area where the tests were carried out. Other areas witnessing more dense traffic might involve active vessel traffic services which might be able to detect abnormal utilization of the message type used for the covert channel. Thus, a practical evaluation of the channel in other geographical areas and involving other AIS transceivers brands is needed to assess the practicality of the covert channel.

6.1.8 Related to Article 8: Communication and Cybersecurity Testbed for Autonomous Passenger Ship

This article is intended to provide a novel introduction to a maritime-themed testbed in a domain that lacks such testing capabilities. The article focused mainly on the existing capabilities and procedures for utilizing the testbed. However, the evaluation approach is limited to a group of use cases. Additionally, although the scope of the testbed is maritime systems, the utility of well-established tools utilized for cybersecurity testing in other domains (e.g. Metasploit) was not thoroughly investigated in this article. This could have led to discovery of novel attack scenarios.

6.1.9 Related to Article 9: Evaluation of a Cyber Risk Assessment Approach for Cyber-Physical Systems

The main observed approach in the literature for evaluating risk assessment methods is the engagement of experts and stakeholders. This article follows the same approach for evaluating the FMECA-ATT&CK risk assessment method based on key characteristics. Among these characteristics is the applicability of the method in different application domains. For this, the utilized use cases are related to the maritime and energy domains. This restricts the measured applicability only to these sectors. Additional use cases are needed to measure the applicability of the method in other domains. Additionally, the number of involved participants in the evaluation was less than what was originally intended which might have reduced the quality of the assessment.

6.1.10 Related to Article 10: Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework

Due to the lack of other cyber risk assessment results for the INS, the accuracy of the obtained results rely on the judgment of the experts conducting the assessment. Additionally, the INS included components that are not compatible with the classification criteria proposed in FMECA-ATT&CK, this hindered the identification and consequently the assessment of risks associated with them. Moreover, due to the component-level nature of FMECA-ATT&CK, risk propagation was not considered.

6.2 Future research

The work in this thesis has merely scratched the surface of the expected challenges related to the communication and cybersecurity aspects of the APS technology and other relevant maritime and cyber-physical systems. This section introduces directions for future work based on the gaps identified during the development of works included in this thesis.

6.2.1 Communication Technologies for the operations of autonomous vessels

The developed communication architecture has been designed to be flexible, modular, and extendable. It can support a wide range of use cases due to its design. The implementation of the proposed communication architecture in future use cases of autonomous passenger transportation is a possible future direction. Also, new technologies are expected to be introduced into maritime operations including the VHF Data Exchange System (VDES) and satellite internet constellations such as Starlink. The utilization of APS technology in diverse transportation solutions is envisioned in the future. This would require investigating the suitability of the new technologies for that purpose. This includes empirical studies to measure the range, data rates, and latency for each technology considering the high bandwidth requirements for the APS technology to be remotely monitored. Moreover, investigating the cyber risks associated with each communication technology is a topic worthy of investigation. Limited research exists regarding the security of VDES and satellite internet constellation.

6.2.2 Maritime Cyber Risk Management

Cybersecurity in the maritime domain in general and specifically in autonomous ships is a recent field of study. The heterogeneity of the maritime domain makes its security a challenging task, some consider cybersecurity as the biggest challenge facing the undergoing digital transformation in maritime [107]. The results of this thesis partially contribute to a vast scope of possible use cases and challenges. The

main identified directions for future work related to cyber risk management can be summarized hereafter:

Integrated safety and security risk management

Toward addressing the recent calls to include cyber risk management within safety management systems, investigating the convergence and integration of safety and security risk management in a systematic and standard-aligned manner is an important direction for future work. Considering a system development perspective, this can start from the early stages of requirement elicitation, system design, and implementation until system disposal. A direction could be to study the existing safety and security procedures with the intent to identify supporting and conflicting relations and manners to combine.

Continuous Cyber Risk Assessment

This thesis addressed the identification and development of a semi-automated cyber risk assessment method aiming to support the continuity of the risk assessment process by reducing efforts and the need for expert judgment. Although the proposed approach has been demonstrated to support this goal, several limitations were identified and have been previously discussed. Future efforts will address those limitations. This includes the consideration of benign and adversarial threats, improved asset classification, additional technology domains, improved safety, financial, and environmental impact estimation criteria, and extended risk mitigation measures. Additionally, the proposed method is suitable during system design. Investigation of the adaptability of the method across the system development life cycle is another possible future direction.

Simulation and emulation toward digital twins

The utilization of simulation and/or emulation platforms has proven its utility in the development of autonomous vehicles and vessels. Future efforts to develop virtual platforms specifically for cybersecurity and safety testing are required. Digital twin technology would have clear advantages in this domain. Investigating the cyber security of such digital twins is an interesting topic for research as well as investigating the utility of digital twins for cybersecurity.

Artificial Intelligence

The domain of AI and big data in maritime is a standalone research domain due to the high degree of authorship concentration and exponential growth of publications in a wide range of applications [108]. The European Commission for technical robustness and safety [109] already paved the way for the development of trustworthy AI. Their guideline suggests three components of trustworthy AI,

namely, lawful AI, ethical AI, and robust AI. Research related to AI and ML considering legal and cybersecurity issues in maritime is not sufficiently explored yet and is identified as a venue that requires more attention [108]. This field requires extensive analysis to increase the trustworthiness of the maritime systems of the future. A possible direction for future work is to investigate the development of trustworthy AI in the maritime domain for a resilient and secure maritime ecosystem toward maintaining the well-being, prosperity, and security of our society. This could include investigating the threats against ML and AI; also known as adversarial machine learning as well as studying the utilization of ML and AI for cybersecurity in a maritime context.

New Sensory Systems

The APS as a technology is envisioned to rely on a wide range of sensory systems including video cameras, radar, lidar, GPS, and others. Additionally, a unique composition of protocols and operating systems has been witnessed in the development of the APS prototype named milliAmper2. This includes the utilization of NMEA protocol, AIS protocol, and Real Time Streaming Protocol (RTSP) side by side with the Robot Operating System (ROS). This thesis has addressed NMEA and AIS protocols in sufficient depth. However, the cybersecurity of other systems and protocols not covered in this thesis is worthy of investigation. This could be conducted through scoping maritime-specific intrusion detection and Security Incident and Event Monitoring (SIEM) systems.

Incident Response Plans

The geographically distributed and dispersed teams' nature of the APS technology and its unconventional mode of operations challenges existing timely safety-critical cyber incident response methods. A future direction could be the improvement of incident and cybersecurity readiness for multi-vessel operations. This includes the technologies and procedures to coordinate distributed teams and technologies across infrastructures (e.g., ferry, remote control center, shore infrastructure) for the development of incident response and recovery plans as well as cybersecurity exercises.

Security of the Supply Chain

The ferry includes components from multiple providers spanning various technologies including traditional ICT, OT, marine technology, as well as ML and AI. These systems are interfaced and integrated to collectively provide the functions desired by the system. Very limited room was left for including cybersecurity considerations during the integration process. An effort to integrate cybersecurity within the supply chain of autonomous maritime systems is an important future direction.

Another aspect is related to the degree of interconnections between the ferry infrastructure as a maritime transportation system with other transportation modes such as sea-going ships and land-based transportation such as buses and trains. This is an envisioned stage for the technology in the coming years. Establishing secure interfaces with the different systems in the different sectors falls within the efforts related to the security of the transportation service supply chain.

6.2.3 Forensics Readiness in the Maritime Sector

The complexity of the maritime domain extended from the sophistication of maritime systems, distributed nature, international voyages, and many other challenges to the forensics readiness in this sector. An interesting future direction is to investigate the readiness of maritime systems and processes for forensics investigations when cyber incidents might be involved. Utilizing the APS as a use case for such research would be a clear advantage as this advances the field toward futuristic autonomous vessels.

Chapter 7

Conclusions

Technological advances and new modes of operation have emerged from the ongoing digital transformation in the maritime sector. As an example, Autonomous Passenger Ships (APS) are being proposed for inland transportation. It is anticipated that the APS technology will operate in an auto-remote mode; autonomous when possible and controlled remotely when necessary. Addressing the technical challenges related to the communications and cybersecurity of the APS is the main premise of this thesis.

APS relies on many interconnected components to transport passengers in urban water channels safely and securely; a communication architecture that connects these components is needed. Answering the question, “what communication architecture should be defined for the APS” constitutes the first objective this thesis is addressing. Additionally, in the wake of recent calls for cyber risk management processes to be implemented in the maritime domain, investigating ways to provide cyber risk management functions for the APS is another need. In this regard, the second and third objectives are defined for this thesis. The second objective is to answer the question, “what risk assessment method is most suitable for assessing the cyber risks of the APS” while the third objective is to answer the question, “What is a suitable approach for managing the cyber risks against the APS”.

To provide answers to the aforementioned questions, while at the same time lacking a clear vision of the APS as a technology, the design science research methodology (DSRM) was followed. Starting with defining the problems, identifying design requirements, then designing and developing artifacts for addressing the problems, and later demonstrating and evaluating the developed artifacts. Also, several system engineering processes related to architecture and design definition, and system analysis were found relevant and were integrated into the different stages of DSRM.

Regarding the first objective, a suitable communication architecture for the APS needs to consider the domain-specific requirements communicated by the different relevant stakeholders. At the same time, since the technology is recent and is subject to changes in relevant regulations and technology, a flexible, mod-

ular, and resilient design is needed. Modularity is demonstrated by defining a group of architectural modules for each expected function. Additionally, flexibility is demonstrated through the proposition of different suitable technologies for each architectural component. Furthermore, resiliency is addressed by defining a suitable hierarchy and redundant components. The proposed architecture was verified concerning addressing the identified requirements and found to be sufficient. Moreover, the ongoing adoption of several aspects of the proposed architecture in the real APS prototype named milliAmpere 2 is strong evidence to support the claims in this thesis.

Regarding the second objective, a suitable risk assessment method for the APS needs to adhere to the domain requirements which includes the consideration of the relevant technology domains including Information Technology (IT) and Operational Technology (OT). In addition to that, the risk assessment should consider several elements of the impact of cyber incidents such as operations and safety in a way that supports continuous risk management processes. This was addressed through the proposition of a new cyber risk assessment approach which is based on the Failure Mode, Effects & Criticality Analysis (FMECA). The approach is comprehensive in its consideration of components, threats, mitigation measures, and impact elements. The approach also supports the need for continuous risk management processes by automating certain tasks through the reliance on the concept of curated knowledge drawn from the encoded knowledge and semantics in the Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework. Moreover, the approach aims to reduce the impact of bias introduced by the reliance on expert judgment and opinions, this is addressed through the utilization of the knowledge encoded in ATT&CK as well as centrality metrics from graph theory. The proposed risk assessment method known as FMECA-ATT&CK has been thoroughly demonstrated and evaluated. It has been demonstrated through application against several target systems and evaluated through the engagement of experts. The results suggest that FMECA-ATT&CK is not only suitable for the identification of risks against the APS, but it is also suitable for a wider range of cyber-physical systems (CPS) including a shipboard Integrated Navigation System (INS), several designs of APS architectures including milliAmpere 2, and a digital substation. This designates FMECA-ATT&CK as a strong candidate semi-automated approach to cyber risk assessment for CPS.

Regarding the third objective, a suitable risk management approach for the APS needs to consider the domain-specific requirements as well as have the ability to address the identified cyber risks. This was addressed through the proposition of a standard-aligned cyber risk management approach named Threat Informed Defense in Depth (TIDiD). TIDiD, aligns two widely adopted cybersecurity strategies, namely, Threat Informed Defense, and Defense in Depth (DiD). The approach extends the curated knowledge in the ATT&CK framework by mapping the mitigation measures with the different DiD elements of cyber defense. This allows for more evolved risk management capabilities. DiD is a commonly observed design pattern in cybersecurity, not only in the maritime domain. It refers

to the utilization of several layers of defense. The ATT&CK framework; a pillar instrument in the TID strategy, allows for mapping the risks with the relevant mitigation measures which support the processes of risk control selection and evaluation. Moreover, several system engineering processes were utilized for the development of a cybersecurity architecture that supports the risk management functions, implementing those processes within the context of the APS has identified several challenges originating from the unconventional mode of operation that is the auto-remote. For instance, the need to include a remote control center in the control loop of the ferry challenges several aspects of the existing DiD practices such as the concept of network zones. These challenges have motivated a deeper exploration of certain threats and relevant mitigation measures. Among the areas which attracted increased focus were anomaly detection and analysis and the utility of the Automatic Identification System (AIS) for establishing covert channels for command and control (C&C). The APS's increased reliance on cyber components and sensor data for the provisioning of navigational functions introduces a wide range of cyber risks. An anomaly and intrusion detection component is deemed necessary and was further explored in this thesis. It was found that specification-based anomaly detection solutions can be suitable for detecting a wide range of relevant anomalies while machine learning approaches might be needed for more advanced threats. On the other hand, the APS inclusion of the insecure AIS technology introduces a threat of covert channels utilized in C&C during the development of cyber attacks. The feasibility of such covert channels has been investigated empirically and found to be feasible with existing technologies. This threat goes beyond the APS technology and affects other traditional maritime components employing AIS in their operations. Finally, the different proposals regarding risk management have been evaluated through various techniques including simulation, checklists, adversary emulation, unit testing, and risk analysis. The proposals' feasibility, suitability, and utility were found sufficient for the APS technology. A wide range of areas deserves further attention, this has been objectively discussed as limitations earlier in the thesis. Still, several elements of the proposed risk management and cybersecurity architecture have been integrated into the development of the milliAmpere 2 ferry and the remote control center. This supports the claims in this thesis regarding the suitability of the proposed solutions.

APS has demonstrated success in trials involving existing communication technologies. The security of the system has been enhanced through the use of existing solutions and processes, including those in this thesis. However, many more areas require additional attention to improve the APS remote monitoring capabilities and cybersecurity posture. This paves the way for future research in the cyber resilience of APS technology and similar maritime technologies relying on information and communication technology (ICT), OT, ML and AI. The convergence of safety and security, the adversarial machine learning threats, as well as incident response and recovery plans are among such future directions.

Bibliography

- [1] *Transport modes*, https://ec.europa.eu/transport/modes_en, Jan. 2019.
- [2] *European defence agency, maritime domain*, <https://eda.europa.eu/docs/default-source/eda-factsheets/2017-09-27-factsheet-maritime>, 2017.
- [3] *Ocean shipping and shipbuilding*, <https://www.oecd.org/ocean/topics/ocean-shipping/>, 2019.
- [4] M. Fruth and F. Teuteberg, 'Digitization in maritime logistics—what is there and what is missing?' *Cogent Business & Management*, vol. 4, no. 1, p. 1411066, 2017.
- [5] D. R. Schallmo and C. A. Williams, 'History of digital transformation,' in *Digital Transformation Now!* Springer, 2018, pp. 3–8.
- [6] L. Heilig, E. Lalla-Ruiz and S. Voß, 'Digital transformation in maritime ports: Analysis and a game theoretic framework,' *Netnomics: Economic research and electronic networking*, vol. 18, no. 2-3, pp. 227–254, 2017.
- [7] *Projects carried out by members of nfas*, <http://bit.ly/NFASProjects>.
- [8] *Autonomous all-electric passenger ferries for urban water transport*, <https://www.ntnu.edu/autoferry>, Jul. 2021.
- [9] G. Havdal, C. T. Heggelund and C. H. Larssen, 'Design of a small autonomous passenger ferry,' M.S. thesis, NTNU, 2017.
- [10] D. Patraiko, 'The development of e-navigation,' *TransNav, International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 1, no. 3, 2007.
- [11] M. IMO, '85/26/add. 1 report of the maritime safety committee on its eighty-fifth session,' *International Maritime Organization, London*, 2008.
- [12] K. An, 'E-navigation services for non-solas ships,' *International Journal of e-Navigation and Maritime Economy*, vol. 4, pp. 13–22, 2016.
- [13] A. Greenberg, *The untold story of notpetya, the most devastating cyberattack in history*. [Online]. Available: <https://bit.ly/MaerskAttack>.
- [14] *Cosco shipping lines falls victim to cyber attack*, Jul. 2018. [Online]. Available: <https://bit.ly/COSCOAttack>.

- [15] K. Tran, S. Keene, E. Fretheim and M. Tsikerdekis, 'Marine network protocols and security risks,' *Journal of Cybersecurity and Privacy*, vol. 1, no. 2, pp. 239–251, 2021.
- [16] M. Balduzzi, A. Pasta and K. Wilhoit, 'A security evaluation of ais automated identification system,' in *Proceedings of the 30th annual computer security applications conference*, 2014, pp. 436–445.
- [17] T. M. S. Committee, *International maritime organization (imo) (2017) guidelines on maritime cyber risk management*, <https://www.imo.org/en/OurWork/Security/Pages/Cyber-security.aspx>, 2017.
- [18] T. M. S. Committee, *Interim guidelines on maritime cyber risk management (msc-fal.1/circ.3/rev.1)*, <https://cutt.ly/6R8wqjN>, 2021.
- [19] *Imo completes regulatory scoping exercise for autonomous ships*, <http://bit.ly/IMOMASS>, May 2021.
- [20] *Nfas - norwegian projects*, <https://cutt.ly/NFAS>, 2021.
- [21] A. Amro, V. Gkioulos and S. Katsikas, 'Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation,' in *Computer Security*, Springer, 2019, pp. 69–85.
- [22] A. Amro, V. Gkioulos and S. Katsikas, 'Communication architecture for autonomous passenger ship,' *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X211002546, 2021.
- [23] A. Amro, G. Kavallieratos, K. Louzis and C. A. Thieme, 'Impact of cyber risk on the safety of the milliampere2 autonomous passenger ship,' in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 929, 2020, p. 012018.
- [24] A. Amro, V. Gkioulos and S. Katsikas, 'Assessing cyber risk in cyber-physical systems using the attck framework,' *ACM Trans. Priv. Secur.*, Nov. 2022, Just Accepted, ISSN: 2471-2566. DOI: 10.1145/3571733. [Online]. Available: <https://doi.org/10.1145/3571733>.
- [25] A. Amro and V. Gkioulos, 'Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth,' *International Journal of Information Security*, Nov. 2022, ISSN: 1615-5270. DOI: 10.1007/s10207-022-00638-y. [Online]. Available: <https://doi.org/10.1007/s10207-022-00638-y>.
- [26] A. Amro, A. Oruc, V. Gkioulos and S. Katsikas, 'Navigation data anomaly analysis and detection,' *Information*, vol. 13, no. 3, 2022, ISSN: 2078-2489. DOI: 10.3390/info13030104. [Online]. Available: <https://www.mdpi.com/2078-2489/13/3/104>.
- [27] A. Amro and V. Gkioulos, 'From click to sink: Utilizing ais for command and control in maritime cyber attacks,' in *European Symposium on Research in Computer Security*, Springer, 2022, pp. 535–553.

- [28] A. Amro and V. Gkioulos, 'Communication and cybersecurity testbed for autonomous passenger ship,' in *European Symposium on Research in Computer Security*, Springer, 2021, pp. 5–22.
- [29] A. Amro and V. Gkioulos, 'Evaluation of a cyber risk assessment approach for cyber-physical systems: Maritime- and energy-use cases,' *Journal of Marine Science and Engineering*, vol. 11, no. 4, 2023, ISSN: 2077-1312. DOI: 10.3390/jmse11040744. [Online]. Available: <https://www.mdpi.com/2077-1312/11/4/744>.
- [30] A. Oruc, A. Amro and V. Gkioulos, 'Assessing cyber risks of an ins using the mitre att&ck framework,' *Sensors*, vol. 22, no. 22, p. 8745, 2022.
- [31] A. Amro, 'Cyber-physical tracking of iot devices: A maritime use case,' in *Norsk IKT-konferanse for forskning og utdanning*, 2021.
- [32] Ø. Rødseth and H. Nordahl, 'Definitions for autonomous merchant ships,' in *Norwegian Forum for Unmanned Ships*, 2017.
- [33] L. Register, 'Cyber-enabled ships: Deploying information and communications technology in shipping-lloyds register's approach to assurance,' *London: Lloyds Register*, <http://www.marinelog.com/index.php>, 2016.
- [34] DNV GL, 'Dnvgl-cg-0264: Autonomous and remotely operated ships,' 2018.
- [35] Ø. Rødseth and H. Burmeister, 'Munin deliverable d10.1: Impact on short sea shipping,' 2015. [Online]. Available: <http://www.unmanned-ship.org/munin/wp-content/uploads/2015/10/MUNIN-D10-1-Impact-on-Short-Sea-Shipping-MRTK-final.pdf>.
- [36] Ø. Rødseth, 'Munin deliverable 4.3: Evaluation of ship to shore communication links,' 2012, pp. 12–31. [Online]. Available: <http://www.unmanned-ship.org/munin/wp-content/uploads/2014/02/d4-3-eval-ship-shore-v11.pdf>.
- [37] C. Ramboll, 'Advokatfirma: Analysis of regulatory barriers to the use of autonomous ships: Final report,' *Danish Maritime Authority, Copenhagen*, pp. 1374–1403, 2017.
- [38] NIST, 'Framework for improving critical infrastructure cybersecurity,' NIST, 2018.
- [39] *United states coast guard - cybersecurity - maritime specific cybersecurity framework profiles*. [Online]. Available: <https://www.dco.uscg.mil/Our-Organization/Assistant-Commandant-for-Prevention-Policy-CG-5P/Inspections-Compliance-CG-5PC-/Office-of-Port-Facility-Compliance/Domestic-Ports-Division/cybersecurity/>.
- [40] Bureau Veritas, 'Guidelines for autonomous shipping,' Bureau Veritas, 2017.
- [41] International Association of Classification Societies (IACS), 'Ur e22 on board use and application of computer based systems - rev.2 june 2016,' International Association of Classification Societies (IACS), 2017.

- [42] Ø. J. Rødseth, B. Kvamstad, T. Porathe and H.-C. Burmeister, 'Communication architecture for an unmanned merchant ship,' in *OCEANS-Bergen, 2013 MTS/IEEE, IEEE*, 2013, pp. 1–9.
- [43] *Small enterprise design profile reference guide - small enterprise design profile(sedp)-network foundation design [design zone for security]*, https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Small_Enterprise_Design_Profile/SEDP/chap2.html, Oct. 2013.
- [44] A. Shostack, *Threat Modeling: Designing for Security*. 2014, vol. Wiley Publishing.
- [45] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington and C. B. Thomas, 'Mitre att&ck: Design and philosophy,' *Technical report*, 2018.
- [46] O. Alexander, M. Belisle and J. Steele, 'Mitre att&ck® for industrial control systems: Design and philosophy,' 2020.
- [47] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, NJ, 1996, vol. 2.
- [48] G. Stergiopoulos, M. Theocharidou, P. Kotzanikolaou and D. Gritzalis, 'Using centrality measures in dependency risk graphs for efficient risk mitigation,' in *International Conference on Critical Infrastructure Protection*, Springer, 2015, pp. 299–314.
- [49] C. León, 'Authority centrality and hub centrality as metrics of systemic importance of financial market infrastructures,' *Available at SSRN 2290271*, 2013.
- [50] A. Akbarzadeh and S. Katsikas, 'Identifying critical components in large scale cyber physical systems,' in *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, 2020, pp. 230–236.
- [51] J. Boyens, C. Paulsen, R. Moorthy, N. Bartol and S. Shankles, 'Nist special publication 800-161: Supply chain risk management practices for federal information systems and organizations,' *NIST. April*, 2015.
- [52] D. Rasmus Nord Jorgensen in Copenhagen, *Bimco: The guidelines on cyber security onboard ships*, <https://iumi.com/news/blog/bimco-the-guidelines-on-cyber-security-onboard-ships>.
- [53] M. Fabro, E. Gorski, N. Spiers, J. Diedrich and D. Kuipers, 'Recommended practice: Improving industrial control system cybersecurity with defense-in-depth strategies,' *DHS Industrial Control Systems Cyber Emergency Response Team*, 2016.
- [54] A. Fielder, T. Li and C. Hankin, 'Defense-in-depth vs. critical component defense for industrial control systems,' in *4th International Symposium for ICS & SCADA Cyber Security Research 2016 4*, 2016, pp. 1–10.

- [55] zvelo, *Fight ransomware with defense in depth*, =<https://zvelo.com/fight-ransomware-with-defense-in-depth/>, Accessed 11.10.2021.
- [56] V. Mavroeidis and S. Bromander, 'Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence,' in *2017 European Intelligence and Security Informatics Conference (EISIC)*, IEEE, 2017, pp. 91–98.
- [57] MITRE, *Threat-informed defense*, <https://www.mitre.org/news/focal-points/threat-informed-defense>, Accessed 05.01.2022.
- [58] *Threat-based defense*, <https://www.mitre.org/capabilities/cybersecurity/threat-based-defense>.
- [59] Y. Zhao, W. Li and P. Shi, 'A real-time collision avoidance learning system for unmanned surface vessels,' *Neurocomputing*, vol. 182, pp. 255–266, 2016.
- [60] F. Goerlandt and K. Pulsifer, 'An exploratory investigation of public perceptions towards autonomous urban ferries,' *Safety science*, vol. 145, p. 105 496, 2022.
- [61] C. A. Thieme, C. Guo, I. B. Utne and S. Haugen, 'Preliminary hazard analysis of a small harbor passenger ferry—results, challenges and further work,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1357, 2019, p. 012 024.
- [62] C. Guo, S. Haugen and I. B. Utne, 'Risk assessment of collisions of an autonomous passenger ferry,' *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X211050714, 2021.
- [63] O. A. Alsos, E. Veitch, L. Pantelatos, K. Vasstein, E. Eide, E.-M. Petermann and M. Breivik, 'Ntnu shore control lab: Designing shore control centres in the age of autonomous ships,' in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2311, 2022, p. 012 030.
- [64] S. Y. Enoch, J. S. Lee and D. S. Kim, 'Novel security models, metrics and security assessment for maritime vessel networks,' *Computer Networks*, vol. 189, p. 107 934, 2021.
- [65] S. Standard, R. Greenlaw, A. Phillips, D. Stahl and J. Schultz, 'Network reconnaissance, attack, and defense laboratories for an introductory cybersecurity course,' *ACM Inroads*, vol. 4, no. 3, pp. 52–64, 2013.
- [66] M. S. Lund, O. S. Hareide and Ø. Jøsok, 'An attack on an integrated navigation system,' 2018.
- [67] S. Papastergiou, E.-M. Kalogeraki, N. Polemi and C. Douligieris, 'Challenges and issues in risk assessment in modern maritime systems,' in *Advances in Core Computer Science-Based Technologies*, Springer, 2021, pp. 129–156.

- [68] J. Pavur, D. Moser, M. Strohmeier, V. Lenders and I. Martinovic, 'A tale of sea and sky on the security of maritime vsat communications,' in *2020 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2020, pp. 1384–1400.
- [69] K. Tam and K. Jones, 'Macra: A model-based framework for maritime cyber-risk assessment,' *WMU Journal of Maritime Affairs*, vol. 18, no. 1, pp. 129–163, 2019.
- [70] C. Hemminghaus, J. Bauer and E. Padilla, 'Brat: A bridge attack tool for cyber security assessments of maritime systems,' 2021.
- [71] Y. Jo, O. Choi, J. You, Y. Cha and D. H. Lee, 'Cyberattack models for ship equipment based on the mitre att&ck framework,' *Sensors*, vol. 22, no. 5, p. 1860, 2022.
- [72] E. Yağdereli, C. Gemci and A. Z. Aktaş, 'A study on cyber-security of autonomous and unmanned vehicles,' *The Journal of Defense Modeling and Simulation*, vol. 12, no. 4, pp. 369–381, 2015.
- [73] E. A. Oladimeji, S. Supakkul and L. Chung, 'Security threat modeling and analysis: A goal-oriented approach,' in *Proc. of the 10th IASTED International Conference on Software Engineering and Applications (SEA 2006)*, Citeseer, 2006, pp. 13–15.
- [74] M. M. Islam, A. Lautenbach, C. Sandberg and T. Olovsson, 'A risk assessment framework for automotive embedded systems,' in *Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security*, 2016, pp. 3–14.
- [75] G. Kavallieratos, S. Katsikas and V. Gkioulos, 'Safesec tropos: Joint security and safety requirements elicitation,' *Computer Standards & Interfaces*, vol. 70, p. 103 429, 2020.
- [76] M. J. Lopes, F. Teixeira, J. B. Mamede and R. Campos, 'Wi-fi broadband maritime communications using 5.8 ghz band,' in *2014 Underwater Communications and Networking (UComms)*, IEEE, 2014, pp. 1–5.
- [77] M. Ludvigsen, S. M. Albrektsen, K. Cisek, T. A. Johansen, P. Norgren, R. Skjetne, A. Zolich, P. S. Dias, S. Ferreira, J. B. de Sousa *et al.*, 'Network of heterogeneous autonomous vehicles for marine research and management,' in *OCEANS 2016 MTS/IEEE Monterey*, IEEE, 2016, pp. 1–7.
- [78] J. Pinto, P. S. Dias, R. Martins, J. Fortuna, E. Marques and J. Sousa, 'The lts toolchain for networked vehicle systems,' in *2013 MTS/IEEE OCEANS-Bergen*, IEEE, 2013, pp. 1–9.
- [79] A. Emam, A. Mtibaa and K. A. Harras, 'Message in a bottle: Extending communication coverage via boat-to-boat wifi communication,' in *Proceedings of the 13th Workshop on Challenged Networks*, 2018, pp. 63–69.

- [80] M. Höyhtyä, J. Huusko, M. Kiviranta, K. Solberg and J. Rokka, 'Connectivity for autonomous ships: Architecture, use cases, and research challenges,' in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, IEEE, 2017, pp. 345–350.
- [81] M. Höyhtyä, T. Ojanperä, J. Mäkelä, S. Ruponen and P. Järvensivu, 'Integrated 5g satellite-terrestrial systems: Use cases for road safety and autonomous ships,' in *Proceedings of the 23rd Ka and Broadband Communications Conference, Trieste, Italy*, 2017, pp. 16–19.
- [82] H. Ferreira, F. Silva, P. Sousa, B. Matias, A. Faria, J. Oliveira, J. M. Almeida, A. Martins and E. Silva, 'Autonomous systems in remote areas of the ocean using bluecom+ communication network,' in *OCEANS 2017-Anchorage*, IEEE, 2017, pp. 1–6.
- [83] A. Zolich, A. Soegrov, E. Vågsholm, V. Hovstein and T. A. Johansen, 'Coordinated maritime missions of unmanned vehicles—network architecture and performance analysis,' in *2017 IEEE International Conference on Communications (ICC)*, IEEE, 2017, pp. 1–7.
- [84] R. Stelzer and K. Jafarmadar, 'Communication architecture for autonomous sailboats,' in *Proceedings of International Robotic Sailing Conference*, 2009, pp. 31–36.
- [85] S.-W. Jo and W.-S. Shim, 'Lte-maritime: High-speed maritime wireless communication based on lte technology,' *IEEE Access*, vol. 7, pp. 53 172–53 181, 2019.
- [86] M. Hoefl, K. Gierłowski, J. Rak and J. Woźniak, 'Netbaltic system—heterogeneous wireless network for maritime communications,' *Polish Maritime Research*, 2018.
- [87] S. N. Rao, D. Raj, V. Parthasarathy, S. Aiswarya, M. V. Ramesh and V. Rangan, 'A novel solution for high speed internet over the oceans,' in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2018, pp. 906–912.
- [88] Y. Xu, S. Jiang and F. Liu, 'A lte-based communication architecture for coastal networks,' in *Proceedings of the 11th ACM International Conference on Underwater Networks & Systems*, 2016, pp. 1–2.
- [89] H.-J. Kim, J.-K. Choi, D.-S. Yoo, B.-T. Jang and K.-T. Chong, 'Implementation of maricomm bridge for lte-wlan maritime heterogeneous relay network,' in *2015 17th international conference on advanced communication technology (ICACT)*, IEEE, 2015, pp. 230–234.
- [90] M.-T. Zhou and H. Harada, 'Cognitive maritime wireless mesh/ad hoc networks,' *Journal of Network and Computer Applications*, vol. 35, no. 2, pp. 518–526, 2012.

- [91] W. Du, M. Zhengxin, Y. Bai, C. Shen, B. Chen and Y. Zhou, 'Integrated wireless networking architecture for maritime communications,' in *2010 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, IEEE, 2010, pp. 134–138.
- [92] A. Karantjias, N. Polemi and S. Papastergiou, 'Advanced security management system for critical infrastructures,' in *IISA 2014, The 5th international conference on information, intelligence, systems and applications*, IEEE, 2014, pp. 291–297.
- [93] C. Grigoriadis, S. Papastergiou, P. Kotzanikolaou, C. Douligeris, A. Dionysiou, A. Elias, K. Bernsmed, P. H. Meland and L. Kamm, 'Integrating and validating maritime transport security services: Initial results from the cs4eu demonstrator,' in *2021 Thirteenth International Conference on Contemporary Computing (IC3-2021)*, 2021, pp. 371–377.
- [94] B. Svilicic, J. Kamahara, M. Rooks and Y. Yano, 'Maritime cyber risk management: An experimental ship assessment,' *The Journal of Navigation*, vol. 72, no. 5, pp. 1108–1120, 2019.
- [95] P. Rajaram, M. Goh and J. Zhou, 'Guidelines for cyber risk management in shipboard operational technology systems,' *arXiv preprint arXiv:2203.04072*, 2022.
- [96] G. Kavallieratos and S. Katsikas, 'Managing cyber security risks of the cyber-enabled ship,' *Journal of Marine Science and Engineering*, vol. 8, no. 10, p. 768, 2020.
- [97] ISO, *Iso 31000:2018 risk management — guidelines*, 2018.
- [98] K. Stouffer, J. Falco, K. Scarfone *et al.*, 'Guide to industrial control systems (ics) security,' *NIST special publication*, vol. 800, no. 82, pp. 16–16, 2011.
- [99] K. Furumoto, A. Kolehmainen, B. Silverajan, T. Takahashi, D. Inoue and K. Nakao, 'Toward automated smart ships: Designing effective cyber risk management,' in *2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, IEEE, 2020, pp. 100–105.
- [100] G. Sabaliauskaite, S. Adepur and A. Mathur, 'A six-step model for safety and security analysis of cyber-physical systems,' in *Int. Conference on Critical Information Infrastructures Security*, Springer, 2016, pp. 189–200.
- [101] I. 6. T. Committee *et al.*, 'Analysis techniques for system reliability-procedure for failure mode and effects analysis (fmea),' 2018.
- [102] T. Skarshaug, 'Developing and using a virtual platform to test cyber security for autonomous vehicles and vessels,' M.S. thesis, NTNU, 2020.

- [103] U. J. Ruud, 'Cyber threats and vulnerabilities in the integrated navigation system,' M.S. thesis, NTNU, 2022.
- [104] A. L. Solli, 'Automated red teams in maritime cybersecurity exercises,' M.S. thesis, NTNU, 2022.
- [105] D. del Riego San Martín, 'It risk assessment automation in healthcare networks,' M.S. thesis, NTNU, 2022.
- [106] *Ntnu trials world's first urban autonomous passenger ferry*, <https://sciencebusiness.net/network-updates/ntnu-trials-worlds-first-urban-autonomous-passenger-ferry>, 2022.
- [107] *Cybersecurity is a top priority for digital transformation*, <https://www.helpnetsecurity.com/2020/09/29/cybersecurity-top-priority-digital-transformation/>, Sep. 2020.
- [108] Z. H. Munim, M. Dushenko, V. J. Jimenez, M. H. Shakil and M. Imset, 'Big data and artificial intelligence in the maritime industry: A bibliometric review and future research directions,' *Maritime Policy & Management*, vol. 47, no. 5, pp. 577–597, 2020.
- [109] E. Commission, C. Directorate-General for Communications Networks and Technology, *Ethics guidelines for trustworthy AI*. Publications Office, 2019. DOI: doi/10.2759/346720.

Part II

Published Research Papers

Paper I

A. Amro, V. Gkioulos and S. Katsikas, 'Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation,' in *Computer Security*, Springer, 2019, pp. 69–85

Connect and Protect: Requirements for Maritime Autonomous Surface Ship in Urban Passenger Transportation

Ahmed Amro¹, Vasileios Gkioulos¹, and Sokratis Katsikas^{1,2}

¹ Norwegian University of Science and Technology, Gjøvik, Norway

ahmed.amro@ntnu.no; sokratis.katsikas@ntnu.no; vasileios.gkioulos@ntnu.no

² Open University of Cyprus, Faculty of Pure and Applied Sciences, Nicosia, Cyprus
sokratis.katsikas@ouc.ac.cy

Abstract. Recent innovations in the smart city domain include new autonomous transportation solutions such as buses and cars, while Autonomous Passenger Ships (APS) are being considered for carrying passengers across urban waterways. APS integrate several interconnected systems and services that are required to communicate in a reliable manner to provide safe and secure real-time operations. In this paper, we discuss the APS context, stakeholders, regulations, standards and functions in order to identify communication and cybersecurity requirements towards designing a secure communication architecture suitable for APS.

Keywords: autonomous ship · communication system · communication security · cyber security

1 Introduction

According to the most recent report from the Norwegian Shipowners' Association, exactly half of the global shipping companies will implement autonomous ships by 2050, while Rolls-Royce aims to operate autonomous unmanned ocean-going ships by 2035 [25]. In this direction, the International Maritime Organization (IMO) started to address the regulatory scope for autonomous ships [8]. Norway is leading the autonomous shipping industry by opening several testing areas for the development of this technology, in addition to the production of Yara Birkland, the world's first all-electric and autonomous cargo ship [27], and other projects aiming to operate autonomous passenger ferries in different locations [5,28]. Many other initiatives all around the globe are taking place towards the development of autonomous ships; for instance, in 2018, Rolls-Royce and a Finnish ferry operator demonstrated the world's first fully autonomous ferry in Finland [26].

The Norwegian Forum for Autonomous Ships (NFAS) has provided definitions for autonomous ships, their context, and functions in [33]. A classification of autonomous maritime systems was suggested, depending on the operational area (underwater or surface), the control mode (remote control or autonomous) and the manning levels (from continuously manned to continuously unmanned).

This paper is targeting a specific autonomous maritime system which is the Maritime Autonomous Surface Ship (MASS) with a specific application for passenger transportation in urban waterways, to which we refer to as Autonomous Passenger Ship (APS). A comprehensive definition for a ship is suggested by NFAS: "a vessel with its own propulsion and steering system, which execute commercially useful transport of passengers or cargo and which is subject to a civilian regulatory framework". Consequently, an autonomous ship is defined as "a ship that has some level of automation and self governance". The typically expected operational mode of autonomous ships that is appropriate for APS is called "autoremove" and refers to a ship operating in a fully autonomous mode with the ability for a human intervention in case of emergency to take over full control of the ship operations [19].

With the increased research in the maritime industry focused at autonomous ships, the technological improvements were directed toward benefiting the development of smart cities through the smart transportation domain. The city of Trondheim which was recently stamped by EU as smart city [10] has opened the Trondheim Fjord as the world's first testing area for autonomous ships [39]. The idea behind the development of a smart city includes suggesting solutions for improving the citizens quality of life [38]. In this direction, the city of Trondheim is considering the application of a new technology i.e the autonomous ferry (*Autoferry*) [1] through the Trondheim canal to improve residents' life as an alternative to a high-cost bridge [40]. In this paper we focus on this new type of autonomous ships that will be used for passenger transportation in urban waterways.

Operating an autonomous passenger ship in a highly congested area is challenging for many reasons. Such a ship is expected to require the development of new technologies, while maintaining security and safety for the surrounding environment, the ship itself, and its passengers. Designing a suitable communication architecture is a crucial factor for safe operations, since improper communications is considered a primary factor for maritime casualties [11]. Additionally, according to ship owners, the most significant challenges for the usage of unmanned ships are rules and regulations, in addition to competence, compatible ports and fairways, and cyber security [27]. Therefore, the APS' communication architecture should satisfy certain requirements, deriving from the applicable rules and regulations and should be compatible with the views of the stakeholders of the APS ecosystem. Accordingly, this paper aims to identify requirements for a secure communication system in the specific case of APS. To this end, we identify the APS's stakeholders and their views and goals; we analyze existing regulations, guidelines and standards governing the design and operation of autonomous vessels; and we consider the functionality that such vessels should have to be able to operate safely.

The remaining of the paper is organized as follows: In Section 2 we review relevant research works. In Section 3 we discuss the APS's context, stakeholders, functions, relevant regulations, standards and guidelines. In Section 4 we present the identified requirements for the APS secure communication architecture. Fi-

nally, in Section 5, we summarize our conclusions and we present directions for future work.

2 Related Work

Several studies targeted the design and development of autonomous vessels. A master thesis proposed a design for a small autonomous passenger ferry that aims to be used for transporting passengers across the Trondheim city canal [22]. Another work proposes a technique for carrying out autonomous vessel steering tasks in coastal waters by implementing an agent system; each agent is deployed to perform specific tasks controlled by an agent platform installed on a computer on shore [24]. Neither of these works discussed communication or cybersecurity in their design proposals. Reliable communication capabilities are considered crucial toward the development of autonomous passenger vessels [22,24]. The literature is rich in various works targeting the communication architecture for autonomous ships, focusing on different operational areas, vessel types, and functional requirements. Furthermore, several navigation solutions known as e-navigation have been introduced by IMO in order to reduce human and traditional machine errors, and improve safety related to navigation on board ships, toward better protection for passengers, crew, maritime systems and the environment [30]. The e-navigation solutions targeted SOLAS (International Convention for the Safety of Life at Sea)-based ships, making them inapplicable to the APS. Nonetheless, a previous work discussed the integration of e-navigation solutions for non-SOLAS manned ships [12].

Moving toward autonomous ships, Maritime Unmanned Navigation through Intelligence in Networks (MUNIN) was a project that targeted the technical aspects in the operation of unmanned merchant vessels, and the assessment of their technical, economic and legal feasibility [31]. The project produced many deliverables, including the ship and communication architecture, remote bridge, autonomous engine room, and shore control center. The MUNIN project also produced a communication architecture for unmanned merchant ships, also suggesting communication and legal requirements to carry out unmanned operations in close to shore areas [32]. The MUNIN communication architecture is expected to influence the design and implementation of the communication architecture for the APS. Bureau Veritas, a member of the maritime classification society, published a document providing guidelines for suggested functions and components in autonomous ships [15]. The document aimed to provide guidelines for achieving the most essential functionality and improved reliability, being helpful in the process of studying related communication and cybersecurity requirements. The document also provided communication requirements for functionality and increased reliability. Although the document focused on satellite communications, which is not relevant for urban passenger transportation, the proposed considerations can be adjusted to radio frequencies in close to shore operations. Although the guidelines exclude ships smaller than 20m, we believe that the suggested guidelines related to communication are relevant for the APS. Addi-

tionally, DNV GL published several documents discussing aspects of autonomous ships. In their position paper they discussed the expected change in navigation, the regulatory scope, safety assurance, and social and ethical assurance [21]. Another related document from DNV GL is the class guidelines for autonomous and remotely operated ships [19]. In this document, DNV GL discussed several aspects including navigation functions, communication functions and cybersecurity considerations.

Several works discuss the lack of a regulatory framework that governs the operation of autonomous ships and suggests solutions to adapt to such technology. The Danish maritime authorities published a report on the regulatory barriers to the use of autonomous ships, suggesting suitable steps toward tackling these barriers, such as creating new laws for autonomous ships or amending existing ones [17]. Another work surveyed relevant regulations that might affect the operational capacity of autonomous ships [23]. The authors discussed regulations like SOLAS, COLREGS (International Regulations for Preventing Collisions at Sea), and others in detail, and pointed out that the regulations in their current form limit the deployment of autonomous ships. The work presented in [23] suggested generic communication requirements in order to satisfy certain regulations such as the availability of delay-free, reliable, fast and secure communication between the ship and control center.

3 The APS ecosystem

3.1 System Context

A general system context for the operation of a MASS as shown in Fig. 1 was suggested by NFAS. A brief description of the context components and their

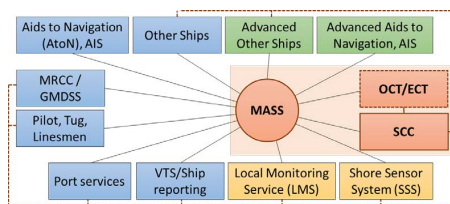


Fig. 1. Context diagram for autonomous ship operation [33].

relevance to the APS is given below:

- **Remote Control Center(RCC)**: The implementation of such controlling entity is common across most works involving autonomous ships. Some refer to this entity as Shore Control Center (SCC), others as Remote Control Center

(RCC); herein we adopt the latter term. An RCC functions as an observer, by monitoring the APS status, but in some cases it might be forced to take control of the ship in order to avoid accidents. For this reason, it was concluded that certain manning requirements are important for the RCC to operate [36]. Additionally, a single or a chain of RCCs might be expected to serve several ships concurrently. The location of the RCC might be on shore or it can reside on-board another vessel (e.g. an escort vessel).

- **Emergency Control Team (ECT)**: a team which is expected to intervene in case of emergencies endangering the passengers or the surrounding environment. For instance, a passenger falling into water, or the ship not responding to remote commands and heading on a collision course.
- **Shore Sensor System (SSS)**: A collection of sensors are expected to be mounted on shore to aid some functions of the APS. For instance, ship automatic docking, charging, and other functions related to passenger embarking and disembarking.
- **VTS/RIS**: Ships are expected to establish contact with Vessel Traffic Services (VTS) for guidance and reporting. Moreover, the European Parliament has defined activities towards establishing harmonized River Information Services (RIS) for inland waterways to facilitate navigation [13].
- **Aids to Navigation (AtoN)**: Collection of systems expected to provide real-time information for the ship navigation system regarding weather, other ships, location awareness, etc. Examples of such systems are the Automatic Identification System (AIS), the Global Navigation Satellite System (GNSS), Radar, Light Detection and Ranging (Lidar), etc.
- **Other Ships**: The APS is expected to communicate with other ships in the area for sharing navigational information using several agreed upon communication systems, such as Very high frequency (VHF), the more advanced VHF Data Exchange System (VDDES) or AIS.
- **Port Services**: Some services, such as electric charging, maintenance, passenger embarking and disembarking, might be provided to the APS at the port or quay.

Other components in Fig.1, such as the Maritime Rescue Coordination Centre (MRCC), Global Maritime Distress and Safety System (GMDSS), and Service vessels (Pilot, tug, etc.) are less relevant to the case of the APS, due to the smaller size of its operational area.

3.2 APS Stakeholders

It is important for the development of the APS communication system to grasp an overview of all the system’s stakeholders and understand their requirements. Several works discussed the stakeholders of autonomous ships; some focused on the regulator’s perspective [17], whilst others provided an overview of all stakeholders from the shipping industry perspective [41]. In the context of APS, we identified seven categories of stakeholders, as shown in Fig. 2. Detailed descriptions of each stakeholder category, their interactions and their interest in the system are provided below:

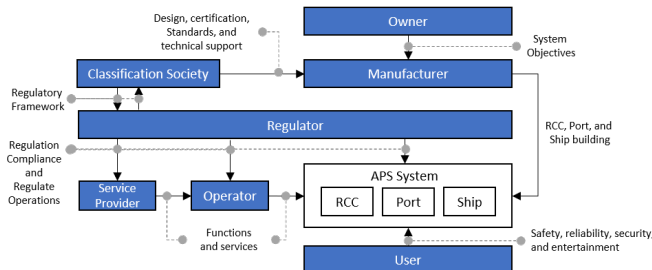


Fig. 2. APS stakeholders and their interactions

- **Owner:** The entire APS or parts of it might be owned by one or several entities. Usually, system owners dictate the objectives to be realized by the manufacturer.
- **Manufacturer:** All entities involved in the design and the implementation of APS, RCC, and port systems and facilities. Such entities are expected to follow standards and requirements related to functionality, reliability, safety, cybersecurity set by the classification society.
- **Classification Society:** Entities that contribute to the maritime domain, including through providing recommendations and suggesting relevant standards for ship manufacturers. The International Association of Classification Societies (IACS) consists of twelve members (including Bureau Veritas and DNV GL) that contribute to the classification design, construction, and rules and standards compliance for more than 90% of the world’s ships. IACS is also recognized by IMO as the principal technical advisor [3].
- **Regulator:** A crucial component for the operation of APS is a relevant civilian legal framework. While such a framework does not exist at the time of writing this paper, its development is an ongoing task carried out by IMO [8], assisted by several other entities [19]. Additionally, the operations of APS are expected to be regulated through ship registration and instructions from several entities such as local maritime authorities and traffic regulators (VTS, RIS, etc.). Ensuring regulatory compliance is another task performed by some regulatory entities.
- **Operator:** All entities responsible for realizing the functions of the different components of the APS ecosystem; these are mainly the RCC, ship, and port. It must be noted that in some cases the system might be operated by its owner.
- **Service providers:** Supporting entities that provide additional functions and services for the system’s operators. Services may include maintenance, connectivity, insurance, technical support, ship movement in and outside water, etc.

- **User:** Passengers constitute an important component of the APS ecosystem. Their safety and well being is the top priority when designing and operating the ship. Passengers expect such a ship to be safe, reliable, secure, and entertaining.

3.3 Regulations, Standards, and Guidelines

As mentioned earlier, the definition of a ship includes a regulatory framework that governs its operation, mainly to ensure safety, security and protection of the environment. Internationally, such responsibility falls upon the IMO, while regional or national regulatory entities are entitled to issue their own regulations within their jurisdiction [21]. Several international regulations need to be considered while moving forward toward autonomous ships. The identified international regulations and their applicability to APS is depicted in Table 1.

Table 1. International Regulations and Standards relevant to APS

Regulations			
Title	Section/Chapter	Scope	APS Applicable
SOLAS	TSM	International voyages	✓
	TSFS		✓
	GMDSS		✓
UNCLOS		Sea	✓
STCW			✓
MARPOL			✓
SAR			✓
COLREG		Sea Connected	✓*
Standards			
IEC 61162	1 (NMEA 0183)	Serial Communication	✓
	3 (NMEA 2000)		✓
	450		Ethernet
IEC 61850	460	Ethernet and Security	✓
	90-4	LAN Engineering	✓
MSC.252	83	Integrated Navigation System	✓
IEC 62443	3-3	Security of Industrial Control Systems	✓
ISO/IEC 27000	27001	Information Security Management Systems	✓
	27002		✓
IEC 62940		Communication between on-board systems and external computer systems	✓

✓*: Require modifications

In the case of APS in urban transportation, the most related regulations are the Convention on the International Regulations for Preventing Collisions at Sea (COLREG) which applies to all vessels operating at sea or waterways connected to the sea and accessed by seagoing vessels [29]. This can apply to APS operating in rivers and canals linked to the sea. An important regulation that affects the core functionality of the autonomous vessels to operate safely at water is Rule 5 in COLREG. The rule basically requires that the ship shall maintain proper lookout by proper means to avoid collision [29]. Considering that 48.9% of 1522 reported maritime accidents in the Republic of Korea between 2013-2017 were related to improper lookout, [41] it is evidently crucial to address this issue in autonomous ships. Additional regulations concerning passenger vessels differ between regions and countries. The European Union enforces

several regulations regarding the cybersecurity of ships and ports like the NIS directive (EU 2016/1148) and the General Data Protection Regulation (GDPR) for processing data of EU citizens, in addition to some other regulations that are related to ships in international voyages. In the Nordic region, each country specifies the passenger vessel types that require an operation certificate. Finland and Norway require all vessels of all sizes to acquire certificates, whilst in Sweden and Denmark certificates are required only for vessels carrying more than 12 passengers. Additionally, all passenger vessels that require certificates must comply with the regulations set by the maritime administration in that country. In Norway, for instance, such administration is the Norwegian Maritime Authorities (NMA) [4]. IACS [2,6,7], DNV GL [19], and Bureau Veritas [15], the most referenced standards that are suggested to be followed are also depicted in Table 1. Additionally, the most referenced guidelines to be considered in providing cybersecurity protections for autonomous ships come from the National Institute of Standards and Technology (the NIST Framework) [37], from IMO in resolution MSC.428(98) (MSC-FAL.1/Circ.3) [16], and from the French National Cybersecurity Agency (ANSSI) [14].

3.4 APS Functions

In order for the APS to operate safely, it must support functions that include navigation, machinery and passenger management, and communications. In this paper we focus on the communication functions and cybersecurity considerations for the APS to perform its intended functions, with an increased focus on navigation. DNV GL discussed the navigation functions that are expected of a vessel in autoremote operation [19]. These are listed below:

- **Voyage Management:** This function includes tasks such as the planning, updating, and recording of voyage data.
- **Condition Detection and Analysis:** This function includes tasks such as proper lookout and situational awareness (e.g. determination of position)
- **Contingency Planning:** A critical safety feature that is expected of any APS is referred to as Minimum Risk Condition (MRC). MRC is a state with the lowest possible risk where the ship should be programmed to enter in case of abnormal situation during operations such as the loss of communication links [19]. MRC can also be referred to as fail-safe condition.
- **Safe Speed:** The human in control or in supervisory mode must receive sufficient information regarding the situational awareness to keep the ship's speed within regulated limits.
- **Maneuvering:** To enable maneuvering for collision avoidance or voyage route change, an effective two way communication to provide sufficient situational awareness for either the autonomous system or the RCC in control to make correct decisions.
- **Docking:** An effective two way communication with the docking stations on board and on the shore (e.g. SSS).
- **Alert Management:** An alerting functionality through a Central Alert Management system (CAM) is crucial to achieve safety.

To realize such functions, a combination of systems are expected to be integrated within the APS. These systems require a certain level of connectivity and cybersecurity protection, which should be provided by the communication functions and protected using cyber security controls.

4 Communication and Cybersecurity Requirements

Based on [15,19] and on our analysis of the APS ecosystem in Section 3, this section presents the extracted communication and cybersecurity requirements of the APS to perform its expected functions (cf section 3.4). These requirements derive from the perspective of each stakeholder (cf section 3.2), and their presentation is organized accordingly.

4.1 Requirements deriving from the regulators' perspective

At the time of writing this paper there exist no specific regulations that govern the operations of autonomous ships. Nevertheless, the main aim of the regulators of APS is to ensure safety, security and environmental protection. This implies that autonomous ships must achieve a level of safety and security that is at least equivalent to that of a traditional ship.

4.2 Requirements deriving from the Classification Society's perspective

Both DNV GL [19] and Bureau Veritas [15] have offered communication and cybersecurity requirements for autonomous ships to operate in compliance with the related regulations, especially COLREG and SOLAS. Bureau Veritas suggested requirements focusing on the functionality and reliability of autonomous ships, whereas DNV GL focused more on safety. An overarching requirement is that *An efficient and secure communication network should be implemented to enable communication between internal and external systems of the autonomous ship.*

In the sequel, we discuss in detail the requirements for (efficient) and (secure) communication in the APS case. Three main communication categories have been identified for the APS to perform its intended functions: 1. External communication including connection with the RCC and external systems and stakeholders; 2. internal communication between on-board ship components; and 3. communication with other vessels in the vicinity. This subsection discuss the communication requirements for each communication category in addition to general requirements that apply across all categories. Additionally, this subsection discusses cybersecurity requirements mapped to the relevant NIST framework function as suggested by Bureau Veritas [15]. Each requirement in this section is titled with a three level coding scheme. The first level is related to the domain (communication (C) or Cybersecurity (S)). The second level is related to the sub-domain. The communication sub-domains are external (X), internal

(N), with other ships (O) or general (G). The cybersecurity sub-domains are identification (I), protection (P), detection (D), response and recovery (R). The third level refers to the relative numbering of the requirement within its category.

Communication Requirements: This subsection discusses external and internal communication requirements, in addition to the communication with other ships and other general communication requirements.

– External Communication

First, *a dedicated physical space must be allocated separately from the controlled vessel*, which can be on the shore or on-board another ship. The required level of reliability, availability, and security of the communication link will increase with increased control of the RCC over the APS, depending on the latter’s autonomy level. Additional communication with off-ship systems is required. Examples of off-ship systems that are leveraged for operational purposes are SSS, AtoN, VTS and RIS communication (cf Section 3.1). Additionally, other systems may require access to the ship’s systems, to provide services such as maintenance, processing insurance claims, etc. Communication with external stakeholders is expected by the APS either by automated systems on the vessel itself, or by the personnel on the RCC. The requirements for the forementioned communication are discussed below:

- **C-X-1:** *The link’s minimum acceptable network latency and maximum required bandwidth should be calculated, documented and implemented.* MUNIN provided minimum accepted requirements of latency and bandwidth [34]. In total 4Mbps accumulated link is considered the minimum link bandwidth for ship to shore communication. The required bandwidth is expected to be larger in the case of APS due to the implementation of new technologies with high data requirements such as the lidar. For instance, the targeted lidar for implementation in the *Autoferry* project [1] requires local transfer rate between 9-20 Mbps. Although the amount of data to be transmitted to the RCC is expected to be much less, in case of an increased control of the RCC over the vessel, the full lidar data might be expected for transmission. Additionally, the accepted latency suggested by MUNIN ranges from 0.05 seconds for ship to ship communication up to 2.5 seconds for HD video.
- **C-X-2:** *A dedicated, permanent and reliable link for emergency push buttons for passengers should exist.* Such button should be used to indicate passenger related emergency and is expected to initiate intervention of the available ECT (cf Section 3.1) in the area or to change the autonomy level to provide the RCC full control of the APS if appropriate.
- **C-X-3:** *The link with the RCC should be fault-tolerant so that it operates at full capacity even in case of failure in a single component*
- **C-X-4:** *Traffic in the link with the RCC should be prioritized according to a pre-defined prioritization policy to enable traffic with higher priority to be forwarded in case of reduced bandwidth.* DNV GL suggested a prioritization policy so that the traffic is prioritized in the following order, from highest to lowest priority: 1. Control messages for emergency (e.g. MRC activa-

tion); 2. commands for remote control of key vessel functions; 3. situational awareness data for remote control of key vessel functions; 4. supervision data; 5. maintenance data.

- **C-X-5:** *The operator should be able to seamlessly switch and distribute different vessel data between the different communication channels without a negative effect on the operations e.g. situational awareness data on one channel, the rest on another.*
 - **C-X-6:** *Communication links should operate according to appropriate QoS requirements and adapt with signal degradation.* The QoS requirements are case dependent based on the implemented systems on board the APS. For instance, a rule could be established that delay sensitive systems (i.e. collision avoidance) should be carried through an appropriate communication channel that provides the lowest delay whereas delay tolerant systems (i.e. HD video) could be channeled through a communication channel with higher but still appropriate delay.
 - **C-X-7:** *The network should integrate monitoring and notification systems for real-time or near real-time link quality analysis, based on data collection and aggregation subsystems which satisfy intrinsic and contextual Quality of Information requirements to support such real-time/near real-time situational awareness and incident response.* The notification functionality is expected to be integrated within the ship's CAM.
 - **C-X-8:** *The operator should have independent troubleshooting capabilities over each one of the communication links.* Troubleshooting one link should not interrupt the operations of another.
 - **C-X-9:** *Communication link with RCC should be established using redundant communication channels, including main and backup channels, preferably using different communication technologies and service providers.* The communication architecture presented by MUNIN was mainly focusing on deep sea operations. This entails the application of satellite communication for carrying ship to shore operations as a primary communication channel; this is different compared to inland or short sea shipping such as the APS, where high communication requirements are needed. In this case, mobile communication or Wi-Fi channels can be primarily used [35].
- Internal Communication
- **C-N-1:** *The Communication network design should comply with the applicable requirements in the relevant standards.* (cf table 1)
 - **C-N-2:** *A Segregated network design should exist to avoid failure cascading.* DNV GL suggested a specific network arrangement that applies network segregation [19]. They suggested that the following systems should not be connected to the same network: 1. Navigation system; 2. Communication system; 3. Machinery control and monitoring system; 4. Safety systems; 5. Control systems that serve redundant vessel services; 6. Auxiliary systems not related to vessel key functions; 7. Other systems from different system suppliers. Suggested network segmentation methods include air-gap, VLAN, firewalls etc.

- **C-N-3:** *A redundant network design should exist with automatic transition/activation/restoration between the main and backup system components.*
 - **C-N-4:** *It should be possible to divert connectivity to local resources upon loss of remote resources. (e.g in case of distributed network or cloud services providing data storage, backup local storage for critical data are expected to be implemented)*
 - **C-N-5:** *Connectivity to several systems on-board, such as passenger management system, alert system (CAM), log book, and local sensors should exist. The passenger management system provides certain services to the passengers on-board such as voice communication, trip status, and internet-access. Local sensors may include weather sensors, positioning sensors and others.*
 - **C-N-6:** *If several wireless communication links are expected to operate closely on-board with a risk of interference, a frequency coordination plan should be made and documented and then tested on board.*
- Communication with other vessels
- **C-O-1:** *The APS should be able to communicate with other vessels. For such communication, line of sight (LOS) communication system mainly based on AIS or digital VHF with range of at least two kilometers should be used. This communication includes position and route advertisement which is essential for safe navigation and collision avoidance.*
- General Communication Requirements
- **C-G-1:** *Important communicated data should be recorded and logged to be analyzed when needed. DNV GL proposed the minimum data that is required to be recorded [19]: 1. The status of the vessel’s key functions including the communication links; 2. Alerts; 3. Manual orders or commands; 4. All input and output data to or from the decision support and automation systems. In case the data is recorded on board, an early alert should be raised in case storage capacity exceeds a certain threshold and it should be possible for it to be transferred to shore.*
 - **C-G-2:** *The network components and equipment should be type-approved in compliance with the related certification policy. For technologies implemented in autonomous vessels to be certified by DNV GL, type approval is discussed in a specified class program for cybersecurity [20]. Type approval according to Bureau Veritas includes compliance with the IEC 61162 standards (all parts) and the MSC.252(83) performance standards.*
 - **C-G-3:** *The transmission protocol in each link should comply with a relevant international standard, for example, 802.11 or 802.15 series for wireless communication.*
 - **C-G-4:** *Wireless data communication should employ an internationally recognized system with the following features: 1. Message integrity including fault prevention, detection, diagnosis, and correction; 2. Device configuration and authentication by permitting the connection only for devices that are included in the system design; 3. Message encryption to maintain mes-*

sage confidentiality; 4. Security management to protect network assets from unauthorized access.

- **C-G-5:** *A coverage-analysis of the different wireless communication systems must be performed in order to determine its effectiveness.* To this end, a wireless communication testbed that simulates or emulates the communication architecture of the APS can be leveraged.
- **C-G-6:** *All protocols and interfaces implemented in the communication links should be documented.*

Cybersecurity Requirements: This section discusses requirements for the cybersecurity of the APS communication system. A recognized framework should be applied to prevent or mitigate cybersecurity incidents, and in this paper we approach and discuss the identified cybersecurity requirements in the context of the NIST framework [37].

– Identification

- **S-I-1:** *An up-to-date cybersecurity management framework should exist to govern the operations of cyber systems. It should include necessary policies, procedures and technical requirements.* According to the IMO resolution MSC.428(98), ship owners/operators must address cybersecurity risks in their management systems [16]. This can be achieved through an Integrated Ship Security and Safety Management System (IS3MS).
- **S-I-2:** *A regularly updated map of the IT installations and the network architecture should be established with a list of the equipment specified by model number and software specified by software version number.*
- **S-I-3:** *Network user accounts should be inventoried with the associated privileges, reflecting actual authorization.*

– Protection

- **S-P-1:** *User access management should exist and support the best practices in secure authentication, avoidance of generic and anonymous accounts, secure password and password change policies.*
- **S-P-2:** *Regular network software updates must be performed, according to an update policy that includes a list of components, responsibilities, means of obtaining and assessing updates, updates verification, and a recovery processes in case of failure.*
- **S-P-3:** *The network should be protected using secure protocols, e.g. encrypted transmission, and/or authentication as appropriate.*
- **S-P-4:** *Protection from malware should be implemented to prevent spreading between systems or network segments.*
- **S-P-5:** *Any personnel who shall access the system should be trained on relevant cybersecurity policies.* It has been determined that a major cause of cybersecurity incidents is the lack of awareness [19].
- **S-P-6:** *Software-based components should go through regular security analysis with suitable update policy.*

– Detect

- **S-D-1:** *Monitoring capabilities should be put in place to detect abnormal events.* Abnormal events such as several log-in failures, or massive data

transfer. Monitoring capabilities might include Intrusion Prevention Systems, Firewalls, etc. Additionally, such monitoring capabilities should adapt to the existence of encrypted traffic through utilizing best practices such as SSL/TLS proxies and/or anomaly detection.

- Response and Recovery
 - **S-R-1:** *An incident response plan should be formulated, including the isolation of infected components and detailed reporting.* First action after the isolation of all infected machines from the network, for each detected incident a feedback should be documented, and lessons learned sessions should be arranged, to improve defensive measures for similar events in the future.
 - **S-R-2:** *Availability of backup facilities for essential information should be made available with a suitable backup plan.*

4.3 Requirements deriving from the Service Providers' perspective

Additional cybersecurity considerations should be given regarding the service providers, especially in the case of them being provided from an external party rather than the systems operators. A list of identified possible service providers categories and their related cybersecurity considerations is given below:

- **Ship Registry:** secure authentication controls should exist for ship certification and revocation of certificates.
- **IT Service Providers:** controls regarding authorization and access control should exist.
- **System installation:** controls to verify proper and secure systems installation according to a defined list of configuration parameters should exist.
- **Maintenance:** access to the system to provide software and/or hardware maintenance services should be controlled, monitored, and verified.
- **Financial services:** controls should exist to protect processes related to passengers payments.
- **Insurance services:** controls should exist to secure access or disclosure of certain data in case of accidents.

4.4 Requirements deriving from the Users' perspective

Essentially, passengers safety should be guaranteed by all means during trips. Communication solutions for passengers to communicate with the ship operators and vice versa should be made available. Additionally, certain regulations exist to protect passengers privacy, for instance in Europe, compliance with GDPR is expected and in Norway there exist regulations including Privacy Law and personal data act that are set forth by The Norwegian Data Protection Authority (Datatilsynet) [9] governing tracking (The use of WiFi, Bluetooth, beacons and intelligent video analytic.), video surveillance and anonymity [18]. So, passengers should be protected against tracking, and their information should be processed with privacy considerations.

5 Conclusion and Future Work

A special type of autonomous ships is the Autonomous Passenger Ship. APSs operating in urban waterways constitute a case of increased interest when it comes to the design and implementation of their communication system. In order to define communication and cybersecurity requirements in this case, we defined and analyzed the APS ecosystem in terms of context, stakeholders, regulations, standards, and functions. By leveraging this analysis, we extracted communication and cybersecurity requirements that need to be satisfied so as the APS may perform its required functions. This work is part of an ongoing project called *Autoferry* [1]. Our future work will design and implement a communication architecture and an IS3MS for the *Autoferry* as a use case of an APS system, according to the requirements defined in this paper.

References

1. Autonomous all-electric passenger ferries for urban water transport. =<https://www.ntnu.edu/autoferry>
2. IACS rec 164 - communication and interfaces - new nov 2018. IACS
3. International association of classification societies. =<http://www.iacs.org.uk/>
4. Nordic boat standard. =<https://www.sdir.no/en/guides/nordic-boat-standard/>
5. Projects carried out by members of nfas. =<http://bit.ly/NFASProjects>
6. IACS rec 158 - physical security of onboard computer based system - new oct 2018. =<http://www.iacs.org.uk/download/8782>
7. IACS rec 159 - network security of onboard computer based systems - new sep 2018. =<http://www.iacs.org.uk/download/8652>
8. IMO takes first steps to address autonomous ships. =<http://bit.ly/IMOAutonomous>
9. Tracking in public spaces. =<http://bit.ly/DatatilsynetTracking>
10. Trondheim blir smartby. =<http://bit.ly/Trondheimkommune>
11. Focus on risks 2018. =<http://bit.ly/sdirRisks2018> (Nov 2017)
12. An, K.: E-navigation services for non-solas ships. *International Journal of e-Navigation and Maritime Economy* **4**, 13–22 (2016)
13. Andrés, S., Piniella, F.: Aids to navigation systems on inland waterways as an element of competitiveness in ulcv traffic. *International Journal for Traffic and Transport Engineering* **7**(1) (2017)
14. ANSSI: Information systems defence and security: France’s strategy (2011)
15. Bureau Veritas: Guidelines for autonomous shipping. <http://bit.ly/BureauVeritas641NI2017> (2017)
16. Committee, T.M.S.: Maritime cyber risk management in safety management systems (2017)
17. Danish Maritime Authority: Analysis of regulatory barriers to the use of autonomous ships. Danish Maritime Authority, Final Report, December (2017)
18. Datatilsynet: The anonymisation of personal data. =<http://bit.ly/DatatilsynetAnonymisation>
19. DNV GL: Dnvg-l-cg-0264: Autonomous and remotely operated ships (2018)
20. DNV GL: Dnvg-l-cp-0231: Cyber security capabilities of control system components (2018)

21. DNV GL – Maritime: Remote-controlled and autonomous ships position paper (2018)
22. Havdal, G., Heggelund, C.T., Larssen, C.H.: Design of a Small Autonomous Passenger Ferry. Master’s thesis, NTNU (2017)
23. Komianos, A.: The autonomous shipping era. operational, regulatory, and quality challenges. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* **12** (2018)
24. Lebkowski, A.: Design of an autonomous transport system for coastal areas. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* **12** (2018)
25. Levander, O., Marine, R.R.: Ship intelligence—a new era in shipping. In: *The Royal Institution of Naval architects, Smart Ship Technology, International Conference proceedings*. pp. 26–27 (2016)
26. MI News Network: Rolls-royce and finferries demonstrate world’s first fully autonomous ferry. =<http://bit.ly/marineinsightRollsRoyce> (Dec 2018)
27. Norwegian Shipowners’ Association: Maritime outlook 2018. Tech. rep., Norwegian Shipowners’ Association (2018)
28. Olsen, S.: Autonom ferge ballstadlandet. =<http://bit.ly/lofotenmatpark>
29. Organization, I.M.: Convention on the international regulations for preventing collisions at sea, 1972 (colregs) (1972)
30. Patraiko, D.: The development of e-navigation. *TransNav, International Journal on Marine Navigation and Safety of Sea Transportation* **1**(3) (2007)
31. Porathe, T., Burmeister, H.C., Rødseth, Ø.J.: Maritime unmanned navigation through intelligence in networks: The munin project. In: *12th International Conference on Computer and IT Applications in the Maritime Industries, COMPIT’13, Cortona, 15-17 April 2013*. pp. 177–183 (2013)
32. Rødseth, Ø., Burmeister, H.: Munin deliverable d10.1: Impact on short sea shipping. =<http://www.unmanned-ship.org/munin/wp-content/uploads/2015/10/MUNIN-D10-1-Impact-on-Short-Sea-Shipping-MRTK-final.pdf> (2015)
33. Rødseth, Ø., Nordahl, H.: Definitions for autonomous merchant ships. In: *Norwegian Forum for Unmanned Ships* (2017)
34. Rødseth, Ø.: Munin deliverable 4.3: Evaluation of ship to shore communication links. <http://www.unmanned-ship.org/munin/wp-content/uploads/2014/02/d4-3-eval-ship-shore-v11.pdf> (2012)
35. Rødseth, Ø.J., Kvamstad, B., Porathe, T., Burmeister, H.C.: Communication architecture for an unmanned merchant ship. In: *OCEANS-Bergen, 2013 MTS/IEEE*. pp. 1–9. IEEE (2013)
36. Rødseth, Ø.J., Tjora, Å.: A system architecture for an unmanned ship. In: *Proceedings of the 13th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT)* (2014)
37. Sedgewick, A.: Framework for improving critical infrastructure cybersecurity, version 1.1. Tech. rep., National Institute of Standards and Technology (2019)
38. Sikora-Fernandez, D., Stawasz, D., et al.: The concept of smart city in the theory and practice of urban development management. *Romanian Journal of Regional Science* **10**(1), 86–99 (2016)
39. SINTEF: Test site opens for unmanned vessels. =<http://bit.ly/sintefTestSites>
40. Skille, A., Lorentzen, S.: Foreslår førerløs passasjerferge i trondheim. =<http://bit.ly/nrkTrondheim>
41. Yoon, I.: Technology assessment - autonomous ships (09 2018). <https://doi.org/10.13140/RG.2.2.36778.88009>

Paper II

A. Amro, V. Gkioulos and S. Katsikas, 'Communication architecture for autonomous passenger ship,' *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X211002546, 2021

Communication Architecture for Autonomous Passenger Ship

Journal of Risk and Reliability (JRR)
XX(X):1-??
© The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

Ahmed Amro¹ and Vasileios Gkioulos¹ and Sokratis Katsikas^{1,2}

Abstract

Novel innovations have been witnessed in the past few years in the field of technology for autonomous vehicles. These have been exploited in various applications in the maritime domain; one such application is the proposal to develop autonomous passenger ships (APS) or ferries for carrying passengers in urban waterways. Such technology requires the integration of several components to support the safe and secure operation of the ferries. In this paper, a communication architecture is proposed, that satisfies pre-established communication requirements and supports autonomous and remotely controlled functions of an APS. The architecture was designed using the Architecture Analysis and Design Language (AADL); this enabled an iterative design process to be followed and allows for future improvements. The proposed architecture is verified by showcasing the role of the different architectural components in addressing the requirements and in supporting the expected functions in a number of operational scenarios based on the expected operations of an APS use case called "Autoferry". Furthermore, the proposed architecture has been evaluated by demonstrating its ability to achieve the expected performance according to the requirements, in simulated experiments using the network simulator GNS3.

Keywords

Autonomous Passenger Ship, communication Architecture, Reliable Communication, Safe Navigation

1 Introduction

Schallmo et al. (1) provide a brief history and some of several existing definitions of digital transformation, summarizing that it as a process which aims at novel value creation, process optimization, enhancement of experience, and establishment of new foundational capabilities. In the maritime domain, as discussed by Heilig et al. (2), digital transformation found its first significant applications within port management and logistics. Soon after, the introduction of innovative Information and Communication Technologies (ICT) was also focused directly toward the ships, aiming to enhance how they are built, operated and maintained.

This process motivated research and innovation activities towards novel and sustainable maritime transport systems, promoting the development of remotely controlled, automated, and autonomous ships. Definitions and advancement towards their attainment can be found in Rødseth et al. (3). The Norwegian Forum for Autonomous Ships (NFAS) currently reports multiple active and completed projects (4) that aim to develop enabling technologies, and also complete platforms. One example of such is the Autonomous all-electric Passenger Ships (APS) for urban water transport (5), from which the work presented in this paper originates and is part of.

The challenges associated with developing an APS, including those specific to the interaction with the environment, and the navigation of the autonomous system, with the primary objective to maintain the safety and security of passengers, systems, and the surrounding environment were presented in an earlier study by Havdal et al. (6).

Multiple initiatives have focused on the development of E-Navigation, also coordinated through the International Maritime Organization (IMO) as described by Patraiko (7). E-Navigation has been defined by the International Association of Lighthouse Authorities (IALA) as "the harmonised collection, integration, exchange and presentation of maritime information aboard and ashore by electronic means to enhance berth-to-berth navigation and related services, safety and security at sea, and the protection of the marine environment" (8). Originally, E-navigation was suggested as an open sea navigation solution. However, as presented by Kwang (9), the examination of the IMO's e-navigation Strategy Implementation Plan (SIP) and of Korea's national SIP for e-navigation reveals that E-navigation services are essential also for inland navigation. Kwan (9) also argued that digital communication services such as Long-Term Evolution (LTE) and Automatic Identification System (AIS) are key enablers for the implementation of E-navigation services. This is also the case for APSs used for inland transportation of passengers.

Supporting E-navigation within the context of the APS raises various system-specific requirements, which have been extracted and analysed by the authors in earlier work

¹Norwegian University of Science and Technology, Gjøvik, Norway

²Open University of Cyprus, Faculty of Pure and Applied Sciences, Nicosia, Cyprus

Corresponding author:

Ahmed Amro, Norwegian University of Science and Technology, Gjøvik, Norway

Email: ahmed.amro@ntnu.no

(10); therein, in addition to communication and security-related requirements, the system context, the involved stakeholders, and the relevant regulations, standards, and guidelines were discussed. The extracted requirements were proposed by the stakeholders with a focus on regulatory compliance, functionality, reliability and safety of autonomous ships. The work in (10) leads to the conclusion that ICT technologies implemented in contemporary ships are not sufficient for autonomous all-electric passenger ferries for urban water transport, given the operational conditions and requirements of the latter.

Accordingly, in this paper, a tailored Communication Architecture is proposed, that aims to satisfy the communication requirements established in (10), and to address the needs of the various stakeholders. The architecture is designed so as to include elements to allow the design and development of a complementary security architecture that will address the security requirements established in (10).

Using modeling and design languages is an observed approach in the literature on autonomous ships. For instance, Rødseth and Tjora (11) referred to the extensive use of Unified Modeling Language (UML) and scenario-based modeling in the MUNIN project to describe functionality. Additionally, the application of system modeling methods during the development of autonomous ship systems has been explored by Basnet et al. (12). Particularly, the authors explored both System Modelling Language (SysML) and Object Process Methodology (OPM) and argued that both methods were suitable to handle the system complexity and communication of system information. Moreover, the Architecture Analysis Design Language (AADL) (13), which can complement SysML, has been proposed for the analysis of critical systems due to its ability to combine information related to hardware, operating system, and code to implement functions. This allows AADL to be applied at advanced stages during the system development (14; 15). This specific capability deemed AADL as the most suitable in this work, especially to support the efforts during the development of the complementary security architecture. Therefore, the presented communication architecture is modeled using AADL for the description and analysis of the architecture in four abstraction layers, namely i) model, ii) service, iii) protocol and interface, iv) implementation. In the paper, the various components, along with their connections and dependencies, are presented, with details on selected aspects across the four abstraction layers. Furthermore, the architecture is conceptually verified against the requirements of (10), and a use case is presented in order to highlight further the functionalities of the various architectural components. Finally, the IP-based network of the architecture is evaluated using the GNS3 simulator, demonstrating capabilities such as increased availability of the internal and external network; and network segregation and traffic prioritization capability.

The remaining of the paper is structured as follows: In Section 2 related work and background information is given. The architecture development methodology is presented in Section 3. In Section 4, the Autoferry use case is described, so as to put the proposed communication architecture into an operational context. In Section 5, we present our proposal for the communication architecture. Section 6 discusses the

verification of the proposed architecture against its design requirements, as well as its applicability to other use cases, including of a larger scale. Finally, Section 7 summarises our conclusions and outlines directions for future work.

2 Background

E-navigation entered the IMO's agenda officially in 2006, with initial work been aided by IALA due to their expertise in the areas of navigational aids and Vessel Traffic Services (VTS). IALA identified three primary objectives for E-navigation, namely the development and provisioning of infrastructure for transferring information onboard ships; between ships; and between ships and onshore stakeholders (7).

As regards autonomous merchant ships, Rødseth et al. (16) proposed four operational modes, namely i) autonomous execution, ii) autonomous control, iii) remote control, and iv) fail to safe. Autonomous execution is the routine operational mode where the ship follows a pre-established set of instructions, transitioning to autonomous control when independently resolving minor problems. Remote control by shore operators is required when occurring circumstances fall outside the predetermined operational envelope. Finally, the fail-to-safe mode is entered when the remote control is necessitated but can not be achieved. E-Navigation is essential across all operational modes and therefore adopted and expected to be supported by the communication architecture proposed in this paper.

Placing these modes in the context of the APS, continuous transmission of real-time telemetry to a Remote Control Center (RCC) is necessitated for remote monitoring during autonomous execution. In case of minor changes, such as changing route or speed for dynamic positioning relative to moving objects, the APS transitions to the autonomous control mode where it keeps performing autonomously but transmits supplementary situational awareness data to the RCC and can expect to receive minor control commands. In case of unresolved hazardous situations such as possibly unavoidable collisions, the APS must transmit data to the RCC for enhanced and complete situational awareness, also receiving real-time control commands from the RCC. Finally, in case of loss of communication with the RCC or if the passenger emergency push button (EPB) is pressed, the APS is expected to initiate an appropriate fail to safe procedure (F2S). Several F2S procedures are expected to be available in the operational envelope, such as calling the nearby Emergency Control Team (ECT) to approach the ship and take control of it, while maintaining a fixed position.

These operational modes can only be achieved using E-navigation services based on reliable communication, with low latency and sufficient bandwidth to accommodate the amount of data generated and transmitted by the sensors. In this direction, many communication architectures have been proposed in the literature for maritime operations including autonomous maritime vessels (16; 17; 18; 19; 20). A reference architecture for crewless merchant ships has been proposed by Rødseth et al. (16) as part of the Maritime Unmanned Navigation through Intelligence in Network (MUNIN) project (21). This architecture, shown in Fig 1 was based on explicit assumptions related to

redundancy, security, network segregation and multiple RCCs, where having redundant communication links is realized by providing a main, a backup, and a dedicated link for rendezvous with ECT. Moreover, this reference architecture suggested an autonomous ship controller (ASC) as an entity that performs the autonomous and remotely controlled operations, in addition to controlling the mapping between the available communication resources and the control mode of the ship.

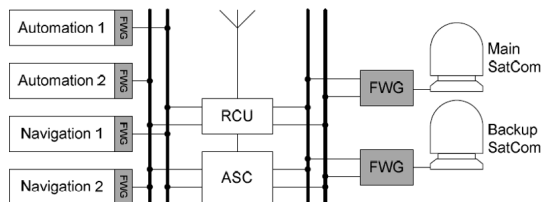


Figure 1. On board Reference Architecture (16)

Zolich et al (17) conducted a survey of communication and networks for autonomous marine systems, including autonomous vessels. The authors discussed the observed communication technologies and protocols used in different applications within the maritime domain. They highlighted that mobile communication technologies, as well as WiFi communication, are widely adopted in applications requiring large throughput and low latency.

In more recent works, Höyhty (18) aimed to address the challenges of navigating in port areas. Such areas are susceptible to accidents due to their increased traffic and limited manoeuvrability space, that raises the need for manoeuvres to be done accurately to avoid damaging the pier area; a similar challenge faces the operations of the APS. Höyhty argued that a reliable communication solution based on mobile communication technologies is needed to address this challenge. To this end, he proposed a high-level communication architecture consisting of satellite and terrestrial components for data transmission. Höyhty also suggested including in the architecture an intelligent entity called *connectivity manager*, which would be responsible for managing connectivity over multiple carriers, prioritizing traffic, cooperating with other ships, etc. Later, Höyhty (19) described an architecture of the connectivity manager as well as its functionalities, such as quality of service measurements, prioritization, and spectrum sharing. The connectivity manager component proposed by Höyhty (18) (19) is considered highly relevant to the case of APS communication, as it addresses a number of the requirements established in (10). Consequently, it has been adopted as a component of the architecture proposed herein.

Another communication architecture, utilizing LTE technology for maritime communications, was proposed by Jo and Shim (20). The authors argued that LTE can be reliably used to increase the range of ship-to-shore communication up to a range of 100 km. However, communication for autonomous operations was not the main focus of their work. Regarding LTE performance, some concerns have been raised by Mir and Filali (22). The authors performed

an evaluation of LTE in comparison to IEEE 802.11p technology. They argued that LTE outperformed IEEE 802.11p in aspects such as reliability, scalability and mobility. However, LTE failed to satisfy delay requirements in networks with high traffic loads.

Stelzer and Jafarmadar (23) have proposed a multi-stage communication architecture for autonomous sailboats. The authors highlighted the benefits of using several technologies to employ in the boat-to-shore link for providing reliable and cost-effective monitoring and control capabilities. In the first stage, a WiFi connection was proposed and exhibited the best performance. In the second stage, GPRS and UMTS cellular connections were made available to increase coverage in a cost-effective manner. For the last stage, satellite communication was suggested and implemented using the Iridium satellite services.

The reference architecture proposed by Rødseth et al (16) was found to be the most relevant to our work since it addressed several of the communication requirements established in (10). This architecture constitutes a significant part of the current state of the art, but it lacks certain elements necessary in the context of the APS case. Differences in operational conditions as well as in functional and non-functional requirements between the APS case and the cases considered in (16) require notable enhancements and modifications, that were made to design the architecture proposed herein. Such enhancements and modifications of significant importance are the protocols and interfacing, the integrated communication technologies, and the provided services. Some other aspects, such as the concept of the connectivity manager, and the utilization of mobile communication technologies rather than satellite communication, were influenced by other works in the literature and have been adapted and integrated into the proposed architecture. The proposed architecture was also influenced by the network engineering and design principles introduced by Cisco Inc., which are related to Hierarchy, Modularity, Resilience and Flexibility. Defining a suitable network hierarchy is a critical factor. Computer networks usually comprise a three-tier hierarchical model consisting of access, distribution, and core. Furthermore, the core and the distribution tiers can be merged into a collapsed tier, which, as described by Paptic (24) can reduce costs and complexity while also increasing redundancy. Additionally, having a modular design is beneficial due to the isolation it provides, and the ability to update or upgrade technologies seamlessly. The resilience principle refers to maintaining an operable status under normal and abnormal conditions, and one way of realizing it is through redundancy, by avoiding single points of failure. Lastly, due to continuous changes in technology, the network design should leave room for flexibility in the choice of technologies.

In our study, AADL and OSATE, an open-source tool that supports it (25), have been utilized for developing and modelling the proposed architecture. AADL is a language that enables early system's architecture analysis, providing a comprehensive set of notations for the description of system components, modes, properties, information flows and events (13). The developed model of the architecture describes all the entities in the APS context. Custom AADL properties were developed in order to

describe several aspects of the architecture, including communication properties of the connections between components, functional requirements, as well as security requirements. Having a model of the communication architecture in AADL enables its structured analysis, and the model is expected to be used in the future for performing threat and risk analyses. The model has been made accessible online ¹.

Notes

1. APS communication architecture AADL model: https://github.com/ahmed-amro/APS-Communication_Architecture.git

3 Methodology

The architecture presented in this article is based on a stable and pre-specified set of requirements, aiming to field early initial operational capabilities concerning communications (main scope of this article) and security (only highlighted here). Accordingly, an adapted pre-specified multistep model of incremental and evolutionary development is utilized, as suggested in (26). The generic system life cycle model (27) has adapted ISO/IEC 15288:2015 (28) and ISO/IEC 24748-1:2010 (29) suggesting a series of steps for the specification, development, operation and retirement of systems. For the purposes of our study and the current Technology Readiness Level of the Autoferry project, this study is focused on the two initial stages, namely:

- **Concept definition:** "Developing the concept of operations and business case; determining the key stakeholders and their desired capabilities; negotiating the stakeholder requirements among the key stakeholders and selecting the system's non-developmental items (NDIs)" (27). These have been primarily addressed earlier at (10), and are further detailed here, in appendix B.
- **System Definition:** "Developing system architectures; defining and agreeing upon levels of system requirements...Performing system analysis in order to illustrate the compatibility and feasibility of the resulting system definition" (27). This is the main contribution of this article, namely presenting the development of the system architecture in section 5, and the initial system analysis in section 6.

The concept definition phase refers both to communications and security, in order to identify and reconcile conflicting objectives and requirements, while the first system definition phase is referring to communications, as presented in the remainder of the article, also carrying security implications given the common concept definition phase. The second system definition phase, focusing purely on security is outside the scope of this article and will be presented in future work. The main contributions and methods used for each phase can be further detailed as follows:

3.1 Concept definition

This topic has been primarily addressed in earlier work. In the current article, we expand upon this, focusing primarily on the pre-established requirements by defining:

1. their prioritization according to the MoSCoW method (Must have, Should have, Could have, Won't have)
2. their classification according to their nature:
 - Quantitative property: a requirement indicating a property that can be described with certain units of measurements, such as latency or bandwidth.
 - Qualitative property: a requirement indicating a property that can be observed but not measured, such as redundant design.
 - Support of capability: a requirement indicating the capacity to perform a certain activity, such as network troubleshooting.
 - Action (i.e. operational activity): a requirement indicating performing a certain activity, such as frequency coordination planing.
3. their corresponding verification criteria at the design and implementation levels, under the limitations with respect to metric quantification due to the phrasing of the requirements by external stakeholders;
4. the architectural components that provide the functions which satisfy the requirements;
5. the verification methods used for system analysis;
6. future work directions.

3.2 System Definition

3.2.1 Development of the system architecture: An overview of the system definition process is depicted in Figure 2. To define the architecture's functions and structure, we initially followed the Goal Tree Success Tree (GTST) functional decomposition framework proposed by Kim and Modarres (30). The functional hierarchy described by the GTST includes a Goal Tree (GT) and a Success Tree (ST). In the GT, three basic levels can be formulated: the goal or functional objectives, functions, and sub-functions. In the ST, the system structure is formulated as a collection of sub-systems utilized to realize the functions identified in the GT.

Firstly, in the GT, the goal functions with relevance to the needed communication architecture have been established in our previous work from the views of the different stakeholders (10)—namely, safe and secure navigation as well as reliable and secure communication. Then, the goal functions are decomposed to functions and sub-functions assuring that the goal functions are achieved. The functions can be described in several ways, such as main and supporting functions. The main functions can be derived from the goal functions, and the supporting functions can be suggested toward the fulfillment of the goal functions. Furthermore, the functions can be decomposed into further sub-functions (more details in Section 4).

Secondly, in the ST, the identification of system structure was influenced by different works in the literature each proposing some design artifacts that deemed relevant to the needed communication architecture. Additionally, joint work with the other project members in the Autoferry project (5) provided guidelines for the proposition of the navigation and machinery systems as well as the emergency modules. The resulted GTST is presented in Figure 3. Then, the interactions of the system functions and the component distribution over the operational context were specified

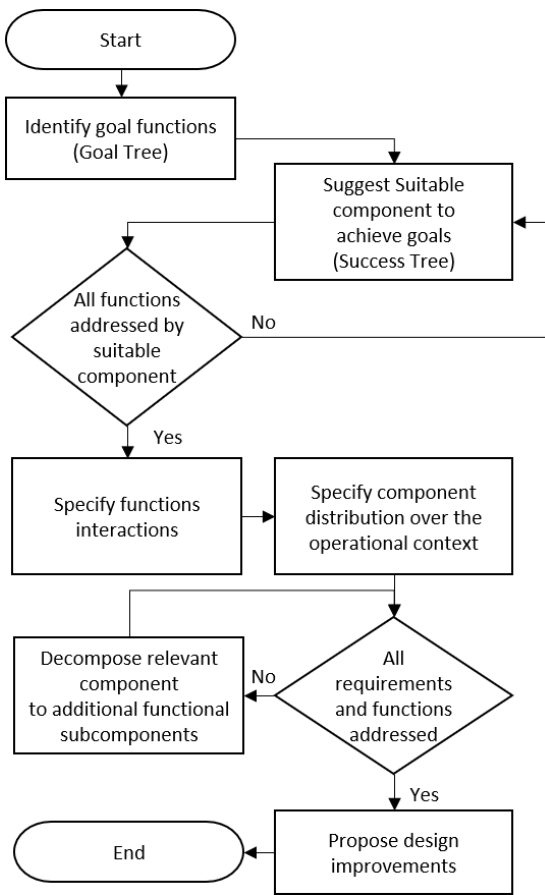


Figure 2. The system definition process

according to the expected operational modes specified in the literature.

Afterward, AADL was utilized in the development of the architecture components by undergoing through several iterations of system decomposition to ensure the realization of the system functions as well as the fulfillment of the established requirements. Moreover, certain design decisions were carried to introduce design improvements related to flexibility, scalability and expandability in a way that doesn't contradict with design requirements. The detailed description of this step is presented in Section 5.

3.2.2 System analysis: Such an analysis is performed not only in the system definition stage but at any stage across the life cycle that engineering or technical decisions are made. This allows for the quantitative and qualitative assessment of available architectural choices, and the affirmation of compliance with the elicited system requirements (31). In this article we focus on the effectiveness analysis of the proposed architecture, utilizing the use case described in section 4 to establish the operational context, and the scenarios presented in section 6 for the conceptual and experimental affirmation of the established requirement verification criteria. The assessment criteria, use case, and scenarios have been selected according

to the anticipated context of use of the system and in accordance with the current Technology Readiness Level of the Autoferry project, which is currently in the early technology development stage. Where operational or technical assumptions have been made due to the current maturity state, these have been captured and documented across sections 5 and 6.

4 Use case

The Autoferry, (Fig 4), will be developed to transport passengers across the Trondheim city canal as an alternative to a high-cost bridge. The operational area of the autonomous ferry is shown in Fig 5a. The ferry goes in both directions across the canal and its route is approximately 110m long. The canal witnesses traffic of mostly small size boats and, occasionally, of kayaks. The ferry will be monitored and able to be controlled by a main RCC stationed at the NTNU campus in Trondheim, at an approximate distance of 1.9 km from the operational area of the ferry, as shown in Fig 5b. A 5G mobile communication infrastructure is being built in the operational area to support the operation of the Autoferry.

The autonomous ferry is expected to carry passengers (max 12 on each trip) from one side of the canal to the other. A number of E-navigation functions is expected to be needed to support different operational modes, similar to the ones proposed by Rødseth et al (16) and previously discussed in Sec 2. Moreover, the ferry is expected to communicate with other vessels in the area, and to offer the necessary traffic services to maintain safe navigation routes according to the requirements established in (10). Furthermore, the ferry will be all electric and is expected to integrate new technologies that are highly interconnected; this makes it susceptible to cyber attacks. Therefore, a communication architecture is needed to support the identified goal functions specified earlier, namely safe and secure navigation; and secure and reliable communication.

During the decomposition of the goal functions, the identification of the main, supporting and system functions and sub-functions for the first goal function was based on the relevant literature and influenced by the stakeholders' viewpoints. Specifically, the notion for the decomposition of navigation functions is influenced by the proposed operational modes by Rødseth et al. (16), functions proposed in the MUNIN project (33) as well as DNV.GL's proposed "autoremove" operational mode (34). For the second goal function, the identification of functions and sub-functions was based on the established communication requirements (10). Each function and sub-function was derived to address a certain requirement until all requirements are addressed.

A logical view of the functions and their interactions is shown in Fig 6. In this figure, "Engine Monitoring and Control" refers to the capabilities to monitor the engines status and control them, "Navigation" refers to the capabilities to receive navigation data, establish situational awareness and define safe routes. "Remote" refers to the capabilities being carried through RCC operators, "Autonomous" refers to the capabilities being carried by the APS itself, while "Emergency Remote" refers to the capabilities being carried by the ECT. "Passenger Safety" refers to the capabilities to initiate emergency signals

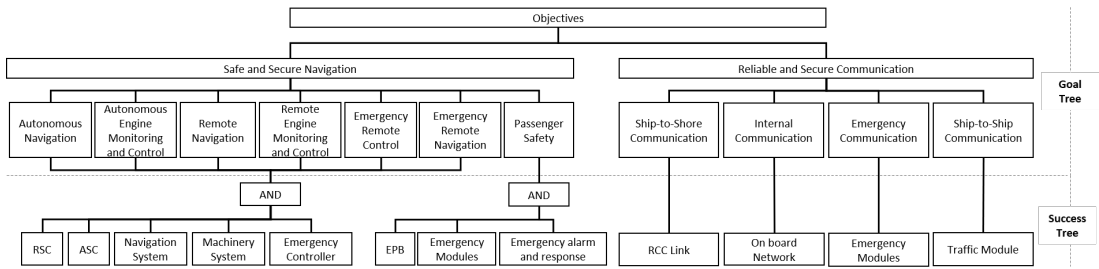
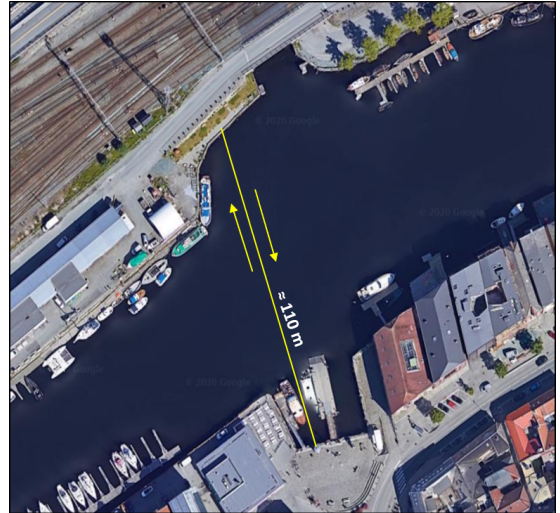


Figure 3. The GTST of the Communication Architecture



Figure 4. APS Use case: the autonomous ferry (Autoferry)



(a) Close view



(b) Wide view

Figure 5. APS Operational Area (Photos by Google Earth(32))

indicating safety-critical event related to passengers. The arrows indicate the direction of networked interactions between the different functions, but they do not capture the transitive interactions between the different functions at the service layer. The functions for safe and secure navigation rely on the communication functions and other supporting functions such as power, security, control and emergency response. Each supporting function is provided by a dedicated system or personnel. On the other hand, the functions for reliable and secure communication enable the navigation functions as well as other functions necessary for network and system management (NSM). A detailed list of the functions that aim to provide safe and secure navigation as well as reliable and secure communication is shown in Table 2 in Appendix A.

5 Communication Architecture

As outlined by Large et al. (35), communication architectures describe both logical and physical interconnections of all the identified elements in an ecosystem from the signal generation to its termination. In this section, the proposed communication architecture is presented, describing the ecosystem of a generic APS, the constituent systems and their subsystems. The architecture is modelled using AADL (13), thus enabling an extended analysis on one hand, and design modifications in the future on the other. In this paper, the modelling and analysis are presented at the

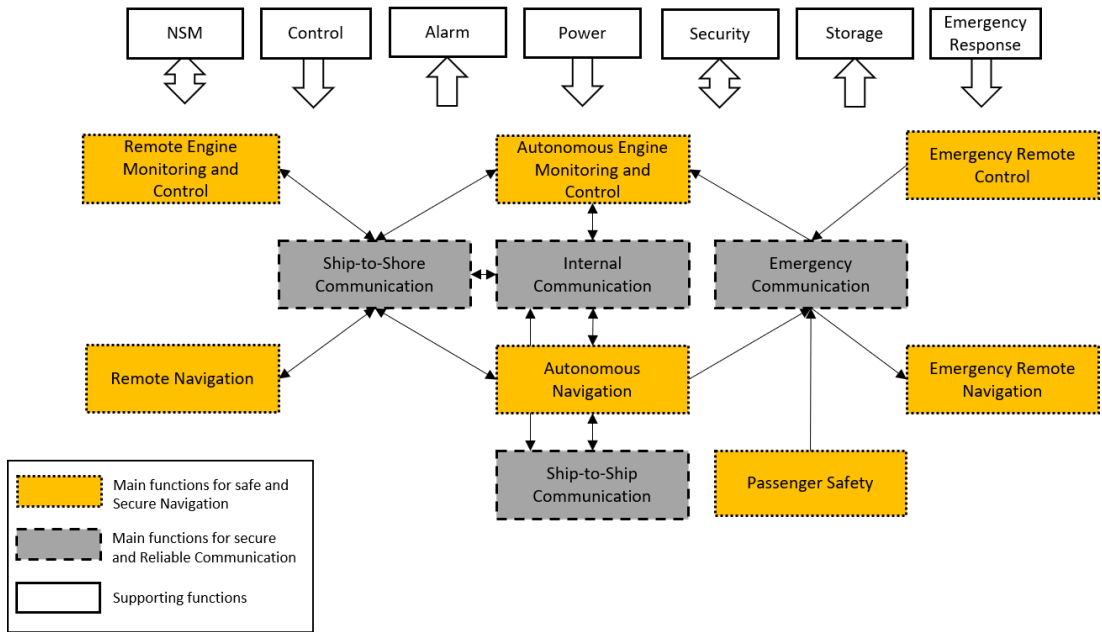


Figure 6. Logical view of the Autoferry functions and their interactions

level of system components, including their properties and interconnections.

5.1 Context View

Initially, the identified system structure reflected in the ST in Figure 3 was distributed across the APS context identified in (10). The distribution was based on the appropriate interactions across the different system functions, which are reflected in Figure 6. Moreover, additional context entities which were not discussed in (10) were suggested to provide design improvements towards realizing the system functions. Namely, mobile network, and Cloud Component (more details in section 5.6 and section 5.5 respectively). The outcome of this initial stage represents the highest level of abstraction with regards to the architecture components, reflected in Figure 7. This view aids the understanding of the various interacting entities in the APS context; explicit details related to each system are discussed in the following subsections.

5.2 APS

The APS itself is the central element of the communication architecture, as it is involved in all the main functions, with the remaining context components supporting its operation. In this section, four subcomponents of the APS onboard architecture are discussed, namely, the onboard network, the Autonomous Ship Controller (ASC), the navigation system, and the machinery system.

The onboard network is responsible for facilitating the different communication functions, while the ASC hosts the logic responsible for autonomous, remote and emergency navigation as well as engine monitoring and

control functions, in addition to other system and network management components. Furthermore, the navigation system is the largest source of data to be traversing through the network and to be processed toward aiding the autonomous, remote and emergency navigation functions. Finally, the machinery system is responsible for supporting the movement ability of the APS and realizing the autonomous, remote, and emergency engine monitoring and control functions. Further discussion for the main sub-components as well as a brief discussion regarding additional expected systems is provided below.

The onboard network architecture presented in Fig 8 utilises two core/distribution tiers for high availability of communication functions, dividing the network into two main segments. The first segment provides ship-to-shore communication functions and limited internal communication functions. It connects the internal system components with the components on the RCC and components hosted on other entities in the context, using high-speed network access. The second segment provides more internal communication functions as well as emergency communication functions through connecting the internal APS systems and subsystems, in addition to integrating some low-speed connections from parts of the context (Aids to navigation, and ECT). Further description for each component in the network architecture is provided below.

5.2.1 The core/distribution tier : This component consists of two parts named A and B, for redundancy and to support network scalability as per the adopted network design principles (36). The main reason for having two parts for the core/distribution tier is to allow the ship network to operate both when communication with the RCC is possible and in the case of communication outage. While the

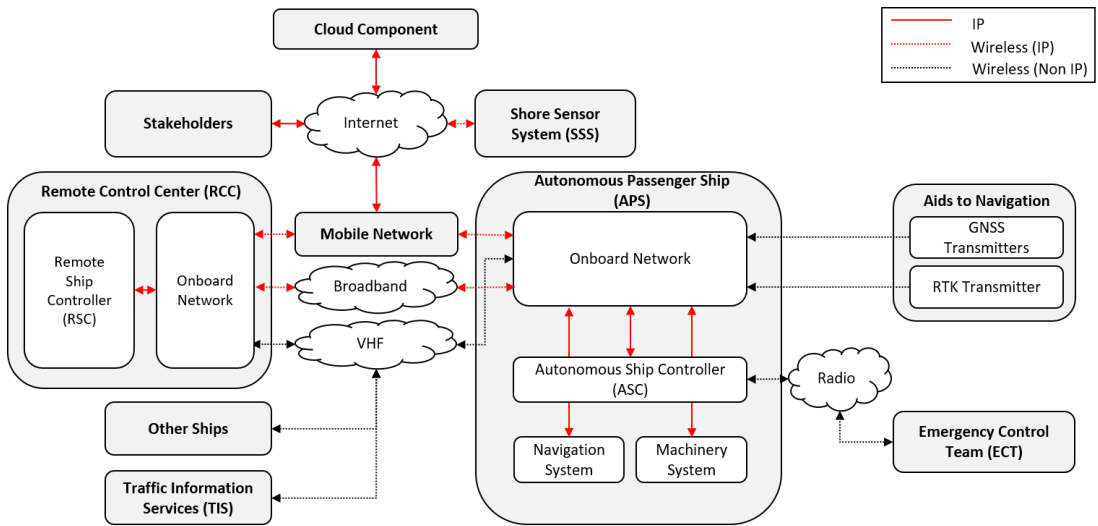


Figure 7. Overview of the APS Context

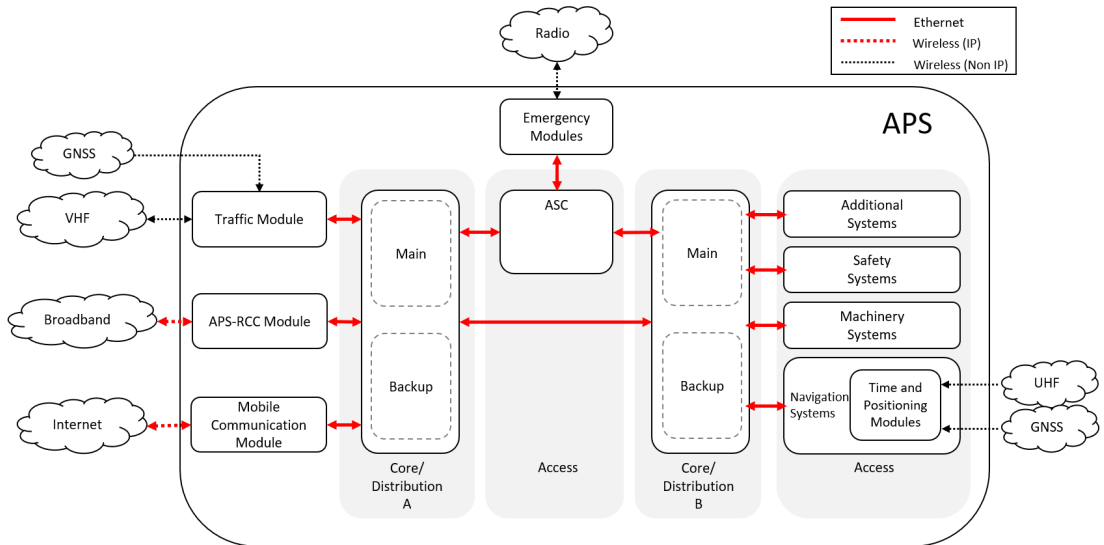


Figure 8. On board Network architecture

communication with the RCC is available, part A will handle the core/distribution related tasks while part B will primarily handle the distribution tasks to the internal networks. In case of loss of communication with the RCC or total failure of part A, part B will take responsibility for the core/distribution tasks, connecting the ASC with the internal networks. Two redundant units are proposed in each part. The units can utilise load sharing according to a load balancing policy, or one of them can be the main and the other stands as a backup. We propose the application of Layer 3 switches with load balancing and inter-VLAN (Virtual Local Area Network) routing capabilities to handle the core connectivity in addition to traffic distribution. This arrangement provides

high availability of external and internal connectivity, in addition to satisfying pre-established requirements related to fault-tolerance, redundancy and capacity. Additionally, appropriate traffic distribution is governed by inter-VLAN routing to satisfy the requirements related to network segregation.

5.2.2 Gateways : A gateway in this architecture operates as a bridge between two networks. Accordingly, the proposed architecture includes several gateways to carry ship-to-shore, ship-to-ship and emergency communication as well as supporting navigation functions. The gateways are represented as modules in the architecture to comply with modular network design principles. Each module represents

a gateway for a communication link without restricting the technology to be used for its implementation.

- **Mobile Communication Module (MC Module):** Provides connectivity to high-speed mobile networks for carrying ship-to-shore functions. 5G is a possible implementation option due to the expected larger bandwidth and lower latency. Additionally, as observed in the literature, LTE and 4G have been evaluated for E-navigation services and showed positive signs with some challenges related to latency. This module will provide the ferry with internet access in order to connect to the cloud component and the RCC.
- **APS-RCC Module:** To satisfy the requirement for redundancy of the link with the RCC, a backup module that provides direct connectivity to the RCC for carrying ship-to-shore communication functions is proposed. Broadband communication technology such as Wi-Fi, Mobile Communication, or the Maritime Radio Broadband (MRB) technology by Kongsberg (37) are considered for implementation of this module since they are internationally recognised wireless communication technologies which are indicated in the requirements. Wi-Fi is suggested if the RCC location is within proximity to the APS while mobile communication can be utilised when out of range of Wi-Fi. However, challenges related to latency are expected when using mobile communication.
- **Traffic Module:** This module aims to provide ship-to-ship communication functions, including communication with Traffic Information Services (TIS) such as the Vessel Traffic Service (VTS), or the River Information Service (RIS). In order to satisfy the ship-to-ship communication requirement, VDES, or AIS are considered for implementation in this module since they provide Line-of-Sight (LoS) communication, and they have been proposed by Kwan (9) as digital communication technologies needed to implement E-navigation services. The traffic module is expected to receive signals from the Global Navigation Satellite System (GNSS) for positioning and timing.
- **Emergency Modules:** Two modules are proposed for providing emergency communication functions. The first module is connected to ASC to provide emergency remote control and navigation functions over radio communication connected with the ECT. The second module, as desired in the requirements, provides a dedicated link over a mobile communication system (e.g. The Universal Mobile Telecommunications System (UMTS)) that allows the transmission of an emergency signal when a passenger onboard the APS press an emergency push button.
- **Time and Positioning Modules:** Two modules are leveraged for supporting the navigation system with positioning and timing data. The expected implementation technologies are GNSS and Real-time kinematic (RTK) receivers. The GNSS receiver provides positioning and timing data while the RTK receiver provides position correction data.

5.2.3 **ASC** : The Autonomous Ship Controller hosts the logic responsible for performing the functions related to

autonomous, remote, and emergency navigation, engine monitoring and control. The proposed architecture of the ASC system includes a main and a backup system, each of which consists of several subsystems. This arrangement can utilise the server virtualization technology to simplify the management of such systems in addition to providing the required high availability. The proposed architecture for the ASC is shown in Fig 9 and is discussed in detail below:

- **ANS:** The Autonomous Navigation System (ANS) hosts the logic to carry the navigation functions in the different operational modes such as collision avoidance, situational awareness, and operational mode alteration. Additional features in the backup unit are the connectivity to the emergency control module to enable emergency remote control by the ECT in case of emergency and loss of ship-to-shore communication with the RCC. Additionally, the backup unit is expected to host the routines for the several Minimum Risk Conditions (MRC) that governs the APS operations under the Fail-to-safe operational mode.
- **AEMC:** The Autonomous Engine Monitoring and Control (AEMC) system hosts the logic for performing engine monitoring and control functions through retrieving engine data and forwarding commands to control the movement of the ship. Similar to the ANS backup unit, the AEMC backup unit is provided with connectivity to the emergency control module.
- **Network and System Management:** is a group of components that host required services. These components include User Access Management (UAM); Connectivity Management (more details below), the ship's Digital Logbook; and additional network and system management entities. Such entities include the Domain Controller; remote access (jump) servers; backup servers; and a system and network documentation repository. The digital logbook is expected to host recording and logging capabilities of important data, as indicated in the requirements.
- **Integrated Ship Safety and Security Management System (ISM3S):** is a group of components that host services related to the safety and security of the ship and provide supportive safety and security functions. Possible components are Security Information Event Management (SIEM), primary Intrusion Detection and Prevention Systems (IDS/IPS), and the ship's Central Alarm Management (CAM) which hosts the systems responsible for safety-related alarms in compliance with the pertinent safety regulations.
- **Connectivity Manager** : The concept of a connectivity manager was proposed by Höyhty et al (18) as an intelligent entity responsible for ensuring the robustness of the ship's communications in any and all environments. Its application was proposed in satellite-terrestrial integration by Höyhty (19). We adopt the notion for the need of an intelligent network management entity in the APS ecosystem due to the increased autonomy, and we propose the APS Connectivity Manager as an autonomous network manager with functions aiming to reduce the need for human

network operators. The proposed Connectivity Manager will be hosted in the ASC network and is expected to provide the identified communication functions (see Appendix A) to satisfy a number of the communication requirements established in (10). Several components are proposed as part of the APS Connectivity Manager; these are discussed in the sequel:

- **Quality of Service Controller (CM-QoS):** This component is responsible for maintaining the required level of service quality. Other than managing the QoS rules through the establishment, enforcement, and management of rules, this module is responsible for enabling traffic prioritization and traffic re-direction. It handles the establishment of a prioritization policy, its communication to the appropriate network devices and any additional tasks related to traffic prioritization. Moreover, this component is responsible for managing the functionality for diverting communication paths within the network, depending on the available communication resources. This can be achieved by monitoring the status of the network links and updating routing and Inter-VLAN routing tables based on the connectivity state to direct traffic from the available sources to the available destinations. Additionally, this controller is responsible for managing the traffic load over the links, based on a pre-established load balancing rule or managed by the operator.
- **Network Monitor and Troubleshooter (CM-NMT):** This component is responsible for collecting the relevant network-related logs, performance indicators, in addition to providing automatic network self-checking and to triggering network troubleshooting by the operator. The component is also responsible for generating, along with the CAM software, the appropriate alarms.
- **Network Software Updater (CM-NSU):** This component is responsible for managing the retrieval, installation, verification of updates and recovery from them in case of failure.
- **Network Segmentation Manager (CM-NSM):** One of the most critical aspects of the proposed architecture is the segregation by design for reliable and secure network operation. This component is responsible for managing the network segregation feature through the establishment and enforcement of the segregation policies, as well as validating their enforcement.
- **Network Security Coordinator (CM-NSC):** The need for a dedicated entity for managing cybersecurity risks was established in (10) and is proposed in this paper, as is also the case with the IS3MS. The communication networks play an essential role in managing cyber-attacks, especially wireless networks. Accordingly, coordination between the connectivity manager and

the IS3MS related to the communication of unexpected events and the enforcement of the various policies is expected.

- **Network Device Backup Controller (CM-NDBC):** This component is responsible for retrieving and maintaining the backups of the network devices, in accordance with a backup policy. Additionally, it provides access to these backups for the other Connectivity Manager components, for example for CM-NI, CM-NSU and others, if needed.

We propose the development of the Connectivity Manager based on the FCAPS network management model (38) with appropriate adjustments, so as to take into account the autonomous operational environment.

5.2.4 Navigation Systems : As discussed earlier, the navigation system is a critical component of the APS that is responsible for collecting the required data for sensing the surroundings and enabling the APS to make informed decisions. No standardised navigation system has hitherto been proposed for the APS. The main design decisions related to the communication network is to avoid traffic congestion due to the transmitted data from the extensive amount of sensors (lidars, radars, video cameras, EO cameras) and to support scalability if the required amount of supporting components is to increase (e.g. more sensors). This could be achieved by applying two solutions. The first solution is to utilise Sensor Processing Units (SPU) to reduce the amount and frequency of transmitting sensor data. However, such a solution has been proven incapable of providing sufficient operational guarantee in case of faults (39), in addition to the expected increased latency that may hinder the control operation (40). The second solution is to connect the SPUs or to distribute sensors across multiple switches so that the traffic flowing to the switch stack in Core/Distribution B is distributed over multiple interfaces. Such an arrangement would add resilience to the navigation system by providing multiple access paths for the sensor data since even in the case that few SPUs or sensor switches failed, the remaining units will still be able to communicate some sensor data to the ASC.

The Global Navigation Satellite System (GNSS) and the inertial measurement unit (IMU) play a crucial role in the ship navigation system to provide accurate and timely positioning and timing data. At the same time, both components rely on external signals received through GNSS and RTK receivers, respectively. GNSS signals are susceptible to various attacks such as spoofing and jamming, and they require additional processing to ensure their security (41). A possible arrangement for the navigation system and its connectivity to the APS internal network is shown in Fig 10a: the navigation system transmits the sensor data as well as the GNSS processed information to the ANS, to aid the navigation functions.

5.2.5 Machinery Systems : Similar to the navigation systems, there is no standardized machinery system proposed for the APS. A possible arrangement of the machinery system and its connectivity to the APS internal network is depicted in Fig 10b. The Dynamic Positioning (DP) system

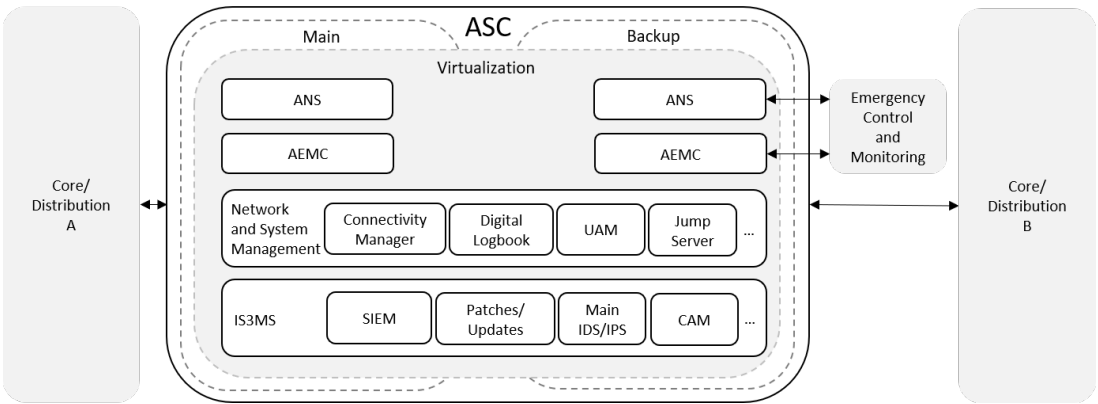
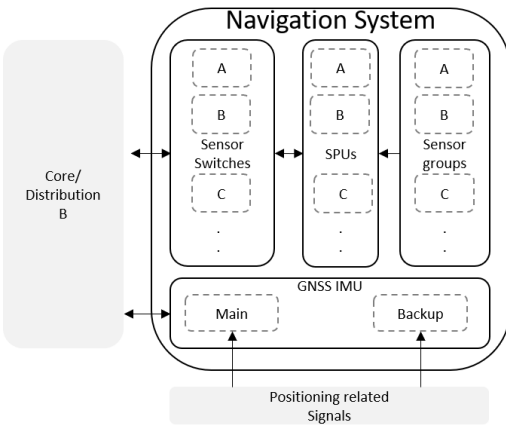
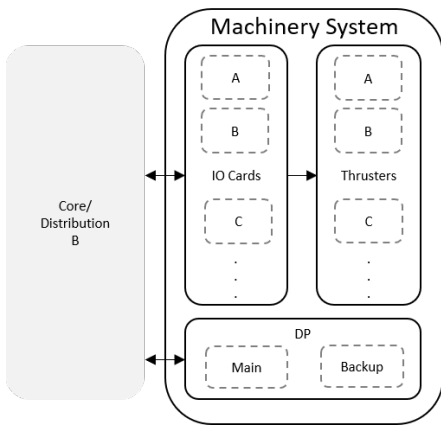


Figure 9. ASC architecture



(a) Navigation System



(b) Machinery System

Figure 10. Simplified Overview of the navigation and machinery systems

is responsible for maintaining the floating structure of the APS in a fixed position as well as on its established route by utilizing active thrusters. Furthermore, for functional redundancy, multiple thrusters are expected to be utilized and be connected to the core/distribution part B through Input/Output (IO) cards (42) which convert packet data (e.g. commands) coming from the control function (e.g. DP system), and send them to the thrusters to navigate the APS. The machinery system, together with the AEMC perform the engine monitoring and control functions through sending engine performance parameters and receiving route specification and control commands from the AEMC.

5.2.6 Additional APS Systems : Additional systems are expected to be attached to the APS internal network to provide supporting functions, as mentioned in Sec 4; these may include safety systems, power management systems, and passenger management and entertainment systems. Specifically, systems dedicated to providing safety functions for the APS and passengers include anchor drop, horn, alarm, lantern and user panel to control such systems manually. The systems are connected to the ship’s network through General IO module, such as a PLC, while further security mechanisms are expected to monitor the commands going toward the machinery and safety systems to protect them from malicious attacks. By having the Core/Distribution part B separated from Core/Distribution part A, the network topology enables the network to be scalable and able to accommodate new systems in the future. Such systems can be added to the network by connecting a system’s gateway or IO Card to the switch stack in Core/Distribution part B and having a dedicated VLAN created for them. Consequently, proper inter-VLAN routing is to take place to route the traffic to and from those systems appropriately.

5.3 RCC

In order to support remote navigation, as well as remote engine monitoring and control functions, it has been determined that a remote controlling entity located in a separate physical location is required. This entity can be hosted onshore or onboard another ship (10). In Fig 11, we propose a possible network architecture for the RCC

that is compatible with the previously proposed network architecture on board the APS. Further description for each component in the network architecture is provided below:

5.3.1 On board network : In this system, we propose a similar arrangement to the one discussed in Sec 5.2 for the Core/Distribution tier and gateways, with the exception that there is no need for two Core/Distribution parts, due to the anticipated small number of integrated systems on the RCC and the fact that the traffic module is only needed if the RCC is hosted onboard another ship.

5.3.2 RSC : The Remote Ship Control (RSC) consists of components responsible for carrying the remote functions in addition to managing the RCC and APS networks, systems and security. The proposed architecture of the RSC system as shown in Fig 12 consists of a main and a backup system, each of which consists of several subsystems. Similar to the ASC, this arrangement can also utilise the server virtualization technology, for the same reasons (refer to the discussion of the ASC architecture in 5.2). Another advantage for the application of server virtualization technology would be to facilitate the migration of RSC from one RCC to another. The systems and subsystems in the RSC and the ASC perform similar tasks. The Remote Navigation System (RNS) and Remote Engine Monitoring and Control (REMC) are expected to perform similar tasks to their equivalent autonomous systems ANS and AEMC, respectively. The main difference is that in the RNS and REMC, the analysis and control can be performed by an operator. Also, the Network and System Management and the IS3MS perform similar tasks, but with a focus on the RCC network, and they support the operations of their corresponding systems onboard the APS.

5.4 ECT

As described in Sec 5.2, there are two emergency-related modules onboard the APS that are utilised to establish emergency communication with the ECT. We propose a simplified architecture of the ECT, as shown in Fig 13a. The Emergency Alarm and Response system is responsible for performing the required tasks when the passenger emergency button on board the APS is pressed, while the emergency signal is received over a mobile communication link. A possible response action could be to raise a vocal and illuminated alarm so that the ECT notice an emergency and make their way toward the APS. The Emergency controller is a separate control system for managing the APS by a human operator through a LoS or short-range communication as suggested by MUNIN (21). In order to perform emergency navigation and engine monitoring and control functions, the emergency controller should be compatible with the backup ANS and AEMC (refer to the ASC architecture in Sec 5.2).

5.5 Cloud Component

Several cloud components such as the Maritime Connectivity Platform (MCP) (43), and DNV.GL's Veracity platform (44) have been utilised in maritime operations. In this work, we propose the utilization of such online applications to provide several functionalities. One goal for the application of a cloud component in the architecture proposed herein

is to facilitate communication among the different APS stakeholders as described in (10); this can be accomplished through a dedicated portal for the APS ecosystem. An additional service could be utilised for the APS and RCC registration and binding in order to establish communication links between them. Additional services can be leveraged in the case of utilizing a cloud service such as online storage to backup the APS's and RCC's digital logbooks and other essential data. Furthermore, the application of a cloud component can facilitate the connection of additional context entities to the APS ecosystem without requiring changes in the proposed architecture. For instance, a Shore Sensor System (SSS) is expected to be implemented to aid the autonomous navigation functions, and a possible communication channel with the APS could be through a cloud service that pulls sensor data from the SSS and pushes this data or a processed version of these data to the APS through the APS MC module. Additional cloud services related to the connectivity manager and IS3MS could be utilised through the cloud component. A simplified architecture for the cloud component is shown in Fig 13b.

5.6 Mobile Network

The operational area of the APS in inland waterways enables the APS ecosystem to utilise the high-speed mobile communication infrastructure. As observed in the literature (Sec 2), many works have proposed, and some have evaluated the application of various mobile communication technologies in the maritime domain and for autonomous vehicles as well. 5G has yet to be evaluated in such applications, but it has been proposed by many. On the other hand, LTE and 4G technologies have shown promising results in previous communication architectures for traditional ship navigation. We propose the utilization of mobile communication technologies (see Fig 14); this satisfies the requirement for minimum bandwidth of 4 Mbps and maximum latency of 1 second for remote control and up to 2.5 seconds for HD video. To achieve flexibility in the architectural design, the mobile communication has been modelled as a forwarder of communication between the connected context entities, i.e. APS, RCC, SSS, and cloud component as well as a gateway to internet access. Thus, we pose no restriction on the communication technology to be employed (4G, LTE, or 5G), and we leave the implementation option to the technology that best satisfies the pertinent requirements. Nonetheless, it must be noted that an appropriate Service Level Agreement (SLA) of the ASP operator with the service provider of the mobile communication should be established, to maintain the required Quality of Service (QoS). We support the suggestion proposed by Höyhty et al (18) regarding the utilization of the mobile edge computing (MEC) technology supported by mobile communication infrastructures such as LTE, 4G and expected to be improved in 5G. The possible implementation of MEC in the APS architecture would move the suggested cloud component into the Mobile Communication infrastructure, which could drastically reduce the latency (see Fig 14).

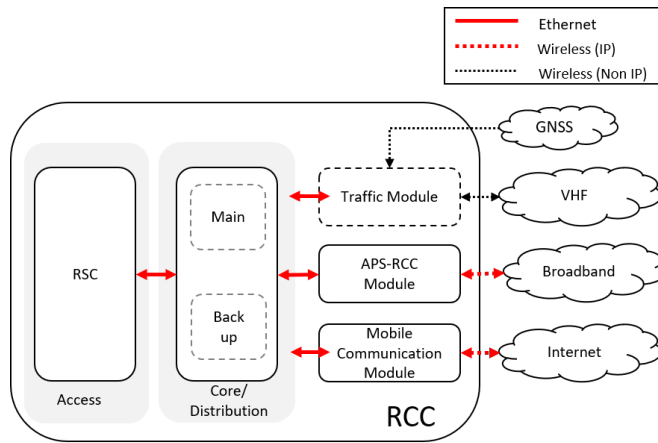


Figure 11. Remote Control Center Network Architecture

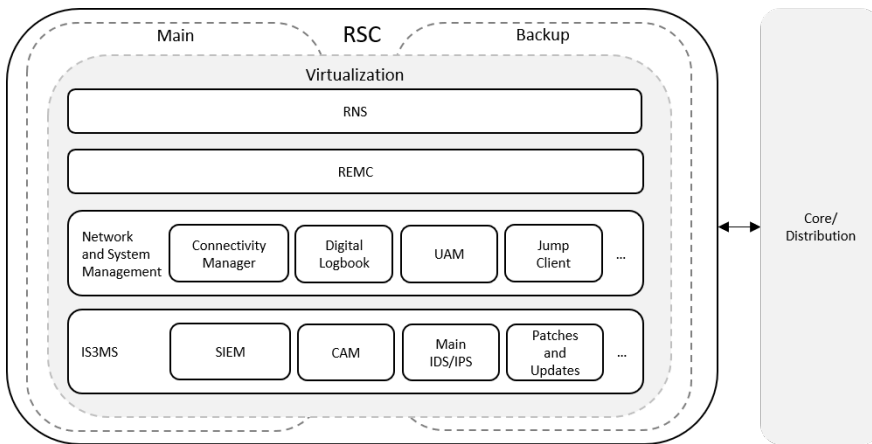


Figure 12. RSC architecture

5.7 SSS

A group of sensors are expected to be mounted on the shore side to aid several functions related to navigation, docking/undocking, passenger loading/unloading, etc. The arrangement of the SSS is still under development, and in this paper, we suggest that the SSS utilises the application of a cloud component to transfer the sensor data to the APS without adding additional communication modules onboard the APS. Such an arrangement enables other APSS approaching the shore to utilise the SSS as well.

6 Verification

In order to verify that the proposed architecture meets the requirements in (10) that relate to communications, a number of operational scenarios were defined. By leveraging these, we showcase how the architecture provides the required functionality. Additionally, the IP network part of the proposed architecture was implemented in a network simulator; this allowed experimentation that also showed the

architecture’s ability to meet the established communication requirements of the APS.

6.1 Verification Scenarios

In this section, a number of operational scenarios that use the communication capabilities of the architecture, and are drawn from the expected operations of the Autoferry (Fig (5)) and influenced by a number of scenarios delivered by the MUNIN project (33) are described.

6.1.1 Traffic Communication: A crucial requirement for the APS is the ability to communicate with the surrounding ships using LoS communication. It is also recommended that the ship follows the guidance provided by traffic services in the area, such as broadcast messages from VTS. At all times, the APS is expected to receive broadcast traffic messages through the traffic module (Section 5.2.2) and use them to determine safe routes. At the same time, the traffic module is used to broadcast the APS status to surrounding ships for

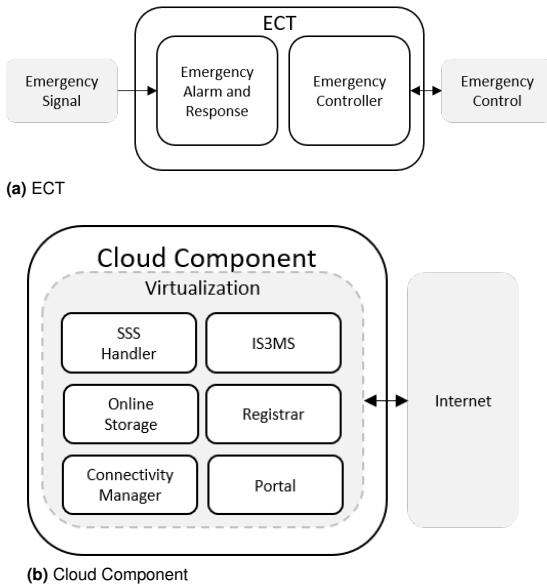


Figure 13. Architectures of ECT, and cloud component

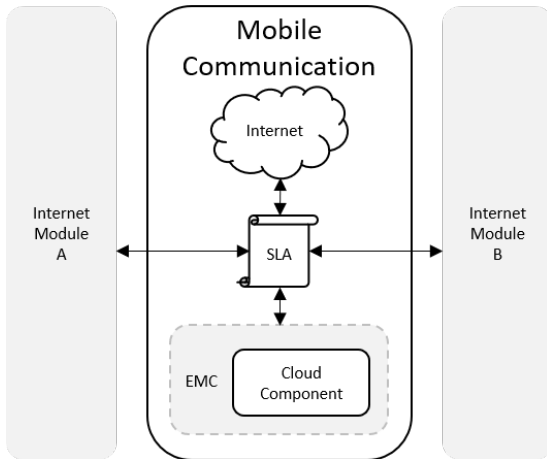


Figure 14. Mobile Communication Utilization

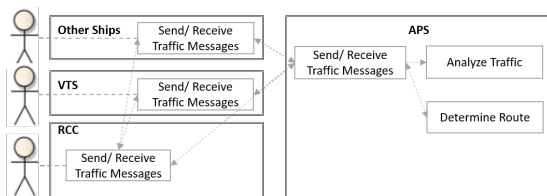


Figure 15. Traffic communication Scenario

safe navigation. Fig 15 shows a scenario of a typical ship-to-ship communication between the APS, other ships and VTS. Moreover, the RCC is expected to communicate with

the surrounding ships if the ship-to-ship communication was not sufficient.

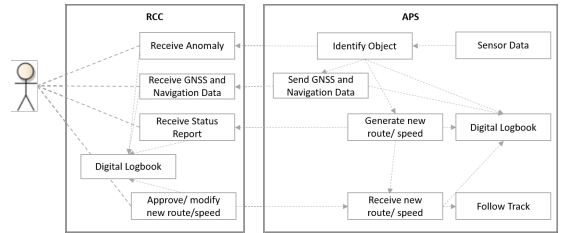


Figure 16. Collision detection and Avoidance Scenario

6.1.2 Collision detection and Avoidance: During the voyage, the APS operates in autonomous execution mode, and this includes the ability of the ship to perform autonomous and remote navigation functions and engine monitoring and control functions. Fig 16 reflects a scenario of expected communication within the APS and with the RCC to carry out those functions. The ANS collects data from different sensors to perform autonomous navigation functions. In case of a detected object that poses a possible collision threat, the ANS generates a new navigation plan that may include another safe route or modified speed. Meanwhile, a summary of the navigation and GNSS data is sent to the RNS for monitoring together with the newly generated navigation plan. The operator in the RCC receives the navigation data and decides whether to approve or modify the ANS plan for collision avoidance. The ANS waits for the RNS approval or modification for a specific time. If received, the ANS adopts the commands and forwards them to the machinery system. Otherwise, the ANS carries on with its own new navigation plan.

In order to successfully realize such a scenario, both internal and external communications are expected to operate reliably according to the established requirements. For instance, high availability of the communication link with the RCC is needed for the remote navigation, engine monitoring and control functions; this is realized through defining the MC module and the APS-RCC module as a redundant pair. Moreover, a fault-tolerant network is required to reliably carry autonomous navigation, engine monitoring and control operations, even in case of a single component failure. This requirement is addressed by the proposal of redundant network design and redundant devices in both of the core/distribution tiers, the ASC, redundant sensors, redundant thrusters, etc.

6.1.3 Loss of Communication: An APS is prone to communication loss. To maintain safe operations, the APS is expected to operate in a fail to safe mode in case of communication loss. Fig 17 describes a scenario showing the APS internal communication and emergency communication with the ECT to handle the loss of communication with the RCC and to maintain safe operations. Initially, as long as the ship-to-shore communication with the RCC is available, the ANS receives continuous updates of several MRC plans which govern the ship operations under fail to safe mode

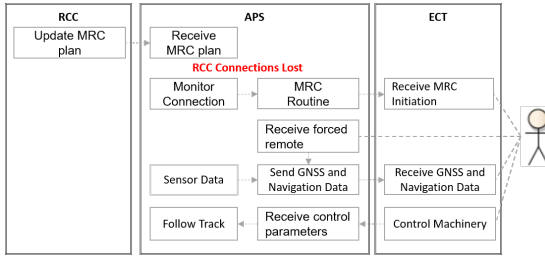


Figure 17. Loss of Communication Scenario

depending on several factors (location, wind, etc). The CM-NMT module that is part of the connectivity manager on the APS has implemented a feature to monitor the connection with the RCC continuously. When the connection is lost, the most appropriate MRC plan is initiated by the ANS, and a signal is sent to the ECT of this action. When the personnel on the ECT receives this signal and have the ability to respond, they send a forced remote signal to the APS forcing it to operate under remote control mode in order to perform emergency navigation and control functions by sending the navigation and GNSS data to the operator on board the ECT; the operator uses these data to navigate the ship safely.

To successfully realize the operation in the described scenario, several requirements need to be met. The requirement related to link quality monitoring and notification is met by the CM-NMT module (Section 5.2.3). Additionally, during the fail to safe mode, it is crucial that the internal network is available for autonomous navigation, engine monitoring and control.

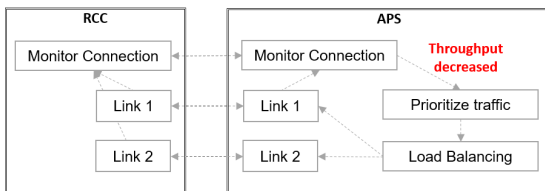


Figure 18. APS and RCC Link Degradation Scenario

6.1.4 APS and RCC Link Degradation: The communication links between the RCC and the APS are prone to quality degradation resulting from the loss of one of them or due to regular link performance issues. Fig 18 describes a possible scenario of the APS ability to deal with link quality degradation. A communication service (CM-NMT) on the APS continuously monitors the quality of the connections with the RCC. The service initiates an alarm of a quality degradation related to decreased throughput, which may affect the quality of the establishment of situational awareness of the RCC operators and may also reduce their ability to intervene in case of emergency. Therefore, the CM-QoSC service prioritizes the traffic based on a pre-established prioritization policy and utilizes a load balancer to distribute the traffic across the multiple links with the RCC

by pushing the traffic with the highest priority in the better link and the traffic with the lower priority in the link with lower quality, if it is still active.

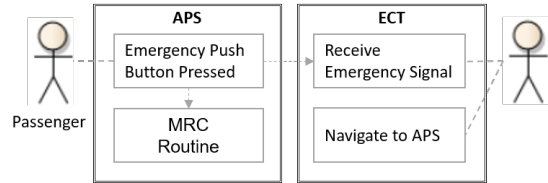


Figure 19. Pressing of Emergency Push Button Scenario

6.1.5 Emergency Push Button: The highest priority in the APS ecosystem is the safety of the passengers. Therefore, one of the main requirements is the establishment of a communication link for an emergency push button (EPB). Fig 19 describes a scenario of the tasks performed when the EPB is pressed. If pressed, due to a situation perceived by passengers as dangerous (e.g. a passenger falls off the ship), passenger safety functions are invoked, including the initiation of the appropriate MRC routine and the transmission of an emergency signal to the ECT. Then, the ECT will navigate to the ship to intervene and perform a rescue operation.

6.2 IP network simulation

In this section, we provide an evaluation of the IP-based network of the proposed architecture by means of experimentation using the GNS3 simulator. GNS3 (Graphical Network Simulator-3) is a software capable of emulating real devices (routers, switches, servers, PCs, etc) using real software images. It allows users the ability to flexibly configure, test, develop their networks without the high cost of real device (45). We implemented and configured the IP-based network in a manner that allows verifying that the proposed architecture meets its design requirements as established in (10).

As shown in Fig 20, both the networks of the APS and the RCC were implemented as well as their interconnections. Moreover, virtual servers were integrated into the network toward implementing the different ASC components (ANS, AEMC, IS3MS etc.). The core/distribution tiers A and B in the APS network as well as the single-tier in the RCC network were implemented using redundant Layer-3 switches. The implemented redundancy protocol is the Gateway Load Balancing Protocol (GLBP). The MC module and the APS-RCC modules were implemented as gateway routers since GNS3 does not directly support wireless communication. However, for the purpose of verifying the architecture, it has been decided that any routing device capable of routing incoming and outgoing traffic through a third network would suffice. The implemented IP routing protocol is Open Shortest Path First (OSPF). Moreover, to satisfy a requirement related to the employed transmission protocol, the implemented transmission protocol is Transmission Control Protocol (TCP) which is compliant with an international standard (46).

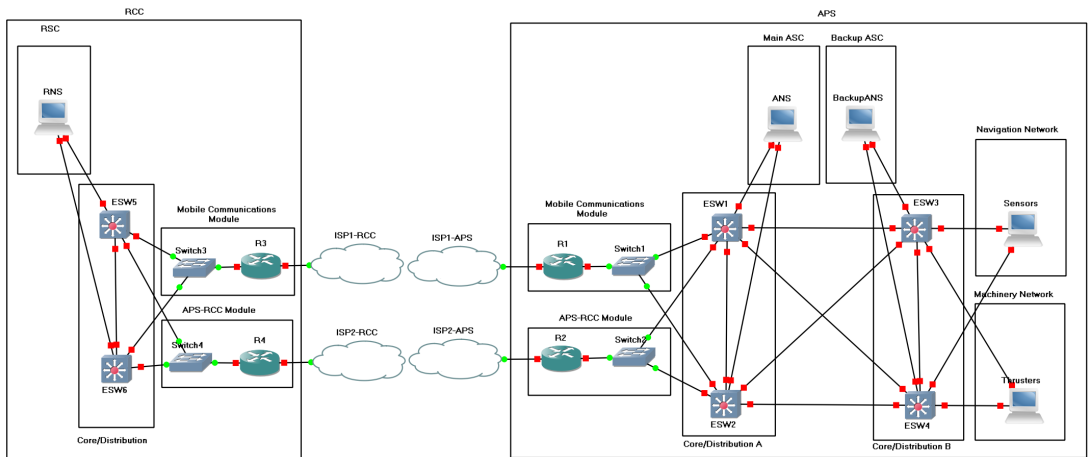


Figure 20. Implementation of the IP network in the proposed architecture using GNS3

Within the capabilities of GNS3, a number of experiments were carried in the simulated architecture to verify that the requirements related to **RCC link redundancy, fault tolerance, network segregation, network redundancy, network troubleshooting, QoS, link quality monitoring and notification, traffic prioritization and traffic redirection** are addressed.

The results can be inferred from Table 1 which depicts the connectivity matrix for all the hosts in the different implemented networks in GNS3. The table reflects the minimum and the maximum number of routing devices allowing traffic between the hosts in each network and the hosts in the other networks. Between each network, there is a maximum number of routing devices. For instance, there exist eight maximum possible routing devices between the main ASC network and the RSC network while only two between the backup ASC and both the machinery and navigation network. For each network pair, the minimum number of nodes required to maintain connectivity was calculated by performing a systematic shutdown of routing devices between them until the connectivity is completely lost without the ability for automatic recovery. For instance, the connectivity between the main ASC network and the RSC network can withstand up to four routing devices to fail (with maximum 1 out of each 2 redundant pair), while the connectivity between the backup ASC and both the machinery and navigation networks can only withstand one routing device to fail. Nevertheless, the results reflect that both internal and external networks are fault-tolerant, and the connectivity with the RCC is redundant. Moreover, the traffic redirection requirement is implicitly satisfied due to the automatic redirection of traffic implemented by the enforcement of weighted load balancing between each network device pair. Moreover, the implementation of network segregation according to the requirement is indicated in Table 1. The network segregation policy has been implemented in this work by creating a virtual LAN (VLAN) for each network; the connectivity between those networks is governed by defining Access Control Lists

(ACLs). For instance, the hosts in the RSC network in the RCC can only access the hosts in the main ASC network, while the hosts in the main ASC can access the hosts in all the networks.

The IP network behaviour in the scenario mentioned in Sec 6.1.3 is simulated during these experiments. When two routers in the pairs (R1 and R2) or (R3 and R4) were shut down, loss of communication between APS and RCC occurred. The "track" feature implemented by the CM-NMT in the Core/Distribution tier A detected that a link had been lost. This feature can be used to notify the ANS to alter the operational mode. Notably, the internal network was not affected by the loss of communication; this enables the ship to perform all the functions that do not require RCC interference. Further, the scenario mentioned in Sec 6.1.4 is also simulated. When one of the routers R1 or R2 was shut down, the "track" feature detected that a link had been lost and the CM-QoS applied a load balancing policy to prioritize traffic, thus meeting the Traffic Prioritization requirement. The link with the RCC was restored automatically using the redundant module, thus showcasing the high availability of the link between APS and RCC.

Moreover, to satisfy the Network Troubleshooting, quality monitoring and notification as well as the operator-triggered traffic redirection requirements, most of the implemented network devices are managed with the Secure Shell (SSH) protocol enabled; this allowed troubleshooting and configuration using SSH within GNS3 from both the RCC and APS networks. Moreover, the notification capability is implemented using the "track" feature that notifies when a link status is changed. The monitoring and troubleshooting capabilities should be hosted in the main ASC network as it has oversight over all other networks, as shown in Table 1. Advanced troubleshooting and monitoring capabilities are targeted for implementation as part of the Connectivity Manager (Section 5.2.3) in future work.

Table 1. Connectivity Matrix for hosts in the implemented networks

Networks	Main ASC	Backup ASC	Nav. Network	Mach. Network	RSC
Main ASC	N/A	2/4	2/4	2/4	4/8
Backup ASC	2/4	N/A	1/2	1/2	0
Nav. Network	2/4	1/2	N/A	0	0
Mach. Network	2/4	1/2	0	N/A	0
RSC	4/8	0	0	0	N/A

For detailed information regarding the verification criteria in the simulation for the targeted requirements, the reader may refer to Table 3 in Appendix B.

6.3 Applicability

The architecture design leveraging the network design principles related to Hierarchy, Modularity, Resilience and Flexibility (see Section 2) enables it to support a wide range of use cases, different than the one targeted in the Autoferry project (see Section 4). In the near future, it is expected to have several APSs operating in the same area. Such operation requires coordination with the RCC and communication among the ships. The proposed architecture can accommodate this operation by virtue of the cloud component: A cloud service can be developed to bind different APSs to the appropriate RCCs. Additionally, other cloud services can be developed to facilitate communication among different APSs. The topology of the scaled network will be a hybrid between centralized star topology with the RCC in the centre managing several APSs, and P2P topology with Ship-to-Ship communication.

Furthermore, avoiding restrictions regarding the choice of the implementation of technology in the gateway modules provides flexibility for various APS routes. For instance, the implementation of high bandwidth, low latency mobile communication technology such as 5G could enable longer routes for the APS.

6.4 Limitations

In this section, we discuss the limitations in the communication architecture design and verification processes and the efforts to overcome them. The limitations are related to the following:

1. Although the APS system functions and expected operations have been previously defined, the APS systems are still under development. This limited the ability to customize design decisions to that would be more suitable in the future APS and confined the architect with best-judgment decisions based on previous experience, discussions with other project

members, and future expectations discussed in the literature.

2. Some requirements lack qualitative or/and quantitative metrics to sufficiently verify their satisfaction in the architecture design. Examples of such limitations:
 - (a) No verification metrics for reliability in C-X-2.
 - (b) No QoS requirements have been defined to verify C-X-6.
 - (c) Measurable metric for redundancy is not provided in C-X-9.

Efforts to overcome this limitation were made by formalizing design-level and implementation-level verification metrics of the requirements as well as suitable verification method as shown in Table 3 in Appendix B.

3. Limited simulation capabilities exist to simulate heterogeneous networks consisting of IP and Non-IP components. Because of this, we were unable to verify the proposed non-IP components using simulation. Therefore, we utilized scenarios to demonstrate the functionality of such components. Furthermore, a testbed that includes the proposed IP and non-IP components in the architecture is undergoing and considered for future work.
4. Additional architecture analysis methods such as technical risk analysis, trade-off studies, cost analysis, usability, dependability, and maintainability analysis were deemed out of the scope of this paper. Nevertheless, a range of these methods, inter alia, are considered in later stages of the architecture development.

7 Conclusions and future work

Many aspects of our modern life are undergoing digital transformation. Autonomous Passenger Ship (APS) is an example of such transformations. APS relies on many interconnected components to carry passengers in urban water channels safely and securely; this requires the definition of a communication architecture capable of connecting all these components. Therefore, a multidimensional design is required to capture the architecture from different perspectives within the APS operational context. Furthermore, with many involved stakeholders in such technology, the communication architecture needs to achieve the goals and satisfy the requirements communicated by these stakeholders.

An adapted and pre-specified multistep model of incremental and evolutionary development was utilized to develop the architecture by following a generic system life cycle model starting with defining the concept and subsequently the system that aims to realize it. In this regard, the Architecture Analysis and Design Language AADL as well as a network simulator were leveraged in the development and analysis of the architecture design.

At the time of writing this paper, there exists no operational APS to fully evaluate the proposed architecture fully. Instead, we relied on a description of an APS use case, namely the Milliampere ferry, as well as a group of operational scenarios to verify the architecture's ability to perform the intended functions. Finally, some

aspects of the architecture have been verified using a network simulator (GNS3); this showed that the architecture meets requirements related to RCC link redundancy, fault tolerance, network segregation, network redundancy, network troubleshooting, QoS, link quality monitoring, and notification, traffic prioritization, and traffic redirection.

The methodology followed for the development of the communication architecture allowed the integration of stakeholders' communicated goals and addressed their requirements in a verifiable manner. Additionally, it allowed the influence and adoption of design artifacts from the literature and relevant best practices and standards in the industry. This allowed the architecture to integrate and suggest features that make it scalable, flexible, and expandable.

As regards directions of future work, note that the work presented in this paper is part of the ongoing Autoferry project (5). An instance of the proposed communication architecture will be implemented to support the operations of a real autonomous passenger ferry (Milliampere). The Milliampere and its supporting systems (Navigation, Machinery, etc) are still under development. Complete evaluation of the proposed architecture will become possible when these systems become available. Until then, the proposed communication architecture will be complemented by a cybersecurity architecture to reduce the risk of cyberattacks. Additionally, a testbed will be developed with real network devices, to allow experimentation with different options for implementing the APC-RCC and the MC modules. The technologies targeted for experimentation are LTE, 4G, 5G and different WiFi versions. Additionally, the testbed will enable further evaluation of the architecture's ability to perform expected functions such as Ship-to-Ship communications, as well as several Internal and ship-to-shore communication functions.

Furthermore, integration between the implemented GNS3 architecture and real situational awareness systems under development by other project members is underway, in addition to adding visual simulation capabilities to the Autoferry. This setup will enable penetration testing of the APS network and of some of its sub-systems, with the ability to observe the result of cyber attacks on the simulated Autoferry. This feature will be useful in studying the effect of cyber attacks on the security and safety of the Autoferry systems.

References

- [1] Schallmo DR and Williams CA. History of digital transformation. In *Digital Transformation Now!* Springer, 2018. pp. 3–8.
- [2] Heilig L, Lalla-Ruiz E and Voß S. Digital transformation in maritime ports: analysis and a game theoretic framework. *Netnomics: Economic research and electronic networking* 2017; 18(2-3): 227–254.
- [3] Rødseth ØJ and Burmeister HC. Developments toward the unmanned ship. In *Proceedings of International Symposium Information on Ships-ISIS*, volume 201. pp. 30–31.
- [4] Projects carried out by members of nfas. <http://bit.ly/NFASProjects>.
- [5] Autonomous all-electric passenger ferries for urban water transport. <https://www.ntnu.edu/autoferry>.
- [6] Havdal G, Heggelund CT and Larssen CH. *Design of a Small Autonomous Passenger Ferry*. Master's Thesis, NTNU, 2017.
- [7] Patraiko D. The development of e-navigation. *TransNav, International Journal on Marine Navigation and Safety of Sea Transportation* 2007; 1(3).
- [8] IMO M. 85/26/add. 1 report of the maritime safety committee on its eighty-fifth session. *International Maritime Organization, London* 2008; .
- [9] An K. E-navigation services for non-solas ships. *International Journal of e-Navigation and Maritime Economy* 2016; 4: 13–22.
- [10] Amro A, Gkioulos V and Katsikas S. Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation. In *Computer Security*. Springer, 2019. pp. 69–85.
- [11] Rødseth ØJ and Tjora Å. A system architecture for an unmanned ship. In *Proceedings of the 13th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT)*. Verlag Schriftenreihe Schiffbau, 2014 Redworth, UK.
- [12] Basnet S, Banda OAV, Chaal M et al. Comparison of system modelling techniques for autonomous ship systems. In *Proceedings of the International Seminar on Safety and Security of Autonomous Vessels (ISSAV) and European STAMP Workshop and Conference (ESWC) 2019*. Sciendo, pp. 125–139.
- [13] Feiler PH, Gluch DP and Hudak JJ. The architecture analysis & design language (aadl): An introduction. Technical report, Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst, 2006.
- [14] de Saqui-Sannes P, Hugues J et al. Combining sysml and aadl for the design, validation and implementation of critical systems. *ERTS 2012* 2012; .
- [15] Kordon F, Hugues J, Canals A et al. *Embedded systems: analysis and modeling with SysML, UML and AADL*. John Wiley & Sons, 2013.
- [16] Rødseth ØJ, Kvamstad B, Porathe T et al. Communication architecture for an unmanned merchant ship. In *OCEANS-Bergen, 2013 MTS/IEEE*. IEEE, pp. 1–9.
- [17] Zolich A, Palma D, Kansanen K et al. Survey on communication and networks for autonomous marine systems. *Journal of Intelligent & Robotic Systems* 2019; 95(3-4): 789–813.
- [18] Höyhty M, Huusko J, Kiviranta M et al. Connectivity for autonomous ships: Architecture, use cases, and research challenges. In *2017 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, pp. 345–350.
- [19] Höyhty M. Connectivity manager: Ensuring robust connections for autonomous ships. In *2019 2nd International Conference on Intelligent Autonomous Systems (ICoIAS)*. IEEE, pp. 86–90.
- [20] Jo SW and Shim WS. Lte-maritime: High-speed maritime wireless communication based on lte technology. *IEEE Access* 2019; 7: 53172–53181.
- [21] Porathe T, Burmeister HC and Rødseth ØJ. Maritime unmanned navigation through intelligence in networks: The munin project. In *12th International Conference on Computer*

- and IT Applications in the Maritime Industries, COMPIT'13, Cortona, 15-17 April 2013. pp. 177–183.
- [22] Mir ZH and Filali F. Lte and ieee 802.11 p for vehicular networking: a performance evaluation. *EURASIP Journal on Wireless Communications and Networking* 2014; 2014(1): 89.
- [23] Stelzer R and Jafarmadar K. Communication architecture for autonomous sailboats. In *Proceedings of International Robotic Sailing Conference*. pp. 31–36.
- [24] Papić S. Achieving optimal redundancy in a small business network. *International Journal of Digital Technology & Economy* 2016; 1(1): 13–23.
- [25] Team SA et al. An extensible open source aadl tool environment (osate). *Software Engineering Institute* 2006; .
- [26] Boehm B and Lane J. DoD Systems Engineering and Management Implications for Evolutionary Acquisition of Major Defense Systems. Technical report, USC-CSSE-2010-500, SERC RT-5 report, March 2010.
- [27] Forsberg K, Turner R and Adcock R. in *SEBoK Editorial Board. 2020. The Guide to the Systems Engineering Body of Knowledge (SEBoK)*, v. 2.2 R.J. Cloutier (Editor in Chief). Hoboken, NJ: The Trustees of the Stevens Institute of Technology. www.sebokwiki.org. BKCASE is managed and maintained by the Stevens Institute of Technology Systems Engineering Research Center, the International Council on Systems Engineering, and the Institute of Electrical and Electronics Engineers Computer Society., chapter Generic Life Cycle Model. March 2010.
- [28] *Systems and Software Engineering - System Life Cycle Processes*. Geneva, Switzerland: International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC)/Institute of Electrical and Electronics Engineers. ISO/IEC 15288:2015. .
- [29] *Systems and Software Engineering, Part 1: Guide for Life Cycle Management*. Geneva, Switzerland: International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC), ISO/IEC 24748-1:2010.
- [30] Kim I and Modarres M. Application of goal tree-success tree model as the knowledge-base of operator advisory systems. *Nuclear Engineering and Design* 1987; 104(1): 67–81.
- [31] Faisandier A, Madachy R and Adcock R. in *SEBoK Editorial Board. 2020. The Guide to the Systems Engineering Body of Knowledge (SEBoK)*, v. 2.2 R.J. Cloutier (Editor in Chief). Hoboken, NJ: The Trustees of the Stevens Institute of Technology. www.sebokwiki.org. BKCASE is managed and maintained by the Stevens Institute of Technology Systems Engineering Research Center, the International Council on Systems Engineering, and the Institute of Electrical and Electronics Engineers Computer Society., chapter System Analysis. May 2020.
- [32] Google earth. <https://earth.google.com/>.
- [33] Munin scenarios. <http://www.mits-forum.org/munin/index.htm>.
- [34] DNV GL. Dnvg1-cg-0264: Autonomous and remotely operated ships 2018; .
- [35] Large D and Farmer J. *Broadband cable access networks: the HFC plant*. Morgan Kaufmann, 2008.
- [36] Pueblas M, Gyurinda S, Strik J et al. Small enterprise design profile reference guide. *CISCO, Capitulo* 2010; 5.
- [37] Maritime broadband radio - mbr. <https://www.kongsberg.com/maritime/products/bridge-systems-and-control-centres/broadband-radios/maritime-broadband-radio>.
- [38] Hegering HG, Abeck S and Neumair B. *Integrated management of networked systems: concepts, architectures and their operational application*. Morgan Kaufmann, 1999.
- [39] Steinbaeck J, Steger C, Holweg G et al. Next generation radar sensors in automotive sensor fusion systems. In *2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, pp. 1–6.
- [40] Gürses E and Akan ÖB. Multimedia communication in wireless sensor networks. In *Annales des Télécommunications*, volume 60. Springer, pp. 872–900.
- [41] Albrektsen SM, Bryne TH and Johansen TA. Robust and secure uav navigation using gnss, phased-array radio system and inertial sensor fusion. In *2018 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, pp. 1338–1345.
- [42] Løvø E. Marine technologies’ control networks analysis and optimization for failsafe operations examined for class certification ; .
- [43] platform Consortium MCC MC. Maritime connectivity platform. <https://maritimeconnectivity.net/>.
- [44] DNV G. Veracity—an open industry platform, 2017.
- [45] Neumann JC. *The book of GNS3: build virtual network labs using Cisco, Juniper, and more*. No Starch Press, 2015.
- [46] Hunt C. *TCP/IP network administration*, volume 2. ”O’Reilly Media, Inc.”, 2002.

Appendices

A Appendix A: Navigation and communication Functions

A comprehensive list of expected functions and sub-functions needed to achieve the goals of the Autonomous Passenger Ship (APS), together with the source or requirement proposing the functions as well as the proposed architectural component to realize them, is depicted in Table 2.

B Appendix B: Verification of Requirements

A detailed analysis of the communication requirements verification has been conducted and is depicted in Table 3 to demonstrate the architecture satisfaction of the communicated requirements. The table describes the addressed requirements, the required verification criteria, the relevant architectural components, efforts made to verify as well as future efforts for evaluation.

Table 2. Functions for safe and secure navigation and secure and reliable communication

Goal Function	Functions	Sub-functions	Proposed/ Requirement	Architectural Components	Goal Function	Functions	Sub-functions	Requirement	Architectural Components	
Safe and Secure Navigation	Autonomous Navigation	Initiate MRC plan		ANS	Secure and Reliable Communication	Main Functions	Establish communication links	C-3,*	Gateways - Core/distribution tiers	
		Receive navigation data					C-N5, C-G-4	CM-NSC		
		Receive requests					C-3,*	CM-ONS		
		Upload new route					C-X-4	CM-ONS		
		Send reports					C-X-4, S-6	CM-ONS		
	Remote Navigation	Receive navigation data	Literature and Stakeholders goals		- ANS - Traffic Module - Navigation System	Secure and Reliable Communication	Main Functions	Establish inter-VLAN routing table	C-X-4, S-6	CM-ONS
		Receive requests						C-X-4, S-6	CM-ONS	
		Send reports						C-N2	CM-SSM	
		Receive traffic messages						C-X-7, C-G-4	CM-NSC	
		Send traffic messages						C-N4	CM-NSC	
	Remote Engine Monitoring and Control	Initiate control parameters		AEMC	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Enforcing of security policies	C-G-4	CM-NSC
		Receive requests						C-N4	CM-NSC	
		Send requests						C-N4	CM-NSC	
		Send Engine Data						C-N4	CM-NSC	
		Receive MRC plan						C-N4	CM-NSC	
Remote Engine Monitoring and Control	Receive navigation data		AEMC	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Collecting devices configuration backup	C-N4	CM-NSC	
	Receive requests						C-N4	CM-NSC		
	Send reports						C-N4	CM-NSC		
	Receive new route						C-N4	CM-NSC		
	Send requests						C-N4	CM-NSC		
Emergency Remote Navigation	Send control parameters		AEMC	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Storing devices configuration backup	C-N4	CM-NSC	
	Receive control parameters						C-N4	CM-NSC		
	Receive requests						C-N4	CM-NSC		
	Send requests						C-N4	CM-NSC		
	Send Engine Data						C-N4	CM-NSC		
Emergency Remote Control	Receive MRC plan		AEMC	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Send device configuration backup	C-N4	CM-NSC	
	Receive navigation data						C-N4	CM-NSC		
	Receive requests						C-N4	CM-NSC		
	Receive new route						C-N4	CM-NSC		
	Send reports						C-N4	CM-NSC		
Passenger Safety Security Network Power Alarm Emergency Response Control	Receive traffic messages		AEMC	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Calculate network logs	C-N5	CM-NSC	
	Send traffic messages						C-N5	CM-NSC		
	Send navigation data						C-N5	CM-NSC		
	Send control parameters						C-N5	CM-NSC		
	Receive control parameters						C-N5	CM-NSC		
Supporting Functions	Initiate emergency signal		C-X-2	- ANS - Navigation System	Secure and Reliable Communication	Main Functions	Calculate performance indicators	C-X-7	CM-NSC	
	Receive emergency signal						C-X-2	CM-NSC		
	Receive traffic messages						C-X-7	CM-NSC		
	Send traffic messages						C-X-7	CM-NSC		
	Send Engine Data						C-X-6	CM-NSC		
Out-of-Scope	Emergency Module		Emergency Module	Emergency Module	Emergency Module	Emergency Module	Manage conflicting QoS rules	C-X-6	CM-NSC	
	Traffic Module						C-X-2	CM-NSC		
	Emergency Module						C-X-2	CM-NSC		
	Traffic Module						C-X-2	CM-NSC		
	Emergency Module						C-X-2	CM-NSC		

Table 3. Detailed Requirement Analysis

Code as in (10)	Requirement Description	Priority*	Related to	Verification Criteria		Architecture Components	Verified by Scenario	Simulation	Evaluation Testbed	Verification and Evaluation method	Plan forward
				Design-level	Implementation-level						
C-X-1	Link bandwidth and latency	S	Quantitative Property	Minimum latency and maximum bandwidth are estimated	A minimum bandwidth of 9Mbps and minimum latency of 0.05 seconds for the ship-to-ship link and 2.5 seconds for the video link	Gateways	-	-	Bandwidth and latency tests	Speed and latency tests will be carried in the testbed across the different implemented gateway links to determine the satisfaction of the requirement.	
C-X-2	Emergency push button	S	Capability	A dedicated component that provides the capability exists	A permanent and reliable link is implemented for the EPB	- EPB - Emergency Module - Emergency Alarm and Response	Y 6.1.5	-	Reliability tests	Several implementation choices are being discussed and efforts to integrate this capability to the testbed will be made	
C-X-3	Fault tolerant RCC link	S	Qualitative Property	Fault-tolerance property is designed for the RCC link	Link still operates at full capacity in case of a failure in a single component.	- MC Module - APS-RCC Module - Core/distribution tier - ASC - RSC	Y 6.1.4	Y	Network performance tests	Several performance tests will be carried in the testbed to evaluate the ability of the RCC link to tolerate different types of failures	
C-X-4	Traffic prioritization	S	Capability	Traffic prioritization capability exists in the proposed component for the RCC link	Traffic prioritization according to a defined policy is implemented	CM-QoS	Y 6.1.4	Y	Network configuration test	After the implementation of the capabilities, several tests will be carried in the testbed to evaluate their effectiveness	
C-X-5	Operator triggered traffic redirection	S	Capability	Traffic balancing capability exists in the proposed component for the RCC link	The operator is able to seamlessly switch and distribute traffic between different communication channels without negative effect		-	Y			
C-X-6	Quality of Service	S	Capability	The capability to provide QoS exists in the proposed components for all external links	The link operates according to a defined QoS metrics		Y 6.1.4	Y			
C-X-7	Link quality monitoring and notification	S	Capability	The design integrates monitoring and notification systems for information quality analysis	Real-time or near real-time quality analysis is implemented based on QoS requirements	CM-NMT	Y 6.1.4, 6.1.3	Y			
C-X-8	Network Troubleshooting	S	Capability	Independent troubleshooting capabilities exist for each link	Troubleshooting can be performed by the operator across all the links.		-	Y			
C-X-9	Redundant RCC Link	S	Qualitative Property	Link with RCC is designed to be redundant having at least main and backup	The link with RCC is maintained even with the loss of one of the links	- MC Module - APS-RCC Module	Y 6.1.4	Y	Network performance test	Several implementation choices will be tested using the testbed to propose the most suitable technologies for the main and backup RCC links	
C-N-1	Standard aligned design	S	Qualitative Property	The design achieves compliance with applicable requirements in relevant standards	N/A	- On board topology - CM-NSM	-	-	Compliance checking	N/A	
C-N-2	Segregated network design	S	Qualitative Property	Segregated network design exists	Failure doesn't propagate across segregated networks		-	Y	Network performance test	The feasibility of the segmentation policy after the implementation of the system components will be evaluated	
C-N-3	Redundant network design	S	Qualitative Property	Redundant network design exists	Automatic transition/activation/restoration between main and backup systems is implemented		Y 6.1.3	Y		The efficiency of the transition/activation/restoration process will be evaluated for each system and its backup component	
C-N-4	Event triggered traffic redirection	S	Capability	Traffic redirection capability exists in the internal network	Traffic redirected upon loss of remote connectivity is achieved for the different APS components in the internal network	CM-QoS	Y 6.1.3	Y	Network configuration test	The effectiveness of the implemented capability will be tested in the testbed	
C-N-5	Network capacity	S	Capability	Internal Network has the ability to connect several systems	Connectivity is achieved for the different APS components in the internal network	On board topology	-	Y	Existence	The ability of the implemented network shall be evaluated	
C-N-6	Frequency coordination plan	S	Action	N/A	Frequency coordination plan is made, documented, and tested	Implementation-level requirement	-	-	Wireless signal interference testing	After the implementation of several wireless technologies, a frequency coordination plan will be constructed	
C-O-1	LoS communication	S	Capability	Ship-to-ship capability exists in the APS	Ship-to-ship communication achieved through LoS communication system (AIS or DVHF) for a range of at least 2km	Traffic module	Y 6.1.1	-		Several tests will be conducted to evaluate the effectiveness of the AIS technology for implementation in the traffic module. Special focus will be given to its security.	
C-G-1	Data recording	S	Capability	Recording and logging capabilities of important data is designed in the APS	Recording and logging of important data is implemented	- Digital Logbook - Online Storage	-	Y	Existence	The recording capability will be implemented and evaluated in the testbed.	
C-G-2	Type approved components	S	Qualitative Property	N/A	All network components and equipment are type-approved	Implementation-level requirement	-	Y	Compliance checking	Efforts will be made to ensure that all tested equipment are type-approved.	
C-G-3	Transmission protocol	S	Qualitative Property	N/A	The transmission protocol for each link is compliant with a relevant international standard		-	Y	Existence	The transmission protocols across all links shall be relevant to an international standard.	
C-G-4	Wireless data communication	S	Qualitative Property	N/A	Wireless data communication across the links employ an internationally recognized system with pre-specified features		-	Y	Performance and security testing	Several performance and security tests will be carried to verify that the implemented wireless communication technology includes the required features.	
C-G-5	Coverage analysis	M	Action	N/A	The effectiveness of the wireless communication systems is determined		-	-	Coverage Analysis	Several tests will be conducted using the testbed to evaluate the coverage of the links in different geographical areas and weather conditions	
C-G-6	Network documentation	S	Action	N/A	Documentation of the implemented protocols and interfaces is performed		-	-	Inventory of System-wide protocols and interfaces	After the implementation of all protocols and interfaces, such documentation will be made available.	

* MoSCoW rule (S: Should, M: Must)

Paper III

A. Amro, G. Kavallieratos, K. Louzis and C. A. Thieme, 'Impact of cyber risk on the safety of the milliamperes2 autonomous passenger ship,' in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 929, 2020, p. 012 018

PAPER • OPEN ACCESS

Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship

To cite this article: Ahmed Amro *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **929** 012018

View the [article online](#) for updates and enhancements.

You may also like

- [Design and Development of Ticket Reservation Information System in Travel Business](#)
E S Soegoto and R Fadillah
- [Assessing Bus Performance Rating in Kajang, Selangor](#)
S Norhisham, A Ismail, L M Sidek et al.
- [Causative Factor Analysis of Passenger Ship Accident \(Fire/Explosion\) in Indonesia](#)
W Mutmainnah, L P Bowo, A Nurwahyudy et al.

Impact of cyber risk on the safety of the MilliAmpere2 Autonomous Passenger Ship

Ahmed Amro¹, Georgios Kavallieratos¹, Konstantinos Louzis² and Christoph A. Thieme³

¹Department of Information Security and Communication Technology (IIK), University of Science and Technology (NTNU), Gjøvik, Norway

²Laboratory for Maritime Transport, School of Naval Architecture and Marine Engineering, National Technical University of Athens, Greece

³Department of Marine Technology (IMT), NTNU, Trondheim, Norway

E-mail: ahmed.amro@ntnu.no, georgios.kavallieratos@ntnu.no, klouzis@mail.ntua.gr, christoph.thieme@ntnu.no

Abstract. The digitalization of the maritime sector is continuously growing, leading to increased automation, such as, the development of autonomous vessels. The Autonomous Passenger Ship (APS) is a characteristic instantiation of this development, aiming to transport people on urban waterways. Although emerging technologies deployed in such APS aim to facilitate the functions and operations of the navigation and communication systems, various safety and security risks are inherent to the communication infrastructure due to their interconnectivity. The aim of this work is to study the safety and cyber security of the communication system of an APS, namely the MilliAmpere2 APS. The six step model (SSM) is utilized to facilitate the joint analysis. The application of the SSM enables, among others, the capturing of relationships between cyber attacks and component failures, the assessment of safety and cyber security countermeasures, as well as, the synergies between them. It has been found that most countermeasures in both categories are reinforcing or are conditionally dependent on each other, while few antagonize each another. These findings will allow for improved design and implementation of integrated safety and security management solutions.

1. Introduction

The emergence of the contemporary and interconnected Cyber Physical Systems (CPSs) in the maritime domain and particularly in the autonomous vessels infrastructure, such as, the Autonomous Passenger Ship (APS), is rapid and continuous. To this end, the safety and security analysis of such systems is needed to ensure the vessel's normal and safe operations. Safety and security are interrelated concepts that may face both commonalities and differences in the analysis process since the former is concerned with accidental events while the latter mainly consider malicious actions taken by adversaries [1]. Particularly, security is concerned with the risks originating from the environment impacting the system and typically addresses malicious risks. Whereas safety deals with risks arising from the system that may affect the environment and addresses purely accidental risks. Safety analysis aims to reduce the risks related to systems, humans, and the environment [2] to an acceptable level.

Security analysis aims to minimize the risk related to confidentiality, integrity, and availability of the operational and functional requirements and therefore the data, information, and services of the system [2]. An extended security analysis may also consider the properties of possession or control, authenticity,



utility, and non-repudiation. Hazards can be defined as conditions or states that may cause harm [3]. The risk associated with the hazards is a measure of uncertainty with respect to outcomes and may be described through risk sources, events and their consequences [4]. From a security point of view, vulnerabilities are system or software flaws that could threaten the system. Vulnerabilities could be also considered as system weaknesses [3].

The aim of this work is to identify weaknesses related to safety and security of a communication architecture proposed for safe and secure navigation for an APS [5] in order to remove them or reduce the risk associated with them. We apply the Six Step Model (SSM) to analyse security and safety risks and study the implications that security poses to safety. Particularly, leveraging the multidimensional matrices provided in the SSM, the functions, structure, failures, safety countermeasures, cyber attacks, and security countermeasures are identified for the communication, navigation and control systems of the APS. Although various approaches exist in the literature for security and safety co-analysis, the SSM has been chosen as the most appropriate for the case of the APS, due to its holistic approach to assess interdependencies. The complexity of the communication systems and the novel technology used in the communication infrastructure can be appropriately studied by the graphical models of the SSM. Further, the SSM facilitates the collaboration of both safety and security experts towards a more comprehensive safety and security analysis. The SSM and its application to the MilliAmper2 is described in detail in Section 4. The methods being employed in this article and previous works have been carried out as initial steps of a risk management process that is part of the Autoferry project [6]. The risk management process is aligned with the guidelines proposed by the International Maritime Organization (IMO) regarding the inclusion of cyber risk management within the Safety Management Systems (SMS) [7].

2. Related Work

Many safety and security methods that have been developed, do not directly consider each other. However, the combination of security and safety analysis is expected to result in identifying synergies regarding interactions, events and conflicting countermeasures. Various works in the literature examined the interrelation between safety and cyber security [2, 8]. Particularly, Lisova et al. [8] conducted a systematic literature review for safety and security co-analysis and thirty- three approaches have been identified. Further, Kavallieratos et al. in [2] conducted a survey in co-engineering approaches for safety and cyber security in cyber physical systems. Various systematic approaches have been proposed in the literature to analyse safety and security. The System-Theoretic Process Analysis (STPA), is developed to facilitate the safety analysis of complex systems considering the control structure of the targeted systems. The STPA is extended to accommodate security considerations, called STPA-Sec [9]. Further, SafeSec Tropos [10] is a co-engineering methodology for safety and security requirements elicitation in CPSs based on STPA and the Secure Tropos methods from safety and security domain respectively.

Safety related cyber attacks for autonomous inland ships are identified and assessed by Bolbot et al. [11]. In particular, by leveraging from a Cyber preliminary hazard analysis (PHA) method and existing systems vulnerabilities, potential cyber attacks that may compromise the vessel's safety along with a set of general countermeasures are examined. Further, Bačkalov [12] studied the safety of autonomous inland vessels. Namely, the key characteristics of the autonomous inland vessels are analyzed considering the corresponding legislation and standards related to the safety of sea-going ships and inland vessels. Kavallieratos et al. [10] analyzed the safety and security of a cyber-enabled ship that could be either autonomous or remotely controlled. By the application of the SafeSec Tropos method, they identified the necessary security and safety requirements for such vessels.

3. Background

This section summarizes the background on the MilliAmpere2 passenger ferry, an instantiation of an APS which is under development as part of the Autoferry project [6].

3.1 System description and context of operation

The MilliAmpere2 is designed to carry up to 12 passengers over the harbor channel in Trondheim, Norway. The APS is characterized by a high degree of autonomy where the navigational and operational requirements are fulfilled by the APS. A supervisor in a land-based control centre (during the first year located on site) is responsible for actions needed in case of emergencies. Autonomous functions include navigation, docking, passenger registration, charging. Therefore, the communication of navigational and status data to the land-based control centre is vital [6].

3.2 Communication architecture

The communication architecture of the APS enables the communication with the environment through a heterogeneous group of different technologies. There are six main communication gateways in the APS. Particularly, two IP based gateways aim to establish ship-to-shore communication links with the RCC by leveraging several implementation solutions such as mobile communication (4G/LTE/5G) and Wireless Local Area Networks (WLAN) technologies. The third gateway is intended for ship-to-ship communication to enable the vessels in the area to communicate with the APS. Automatic Identification System (AIS) is proposed for the implementation of this module. The fourth gateway is intended to carry emergency communications for the control and navigation of the APS in case of lost communication with the Remote Control Center (RCC) while the fifth and sixth gateways are utilized to receive signals for real time kinematic (RTK) and global navigation satellite system (GNSS).

The internal network architecture of the APS is designed to include redundant communication paths, segregated sub-network, and secure communication. A centralized monitoring and controlling group of servers called the Autonomous Ship Controller (ASC) is interfaced with two network traffic Core and Distribution tiers (C/D), each consisting of main and backup switches with IP routing capabilities. The former tier (C/D A) connects the external gateways with the servers in the ASC, while the latter (C/D B) connects the secondary (i.e., backup) servers in the ASC with the internally segregated sub-networks. Moreover, a centralized component named the connectivity manager is responsible for performing network management functions by configuring and monitoring the network devices in addition to additional functions related to security. The detailed communication architecture along with the corresponding functions are described by Amro et al. in [5].

3.3 Navigation and Machinery Systems

The navigation system is comprised of components able to collect environmental data, establish situational awareness based on the sensing data, and determine safe navigation routes. The navigation system components are arrays of sensors of different types (EO cameras, video cameras, Lidars, and Radars), in addition to RTK GNSS. All these components send their data through the ship internal network to the Autonomous Navigation System (ANS) that hosts the logic to perform autonomous navigation functions, as well, as support remote navigation by the RCC. Moreover, a machinery system implements maneuvers according to the determined route from the ANS. The machinery system consists of a Dynamic Positioning (DP) system, and thrusters, interfaced through Input/Output (I/O) cards. The machinery system is controlled by an Autonomous Engine Monitoring and Control (AEMC) system which host the logic to monitor engine data and determine the appropriate control parameters. Both the ANS and AEMC are hosted in the ASC servers zone.

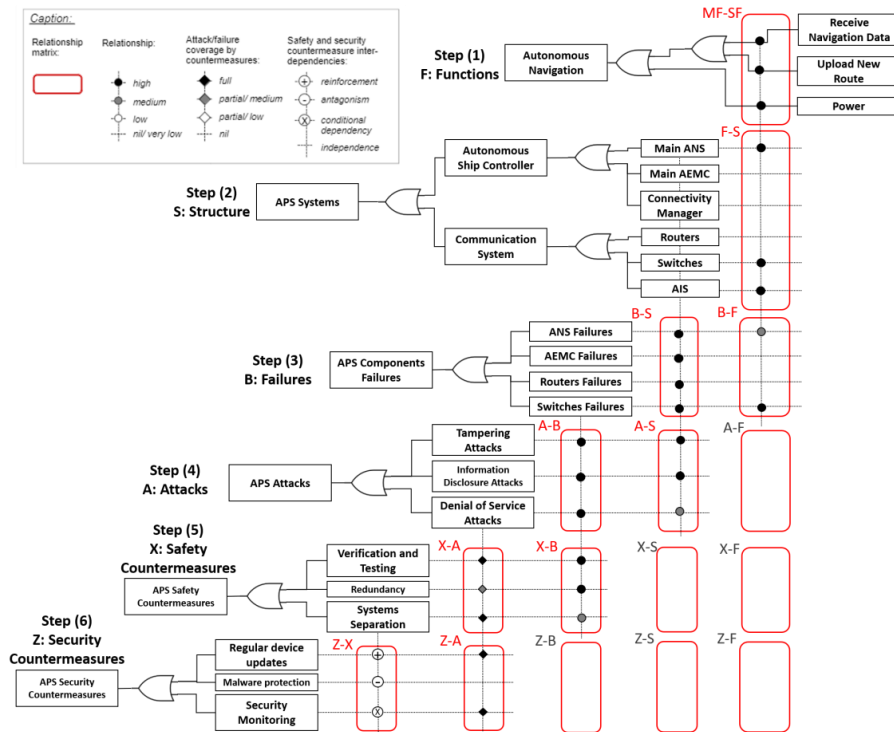


Figure 1. Overview of the applied Six Step Model steps (Adapted from [15])

3.4 Preliminary Hazard Analysis

A PHA for the MilliAmpere2 is presented in [13]. A PHA is a structured hazard identification method, which is guided through keywords that represent possible hazards and their sources [3]. The PHA for the APS was conducted in two sessions with the participation of several experts from several relevant domains. This analysis builds the foundation for identification of safety related issues in this article.

4. The six-step model

This Section summarizes the adopted six-step model (SSM) for the safety and security analysis of the communication, navigation and control systems of the APS. The SSM was proposed by Sabaliauskaite et al. [14] to analyze, both, safety and security aspects of CPS. Furthermore, the applicability of the SSM to and security issues of an autonomous vehicle is examined in [15]. An overview of the SSM that is followed in this work, is depicted in Figure 1. The six steps of the SSM model were performed disregarding some matrices, since their analysis was not considered relevant for the scope of the target analysis.

4.1 Step 1: Functions

The first step of the SSM describes the system's functions. Main and supporting functions are differentiated and their relationships are established. The main functions are cross referenced with the supporting, secondary functions of the system in the MF-SF matrix (Main-functions, Supporting Functions). The matrix provides the basis for further system analysis. Four types of relationships are distinguished, high, medium, low, very low/ none. The relationships description is adopted from [16]. High relationships characterize high dependency of the main function on the supporting function for

proper operation. Medium means that the main functions might be dependent on the supporting functions to operate properly in some operational modes while low relationship means that the main functions are rarely dependent on the supporting functions. Finally, nil/ very low relationships mean that no dependencies on the supporting functions are identified.

4.2 Step 2: Structure

The second step identifies the relationships between the APS components and the APS functions in the S-F matrix (structure: functions). The APS components are identified by decomposing the systems to the appropriate level of analysis. This includes main and supporting systems. This analysis is necessary to determine the relationships between function failures and physical component failures in step 3 (Section 4.3). The rating scheme includes high, medium, low, and nil/ very low levels. A high relationship indicates that the component is highly important for the realization of the function under analysis. A medium relationship indicates that the component might be needed to realize the function in certain operational modes. A low relationship indicates that the component might be needed to realize the function in very specific and rare cases. Nil/ very low is assigned to pairs that have no relationship with each other.

The components are prioritized for threat modeling in Step 4 (Section 4.4) according to their highest effect on the main system functions, considering the relationships studied in Step 2. The scores of the components under analysis are calculated for all the APS's components, taking into account the assigned relationships of each component with the specified system functions following Equation 1. The number of high relationships with systems functions is denoted as "h". Further, the number of medium and low relationships are denoted as "m" and "l" respectively. The components that gathered scores above the average are considered for analysis.

$$Score = 5h + 3m + l \tag{1}$$

4.3 Step 3: Failures

The aim of the third step is twofold; firstly the system failures of main components are identified and secondly the failures' impact on the system's functions are determined. In this step, the B-S matrix (failure: structure) is created. In this matrix, component failures were identified and assessed from the available PHA report [13]. The failures were generalized and the results from the report were used as input to rate the dependencies. The failures are assessed in relation to the system's functions in order to assess the severity of failures on the system's function execution. The information is recorded in the B-F matrix (failures: functions). The rating scheme used for the B-S and B-F matrices included high, medium, low, and nil/very low levels. These indicate the strength of the impact of a failure to either the operation of a component or the implementation of a system function. For instance, a high relationship between a failure and a component means that the latter will most probably not be able to operate, whereas the same level between a failure and a function means that the latter will be potentially severely impaired.

By leveraging the PHA [13] the Function Failure Impact Factor (FFIF) is determined. The FFIF represents the expected impact on the execution of a function that is weighted by potential consequences of the failures. The loss of each function was associated with potential consequence categories that included the following in ascending order of severity: loss of operational/ performance data, loss of remote monitoring and control, and loss of control/ drifting/ grounding/ collision/ injuries/ fatalities. Having calculated the failures' relationships with the system functions (Relationship_{i,j} (Failure_i, Function_j)), the overall impact score for each component failure (i) is calculated using equation 2 where N represents the total number of system functions. Failures with Impact score above the average have been forwarded for analysis in Step 4 (Section 4.4).

$$Impact_i = \sum_{j=1}^N Relationship_{i,j} * FFIF_j \tag{2}$$

4.4 Step 4: Attacks

The assessment of cyber attacks is performed in Step 4 by utilizing the STRIDE method. STRIDE stands for Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service and Elevation of Privilege [17]. The method enables the analysis of complex systems and environments similar to the APS [18, 19].

Potential security threats along with the corresponding attack scenarios are identified and analyzed for the APS's components considering STRIDE. The security analysis takes into account external and internal attackers. The former are able to conduct the attack remotely while the latter perform the attack by infecting system components. Further, malicious passengers are considered as potential adversaries that may attack the APS. The following relationship matrices are generated; matrix A-B (attacks: failures), and matrix A-S (attacks: structure).

The A-S matrix is analyzed through the application of STRIDE. The attacks considered in matrix A-B are a subset of the attacks that are analyzed in the A-S matrix. Using the analysis performed in step 2 (Section 4.2) only attacks against components with most effect on the system's functions have been considered. Therefore, the attacks that are characterized by the highest impact on the main system functions are identified. Further, the A-B matrix includes attack scenarios that violate the security objectives of authenticity, integrity, non-repudiation, confidentiality, availability, and authorization. These objectives aim to ensure the security of the components in the maritime environment [10] and therefore of the APS's infrastructure. The most critical security objectives related to the functions of the APS are integrity, confidentiality, and availability. These ensure the security and reliability of the communication systems.

The A-B matrix reflects the relationships between the prioritized attacks and prioritized failures. The relationships are categorized as high, medium, low, or nil/ very low. A high relationship means that the attack is expected to directly lead to the failure with high possibility. A medium relationship means that the attack triggers the failure with moderate possibility. A low relationship means that the attack may lead to the failure with low probability, while a nil/ very low relationship is considered if no connection between between the attack/ failure pair can be identified.

4.5 Step 5: Safety Countermeasures

Safety countermeasures are identified considering the failures and the attacks in the fifth step. The reasoning for identifying the potential safety countermeasures is based on a high-level consideration for the system design and development process, and strategies to mitigate potential risks. The safety countermeasures include measures that need to be designed into the system (e.g., integrity checks, error handling), measures during commissioning (i.e., testing and verification), and operational measures (e.g., maintenance policies for hardware and software components, minimum risk condition). The matrices assessed in this step are matrix X-B (safety countermeasures: failures), and matrix X-A (Safety countermeasures: Attacks). The X-A matrix enables the identification of synergies between safety and security issues. For X-B and X-A the assessment considered four distinct degrees; full, partial medium, partial low, and nil. A full degree removes the failure or attack and its associated consequences to a large degree or completely. A safety measure assessed as partial medium eliminates the consequences or reduce the consequences to a large degree. A partial low assessment implies conversely a minor reduction in the frequency of occurrence or a minor reduction in the expected consequences. Nil degree describes that no improvement from this measure is expected, or that it has been already implemented.

4.6 Step 6: Security Countermeasures

In this step, the relationship matrices being analyzed are matrix Z-X (Security countermeasures: safety countermeasures) and matrix Z-A (security countermeasures: attacks). The remaining matrices, Z-B (security countermeasures: failures), Z-S (security countermeasures: structure), and Z-F (security countermeasures: functions) were not analyzed. The identification of security countermeasures is needed for the analysis in this step. The considered countermeasures are based on previously established cyber security requirements for the communication architecture [20].

The Z-A matrix described the effectiveness of a countermeasure in mitigating cyber attacks. To this end, three relationship categories; 1) the countermeasure leads to fully mitigating the attack (f), 2) partial mitigation(p), 3) nil (no mitigation or not relevant). By applying equation 3 a score is calculated for each countermeasure to indicate the effectiveness against the attacks based on the analysed relationship.

$$Score = 5f + p \quad (3)$$

The Z-X matrix captures the dependencies between safety and security countermeasures, which is crucial to the study of the synergies between these different sets of countermeasures. In particular, the effects that the security countermeasures may have on the safety countermeasures are represented through four types of relationship as defined in [14]. These are: 1) Reinforcement, 2) Antagonism, 3) Conditional dependency, and 4) Independent.

5. Results

5.1 Step 1: Functional analysis

The proposed communication architecture enables the MilliAmpere2 to perform several functions related to navigation, control, communication, and safety. Figure 2 shows an overview of the functions supported by the communication architecture and reflects their relations with the APS components previously discussed in Section 3 as well as among themselves. These relations highly influenced the analysis in Step 1 and 2 in the SSM model. The main navigation functions provide the situational awareness of the APS for the determination of safe routes. The "Engine Monitoring and Control" functions describe the monitoring and control of the APS's thrusters. Furthermore, autonomous functions are performed by the APS. Remote functions are carried out by the RCC, and emergency functions are executed by the Emergency Control Team (ECT). Further functions are needed to initiate emergency signals by passengers referred to as "Passenger Safety" functions. They are needed to indicate the occurrence of safety-critical events (e.g., passenger falling overboard). Therefore, these functions will only be found in APS or manned autonomous ships, and not in unmanned autonomous ships, since they will not be required. For the purpose of this paper, only emergency functions with respect to passenger communication with the RCC and the emergency services are considered, due to the focus on the communication system. The main communication functions are categorized considering their individual role. The "Ship-to-shore communication" function provides the required connectivity between the ship and the RCC. The "Internal communication" functions provide the needed connectivity between the different components onboard the ship. The "Emergency communication" functions provide the needed connectivity with the ECT. These functions depend on several supporting functions such as power, security, and network system management (NSM).

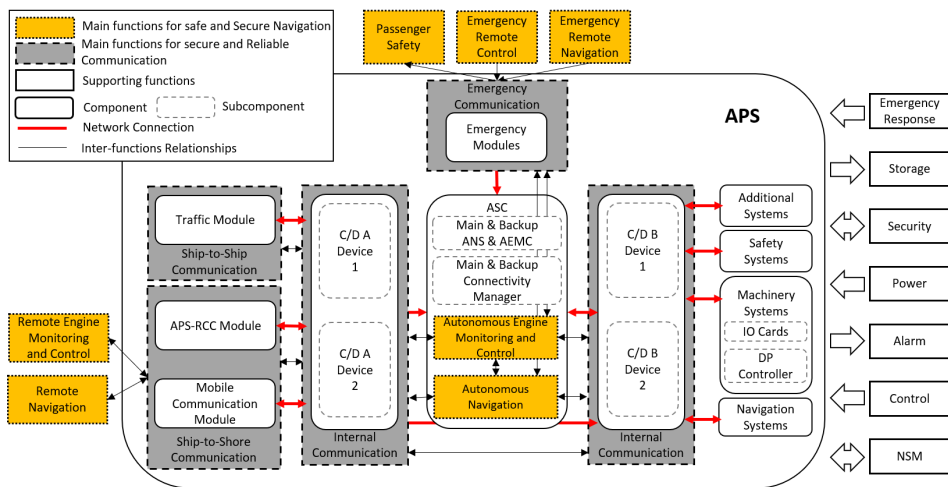


Figure 2. Overview of the relationships between the communication architecture functions and structure (Adapted from [5])

5.2 Step 2: Component assessment

The aforementioned scoring scheme in Step 4.2 was used to identify the components with the highest influence on the system. The components and the number of identified high, medium and low relationships with system functions are shown in Table 1. It can be observed that the components related to communication related have a relatively higher effect over system functions which is logical since they are responsible for the information exchange needed for most functions.

Table 1. Relationships assessment of the components with the most effect on the main system functions

Component	Relationships		
	High	Medium	Low
C/D A Device 1 and 2	30	1	10
Mobile Communication Module	23	1	13
APS-RCC Module	23	1	13
Connectivity Manager	24	1	0
DP controller	9	19	1
C/D B Device 1 and 2	16	3	10
Backup Connectivity Manager	18	1	0
IO cards	10	14	0
Main ANS	15	4	1
Backup ANS	11	4	1
Emergency Module 1	11	1	10

5.3 Step 3: Assessment of failures

As an outcome of step 3, the failures having highest impact according to equation 2 were identified. The failures associated with the connectivity manager are found to have the highest impact score, with C/D part A devices next, and the DP system after them. We argue that the results are plausible since the failures with highest scores would indeed affect the entire APS. For instance, failures in the connectivity manager and network devices would lead to disruption in the information flow within the APS which would lead to total loss of certain functionalities. Additionally, the AEMC, DP, together with the I/O cards and the thrusters highly affect the APS, in case of a component failure. Some may even lead to a blackout affecting the whole APS.

5.4 Step 4: Assessment of attacks

Through the application of the STRIDE methodology several attack scenarios were identified as well as their relevance to the system components. Those attacks were analyzed against the failures discussed in Section 5.3. The results reflect which attacks could cause additional failures, as well as, which failures increase the vulnerability of the system to cyber attacks. It was found that attacks against both, the main and backup, connectivity manager components could cause failures with highest impact, followed by the C/D switches. Additionally, it was observed that denial of service attacks cause higher impact failures than others, followed by tampering, while information disclosure attacks have much lower effect on failures. Moreover, susceptibility to cyber attacks is mostly enhanced by failures in the connectivity manager, ANS, and AEMC components.

5.5 Step 5: Assessment of safety measures

Eleven safety countermeasures were identified. The countermeasures are summarized and described in Table 2. Safety countermeasures were identified from the PHA [13] and common risk mitigation strategies, such as laid out in [3]. For the purpose and scope of the article, the safety countermeasures are generic in nature and the specific implementation for the components needs to be defined and described in the further design process.

Table 2. Identified safety countermeasures

ID	Mitigation measure	Description
CSaf1	Choice of communication protocol	Selecting protocols and bus systems that are robust and suitable for the purpose of communication between the components.
CSaf2	Verification and testing	The component should be tested and its function and behavior verified during different phases of the development process.
CSaf3	Monitoring and trouble shooting through shore operator	The shore operator monitors the system behavior and engages in problem trouble shooting if a problem with the ferry occurs. The ferry design needs to accommodate these trouble shooting abilities.
CSaf4	Component redundancy	A second similar component is introduced in the system design to take over functionalities in case of the failure occurring.
CSaf5	Separate hardware components	The component has its own dedicated computing hardware to run on.
CSaf6	Go to a safe state	A safe state is defined for a failure and will thereby mitigate the consequences of this failure. The ferry needs to be designed such that the safe state can be reached in the failure condition.
CSaf7	Self and status tests of the component	The component must be able to test for correct operation and functioning.
CSaf8	Choice of computing hardware	Sufficient powerful computing hardware needs to be chosen to fulfil the components purpose even under high load conditions.
CSaf9	Cross validation of data inputs for sensor data	The components using sensor data is crosschecked with other data for plausibility. Implausible and invalid data should be rejected.
CSaf10	Hardware maintenance and cleaning policy	A maintenance plan defining preventive and corrective maintenance, including cleaning for the hardware components.
CSaf11	Software maintenance policy	A maintenance policy for the software describing the policy for bug fixing and updating software, and associated tasks.

The assessment of the effectiveness of each safety countermeasure is based on its assumed degree of elimination or mitigation of a failure. However, due to the generic nature of these, their impact cannot be definitively assessed since the implementation of the measure for each component was not specified in detail. For the same reason, it cannot be assumed that all failures are removed from the system. The safety countermeasures address all the failures, as well as, most of the attack scenarios. Only hardware

maintenance is not addressing any of the cyber attacks. Most failures and attacks can be addressed through testing and verification efforts (CSaf2) and Self and status tests (CSaf7). Monitoring through an operator (CSaf3) can assist to some degree in identifying safety and security related issues. However, adequate procedures need to be established in order to troubleshoot efficiently and react appropriately. Cross validation of data (CSaf9) and a hardware maintenance policy (CSaf10) is mainly relevant for the sensors and the actuators.

5.6 Step 6: Assessment of security measures

The analysis performed in the Z-A matrix is utilized to assess the countermeasures coverage of the identified attacks. By using equation 3, as depicted in Table 3 various countermeasures are found to mitigate either fully or partially several attacks, such as the application of secure network protocols, and the preparation of incident response plans. The analysis facilitates the prioritization of the countermeasures implementation. It can be observed that the implementation on secure network protocols such as Transport Layer Security (TLS) and Virtual Private Network (VPN) would be most effective for attack mitigation, followed by well planned incident response procedures, and security monitoring. Moreover, it was observed that cyber security training for operators is not of high priority which is logical due to the reduced human involvement in the direct operation of the APS. For the operation of the RCC, the cyber security training may still be of importance.

Table 3. Outcome of the assessment of cyber security countermeasures

ID	Countermeasures	Attacks mitigation		
		Fully	Partially	None
CSec1	Secure network protocols	14	2	2
CSec2	Incident response plans	7	10	1
CSec3	Security monitoring for detecting malicious and abnormal incidents	8	8	2
CSec4	User access management system	7	7	4
CSec5	Regular device updates	3	9	6
CSec6	Detailed map of IT and network equipment and software	3	8	7
CSec7	Cyber security management framework	2	9	7
CSec8	Regular software security analysis (penetration testing)	1	8	9
CSec9	Backup facilities	4	2	12
CSec10	Periodic inventory of user accounts and their associated privileges	2	5	11
CSec11	Malware protection	0	8	10
CSec12	Cyber security training	1	6	11

In this work the impact of security risk on safety is examined and analyzed. To this end, Table 4 depicts the relationship between the safety and security countermeasures. By leveraging the information depicted in Table 4 most of the security countermeasures are independent (60 relationships) while 40 relationships were assessed as enhancing the existing safety countermeasures. For instance, the cyber security management framework facilitates and strengthens the corresponding safety countermeasures. Additionally, twenty seven countermeasure relationships are characterized as conditionally dependent. Only five relationships between safety and security countermeasures are characterized as antagonism. Namely, the need to separate system components (CSaf5) and specify certain choices of hardware (CSaf8) for safety countermeasures may complicate the implementation of suitable security monitoring solutions (CSec3).

6. Discussion

The SSM analysis identifies the relationships between components and functions and provide a holistic view for the system under analysis. Therefore, the interplay of safety and security is examined in detail, considering the different viewpoints that are provided by the corresponding matrices. Overall the SSM provides an appropriate analysis of a system under development, in the initial steps of the design process where the functional and operational requirements are not defined in detail. Our analysis shows that the SSM provide results that may help to prioritize identified safety and security issues for more detailed analysis.

The analysis of an abstract system architecture is facilitated by leveraging the SSM. The method extracts rigorous and valid results in high organizational and operational levels. However, the SSM scales badly with increasing system complexity, due to the state-space growth in the SSM matrices. The SSM would benefit from additional guidance on how to represent and model the system and the dependencies among its components in a standardized way, similarly to how the STPA defines the safety control structure as a way to model the system. This would help in better determining, for example, common cause failures and assessing the impact of failures on the system structure and functionality. Additionally, further guidance on ranking the relationships described in each SSM matrix and steps for the failure and attack prioritization are needed. This may reduce the effect of subjective expert assessments that may not be justified in a transparent and reproducible manner.

Table 4. Relationships between safety and security countermeasures

Safety	Security											
	CSec7	CSec6	CSec10	CSec4	CSec5	CSec11	CSec1	CSec12	CSec8	CSec3	CSec2	CSec9
CSaf1	●	○	○	○	○	○	○	○	○	○	○	○
CSaf2	●	●	○	○	○	○	○	○	○	○	○	○
CSaf3	●	○	○	○	○	○	○	○	○	○	○	○
CSaf4	●	○	○	○	○	○	○	○	○	○	○	○
CSaf5	●	○	○	○	●	○	○	○	○	○	○	○
CSaf6	●	○	○	○	○	○	○	○	○	○	○	○
CSaf7	●	○	○	○	○	○	○	○	○	○	○	○
CSaf8	●	○	○	○	○	○	○	○	○	○	○	○
CSaf9	○	○	○	○	○	○	○	○	○	○	○	○
CSaf10	○	○	○	○	○	○	○	○	○	○	○	○
CSaf11	○	○	○	○	○	○	○	○	○	○	○	○

Legend: ● Reinforcement, ● Antagonism, ○ Conditional dependency, ○ Independent

Common cause failures, emerging system behavior, and multiple system failures are hard to include in the assessment. This may reduce the ability to identify interdependencies of failures. The identification of safety countermeasures is performed on a high level and detailed risk countermeasures could be developed early in the design phase. Applying the SSM process later, in the detailed design phases, the identified design changes and risk mitigation measures may come with a high cost. Guidelines on how safety and security measures may be identified with the SSM are also desirable.

7. Conclusion

In this work, a joint safety and security analysis of the MilliAmpere2 Autonomous Passenger Ship (APS) has been conducted. The Six-Step-Model (SSM) was applied to capture the different analysis viewpoints, namely, APS functions, structure, failures, cyber attacks, and safety and cyber security countermeasures. The main goal of the analysis was to infer the effect of cyber threats on safety, as well as, the interrelations between safety and security countermeasures, for design and implementation of integrated safety and security countermeasures.

Several conclusions can be drawn from the application of the SSM. It was found that the connectivity manager has most effect on the system functions. It could cause most failures, and is among the most susceptible to cyber threats. Secure network protocols, incident response plans and security monitoring were identified as the most important security countermeasures to be implemented. Moreover, safety and cyber security countermeasures have been found to be mostly compatible. Some measures are contradictory, which is very helpful to know during the design and implementation of both. Further work is needed to establish a security architecture for the APS that considers interrelations with safety. The outcome of this paper is expected to influence the design and implementation of security countermeasures to be adopted in the target security architecture as well as the undergoing design and implementation of the connectivity manager to mitigate its discovered threats.

Acknowledgements

Thieme acknowledges the support by the Norwegian Research Council through the UNLOCK project, project number 274441.

References

- [1] L. Pi`etre-Cambac`ed`es and C. Chaudet. The sema referential framework: Avoiding ambiguities in the terms “security” and “safety”. *Int. Journal of Critical Infrastructure Protection*, 3(2):55–66, 2010.
- [2] G. Kavallieratos, S. Katsikas, and V. Gkioulos. Cybersecurity and safety co-engineering of cyberphysical systems—a comprehensive survey. *Future Internet*, 12(4):65, 2020.
- [3] M. Rausand. *Risk Assessment: Theory, Methods, and Applications*. John Wiley & Sons, Hoboken, New Jersey, USA, 1st ed. edition, 2013.
- [4] T. Aven and O. Renn. On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12(1):1–11, 2009.
- [5] A. Amro, V. Gkioulos, and S. Katsikas. Communications architecture for autonomous passenger ship. Submitted for review to *Journal of Risk and Reliability (JRR)*.
- [6] NTNU Autoferry. Autoferry - Autonomous all-electric passenger ferries for urban water transport, 2018.
- [7] The Maritime Safety Committee. International maritime organization (imo) (2017) guidelines on maritime cyber risk management. <http://bit.ly/IMORiskManagement>.
- [8] E. Lisova, I. S`ljivo, and A. C`au`sevi`c. Safety and security co-analyses: A systematic literature review. *IEEE Systems Journal*, 13(3):2189–2200, 2018.
- [9] W. Young and N. Leveson. Systems thinking for safety and security. In *Proceedings of the 29th Annual Computer Security Applications Conference*, pages 1–8, 2013.
- [10] G. Kavallieratos, S. Katsikas, and V. Gkioulos. Safesec tropos: Joint security and safety requirements elicitation. *Computer Standards & Interfaces*, 70:103429, 2020.
- [11] V. Bolbot, G. Theotokatos, E. Boulougouris, and D. Vassalos. Safety related cyber-attacks identification and assessment for autonomous inland ships. In *Int. Seminar on Safety and Security of Autonomous Vessels (ISSAV)*, 2019.
- [12] I. Ba`ckalov. Safety of autonomous inland vessels: an analysis of regulatory barriers in the present technical standards in europe. *Safety science*, 128:104763, 2020.
- [13] C. A. Thieme, C. Guo, I. B. Utne, and S. Haugen. Preliminary hazard analysis of a small harbor passenger ferry-results, challenges and further work, 2019.
- [14] G. Sabaliauskaite, S. Adepu, and A. Mathur. A six-step model for safety and security analysis of cyber-physical systems. In *Int. Conference on Critical Information Infrastructures Security*, pages 189–200. Springer, 2016.
- [15] G. Sabaliauskaite and Jin Cui. Integrating autonomous vehicle safety and security. In *Proceedings of the 2nd Int. Conference on Cyber-Technologies and Cyber-Systems (CYBER 2017)*, Barcelona, Spain, pages 12–16, 2017.
- [16] G. Sabaliauskaite and J. Cui. Integrating Autonomous Vehicle Safety and Security. *CYBER 2017 : The Second Int. Conference on Cyber-Technologies and Cyber-Systems, (level 0):75–81*, 2017.
- [17] A. Shostack. *Threat Modeling: Designing for Security*, volume Wiley Publishing. 2014.
- [18] G. Kavallieratos, S. Katsikas, and V. Gkioulos. Cyber-attacks against the autonomous ship. In *Computer Security*, pages 20–36. Springer, 2018.
- [19] G. Kavallieratos, V. Gkioulos, and S. K. Katsikas. Threat analysis in dynamic environments: The case of the smart home. In *15th Int. Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 234–240. IEEE, 2019.
- [20] A. Amro, V. Gkioulos, and S. Katsikas. Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation. In *Computer Security*, pages 69–85. Springer, 2019.

Paper IV

A. Amro, V. Gkioulos and S. Katsikas, 'Assessing cyber risk in cyber-physical systems using the attck framework,' *ACM Trans. Priv. Secur.*, Nov. 2022, Just Accepted, ISSN: 2471-2566. DOI: 10.1145/3571733. [Online]. Available: <https://doi.org/10.1145/3571733>

Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework

AHMED AMRO, VASILEIOS GKIOULOS, and SOKRATIS KATSIKAS, Norwegian University of Science and Technology, Norway

Autonomous transport receives increasing attention, with research and development activities already providing prototype implementations. In this article we focus on Autonomous Passenger Ships (APS), which are being considered as a solution for passenger transport across urban waterways. The ambition of the authors has been to examine the safety and security implications of such a Cyber Physical System (CPS), particularly focusing on threats that endanger the passengers and the operational environment of the APS. Accordingly, the article presents a new risk assessment approach based on a Failure Modes Effects and Criticality Analysis (FMECA) that is enriched with selected semantics and components of the MITRE ATT&CK framework, in order to utilize the encoded common knowledge and facilitate the expression of attacks. Then, the proposed approach is demonstrated through conducting a risk assessment for a communication architecture tailored to the requirements of APSs that were proposed in earlier work. Moreover, we propose a group of graph theory-based metrics for estimating the impact of the identified risks. The use of this method has resulted in the identification of risks and their corresponding countermeasures, in addition to identifying risks with limited existing mitigation mechanisms. The benefits of the proposed approach are the comprehensive, atomic, and descriptive nature of the identified threats, which reduce the need for expert judgment, and the granular impact estimation metrics that reduce the impact of bias. All these features are provided in a semi-automated approach that reduces the required effort and collectively are argued to enrich the design-level risk assessment processes with an updatable industry threat model standard, namely ATT&CK.

CCS Concepts: • **Security and privacy** → **Systems security**.

Additional Key Words and Phrases: Risk Assessment, Safety and Security, Cyber-Physical System, Autonomous Ship, MITRE ATT&CK, FMECA

ACM Reference Format:

Ahmed Amro, Vasileios Gkioulos, and Sokratis Katsikas. 2021. Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework. *J. ACM*, (2021), 35 pages. <https://doi.org/10.1145/1122445.1122456>

1 INTRODUCTION

Interest in automated and autonomous ships has increased in the last years with many ongoing projects in this domain driving the industry into a major transformation [21]. An instance of this trend is the project targeting the development of an Autonomous Ferry (Autoferry) for transporting passengers autonomously across the Trondheim city canal in Norway. We classified the Autoferry earlier as an Autonomous Passenger Ship (APS) in [7], and the reader can find detailed information about the project in [46]. This APS is expected to be fully autonomous, with remote navigation and control capabilities enabled by a heterogeneous communication architecture which we specified earlier [8]. This communication architecture fully supports the autonomous operation of e-navigation,

Authors' address: Ahmed Amro, ahmed.amro@ntnu.no; Vasileios Gkioulos, vasileios.gkioulos@ntnu.no; Sokratis Katsikas, sokratis.katsikas@ntnu.no, Norwegian University of Science and Technology, Gjøvik, Norway, 2815.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

0004-5411/2021/-ART \$15.00

<https://doi.org/10.1145/1122445.1122456>

which is defined by the International Maritime Organization (IMO) as "*the harmonized collection, integration, exchange, presentation and analysis of maritime information onboard and ashore by electronic means to enhance berth to berth navigation and related services, for safety and security at sea and protection of the marine environment*" for which the reader can find more information in [19]. However, introducing automation along with e-navigation increases the likelihood of cyber attacks because of the required increased connectivity and decreased human supervision.

Broadly, cyber attacks that target the maritime domain are increasing both in numbers and severity, also enabled by the aforementioned ongoing digital transformation. Such attacks target all the segments of maritime infrastructure, including the ships, ports, and shipping companies. Some notable and well-known examples of such incidents include the attack against the COSCO shipping company [2], the Austal naval shipbuilder [3] and the notoriously disruptive attack against the Maersk shipping company [25]. Ships themselves have also been targets of attacks, since, arguably, attacks against them are of comparatively low complexity [53], and past incidents targeting their Global Positioning System (GPS) [28] and communication technologies [63] are indicative of both their feasibility and potential impact.

In the case of the APS, its security risks can directly or indirectly endanger the safety of the passengers and affect the operational environment. Potential risks can arise if the ship's remote control capabilities are hijacked, and the ship is directed towards a collision with the surrounding environment or other ships. Therefore, risk management of the APS communication architecture must be implemented in order to increase the trustworthiness, security, and resilience of the integrated systems. Risk management comprises several processes with risk assessment at the core, as discussed in ISO 31010 [16] and ISO 27005 [23]. Furthermore, the relationship between safety and security in the risk management of Cyber-Physical Systems (CPS) like autonomous ships requires additional attention in order to ensure the safety of people and the systems themselves.

In this paper, the Failure Modes Effects and Criticality Analysis (FMECA) [15] has been chosen for conducting a risk assessment for the APS communication architecture. FMECA was conducted to identify risks and suggest mitigation methods to support the efforts towards developing a security architecture for the APS. In order to overcome limitations in existing risk assessment methods (discussed in Section 2), we propose an approach for utilizing the common knowledge encoded within the MITRE ATT&CK framework [58] within the FMECA process. Additionally, we introduce a group of impact estimation metrics that can be calculated from the target system model by utilizing concepts from graph theory [62]. The results reflect the comprehensive nature of the proposed approach in addition to the utility of the suggested metrics in reducing the effect of biased analysis and the need for expert judgment. Finally, several risks have been identified and considerations for mitigation methods have been proposed.

The remainder of this paper is structured as follows. In Section 2, our rationale for the proposition of the risk assessment approach is discussed, in addition to a comparison with relevant works. Additionally, a brief description of an Autonomous Passenger Ship (APS) is provided. The APS constitutes a CPS use case that is utilized to evaluate the proposed approach. Further, Section 2 includes a description of the ATT&CK framework and graph theory to facilitate later discussions throughout the paper. Then, Section 3 describes a group of related risk analysis works that influenced or that are comparable to the work in this paper. After that, Section 4 presents the proposed risk assessment approach for CPS and Section 5 presents an evaluation of the proposed approach using the case of the APS. Section 6 then presents the results of the conducted risk assessment of the APS, highlighting the benefits of the proposed approach. In Section 7 we discuss certain limitations and give recommendations for future work. Finally, Section 8 summarizes our conclusions.

2 BACKGROUND

2.1 Motivation and comparison with existing approaches

The work in this paper is motivated by the need to introduce cyber risk management activities into the maritime domain and specifically the MilliAmpere2 APS use case described in Section 2.2 toward the proposition of a suitable cybersecurity architecture. IMO urged the different maritime industry stakeholders to include cyber risk management into their safety management systems. The resolution suggested some guidelines and requirements for cyber risk assessment and management [18]. This includes the consideration of different cyber technology domains such as IT and OT (Sections 2.1.1 and 2.1.2 in [18]), the consideration of operational, safety and security impacts (Section 1.1 in [18]), and the need for continuous risk assessment and management (Sections 3.3 and 3.5 in [18]). Because of the different technology domains found in the maritime environment, we considered the application of ATT&CK (more details in Section 2.3). ATT&CK includes different technology domains within its threat model, specifically, enterprise, mobile, and industrial control systems (ICS). All were found to be relevant to the APS use case.

Regarding impact estimation for risk calculation, as shown in Table 1, several approaches for estimating the impact of cyber attacks in CPS were observed in the literature. The majority of the studied literature estimated the impact severity through the four elements described in the SAE J3061 Ground Vehicle Standard [17], namely, safety, financial, operational, and privacy/legislative. Other works considered the impact from the perspective of the breached security goals, namely, confidentiality, integrity, and availability, similar to the Common Vulnerability Scoring System (CVSS) criteria. Macher and Armengaud [37] utilized the DREAD impact model, Bolbot et al [12] utilized three impact elements, namely, safety, environmental, and financial, while Tam and Jones [60] employed a novel model named MaCRA. However, we argue that the observed impact estimation approaches fail to clearly capture the impact of all observed adversarial tactics and techniques in ATT&CK including, command and control, defense evasion, discovery, initial access, lateral movement, persistence, and credential access. Therefore, we considered utilizing the impact model proposed in the SAE J3061 standard and extending it to capture the security-related impact that is not captured in the other approaches.

The continuous risk assessment and management requirement has motivated us to reduce the efforts associated with conducting the risk assessment process. Table 1 reflects the sources of knowledge that are utilized in the different survey approaches. Expert judgment constitutes the main source for threat identification, risk calculation, and the proposition of mitigation methods. In this regard, our approach employs the concept of curated knowledge and utilizing available threat-related information in the ATT&CK repository. Similarly, Sheehan et al [54] employed the National Vulnerability Database (NVD) as a source for updatable software-related threat identification methods. However, we argue that the ATT&CK threat model is more appropriate in the design phase in comparison to NVD which relies on specific software and hardware information for the identification of relevant vulnerabilities.

In addition to the requirements communicated by IMO in [18] we argue that the observed risk analysis approaches in the literature are not comprehensive enough in their consideration of threats. As depicted in Table 1, the most observed threat identification method in the studied literature is STRIDE; Spoofing, Tampering, Repudiation, Information disclosure, Denial of service and Elevation of privileges [55]. Monteuis et al [44] however, have extended the STRIDE approach to include two additional threat categories, namely, linkability (which violates privacy) and confusion (which violates trustworthiness). In this regard, we considered the utilization of the ATT&CK framework which provides additional attack description information that STRIDE simply does not provide.

STRIDE-based methods consider only six threat categories, some extended them to eight. On the other hand ATT&CK suggests more than 600 attack techniques across the several technology domains. Therefore, our threat identification approach is more descriptive and comprehensive. In addition to that, ATT&CK is constantly updated, thus providing an updatable feature to our risk assessment approach.

Still, ATT&CK is not a risk assessment framework. Therefore, we have considered several approaches toward the proposition of the most suitable risk assessment method. We referred to the IEC 31010:2019 [16] standard for risk assessment techniques. The standard provides detailed descriptions and comparisons among the most observed techniques employed within the different risk assessment steps. In our quest, we considered the scope, time horizon, requirements for specialist expertise, and the amount of effort required to apply the risk assessment techniques. Our scope of the risk assessment in CPS such as the APS includes components, equipment, and processes. The time horizon should be flexible, also the need for specialist expertise and amount of effort required should be at most moderate, to support the satisfaction of the requirement for continuous risk assessment and management as well as to reduce the effect of biased assessment associated with expert judgments. The aforementioned criteria have led us to FMECA. Moreover, the standard highlights the applicability of FMECA in the different steps in the risk assessment process, namely, risk identification, consequence, likelihood, risk estimation and risk evaluation. Afterwards, in each step of the FMECA process, we aimed to integrate the most suitable technique, while considering the requirements for the risk assessment process mentioned previously and by utilizing relevant artifacts from the literature. We suggested the utilization of a Preliminary Hazard Analysis (PHA) or a hazard and operability study (HAZOP) for the estimation of safety and financial impact based on the previous works by Bolbot et al [12] and Thieme et al [61]. Both works utilized these approaches for estimating the safety impacts of cyber attacks in different maritime use cases. Also, we suggested the utilization of the CVSS exploitability metrics for likelihood estimation based on its common adoption in the literature as depicted in Table 1 (more details in Section 4.6) and its suitability for our approach.

2.2 Communication Architecture of an Autonomous Passenger Ship

The Autoferry project [46] aims to develop an APS prototype named the MilliAmpere2; an autonomous ferry capable of carrying 12 passengers across the Trondheim city canal, proposed as an alternative to a high-cost bridge [29]. The ferry will operate autonomously with a human operator in a Remote Control Centre (RCC) monitoring its operations and with the capability to intervene at any moment.

We have designed a communication architecture for the APS [8] that enables it to communicate with its operational context. The architecture enables the APS to carry out a group of functions, including autonomous and remote navigation and control. Navigation functions rely on collecting sensing information from the surrounding environment through arrays of sensors including lidars, radars, Infra-Red, and video cameras which interface using several Sensor Processing Units (SPU) and sensor switches. Then an Autonomous Navigation System (ANS) achieves situational awareness by leveraging sensor data to determine safe routes.

Additionally, the APS relies on real-time kinematics and the Global Navigation Satellite System (GNSS) for positioning. Moreover, the APS has the ability to carry control functions and maneuvers using a machinery system that includes a Dynamic Positioning (DP) system and active thrusters interfaced through Input/Output cards. The machinery system is supervised by an Autonomous Engine Monitoring and Control (AEMC) system. The APS is also equipped with an emergency push-button for initiating the emergency protocol, according to which a nearby Emergency Control Team (ECT) is expected to intervene when needed.

Table 1. Comparison between our approach and the surveyed works

Work	Technology Domains	Threat Identification			Risk Calculation Model			Risk Calculation	Knowledge Source			Attack Sequence Approach
		Attacker Profile Model	Threat Categories	Attack Categories	Impact	Likelihood	Mitigation Proposition		Threat Identification	Risk Calculation		
Our approach	IT/OT/ Mobile	ATT&CK Groups*	ATT&CK Tactics	ATT&CK Techniques	SAE J3061 (1) + Staging	CVSS Exploitability	Risk Priority Number + Risk criteria	ATT&CK Tactics and Techniques	Graph Theory + Expert Judgment	*		
[44]	IT/OT	Knowledge, Expertise, Equipment	STRIDE + LC	Alter, Listen, Disable, Forge	SAE J3061 (1) + Controllability	Capability: elapsed time, opportunity	Risk Matrix	Expert Judgment	Expert Judgment	Attack tree		
[37]	IT/OT	-	STRIDE + HARA (2)	-	DREAD (3)	Knowledge, resources	Risk Matrix	Expert Judgment	Expert Judgment	-		
[31]	OT	-	STRIDE	-	SAE J3061 (1)	Expertise, knowledge, opportunity, equipment	Risk Matrix	Expert Judgment	Expert Judgment	-		
[10]	OT	-	STRIDE	-	SAE J3061 (1)	Model 1: expertise, knowledge, opportunity, equipment. Model 2: CVSS Exploitability	Bayesian Network	Expert Judgment	Expert Judgment	-		
[54]	IT/OT	-	Software vulns. in NVD	-	CVSS impact: Confidentiality, Integrity, Availability	CVSS Exploitability	CVSS Base Score + Risk Matrix	Expert Judgment	NVD	NVD + Expert Judgment		
[33]	OT	-	STRIDE	-	Predefined criteria similar to SAE J3061 (1)	Predefined criteria considering capability, motivation, controls, available exploits, reachability, vulns., knowledge, authentication	Risk Matrix	Expert Judgment	Expert Judgment	Expert Judgment		
[38]	IT/OT	-	STRIDE	-	CVSS Impact	CVSS Exploitability	CVSS Base Score	Expert Judgment	Expert Judgment	Attack tree		
[12]	IT/OT	Technological level, Ease-of-exploit (EoE)	Predefined list	Predefined list	Safety, environmental, financial	Technological level, activity level, interest level, EoE, exposure level, vulns level, frequency index	Risk Matrix	Expert Judgment	Expert Judgment	Expert Judgment		
[60]	OT	EoE	MacCRA	Damage, theft, denial of service, misdirect, obfuscate	MacCRA (Vulns.+Ease of Exploit+Reward) Vulns: Attack vector, vulns effects Ease of Exploit: Attack agent, type, target type, resources, controls Reward: Attack agent type, target type, attacker goal, target effect	MacCRA	MacCRA	Expert Judgment	Expert Judgment	Expert Judgment		

(1) SAE J3061 impact elements: Safety, Privacy, Financial, Operational
 (2) HARA: Hazard Analysis and Risk Assessment
 (3) DREAD impact elements: Damage Potential, Reproducibility, Exploitability, Affected Users, Discoverability
 * Not included in the current work but planned in future work. Check Section 7 for more details.

Moreover, a set of heterogeneous communication modules and components are proposed and integrated within the onboard network as shown in Figure 1, to satisfy the communication requirements of various stakeholders, as discussed in [7]. Additionally, the architecture supports carrying out the autonomous and remote functions through a group of communication functions including ship-to-shore, ship-to-ship, internal, and emergency communications. Ship-to-shore communication enables the APS to communicate with the RCC through two IP-based redundant communication modules: the Mobile Communication Module (MCM) and the APS-RCC Module. The technologies for implementing these modules are expected to be LTE/4G/5G and Wi-Fi, respectively. Ship-to-ship communication is facilitated through a traffic module such as an Automatic Identification System (AIS), while internal communication is enabled through two Core/Distribution tiers (C/D part A and part B) each implemented using two redundant layer-3 switches. Emergency communication relies on two modules. The first emergency module (Emergency module 1) facilitates communicating with the ECT to perform emergency navigation functions, while the other module (Emergency module 2) is used to transmit emergency signals to the ECT when the emergency push button is pressed by passenger. Finally, an intelligent entity named Connectivity Manager performs autonomous and remote network management functions.

A centralized component named Autonomous Ship Controller (ASC) resides in the center of the APS network which hosts the primary and backup servers hosting the ANS, AEMC, Connectivity Manager, as well as other components for the system, network, and security management.

A group of systems resides in the RCC network for remote navigation functions, control functions, and additional ship-to-shore communication functions. The network modules and devices are equivalent to the ones on board the APS. On the other hand, the Remote Ship Controller (RSC) hosts the Remote Navigation System (RNS), Remote Engine Monitoring and Control System (REMS), in addition to other components for the remote system, network, and security management. More details regarding the communication architecture can be found in [8].

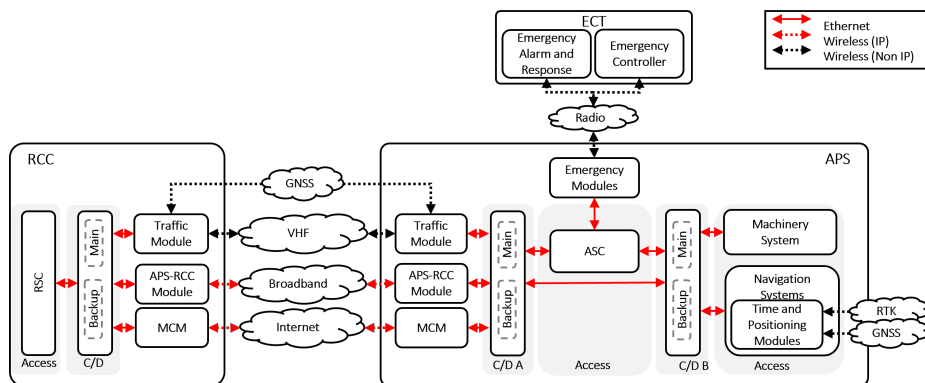


Fig. 1. APS Communication Architecture, Inspired from [8]

In this paper, a risk assessment approach is first presented through the application of a FMECA enriched with the MITRE ATT&CK framework. Then, the proposed approach is evaluated using the communication architecture of the APS.

2.3 MITRE ATT&CK framework

An increase in adopting the ATT&CK framework proposed by MITRE [58][5] is observed in both academia and industry. ATT&CK, which stands for Adversarial Tactics, Techniques, and Common

Knowledge, is a curated knowledge base that models the behavior of cyber adversaries. It provides a common taxonomy in describing the different phases of the adversary attack life-cycle. Among the most important features of ATT&CK, that distinguishes it from other threat models, is the abstraction level in describing adversarial tactics and techniques. High-level models observed in the literature such as STRIDE [55] and the cyber Kill Chain [40], fail to effectively reflect the granularity of actions that adversaries can take, how they relate to one another, their consequences related to adversarial objectives, their correlation with mitigation methods and data sources, and their targeted platforms and systems [58]. We argue that these particular features qualify ATT&CK as an appropriate engine for conducting comprehensive and logically sound risk assessment, utilizing the systematically encoded expert knowledge to reduce effort and inconsistencies during risk estimation. The ATT&CK adversarial model comprises, among others, a group of essential terms, Tactics, Techniques, and Procedures. Tactics represent the adversarial objective of the attack, techniques represent the adversarial method for realizing an objective, while procedures represent the actual software utilized to run the technique to realize the tactic. The framework is organized as a group of matrices for different technology domains, enterprise, mobile, ICS, containers and adversarial machine learning. Each matrix holds the relationships between tactics, techniques, procedures, mitigation methods, and others.

In the context of the APS, since it comprises a collection of Information Technology (IT) and Operational Technology (OT) components, the comprehensive nature of ATT&CK was of particular utility to identify relevant threats for APS's heterogeneous components. Moreover, the well-established relationships in ATT&CK were found to be logically compatible with FMECA. A detailed discussion on applying ATT&CK in the FMECA process is presented in Section 4.

2.4 Graph Theory

Several works have highlighted the utility of graph theory [62] in analysing interconnected infrastructures [57] [35] [4]. Graphs are mathematical structures used to model the relationships between distinct objects [62], while the abstraction of graphs enables them to model a wide range of relationships including networked systems [4], connected organizational structures [35], and other types of related objects. A graph consists of nodes, each representing an object that is involved in a relationship with other objects, while these relationships are represented with edges connecting the related nodes. A group of formal measures has been proposed to analyze the graph, including the centrality measures. These measures can be utilized to estimate the relative influence of a node in the graph. Several centrality measures exist such as closeness centrality, degree centrality, Eigenvector centrality, and many others [57]. The aggregation of all centrality measures has been found to identify nodes with the highest influence over the graph [57]. On the other hand, the Outbound Degree Centrality (ODC) (i.e. out-degree centrality) of a node reflects the number of its neighbors. Nodes with the highest ODC within a graph are called "cascade initiating nodes" [36] and must be prioritized when examining mitigation controls [57]. In this paper, we rely on ODC and the semantics of the combined centrality measures proposed by Stergiopoulos et al [57] during the estimation of the operational and the security related impacts (Section 4.4) for the examined use case.

3 RELATED WORK

In this section, we will discuss, in detail, related works that share considerable similarities with our approach. In the automotive domain, Islam et al [31] argued that a risk assessment be performed in the requirements elicitation phase of the development life-cycle, in order to guide countermeasure integration in the design and development phase. The authors proposed a risk assessment framework that is aligned with known automotive processes related to functional safety and usability.

The authors proposed a novel approach to calculate semi-quantitative risk by applying STRIDE for threat modeling, attacker expertise, required knowledge, equipment and window of opportunity for likelihood estimation, and the common impact elements safety, privacy, financial, and operation. The authors proposed the application of weights to adjust the impact estimates based on the organization's needs. In this paper, we utilize the concept of weights (i.e. factors) from the framework of [31] to adjust the impact assessment of risks according to the followed risk management strategy. Sheehan et al [54] proposed a risk classification framework based on Bayesian networks to evaluate the security level of connected and autonomous vehicles. The authors utilized the software vulnerabilities in the National Vulnerability Database (NVD) for an updatable threat identification approach. Then, they employ Bayesian networks for estimating the likelihood and impact of threats, following the CVSS approach toward calculating the risk. Expert judgment is integrated into the proposed framework for deriving the structure of the Bayesian networks and estimating several risk variables. Our proposed approach in this paper shares similar features with [54] regarding the utilization of CVSS, the integration of expert judgment, the updatability of risk scores as well as the accommodation of existing mitigation techniques into the risk calculation. In contrast, we consider more comprehensive attack techniques and granular impact estimation parameters.

Compared to the domain of autonomous cars, fewer works have addressed risk assessment for autonomous ships. Kavallieratos et al [33], proposed a multilayer architecture for the information and communication technology systems in cyber-enabled ships which include autonomous ships. The authors then applied the STRIDE threat modeling method to identify potential threats. Then the associated risks were assessed using risk matrices following risk estimation criteria inferred from the work by Jelacic et al [32]. The risk estimation criteria consider safety, operations, economic, information leakage, and reputation impact elements. Moreover, the criteria consider attackers' capability, motivation, and knowledge, in addition to existing countermeasures and exploits as well as component reachability. Additionally, Tam and Jones [59] proposed a model-based risk assessment framework called MaCRA [60] and applied it on three futuristic ships with different applications and levels of autonomy. The process started with applying the MaCRA threat assessment framework and then the risk assessment process. The threat assessment considered the different attackers' profiles, their goals, and available resources. Moreover, the ships' vulnerabilities related to the expected technologies and the expected impact of the vulnerabilities have been considered. In the risk assessment process, five-tier values were applied to quantify the risks associated with the identified threats and the risk level was presented through two values, namely Ease of Exploitation (Likelihood) and Attackers reward (Impact). Bolbot et al [12] proposed a cyber risk assessment method for ship systems based on a Cyber-Preliminary Hazard Analysis. The method was applied for conducting a risk assessment and providing design enhancement of the navigation and propulsion systems of inland waterways autonomous vessels. The risk assessment considers attacker groups, system vulnerabilities, attack likelihood, consequences, and existing barriers. The likelihood estimation considers component reachability (i.e. connectivity), attack complexity, attacker group motivation, capabilities, activity level, ease of exploitation, the absence of barriers. The impact estimation considers safety, environmental, and financial consequences.

An application of ATT&CK in the risk analysis of digital substations is presented by Khodabakhsh et. al [34]. The authors utilized the ICS matrix in ATT&CK to identify possible attack paths in a system of digital substations, assessed their potential impact regarding confidentiality, integrity, and availability (CIA), and finally, proposed a group of suitable countermeasures.

In contrast to the related works presented in this section, we propose a comprehensive and systematic approach for identifying relevant attacks against components in CPS architectures, considering three technology domains in ATT&CK, namely ICS, enterprise, and mobile. Additionally,

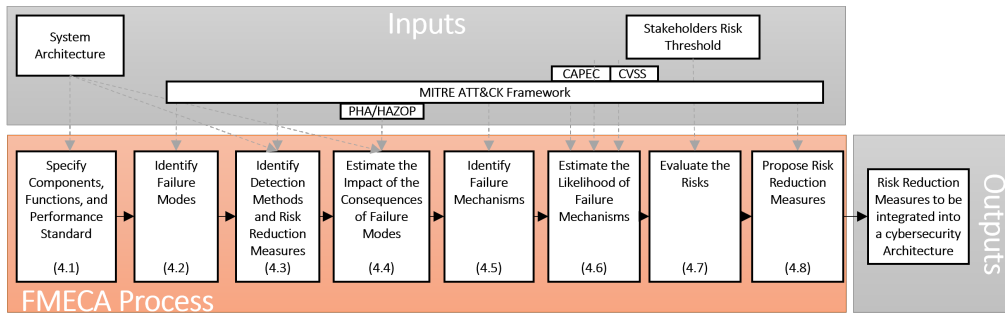


Fig. 2. Overview of the proposed FMECA-based approach showing the integrated information sources

we propose a granular and comprehensive impact estimation approach considering operational, safety, financial, and system and information security-related impacts.

4 THE PROPOSED RISK ASSESSMENT APPROACH

The proposed risk assessment approach is based on a design-level FMECA [15]. A Bottom-up approach is suggested that requires knowledge of low-level components. A FMECA process consists of three main phases, namely planning the analysis, performing it, and finally documenting it. The overall risk assessment process highlighting the utilized information sources is depicted in Figure 2. In the planning phase the objectives, and scope of analysis, as well as the considered scenarios, are identified. The ATT&CK framework aids the FMECA process by describing attack techniques and tactics in a manner that can be used to induce failure scenarios [5]; this feature is of particular utility to the analysis and communication of the identified risks. Additionally, the criteria for the treatment of failure modes should be defined according to the followed risk management strategy. Afterwards, the analysis is performed, and the detailed steps for performing a FMECA utilizing ATT&CK are presented in the subsequent sections. Finally, a FMECA report is generated, which gives detail on the analysis process.

4.1 Specify Components, Functions, and Performance Standard

An architecture model describing the different architectural components, their functions, and interconnections is a vital element of the risk assessment process. The components addressed in this analysis are the functional components, including software and hardware elements. The components are classified as Information Technology (IT), Operational Technology (OT), Wireless, and/or a combination of multiple categories. Wireless components include mobile devices and/or devices with wireless services. The classification of components is conducted based on the criteria shown in Table 2. Additionally, operational modes are proposed to be identified and considered to improve the analysis, particularly when they inflict a change in the system state (set of components and their connections). Moreover, the performance standard for each component function should be specified, to define what constitutes a component failure.

4.2 Identify Failure Modes

A failure mode is defined as a manner in which a failure occurs [15]. In this study, the security of CPSs is analyzed. So, the security failure modes are considered. In system security engineering, a system security failure is defined as "not meeting the security-relevant requirements, objectives,

Table 2. Components Classifications

Class.	Description
IT	Components that are hosted on a traditional IT system such as multipurpose computers or network devices.
OT	Components that are involved in monitoring and controlling functions.
Wireless	Components that are connected to a mobile network or communicate with an external infrastructure such as Aids to Navigation to acquire location-related information in the maritime domain.
IT/OT	Dual-homed components that are hosted on a traditional IT system and are involved in monitoring and controlling functions.
IT/OT/Wireless	Components that are classified as IT/OT and are connected to a mobile network or communicate with an external infrastructure

Table 3. Malicious failure modes according to ATT&CK

Class	Failure Mode	Failure effect
IT, OT, and Wireless	Initial Access	entry to the network.
	Collection	gathering data of interest
	Command and Control	communicating with other compromised components in the network to control them.
	Defense Evasion	avoiding detection.
	Discovery	discovering the environment.
	Execution	running malicious code.
	Impact	impacting the data and/or components.
	Lateral Movement	moving between components within the environment.
	Persistence	maintaining a foothold in the environment.
IT and Wireless	Privilege Escalation	increasing privilege.
	Credential Access	discovering account names and passwords.
	Exfiltration	stealing data
OT	Impair Process Control	impacting the control processes
	Inhibit Response Function	impacting the safety, protection, and monitoring functions from responding.
Wireless	Network Effect	impacting the network traffic.
	Remote Service Effect	impacting components remotely.

and performance measures, to include exhibiting unspecified behavior, exhibiting unspecified interactions, or producing unspecified outcomes, where there is security-relevance" [51]. Earlier works have discussed security failure modes considering the CIA triad [9] and [34]. We propose to go beyond that and consider, for each component, a broader range of security failure modes. For this, we utilized the ATT&CK framework. We argue that the failure modes which are referred to as Tactics in ATT&CK (i.e. kill chain phases) are more comprehensive than the high-level failure modes classification according to the CIA triad since the tactics in ATT&CK involve failing more than one of the CIA attributes or none. The considered security failure modes for each component class are depicted in Table 3.

4.3 Identify Detection Methods and Risk Reduction Measures

In this step, the existing detection and risk reduction measures (i.e. controls) are identified and analyzed. These controls affect the *Detectability* estimation value when estimating the risks of failures. This value constitutes the probability of the attack being detected or mitigated. The calculation of this value is conducted as follows:

4.3.1 The Failure-Mitigation Table (FMT). The FMT is constructed, which captures the possible mitigation methods for each considered failure mechanism as well as their expected efficiency. The ATT&CK framework was consulted for this purpose. A list of all techniques in the different ATT&CK matrices in the first and second column and their suggested mitigation methods in the third column were pulled from the online repository to populate the FMT. The fourth column captures the efficiency of the mitigation method (M) against that failure mechanism (FM) ($Efficiency_{FM,M}$); this value does not exist in the current ATT&CK knowledge base and therefore should be estimated. A typical measurement scale for detectability rating is provided in the FMECA standard [15]; the example is for a wind turbine. A sample of the FMT is depicted in Table 4. In the complete version of the FMT, a single technique could be mitigated by several mitigation methods. Similarly, a single mitigation method could be used to mitigate several techniques.

Table 4. An FMT Sample reflecting some techniques, their suggested mitigation methods, and their estimated efficiency.

Matrix	Technique	Mitigation	Efficiency
ICS	Change Program State	Access Management	0.5
Enterprise	Commonly Used Port	Network Intrusion Prevention	0.5
Mobile	Remote File Copy	Application Vetting	0.5

$$Detectability_{FM,C,M} = Coverage_{M,C} \times Efficiency_{FM,M} \quad (1)$$

4.3.2 The Component-Mitigation Table (CMT). The Component-Mitigation Table (CMT) is constructed, which captures the coverage of mitigation methods for each component. For this, the CMT is populated with the mitigation methods in ATT&CK and their coverage for the existing components in the architecture. What is specifically meant by "coverage" is different from one mitigation method to another. But, in this paper, a component is said to be influenced by a mitigation method if the component in the proposed architecture is subject to an architectural decision that enforces the mitigation method. For instance, if the architecture is designed with network segmentation, all components in the isolated network segments are said to be covered by the network segmentation mitigation method. The CMT structure consists of the operational modes in the first column, the mitigation methods in the second, while the architectural components are spread across the remaining columns, and the coverage of the mitigation methods (M) for each component (C) (covered:1 or not:0) as the values ($Coverage_{M,C}$). A sample of the CMT is depicted in Table 5.

Table 5. A CMT Sample reflecting the coverage of some components by the mitigation methods

Op-Mode	Mitigation	Component A	Component B
All	Access Management	0	1
All	Network Segmentation	1	0

4.3.3 Calculating the Detectability. The value of *Detectability* for each failure mechanism of a specific component is calculated based on whether or not the component is covered by a mitigation method suggested for the specific failure mechanism (as indicated in the CMT) and its mitigation efficiency (as indicated in the FMT). The *Detectability* of the failure mechanism (FM) for a component (C) when considering the coverage of mitigation method (M) is calculated using equation 1.

4.4 Estimate the Impact of the Consequences of Failure Modes

The tactics in ATT&CK are terms that describe the desired outcome of attacks by attackers. This terminology allows the utilization of ATT&CK tactics as a classification of consequences for their corresponding techniques. Additionally, certain techniques have unique consequences; these techniques are grouped for each matrix within special tactic categories, namely, impact, network effect, or remote service effects. In our approach, we propose the consideration of all the tactics across all the relevant matrices in addition to the techniques under the special tactic categories. The impacts of these tactics and techniques are estimated based on the four most observed elements of impact of threats against CPS, namely, safety, financial, operational, and information criticality (e.g. privacy) [44] [31] [10] [33]. Nevertheless, some tactics have no impact according to the observed impact model in the literature; they rather have a security related impact that affects the security of

connected nodes. For instance, a single successful technique aiming to achieve defensive evasion, which is the most observed ATT&CK tactic in 2019 [26] and second-most in 2020 [48], has no immediate impact on safety, privacy, financial or operations. But, it will support the attacker's efforts to stage future attacks. Therefore, we propose a fifth impact element named "Staging" to capture the impact of techniques that facilitate the staging of future attacks. The proposed process for estimating the impact of consequences is as follows:

4.4.1 Failure-Mode-Consequences Table (FMCT). The FMCT is constructed. It captures the mapping between each failure mode and its expected consequences across the entire system, expressed through the five impact elements. The FMCT can differ among different target systems. A sample FMCT is depicted in Table 6. For each matrix in ATT&CK, each failure mode was analyzed and the related impact elements were determined. For instance, all techniques under the collection tactic aim to collect information from the compromised target. The consequences of this attack can be assessed with relevance to the information criticality of the component with regards to the hosted process (i.e. intellectual property), information (i.e. confidentiality and privacy), and location information, therefore, an attack enabling collection will only impact information criticality of the component. As another example, the consequences of a successful attack with a "Loss of Control" failure mode will impact the ICS operations. The impact value can be estimated with relevance to the criticality of the component to the control functions it is involved with, in addition to the possible safety and financial impacts. Based on this mapping, five values are specified for each failure mode, namely Safety Factor (SF), Financial Factor (FF), Information Criticality Factor (ICF), Operational Factor (OF), and Staging Factor (StF). A zero value reflects that no consequence is expected, while a positive value reflects the magnitude of the consequences. The implementation of this approach was influenced by the work of Islam et. al [31]. Based on the risk management strategy, the factor values can be controlled to reflect the priority of impact elements on the final risk value. For instance, the stakeholder concerns may prioritize safety as the greatest concern while considering privacy the lowest. In this case, the values of SF and ICF could be controlled to reflect that priority by increasing SF and decreasing ICF with appropriate proportions, based on the stakeholders' concerns.

4.4.2 The Component-Criticality-Scores Table (CCST). The CCST is constructed. it captures the impact scores for each component that correspond to the previously identified impact elements. The estimation criteria for each score are explained below:

- **Safety and Financial criticality (SC and FC):** safety and financial impact scores for each component can be elicited through a Preliminary Hazard Analysis (PHA) or a hazard and operability study (HAZOP). An example of a set of estimation criteria is depicted in Table 7. The concept of this approach has been observed in hazardous waste management, a hazardous waste index is assigned to waste reflecting the level of safety procedures that are required in its handling, storage, transportation, and treatment [27]. The maximum possible safety and financial consequence values deduced from the PHA or HAZOP analysis for each component are recorded as the corresponding safety and financial criticality. For instance, if a component failure has been estimated to cause a catastrophic safety and financial consequence, the safety and financial impact scores for that component will be the maximum, (i.e. 1). After the analysis of the failure modes in ATT&CK, the nature of possible safety impacts is all similar, possibly leading to a life-threatening incident. On the other hand, the nature of financial impact was found to be different for a single failure mode, namely, Carrier Billing Fraud, a technique in the mobile matrix that could cause a financial impact in the form of unexpected billing for SMS-enabled devices should they exist within the CPS.

Table 6. Mapping of failure modes and their consequences in the FMCT

Mobile failure modes	O	S	I	F	ST	IT failure modes	O	S	I	F	ST	ICS failure modes	O	S	I	F	ST
Collection			1			Collection			1			Collection			1		
Command and Control					1	Command and Control					1	Command and Control					1
Defense Evasion					1	Defense Evasion					1	Defense Evasion					1
Discovery					1	Discovery					1	Discovery					1
Execution	1			1	1	Execution	1		1	1	1	Execution	1	1		1	1
Exfiltration			1		1	Exfiltration			1		1	Theft of Operational Information			1		1
Initial Access					1	Initial Access					1	Initial Access					1
Lateral Movement					1	Lateral Movement					1	Lateral Movement					1
Persistence					1	Persistence					1	Persistence					1
Credential Access					1	Credential Access					1	Damage to Property	1	1	1	1	
Data Encrypted for Impact	1	1	1	1	1	Data Encrypted for Impact	1	1	1	1	1	Denial of Control	1	1			
Privilege Escalation					1	Privilege Escalation					1	Denial of View	1	1		1	1
Carrier Billing Fraud				1		Account Access Removal	1	1		1	1	Impair Process Control	1	1		1	
Clipboard Modification	1					Data Destruction	1	1	1	1		Inhibit Response Function	1	1		1	
Delete Device Data	1		1	1		Data Manipulation	1	1	1	1		Loss of Availability	1	1	1	1	1
Device Lockout	1					Defacement	1	1	1	1	1	Loss of Control	1	1		1	
Downgrade to Insecure Protocols	1		1		1	Disk Wipe	1	1	1	1		Loss of Productivity and Revenue	1	1		1	
Eavesdrop on Insecure Network Communication			1		1	Endpoint Denial of Service	1	1		1		Loss of Safety	1	1		1	
Exploit SS7 to Redirect Phone Calls/SMS	1					Firmware Corruption	1	1		1		Loss of View	1	1		1	1
Exploit SS7 to Track Device Location			1			Inhibit System Recovery	1	1		1		Manipulation of Control	1	1		1	
Generate Fraudulent Advertising Revenue						Network Denial of Service	1	1		1		Manipulation of View	1	1		1	1
Input Injection	1					Resource Hijacking	1	1		1	1						
Jamming or Denial of Service	1	1			1	Service Stop	1	1		1							
Manipulate App Store Rankings or Ratings						System Shutdown/Reboot	1	1		1							
Manipulate Device Communication	1	1	1	1													
Modify System Partition	1																
Obtain Device Cloud Backups			1		1												
Remotely Track Device Without Authorization			1														
Remotely Wipe Data Without Authorization	1	1	1	1													
Rogue Cellular Base Station	1				1												
Rogue Wi-Fi Access Points	1				1												
SIM Card Swap	1				1												
SMS Control	1				1												

O: Operational || S: Safety || I: Information Criticality || F: Financial || ST: Staging

Table 7. SC and FC estimation criteria [61]

Safety Criticality (SC)	Description	Financial Criticality (FC)	Description
None	No injuries	None	No damage to equipment or other property
Minor (0.25)	Single and/or minor injuries	Minor (0.25)	Local equipment damage, small damage to other property, or minor loss of income.
Significant (0.5)	Multiple minor injuries and/or severe injury	Significant (0.5)	Damage to CPS, to other property, or significant loss of income.
Severe (0.75)	Single fatality and/or multiple severe injuries	Severe (0.75)	Severe damage to CPS, other properties, or loss of income equivalent to several days of operation.
Catastrophic (1)	Multiple fatalities and severe injuries	Catastrophic (1)	Loss of CPS or other properties.

- Operational criticality (OC): for assessing the operational impacts, several architectural views are created, to calculate several impact values utilizing metrics from graph theory and multidimensional networks [20]. The ORA software [13] [6] is an example of existing

software that can be utilized to draft the architecture views and provide metrics that are used to calculate the different OC metrics.

After the analysis of the failure modes in ATT&CK, we have observed that certain failure modes could affect the overall performance of CPS, others could only affect the control or monitoring functions. Therefore, three operational impact metrics are calculated for each element. A description of each impact metric is given below:

- **Overall operational impact (OOI):** The aggregated centrality measures of the components in the entire network structure are calculated and scaled by creating a graph representing the expected connectivity between the architectural components and their operational context. Each node represents a component (hardware, or software), while each edge represents a network connection (wired or wireless) as well as an expected application-level connection.
- **Impact to the control functions (I2CF):** for each system state (refer to Section 4.1), a graph is created to represent the connectivity between components involved in the control functions. The aggregated centrality measures are then calculated and scaled for each component.
- **Impact to the monitoring functions (I2MF):** similar to previous, but for the monitoring functions.

Finally, the value of the OC metric is calculated differently, based on the considered failure mode. $OC = I2CF$ for the Manipulation of Control, Loss of Control, Denial of Control, and Impair Process Control failure modes, $OC = I2MF$ for the Denial of View, Loss of View, and Manipulation of View failure modes, while $OC = OOI$ for all other failure modes.

- **Information criticality (IC):** this metric captures the criticality of the component concerning possible privacy or/and confidentiality violations. The confidentiality of data stored, processed, or communicated within the CPS network could involve location information. Also, concerns could exist to preserve the intellectual property of processes hosted within the CPS components. After the analysis of the ATT&CK failure modes, three possible impacts have been identified related to information criticality, namely, the attackers might be able to collect sensitive data (e.g. violates users privacy), to collect data that violate the intellectual property, or to collect location information. Therefore, three possible metrics could be estimated based on the failure mode:
 - **Data Criticality (DC):** this metric captures the importance of data hosted or processed in a component. It is measured for each component according to its involvement in the processing and storage of sensitive data.
 - **Intellectual Property Criticality (IPC):** this metric captures the component criticality regarding the hosting of processes with intellectual value.
 - **Location Information Criticality (LIC):** this metric captures the component criticality regarding the involvement with location information and the sensitivity of such information. If the system under analysis is involved in a location-sensitive use case, this metric could be of value and should be estimated according to the use case specifications. Two failure modes can be estimated using this metric, namely, Exploit SS7 to Track Device Location, and Remotely Track Device Without Authorization.

Based on the risk management strategy, the importance of each metric could differ. Therefore, three factors are proposed to control the prioritization of the information criticality metrics, namely, DC_F , IPC_F and LIC_F ; the values of these factors range from 0 to 1. Finally, the information criticality (IC) of component (C) is calculated using Equation 2. It has been found that in only a single failure mode, namely the Collection failure mode, all three metrics could be of relevance. The values of these metrics for each component can be estimated through

$$IC_C = \frac{(DC_F \times DC_C) + (IPC_F \times IPC_C + (LIC_F \times LIC_C))}{(DC_F + IPC_F + LIC_F)} \quad (2)$$

$$OCC_C = \frac{OC_C + SC_C + IC_C + FC_C}{4} \quad (3)$$

the implementation of an early Privacy Impact Assessment [14] or a Data Protection Impact Assessment [11].

- Staging Criticality (StC) : this metric captures the impact of a failure mode that enables the staging of future attacks. We have observed that the impact of some of the considered failure modes is not captured using the previously mentioned impact elements. These failure modes include command and control, defensive evasion, discovery, initial access, lateral movement, persistence, privilege escalation, and credential access (refer to Table 6). Yet, these failure modes are critical to the security status of a system. Other security impact elements such as confidentiality, integrity, and availability are captured directly or indirectly in other impact elements. For instance, the confidentiality impact is captured directly in the information criticality, while the integrity and availability impacts, should they exist, are captured indirectly in several impact elements: if the integrity and/or availability of information or process are not preserved, information, safety, financial, and operational impacts might occur. We propose that the StC metric can be estimated using two metrics, namely Outbound Degree Centrality (ODC), and Overall Component Criticality (OCC). Details regarding both metrics are presented below:

- Outbound Degree Centrality (ODC): Some failure modes, if materialized, enable the attacker to move to or communicate with other components in the network; the impact of this ability increases with higher ODC of the component. The more connected the node to its neighbors the higher the staging impact. Moreover, regarding the credential access failure mode, the discovered credentials can be utilized in other components not connected to the compromised component, even outside the compromised network. The impact of this case is not captured in this specific metric. Nevertheless, we argue that this metric provides a logical estimate of the impact of this failure mode within the compromised network.
- Overall Component Criticality (OCC): the persistence, defense evasion, and privilege escalation failure modes do not directly impact the attacked node or other nodes. So, we propose the utilization of the combined impact metrics to capture the staging impact of these three failure modes. We argue that the impact of a successful attack aiming to inflict these failure modes can be measured by the combined criticality (OC, SC, IC, FC) of the attacked component, using equation 3. In a study of adversarial behavior, a Unified Kill Chain similar to the ATT&CK framework was studied [50], which showed that persistence, defense evasion, and privilege escalation, occur most frequently among the observed attack paths. Therefore, it is highly likely that attackers applying techniques aiming to achieve these failure modes are aiming to inflict additional impact to the network. Since the future impact cannot be known, a reasonable estimate can be reached by considering all possible impact elements in the estimation.

Finally, the value of the StC metric is calculated based on the considered failure mode. $StC = OCC$ for the persistence, defense evasion, and privilege escalation failure modes while $StC = ODC$ for the other failure modes.

$$Impact_{F,C} = (SF_F \times SC_C) + (FF_F \times FC_C) + (ICF_F \times IC_C) + (OF_F \times OC_C) + (StF_F \times StC_C) \quad (4)$$

The CCST should include scores for each component in the different operational modes. Some scores may not change across the different operational modes such as the IC, while others such as the I2CF, and I2MF, are more likely to change.

4.4.3 The Failure-Mode-Metric Table (FMMT). The FMMT is constructed specifying the metrics used to estimate the impact of each failure mode. A sample of the FMMT is depicted in Table 8. The FMMT reflects the mapping in the FMCT with additional information reflecting the metrics utilized to estimate the impact elements. For instance, the FMCT shown in Table 6 specifies that the Denial of Control failure mode is expected to cause only operational and safety consequences with impact factor of 1 for both; then the FMMT specifies that the operational and safety impacts are estimated using the I2CF and SC metrics, respectively.

Table 8. An FMMT Sample reflecting some failure modes and their proposed impact estimation metrics

Matrix	Failure Mode	OC	SC	IC	FC	StC
Mobile	Data Encrypted for Impact	OOI	SC	IC	FC	ODC
Mobile	Persistence					OCC
Mobile	Exploit SS7 to Track Device Location			LIC		
ICS	Denial of Control	I2CF	SC			
ICS	Denial of View	I2MF	SC		FC	ODC
ICS	Damage to Property	OOI	SC	IC	FC	
Enterprise	Privilege Escalation					OCC
Enterprise	Defense Evasion					OCC

4.4.4 Impact Calculation. Finally, the impact of the failure mode (F) for a component (C) is calculated using equation 4. The impact factors for the failure mode are retrieved from the FMCT. The metrics utilized to indicate the impact estimation for that failure mode are retrieved from the FMMT, while the impact scores for each metric for the component are retrieved from the CCST.

Considering the well-established relationship between the failure modes (i.e tactics) and failure mechanisms (i.e. techniques) in ATT&CK, the estimated impact of each failure mode is considered the same for all failure mechanisms that could cause it. For instance, the collection failure mode could be achieved using more than 17 failure mechanisms (e.g. Automated Collection and Data from Information Repositories). The impact of all of them for the same component is considered the same at the design stage. In future stages in the development life cycle, when a more detailed classification of the hosted information on each component is made available, more granular impact estimation for each failure mechanism would be possible.

4.5 Identify Failure Mechanisms

A failure mechanism is defined as a process that leads to failure [15]. In this paper, the security failure mechanisms are considered. In the remainder of this paper, we refer to failure mechanisms as cyber attacks or techniques interchangeably. The identification of relevant attacks during the risk assessment through the utilization of checklists, classifications, and taxonomies is considered a comprehensive approach, in addition to promoting a common understanding of risk and reducing the need for special expertise [16]. For these reasons, we relied on the ATT&CK framework as the approach for the identifying attacks.

Due to the heterogeneous nature of CPSs, the nature of cyber-attacks is expected to be different. Accordingly, we utilized the *Techniques* and *sub-techniques* in the multiple matrices of ATT&CK, namely, the Enterprise matrix for the IT components, the Mobile matrix for wireless components,

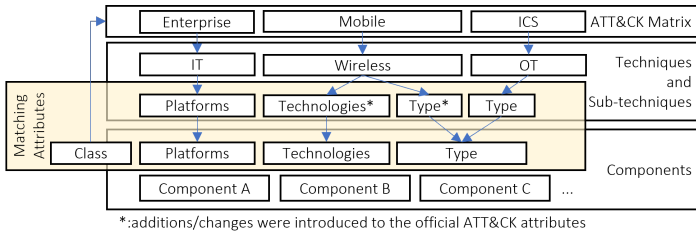


Fig. 3. Identification of relevant techniques per component from the ATT&CK matrices

and the ICS matrix for OT components. Certain components can be classified as a combination of multiple classifications, therefore the attack surface for such components is expected to be broader. The relevant attacks are derived from multiple relevant matrices. For instance, a data historian component is expected to be hosted in industrial control systems, such a system is classified as a dual-homed data historian in ATT&CK, which means that it is both an IT and OT component. This means that the data historian component can be susceptible to both IT-based attacks in the Enterprise matrix as well as to OT-based attacks in the ICS-matrix. The process for identifying relevant attacks for each component is highlighted in Figure 3.

Initially, the Techniques-Description Table (TDT) is constructed. All techniques and sub techniques from the relevant matrices are pulled from the official MITRE ATT&CK online repository [41]. The technique-specific attributes utilized from ATT&CK that are relevant in this step are the platform, for the enterprise techniques; and the type for the ICS techniques. The mobile matrix was developed mainly for mobile devices operating Android or IOS. We have studied the techniques in the mobile matrix and we argue that they can be applied to wireless components hosted in CPSs. To this end, we propose modifications to the mobile matrix to enable the description of attacks outside the scope of traditional mobile devices, by defining the "type" and "technologies" attributes. Additional attributes in the TDT which will be utilized in later steps such as the "Permission Required (PR)", "Kill Chain Phase (Tactics)", and others are retrieved from ATT&CK. Samples of the TDTs are depicted in Table 9.

Table 9. TDT Samples reflecting some techniques and their attributes

Enterprise TDT	Technique	Platform	Tactics	CVSS	PR	AC	UI	AV	
	Network Device CLI	Network	Execution	1.835	0.62	0.77	0.85	0.55	
ICS TDT	Technique	Type	Tactics	CVSS	PR	AC	UI	AV	
	Alarm Suppression	Field Controller/RTU/PLC/IED	Inhibit Response Function	2.221	0.85	0.44	0.85	0.85	
Mobile TDT	Technique	Type	Technologies	Tactics	CVSS	PR	AC	UI	AV
	Jamming or Denial of Service	Any	Cellular, Wi-Fi	Network Effects	3.887	0.85	0.77	0.85	0.85

Secondly, the Component-Description Table (CDT) is constructed from the architecture description. All components in the architecture are tagged with appropriate attributes that allow accurate matching with the relevant attacks. The component-specific attributes are the "Name", "Class", "Platform", "type", "MobileType", and "technologies". The "Class" attribute specifies the component classification (following the criteria specified in Table 2) to enable its matching with techniques from relevant matrices. The "MobileType" attribute specifies the type of mobile device ("Application-based", or not) while the "technologies" attribute specifies the attached technologies with the component (e.g. Wi-Fi, Cellular, Bluetooth, etc.). The "type" and "platform" attributes specify the type corresponding to the ICS asset classification and the platform corresponding to the enterprise platform attribute, respectively. A sample of the CDT is depicted in Table 10. We argue that identifying several failure mechanisms for each failure mode for each component would

support the efforts toward the proposition of risk reduction measures. Since the same failure mode could be triggered in several manners, each has a different mitigation method.

Table 10. A CDT Sample reflecting some components and their attributes

Name	Class	Type	Platform	MobileType	Technologies
Component A	OT/IT	Engineering Workstation	Linux	N/A	N/A
Component B	Mobile	Network	N/A	Non-App	Cell
Component C	OT/IT/Mobile	Network	Network	App-Based	Cell

4.6 Estimate the Likelihood of Failure Mechanisms

The likelihood estimation is proposed to be conducted by utilizing the exploitability score defined in the CVSS [30]. Several works have utilized CVSS during risk assessment to evaluate risks associated with threats rather than vulnerabilities, [38], [54], [10]. We argue that this approach is of great value for the security-by-design approach, since the implementation-level vulnerabilities are unknown during the system design, but the designer should at some time during the development life cycle consider the risk of such vulnerabilities and plan controls to mitigate the risk as early as possible to reduce the cost of remediation.

The base exploitability score in CVSS is calculated using four elements. A description of these elements and their values is depicted in Table 11. The official CVSS guidelines [22] describe the calculation of the exploitability score for vulnerabilities assuming some sort of online, physical, or logical access to the vulnerable component. In this paper, we have modified the description of the "Network" Attack Vector (AV) from the official CVSS guidelines [22] to enable the calculation of the exploitability score for off-line attacks existing in the ATT&CK framework, such as the supply chain compromise techniques, since such techniques could be performed way before the component is operational and no direct access to the component is required. Therefore, we propose that such group of attacks is assigned the highest AV value which corresponds to "Network" in the exploitability score. All the remaining descriptions are followed as the official CVSS guidelines suggest. The Scope metric proposed in the CVSS scheme is ignored since its effect is measured in the impact analysis conducted in this paper through the proposition of the "Staging" impact element. Finally, the likelihood value for each Failure Mechanism (FM) (i.e. attack) is calculated according to Equation 5. We decided to use the same equation for likelihood estimation as the one suggested for the exploitability score calculation in CVSS specified in [22], to make it compatible with this widely used approach to facilitate the analysis and the communication of results.

$$Likelihood_{FM} = 8.22 \times AV \times AC \times PR \times UI \quad (5)$$

Table 11. Exploitability elements, their values, and description.

Exploitability element	Metric	Value	Description
Attack Vector (AV)	Network	0.85	The attack can be carried out remotely and not bound to the local network, such as the internet. Also, if the attack does not require direct connectivity.
	Adjacent	0.62	The attack is bound to the network stack to logically adjacent topology. Such as local IP subnet, Bluetooth connection, or GNSS transmission.
	Local	0.55	The attack can be carried locally on the target component.
	Physical	0.2	The attack requires physical action upon the component.
Attack Complexity (AC)	Low	0.77	The attack requires a low level of combined skills and resources
	High	0.44	The attack requires a considerable level of skills and/or resources.
Privileges Required (PR)	None	0.85	The attack requires no authorization upon initialization to be successful.
	Low	0.62	The attack requires user-level privileges
	High	0.27	The attack requires high privileges(e.g. administrator)
User Interaction (UI)	None	0.85	No user interaction is needed to successfully launch the attack
	Required	0.62	The attack requires an action to be taken by the user.

$$RPN_{FM,C,F} = Likelihood_{FM} \times Impact_{F,C} \times Detectability_{FM,C,M} \quad (6)$$

The next step is to estimate the appropriate values of the exploitability elements for each attack in the ATT&CK framework and record the estimated values in the appropriator TDT (refer to Section 4.5). The number of analyzed attacks are 525, 86, and 81 in the enterprise, mobile, and ICS matrices respectively. This step has been performed by analyzing the descriptions of the attacks with the provided attributes. The analysis went as follows:

- If the attack has a CAPEC [45] attack pattern associated with it, the available attributes in the CAPEC page are retrieved. Some attack patterns have a description of the typical likelihood, resources, and skills required. The missing attributes were estimated based on the provided description. If the attack has more than a single CAPEC pattern associated with it, the maximum likelihood is considered and recorded in the TDT.
- Some attacks have a Common Vulnerabilities and Exposures (CVE) entry associated with them; in this case, the attributes were retrieved from the CVE page. The exploitability value of the CVE version 3.0 was retrieved when available, otherwise the version 2.0 value was retrieved; this lacks the User Interaction metric. Also, when more than one CVE entry is associated with the attack, the highest exploitability score is considered and recorded in the TDT.
- The values for the Privileges Required are provided in the ATT&CK framework techniques headers as attributes. Some attacks have several possible required privileges based on the possible mechanisms to launch the attack; the lowest possible privilege is considered and recorded in the TDT.
- The values for the Attack Vector were estimated using the description and utilizing the data sources attribute in the techniques headers. For instance, an attack that can be detected using a "packet capture" data source was assumed to have at least an Adjacent AV. Moreover, if the technique has an attribute "Remote Support: Yes" this means that it has a Network AV. The estimated value is then recorded in the TDT.
- The values for the Attack complexity were estimated using the description of each technique and then recorded in the TDT.

The final state of the TDT for each technology domain (i.e. matrix) including the estimated likelihood values are provided in our GitHub repository for this work ¹. We provided comments when possible to highlight the assumptions behind the estimate and/or the source providing the estimate. We consider this as another contribution of this paper. Since these tables are architecture-independent, they can be considered as encoded knowledge that can be utilized in future risk assessment tasks for a wide range of CPSs, to reduce the efforts and required skills.

4.7 Evaluate the Risks

The risk value is acquired through the calculation of a Risk Priority Number (RPN) as suggested in the FMECA standard [15]. The calculation of RPN for each failure mechanism (FM) against a component (C) resulting in a certain failure mode (F) is performed according to Equation 6. A qualitative rating can then be elicited based on the distribution of the risk values. The distribution of the likelihood and the detectability value is always between (0.12 - 3.89) and (0 - 1) respectively. On the other hand, the distribution of the impact value depends on the criteria chosen for the impact factor values in Equation 4.

¹https://github.com/ahmed-amro/APS-Communication_Architecture/tree/master/RPNMI

A tool has been developed in this work to aid the calculation of the RPNs for the attacks against CPS architectures and suggest relevant mitigation methods. The tool implements the RPN Calculation and Mitigation Identification (RPNMI) algorithm summarized in Algorithm 1. Initially, all the tables described previously should be constructed and made available as inputs in addition to a list with the operational modes. Then, for each component specified in the CDT the relevant attacks specified in the TDT are retrieved to populate an attack list specifying the list of attacks for each component (refer to Section 4.5). Then, the likelihood of each attack in the attack list is calculated from the TDT (refer to Section 4.6), its impact is calculated based on its associated tactic according to ATT&CK using the FMCT, CCST, and FMMT (refer to Section 4.4), its detectability is calculated using the FMT and CMT (refer to Section 4.3), its RPN is calculated using Equation 6, and its suggested mitigation methods are retrieved from the FMT. Finally, the tool produces all the components' attack lists in each operation mode with RPN and mitigation methods for each attack.

Algorithm 1 RPN Calculation and mitigation identification (RPNMI) algorithm

```

1: procedure RPNMI(OPModes, TDT, CDT, FMCT, CCST, FMMT, FMT, CMT)
2:   for each component in CDT do
3:     AttackList  $\leftarrow$  IdentifyRelevantAttacks(CDT, TDT)
4:     for each Operational Mode in OPModes do
5:       for each attack in AttackList do
6:         Likelihood  $\leftarrow$  CalculateAttackLikelihood(TDT)
7:         Impact  $\leftarrow$  CalculateAttackImpact(FMCT, CCST, FMMT)
8:         Detectability  $\leftarrow$  CalculateAttackDetectability(FMT, CMT)
9:         RPN  $\leftarrow$  Likelihood  $\times$  Impact  $\times$  Detectability
10:        MitigationList  $\leftarrow$  GetAttackMitigation(FMT)
11:       end for
12:     end for
13:   end for
14:   return AttackLists, RPNs and MitigationLists
15: end procedure

```

4.8 Propose Risk Reduction Measures

Finally, after the identification of risk values, the last step in the performing phase of a FMECA analysis is the proposition of risk reduction measures for each failure or failure mode. The ATT&CK framework provides a list of suggested mitigation and detection methods for each technique. Algorithm 1 produces a list of the mitigation methods for each identified attack against components with the risk information to facilitate later analysis to prioritize the integration of the mitigation methods into a security architecture.

5 RISK ASSESSMENT FOR AN AUTONOMOUS PASSENGER SHIP

In this section, we present the details of the tool-assisted application of the proposed approach for the APS use case. The main objective is to assess the risks of cyber threats against a communication architecture for an APS to aid the efforts in managing those risks through the development of a security architecture. The analysis aims to identify cyber risks considering scenarios with malicious intent causing failures in APS components. All scenarios are induced from the description of techniques and tactics in ATT&CK. Additionally, the criteria for the treatment of failure modes are based on the stakeholders' requirements. Safety and reliability are the main topics of concern. Additional topics of concern are the privacy of APS users, financial impact, and the security of the components and their communications. Afterwards, the analysis is performed, and the detailed steps for performing a FMECA utilizing ATT&CK are presented in the subsequent sections. Finally, a FMECA report is generated detailing the analysis process. This section constitutes a summarized

report of the conducted FMECA process and is intended to demonstrate the utility of the proposed approach.

5.1 Specify Components, Functions, and Performance Standard

The targeted components are inferred from the developed model of the communication architecture developed using Architecture Analysis and Design Language (AADL) [1]. The components addressed in this analysis are the functional components that include software and hardware elements. The components are classified as Information Technology (IT), Operational Technology (OT), Wireless, and/or a combination of multiple categories. The classification of components is conducted based on the criteria shown in Table 2. A brief description of each element is presented in Section 2.2 while a detailed discussion on them can be found in [8].

The proposed architecture supports several main functions, including autonomous, remote, and emergency navigation and control, in addition to internal, Ship-to-Shore, Ship-to-Ship, and emergency communication. Each component is involved in one or more system functions. A mapping between the system elements and the system functions has been provided using the goal tree success tree approach [52] the results of which are presented in the communication architecture definition in [8]. Moreover, the Operational Modes (OM) of the APS have been considered during the risk assessment process. The proposed architecture of the APS supports four operational modes, namely, Autonomous Execution (OM-AE), Autonomous Control (OM-AC), Remote Control (OM-RC), and fail-to-safe (OM-F2S). Each component is utilized in one or more operational modes. It has been identified that the overall APS system structure can be in one of two states (set of components and their connections); the first state operates in the three operational modes (OM-AE, OM-AC, and OM-RC) while the second state operates in the fourth operational mode (OM-F2S). This allowed for a more granular risk assessment.

Nevertheless, for the use case employed in this work, the results reflect no considerable difference in the risk values when considering the operational modes. However, we argue that for more advanced systems in which the components' interconnections could differ considerably across different operational modes, considering risk assessment with an operational mode perspective could reveal unexpected risk values.

The performance standard is based on the system security engineering definition of security failure; any violation of one of the established requirements and/or objectives of for the APS components constitutes a system security failure.

5.2 Identify Failure Modes

All the failure modes in ATT&CK (refer to Section 4.2) were considered relevant and have been considered, except two from the mobile matrix, namely, Generate Fraudulent Advertising Revenue, and Manipulate App Store Rankings or Ratings. All the other failure modes; should they occur, violate one or more of the stakeholders' concerns communicated as requirements and objectives in our earlier work [7].

5.3 Identify Detection Methods and Risk Reduction Measures

Considering that the system under analysis is still under development, no detection methods have yet been integrated. On the other hand, some controls have been proposed and included in the architecture description to satisfy previously established requirements. These are, Out-of-Band Communications Channel, Network Segmentation, and Redundancy of Service. Based on this the CMT (refer to Section 4.3) is constructed describing the coverage of the architectural components with regards to the mitigation methods.

5.4 Estimate the Impact of the Consequences of Failure Modes

The estimation of the impact values of failure modes for the APS architecture is conducted as follows:

5.4.1 The Failure-Mode-Consequences Table (FMCT). It is constructed for the APS use case considering the entire communication architecture as a System-of-Systems. The constructed FMCT is depicted in Table 6.

5.4.2 The Component-Criticality-Scores Table (CCST). It is constructed as follows:

- Safety and Financial criticality (SC and FC): safety and financial impact scores for each component were elicited from previously conducted Preliminary Hazard Analysis (PHA) for an APS use case [61].
- The different Operational Criticality (OC) scores, namely, the OOI, I2CF, and I2MF where calculated as described in Section 4.4. Three architecture views were developed using the ORA software. The OOI metric for each component is calculated using the combined centrality measures provided by the ORA software after modeling the entire APS network. The I2CF metric for each component is calculated in a similar manner, but only the components involved in the control functions (Autonomous, remote, and emergency control Section 5.1) were modeled. Finally, the I2MF metric for each component is similarly calculated, but only the components involved in the monitoring functions (Autonomous, remote, and emergency navigation Section 5.1) were modeled.
- The Information Criticality (IC) scores were estimated based on the communicated stakeholders' concerns. A specific requirement has been established to protect passengers' privacy from tracking and surveillance [7]. Also, concerns related to the preservation of intellectual property of processes hosted within the ship components in the Autoferry project [46] have been expressed. The estimation criteria for the IPC and DC metrics are shown in Table 12. The location information has been deemed to be of no impact (zero value) because the APS is utilized for passenger transportation in a fixed operational area.

Table 12. IPC and DC estimation criteria

Data Criticality (DC)	Description	Intellectual Property Criticality (IPC)	Description
None (0)	The component does not store or process sensitive passenger data (e.g. GNSS System)	None (0)	The component host processes with no intellectual property value
Low (0.33)	The component only forwards encoded sensitive passenger data (e.g. network device).	Low (0.33)	The component host processes with low intellectual property value (Common proprietary software) (e.g. network devices)
Medium (0.66)	The component performs the processing of sensitive passenger data (e.g. video camera).	Medium (0.66)	The component host processes with medium intellectual property value (Rare proprietary software)(e.g DP system)
High (0.99)	The component stores sensitive passenger data (e.g. data historian).	High (0.99)	The component host processes with high intellectual property value (Innovative proprietary software) (e.g. ANS)

- The Staging Criticality (StC) scores were calculated as follows:
 - The ODC scores for each component were calculated by the ORA software after modeling the entire APS network.
 - The OCC scores for each component were calculated using all the previously estimated criticality scores according to equation 3.

5.4.3 The Failure-Mode-Metric Table (FMMT). The FMMT is constructed reflecting the metrics that are needed to estimate each impact of each failure mode.

5.4.4 Impact Calculation. The final impact values for each failure mode of each component are calculated by means of equation 4, by utilizing the developed tool.

5.5 Identify Failure Mechanisms

The identification is conducted by utilizing the approach that identifies the relevant attacks for each component using attribute matching as described in Section 4.5. Initially, the TDT table is constructed by retrieving the techniques from the ATT&CK repository. Then the CDT is constructed by consulting the architecture description in [8]. Future work may attempt to perform automatic construction of the CDT table from a formal architecture description provided through an architecture description language such as AADL. Nevertheless, in this work, the CDT is manually constructed in a Comma Separated Value (CSV) format.

5.6 Estimate the Likelihood of Failure Mechanisms

The likelihood for each technique is calculated by means of equation 5, using the available information in the TDT. The result is added to the TDT in the "CVSS" column (refer to Table 9).

5.7 Evaluate the Risks

Afterwards, the tool calculates the RPN of all attacks relevant to all components using the RPNMI algorithm described in Section 4.7. Since the impact factors (refer to Section 4.4) are all chosen to be 1, a qualitative rating of the RPN can be calculated according to the following criteria: low risk rating (0 - 4.86), medium risk rating (4.87 - 9.72), high risk rating (9.73 - 14.58), and critical risk rating (14.59 - 19.44).

5.8 Propose Risk Reduction Measures

The tool additionally provides the mitigation methods suggested for each technique, based on the FMT table (refer to Section 4.3). Therefore, the suggested mitigation methods for each failure mechanism were identified and collected to analyze the most needed mitigation methods to support the effort in the development of a security architecture for the APS.

6 RESULTS AND EVALUATION

In this section, a summary of the results of the risk assessment process are presented to demonstrate the granular and comprehensive outcome of the proposed approach. Additionally, we the evaluation of the different elements of the approach.

After conducting a risk analysis of attacks against 39 different components in the APS architecture, we present an overview of the highest identified risks across the different failure modes, the most observed failure modes, failure mechanisms, and required mitigation methods. Concurrently, we discuss the utility of our approach regarding each outcome and the argued differences compared to other approaches. Then, an evaluation of the proposed metrics for estimating operational and staging impacts is presented.

Since no considerable difference has been identified in the risk values among the different operational modes, all the presented results are related to risks specifically identified for three operational modes, namely, OM-AE, OM-AC, and OM-RC; as they all maintain a unified system state, the risk values are the same among all of them.

The developed tool and the raw results can be found in our shared GitHub repository². Additionally, we have shared the populated tables discussed throughout the paper. These tables were utilized for the risk assessment of the APS communication architecture.

²https://github.com/ahmed-amro/APS-Communication_Architecture/tree/master/RPNMI

6.1 Overview

The comprehensive outcome of our proposed approach is demonstrated in Figure 4. The figure depicts the identified techniques with the highest risks across the different failure modes or tactics. Firstly, the utility of the inclusion of the different ATT&CK matrices is demonstrated through the identification of different risks from all of them. For instance, the manipulation of communication is a risk against wireless technology suggested in the mobile matrix in ATT&CK. In the APS, an expected implementation of the traffic module is an AIS. The result seems consistent with what is observed in the literature since AIS has been deemed susceptible to spoofing attacks by several works [12, 33]. However, in our approach, this risk is identified without the need for expert judgment as it is encoded common knowledge. Similarly, the suggested mitigation methods are drawn from the encoded common knowledge in ATT&CK and are in alignment with an observed direction for improving the security of AIS through encryption as suggested by Goudossis and Katsikas [24]. Also, the consideration of 16 failure mode categories improves the risk and countermeasures description. For instance, Auditing is a proposed countermeasure for the "Modify parameter" technique to impair process control. This granular description of the threat also suggests another descriptive scope for auditing, which is to include technologies and processes that allow the investigation of the modification of parameters sent to the DP controller component. Figure 4 also highlights the logical identification and estimation of risks across the different tactics. For instance, considering the digital logbook for the collection of information is very reasonable as it is the component with the highest information criticality: its function is to log and store information from most components, including passenger-related information. Still, the collection itself constitutes low risk since the collected information is still within the same component. Then, considering the Backup ANS for exfiltration through other network media is also very reasonable as it is a more connected component and is expected to include several communication technologies. Also, exfiltration does constitute a higher risk than collection since it also can affect other components if the exfiltration includes system credentials. Another example is the risks associated with the Sensor Processing Units (SPU). These components are proposed to aggregate the different sources of sensor data before forwarding them to the Autonomous Navigation System (ANS). The estimated risks of targeting this component through the execution of a modified program as well as denial of service are very reasonable. Since both would inhibit the monitoring functions of the APS and disable its ability to establish situational awareness which could lead to hazardous consequences (e.g. collision).

Overall, we argue that no other observed approach in the literature provides such a detailed level of risk identification, estimation, and proposition of countermeasures that can be utilized from the design stage. Also, the approach can be conducted in a semi-automated manner based on an updatable source such as ATT&CK, which results in reduced risk assessment effort.

6.2 Failure Modes

The risks associated with each failure mode have been analyzed. The results suggest that the most critical failure modes are related to adversaries aiming to inflict an impact, execute malicious software, remotely affect the APS services and inhibit the response functions of the APS. Additionally, exfiltration of sensitive passenger information or information with intellectual property as well as affecting the APS network constitute medium risks. The remaining failure modes constitute only low risks according to the followed risk management strategy.

The risks of some techniques such as Defense Evasion, Credential Access, Discovery, etc. have been estimated "Low" due to the risk management strategy followed in the estimation of the impact of failure modes which is captured in the FMCT (refer to Sec 4.4). The strategy considers all the elements of impact as equals; this rendered the impact of these failure modes as low as they only

Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework

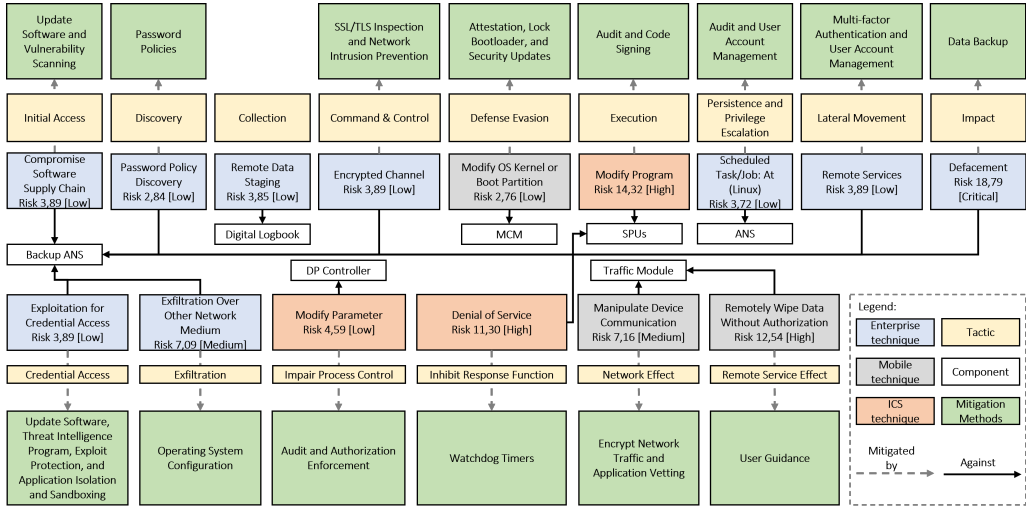


Fig. 4. An overview of the identified highest risk techniques for each tactic, their calculated risk, targeted component, and suggested mitigation method

affect the staging impact element (See Table 6). A different security-focused risk strategy that increases the value of the staging factor (i.e. StF) in the impact estimation could have been adopted; such a strategy would generate different results and it is to be expected in future work.

Compared to other works, we argue that our risk assessment methodology provides a granular description of failure modes. Other methods provide a comparatively less meaningful description of the attackers' objectives in the assessed system. For instance, the persistence, command and control, and defense evasion failure modes are not straightforwardly mapped to the STRIDE threat categories or the CIA objectives. Considering the popularity of such failure modes in the current threat landscape, a methodology that addresses them is required. Moreover, the inclusion of failure modes from the different technology domains (i.e. ATT&CK matrices) provides meaningful context to the failure modes. For instance, the high risk identified due to the remote service effects failure mode highlights the risks associated with the inclusion of wireless technology while the inhibit response function highlights risks due to the inclusion of OT for monitoring functions. We argue that this provides an improved threat communication feature of our approach.

6.3 Failure Mechanisms

The Failure Mechanisms (i.e. techniques) with estimated critical and high risks belong to the "impact" failure mode. Regarding critical risks, an attacker could severely impact the entire APS and its operational area, possibly leading to damage and life-threatening hazard against passengers through the ANS and Backup ANS that are responsible for the navigation functions. Damage could occur through several forms, an example would be similar to the Polish teen incident in which he derailed the city tram system [56], an attacker controlling the main or backup ANS systems could enforce unsafe routes and/or drop the reporting of warnings to the operator to avoid intervention. Moreover, a surprising risk was identified: the possibility to inflict impact through a defacement attack. Defacement attacks usually target web applications by modifying the distributed content [42]. A possible implementation of the control and monitoring functions could be through web services [39, 49]. Therefore, a defacement technique could manipulate or impair the control and monitoring

functions to a severe degree. Regarding high risks, a group of attacks has been identified, that aim to inflict impact through a group of denial of service techniques, namely, reflection amplification, direct network flood, service and OS exhalation flood, or endpoint and network denial of service. Additionally, other attacks could inflict impact through manipulation of transmitted data, scheduled execution, scripting, and project file infection. Two surprising techniques that are not common against ships are "Resource Hijacking", and "Remotely Wipe Data Without Authorization". Resource hijacking is a widely observed technique that attackers carry out to exploit system resources to validate transactions related to cryptocurrency networks [43]. Such a technique could impair the target system by reducing its performance, the effect could be amplified if the component is involved in time-critical functions which is the case for some of the APS components such as the Automatic Navigation System (ANS). Also, adversaries could Remotely Wipe Data Without Authorization for components involved in the control and monitoring functions; this attack could be in the form of ransomware. The Mobile Communication Module (MCM) components are expected to be routers that are not immune against ransomware attacks [47], while the traffic modules are expected to be Automatic Identification Systems (AIS). Until the time of writing this paper, no ransomware attack has been found to attack this specific implementation solution but, at the design stage, the implementation solution is unknown, and therefore, such an attack could be of relevance and therefore should be considered.

Other attacks observed in the maritime domain such as jamming, data encryption for impact (i.e. ransomware), and exfiltration have been found to have a medium impact. These results were not surprising, for different reasons. Jamming attacks have been considered since the system concept definition phase in the Autoferry project; therefore, design solutions were introduced to mitigate their effect through redundant functional components such as other sensors. Thus, the operational effect of the jamming attack is reduced and this is reflected in the risk value. Furthermore, network segmentation is one of the mitigation methods against ransomware and exfiltration attacks. The APS network has been designed with network segmentation for most of the components. This affects the *Detectability* value of the risks against them and thus reduces their risk values. We argue that these results reflect the accuracy of our proposed process.

6.4 Mitigation methods

An outcome of our proposed approach is a list of the most needed mitigation methods, drawn from the suggested mitigation methods by ATT&CK for the identified risks. The mitigation methods against critical risks include data backup, mechanical protection layers, safety, instrumented systems, network allow lists, out-of-band communication channels, and redundancy of service. These controls are considered to be prioritized during the security architecture design. Two special categories of mitigation methods should receive additional focus, namely, "Mitigation Limited or Not Effective" and "Do Not Mitigate". ATT&CK classifies the mitigation method for certain techniques as difficult to mitigate since they are based on the abuse of system features; yet, some of them have proposed suggestions for detection that should be considered in the security architecture design. The identified techniques with high and medium risks that belong to such category include Resource Hijacking, Account Access Removal, Automated Exfiltration, Jamming or Denial of Service, and System Shutdown or Reboot. Therefore, future efforts will be dedicated to suggesting mitigation methods for these techniques in the APS and integrating them within the security architecture. Moreover, two defense evasion techniques, namely Execution Guardrails, and Environmental Keying are proposed not to be mitigated, since the mitigation methods could lead to increasing the risk of compromise.

6.5 Evaluation for the proposed metrics

In this section, we present the results of the analysis conducted to evaluate the proposed staging and operational impact metrics. During the discussion in the coming sections, we will utilize the values in Table 13, which depicts a snapshot from the CCST (Section 4.4) holding the values of the impact metrics for the AEMC, ANS, Backup ANS, and GNSS components.

Table 13. Snapshot from CCST reflecting the criticality scores of the highlighted components

Op-Mode	Component	OC			SC	IC	FC	StC	
		OOI	I2CF	I2MF				ODC	OCC
OM-AE,	AEMC	0.95	1	0	1	0.49	1	0.88	0.86
OM-AC,	ANS	1	0.27	1	1	0.82	1	0.77	0.95
and	Backup ANS	0.99	0.074	0.99	1	0.82	1	1	0.95
OM-RC	GNSS IMU	0.45	0	0.56	1	0.16	1	0.11	0.65

6.5.1 The granularity of operational impact estimation. In this section, we highlight the results of our proposed application of the different operational criticality metrics, namely the Overall Operational Impact (OOI), Impact to the Control Functions (I2CF), and Impact to the Monitoring Functions (I2MF) (refer to Section 4.4). To demonstrate the effect of the application of these metrics on the risk estimation, Table 14 depicts the utilized metrics in the calculation of the impact score for three different failure modes against the ANS and AEMC components. According to Table 6 The "Impair Process Control" failure mode has positive SF, FF, and OF, which means that it is only expected to cause incidents with safety, financial and operational impacts. The safety criticality (SC), and financial criticality (FC) are all estimated using the same metrics for the three failure modes. The operational criticality (OC) on the other hand, can be estimated using either the OOI or the I2CF. Considering the two components, the impact of this failure mode when using the OOI metric has a negligible difference: 2.95, and 3 for the AEMC and the ANS respectively. But, when using the I2CF the difference is noticeable: 3 and 2.27 for the AEMC and the ANS respectively. Since the AEMC is heavily involved in the control functions while the ANS has less involvement, we argue that the I2CF metric reflects a more reasonable estimate of the operational impact than the OOI metric for this failure mode and similar ones involved in the control functions. The other failure mode, namely the "Manipulation of View", has positive SC, FF, OC, and StC, meaning it can cause incidents with safety, financial, operational, and staging impacts. The OC metric can be estimated using either the OOI or the I2MF metrics. The difference in the impact values is negligible when using the OOI metric: 3.84, and 3.78 for the AEMC and the ANS respectively. But, when using the I2MF metric, the difference is noticeable: 2.89, and 3.78 for the AEMC and the ANS respectively. Since the ANS is heavily involved in the monitoring functions while the AEMC has much less involvement, we argue that the I2MF metric accounts for a more reasonable estimate of the operational impact than the OOI metric for this failure mode and similar ones involved in the monitoring functions. Finally, the "Resource Hijacking" failure mode could impact the entire system functions certainly not only the control and monitoring functions. Considering that ANS and AEMC are centralized components involved in several functions other than monitoring and control and both have very similar combined centrality measures (0.95 and 1 for the AEMC and ANS respectively), the OOI metric captures a reasonable estimate of the operational impact for this failure mode should it occur for any of these components.

The majority of observed risk assessment methods are qualitative as they utilize expert judgment for the estimation of the operational impact. This increases the required effort for conducting risk assessment and subjugates the assessment to bias. However, our proposed metrics reduce these shortcomings by relying on a graph-based model of the system for providing a more granular quantitative estimate of the operational impact.

Table 14. Estimation of failure mode impact using different OC metrics

Failure Mode	OC Metric	Componet	OC			SC	IC	FC	StC ODC	Impact Value
			OOI	I2CF	I2MF					
Impair Process Control	OOI	AEMC	0.950538225	-	-	1	-	1	2.950538225	
	I2CF		-	1	-				-	
	OOI		1	-	-				3	
	I2CF	ANS	-	0.272978267	-				2.272978267	
	OOI		0.950538225	-	-				0.8888889	3.839427125
Manipulation of View	I2MF	AEMC	-	-	0	-	2.8888889			
	OOI	ANS	1	-	-	0.7777778	3.7777778			
	I2MF		-	-	1	-	3.7777778			
	OOI		AEMC	0.950538225	-	-	0.8888889	3.839427125		
Resource Hijacking	OOI	ANS	1	-	-	0.7777778	3.7777778			

6.5.2 The granularity of staging impact estimation. As discussed in Section 4.4, the staging impact element estimates the ability of the attacker to stage future attacks which are mainly influenced by the position and criticality of the attacked component in the system network. Table 15 shows the estimates of the impact value of the group of failure modes that do not have any other impact than the staging impact. Also, an example of a failure mechanism for each failure mode is presented. The Backup ANS component is among the most connected components in the network, having the highest Outbound Degree Centrality (ODC) measure. This provides attackers with several options for traversing the network for staging other attacks. Moreover, it is a critical component, having among the highest Overall Component Criticality (OCC). Failing to eliminate persistence, defense evasion and privilege escalation failure modes on this specific component could initiate critical future risks, thus the staging impact is higher. On the other hand, the GNSS IMU system is much less connected and has among the lowest ODC. This limits the attacker's ability to traverse the network. Also, its OCC measure is estimated to be less than that of the Backup ANS, as it is less involved in the overall operations and hosts less critical information. Therefore, its staging impact estimates for the failure mechanisms shown in Table 15 are also less. The results suggest that the ODC and OCC metric provides reasonable estimates of the proposed staging impact.

We argue that other risk assessment methods observed in the literature might overlook the impact of certain failure mechanisms in ATT&CK. For instance, using the legitimate VNC software for lateral movement is not expected to have any safety, financial, privacy, or operational impact on the target component, nor does it inflict an immediate impact on confidentiality, integrity, or availability. However, it aids attackers during the staging of cyber attacks and our approach provides a granular estimation of the impact of such activities. We argue that this impact element is of critical value to the cybersecurity posture as it aids the identification of the most critical risks related to the ability of adversaries to stage attacks.

7 LIMITATIONS, DISCUSSION AND FUTURE WORK

Below, limitations in the proposed approach are discussed with possible improvements to be addressed in future work:

- Traditional FMECA only enables the identification of single failure modes [16]. Nevertheless, the relationships between different failure modes are communicated through the kill chain concept embedded in ATT&CK. The latest version of ATT&CK provides detailed information regarding software (i.e. malware) and threat groups employing the different tactics and techniques. Specifically, 638 software and 129 threat groups are present in the enterprise and mobile matrices in addition to 19 software and 9 groups in the ICS matrix. This information is expected to be utilized as models for propagating threats in the system under analysis. Additionally, the expected paths (i.e. links) in the network utilizing graph theory are also planned to be employed to achieve comprehensive coverage of threat propagation paths.

Table 15. Estimation of failure modes impact using the StC metrics

Failure Mechanism	Description	Failure Mode	StC Metric	OC	SC	IC	FC	Backup ANS	GNSS IMU
								Impact Value	Impact Value
VNC	Attackers may use this remote access software to access other components in the network	Lateral Movement	ODC					1	0.111
DNS	Attackers may communicate using DNS protocol to avoid detection	Command and Control							
Drive-by Compromise	Attackers may obtain initial access using a downloaded malicious payload (e.g. driver)	Initial Access							
ARP Cache Poisoning	Attackers may use this technique to collect and/or relay data such as credential	Credential Access							
Remote System Discovery	Attackers may discover connected systems in the network	Discovery	OCC					0.956	0.654
Valid Accounts	Attackers may create new accounts or use existing ones to keep a foothold in the network	Persistence							
Obscured Files or Information	Attackers may alter files in a manner to make them hard to discover	Defense Evasion							
At (Linux)	Attackers may exploit this scheduling tool to run a process using the privilege of a specified account	Privilege Escalation							

However, the correlation between the different ATT&CK techniques across the different kill chain phases in addition to the suitable methods for estimating the collective likelihood and impact values are yet unresolved issues. Therefore, future work will focus on the correlation between attacks, causing different failure modes to generate attack scenarios composed of coherent steps similar to the concept of attack trees.

- The Checklist risk identification approach is said to lack the ability to identify new attacks [16]. We argue that the comprehensive nature of the tactics and techniques in ATT&CK reduces the effect of this limitation.
- Some components might be covered by multiple mitigation methods. In this paper, for simplicity, the detectability value is estimated based on whether a component is covered by (at least) one mitigation method or not. Future work can investigate how this value is affected when multiple mitigation methods contribute to the coverage.
- We relied on the literature for choosing the ODC to estimate the staging impact. Nevertheless, we have considered other centrality measures, such as the Authority Centrality to estimate the staging impact. However, it is outside the scope of this paper to compare the utility of different centrality measures. In future work, a comparative study could be conducted to do so.
- The mapping between failure mode and consequence reflected in the FMCT (refer to Section 4.4) is constructed after manual analysis of the description of each failure mode in ATT&CK as such, it is subject to bias and therefore should be reconstructed for other use cases with considerations for reducing biased judgment. The IEC 31010 standard [16] provides guidelines for eliciting stakeholders' and experts' views while reducing bias.
- The ODC metric in the staging impact estimation may overlook the fact that in some attacks, the attacker only requires a single point of access to stage future attacks (Low ODC value). However, we argue that the higher the possible points of access from a component to other components, the higher the impact value contributing to the risk.
- The proposed metrics for estimating safety and financial impacts require a prior PHA/HAZOP or similar analysis. Future work may attempt to provide more granular quantitative estimates induced from the architecture description.
- A comparative analysis of our proposed approach is suggested for future work. This includes the engagement of independent experts to isolate and prevent biases introduced by the authors, as well as considering several use cases that could help quantify performance limitations across the various methods.

7.1 Approach Adaptability

Our proposed approach can be applied in different CPS use cases. Also, it is capable of assessing new risks matching the up-to-date threat landscape due to its reliance on the ATT&CK framework. Figure 5 depicts a flowchart for applying the approach at different periods of time, in different use cases, or when the same use case is updated or modified.

When the approach is to be applied against a different system or a modified version of the same system. The system components are classified according to the criteria in Table 2. If some components cannot be classified (e.g. docker containers), the risks associated with them will not be assessed. Then, the relevant failure modes from Table 3 are identified. Then, the CMT is updated to map the relationships between the existing controls and the system components. However, the controls are limited to those in the ATT&CK framework. Therefore, use cases with some controls that do not exist in the ATT&CK framework (e.g. Email Protection) will suffer inaccurate results. Afterward, the FMCT is updated to map the relationships between the failure modes and the consequences according to the defined impact model. Consequently, the FMCT is updated to define the required metrics for the entries in the FMCT. After that, the CCST is updated to specify the criticality values of the components in the system. This is done by estimating the safety, financial, operational, information, and staging criticality metrics using their appropriate approaches discussed in Section 4.4.2. Later, the CDT is updated with the components properties, namely, classification, type, platform, mobile type, and technologies. Then, the risk threshold needs to be defined.

When the a new version or an update to ATT&CK is released. This can be done by updating the TDT by fetching the techniques' information from the ATT&CK online repository and update their CVSS metrics. Then, updating the FMT by fetching the information of the techniques' mitigation methods from the online repository and updating the effectiveness estimation. Later, the FMCT is updated after identifying the new failure modes and defining their expected consequences according to the defined impact model. Consequently, the FMCT is updated to define the required metrics for the new entries in the FMCT.

When attacks techniques and defenses evolve effecting the CVSS or effectiveness estimations in the TDT or CMT respectively. For instance, a new released exploit for an attack technique with different attack complexity. This changes the respective CVSS score of that attack.

Finally, when all the aforementioned conditions are considered and processed, the RPNMI algorithm can be launched to generate updated results.

8 CONCLUSION

A semi-quantitative risk assessment approach is proposed in this paper following a Failure Modes Effects and Criticality Analysis (FMECA) and utilizing the ATT&CK framework. This approach provides a comprehensive risk assessment while reducing the need for expert judgment. Additionally, the approach addresses the heterogeneous nature of CPSs and provides attack descriptions that are relevant for different categories of components. Further, the approach, in addition to identifying the required mitigation methods, can identify areas of concern which the system under analysis can be susceptible to and only limited mitigation methods are yet available. Moreover, the approach allows for the updatability of the risk values through updating input values to reflect the current threat landscape.

The proposed impact estimation metrics are demonstrated to provide a reasonable estimate of the different impact elements, namely operational, information criticality, and security-related impact. Additional efforts are required to provide metrics that are capable of estimating safety and financial impacts.

- [8] Ahmed Amro, Vasileios Gkioulos, and Sokratis Katsikas. 2021. Communication architecture for autonomous passenger ship. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* (2021), 1748006X211002546.
- [9] Arben Asllani, Alireza Lari, and Nasim Lari. 2018. Strengthening information technology security through the failure modes and effects analysis approach. *International Journal of Quality Innovation* 4, 1 (2018), 5.
- [10] Ali Behfarnia and Ali Eslami. 2018. Risk assessment of autonomous vehicles using bayesian defense graphs. In *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. IEEE, 1–5.
- [11] Felix Bieker, Michael Friedewald, Marit Hansen, Hannah Obersteller, and Martin Rost. 2016. A process for data protection impact assessment under the european general data protection regulation. In *Annual Privacy Forum*. Springer, 21–37.
- [12] Victor Bolbot, Gerasimos Theotokatos, Evangelos Boulougouris, and Dracos Vassalos. 2020. A novel cyber-risk assessment method for ship systems. *Safety Science* 131 (2020), 104908.
- [13] Kathleen M Carley and Jürgen Pfeffer. 2012. Dynamic network analysis (DNA) and ORA. *Advances in Design for Cross-Cultural Activities Part I* (2012), 265–274.
- [14] Roger Clarke. 2009. Privacy impact assessment: Its origins and development. *Computer law & security review* 25, 2 (2009), 123–135.
- [15] IEC 60812 Technical Committee et al. 2018. Analysis techniques for system reliability-procedure for failure mode and effects analysis (FMEA). (2018).
- [16] IEC 31010 Technical Committee et al. 2019. 31010: Risk management–Risk assessment techniques. (2019).
- [17] SAE J3061 Vehicle Cybersecurity Systems Engineering Committee et al. 2016. Cybersecurity Guidebook for Cyber-Physical Vehicle Systems. *SAE International* (2016).
- [18] The Maritime Safety Committee. [n.d.]. INTERIM GUIDELINES ON MARITIME CYBER RISK MANAGEMENT (MSC-FAL.1/Circ.3/Rev.1). <https://cutt.ly/6R8wqjN>.
- [19] Mirko Čorić, Anita Gudelj, Zvonimir Lušić, and Sadko Mandžuka. 2019. E-Navigation Architecture Overview and Functional Connection Analysis. *NAŠE MORE: znanstveno-stručni časopis za more i pomorstvo* 66, 3 (2019), 120–129.
- [20] Manlio De Domenico, Albert Solé-Ribalta, Elisa Omodei, Sergio Gómez, and Alex Arenas. 2015. Ranking in interconnected multilayer networks reveals versatile nodes. *Nature communications* 6, 1 (2015), 1–6.
- [21] Okan Duru. [n.d.]. The Future Shipping Company: Autonomous Shipping Fleet Operators. <https://bit.ly/MaritimeFuture>
- [22] FIRST. 2019. Common Vulnerability Scoring System version 3.1: Specification Document. (2019).
- [23] International Organization for Standardization. 2018. *Information Technology. Security Techniques. Information Security Risk Management: ISO/IEC 27005: 2018*. International Organization for Standardization.
- [24] Athanassios Goudossis and Sokratis K Katsikas. 2019. Towards a secure automatic identification system (AIS). *Journal of Marine Science and Technology* 24, 2 (2019), 410–423.
- [25] Andy Greenberg. [n.d.]. The Untold Story of NotPetya, the Most Devastating Cyberattack in History. <https://bit.ly/MaerskAttack>
- [26] INSIKT GROUP. 2020 (accessed December 2, 2020). *Defense Evasion Dominant in Top MITRE ATTCK Tactics of 2019*. <https://www.recordedfuture.com/mitre-attack-tactics/>.
- [27] JP Gupta and B Suresh Babu. 1999. A new hazardous waste index. *Journal of hazardous materials* 67, 1 (1999), 1–7.
- [28] David Hambling. 2017. Ships fooled in GPS spoofing attack suggest Russian cyberweapon. <https://bit.ly/GPSAttack>
- [29] Gina Havdal, Christina Torjussen Heggelund, and Charlotte Hjelmseth Larssen. 2017. *Design of a Small Autonomous Passenger Ferry*. Master’s thesis. NTNU.
- [30] Siv Hilde Houmb, Virginia NL Franqueira, and Erlend A Engum. 2010. Quantifying security risk level from CVSS estimates of frequency and impact. *Journal of Systems and Software* 83, 9 (2010), 1622–1634.
- [31] Mafijul Md Islam, Aljoscha Lautenbach, Christian Sandberg, and Tomas Olovsson. 2016. A risk assessment framework for automotive embedded systems. In *Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security*. 3–14.
- [32] Bojan Jelacic, Daniela Rosic, Imre Lendak, Marina Stanojevic, and Sebastijan Stoja. 2017. STRIDE to a secure smart grid in a hybrid cloud. In *Computer Security*. Springer, 77–90.
- [33] Georgios Kavallieratos, Sokratis Katsikas, and Vasileios Gkioulos. 2018. Cyber-attacks against the autonomous ship. In *Computer Security*. Springer, 20–36.
- [34] Athar Khodabakhsh, Sule Yildirim Yayilgan, Mohamed Abomhara, Maren Istad, and Nargis Hurzuk. 2020. Cyber-risk identification for a digital substation. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*. 1–7.
- [35] Carlos León. 2013. Authority centrality and hub centrality as metrics of systemic importance of financial market infrastructures. *Available at SSRN 2290271* (2013).
- [36] Eric Luijff, Albert Nieuwenhuijs, Marieke Klaver, Michel van Eeten, and Edite Cruz. 2008. Empirical findings on critical infrastructure dependencies in Europe. In *International Workshop on Critical Information Infrastructures Security*.

- Springer, 302–310.
- [37] Georg Macher, Eric Armengaud, Eugen Brenner, and Christian Kreiner. 2016. Threat and risk assessment methodologies in the automotive domain. *Procedia computer science* 83 (2016), 1288–1294.
- [38] Bharat B Madan, Manoj Banik, and Doina Bein. 2019. Securing unmanned autonomous systems from cyber threats. *The Journal of Defense Modeling and Simulation* 16, 2 (2019), 119–136.
- [39] Markus Mathes, Christoph Stoidner, Steffen Heinzl, and Bernd Freisleben. 2009. SOAP4PLC: Web services for programmable logic controllers. In *2009 17th Euromicro International Conference on Parallel, Distributed and Network-based Processing*. IEEE, 210–219.
- [40] Ioan-Cosmin Mihai, Stefan Pruna, and Ionut-Daniel Barbu. 2014. Cyber kill chain analysis. *Int'l J. Info. Sec. & Cybercrime* 3 (2014), 37.
- [41] MITRE. 2020 (accessed December 2, 2020). *Cyber threat intelligence repository expressed in stix 2.0*. <https://github.com/mitre/cti>.
- [42] MITRE. 2020 (accessed December 8, 2020). *Defacement Technique*. <https://attack.mitre.org/techniques/T1491/>.
- [43] MITRE. 2020 (accessed December 8, 2020). *Resource Hijacking*. <https://attack.mitre.org/techniques/T1496/>.
- [44] Jean-Philippe Monteuiis, Aymen Boudguiga, Jun Zhang, Houda Labiod, Alain Servel, and Pascal Urien. 2018. Sara: Security automotive risk analysis method. In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security*. 3–14.
- [45] Thomas L Nielsen, Jens Abildskov, Peter M Harper, Irene Papaconomou, and Rafiqul Gani. 2001. The CAPEC database. *Journal of Chemical & Engineering Data* 46, 5 (2001), 1041–1044.
- [46] NTNU Autoferry. 2018. Autoferry - Autonomous all-electric passenger ferries for urban water transport. <https://www.ntnu.edu/autoferry>
- [47] Emmanuel Olaniyi. 2020 (accessed December 8, 2020). *Ransomware and Routers - Spectranet and Smile Ransomware Hit*. <https://www.cybersecfill.com/ransomware-and-spectranet/>.
- [48] Charlie Osborne. 2020 (accessed December 2, 2020). *Code execution, defense evasion are top tactics used in critical attacks against corporate endpoints*. <https://bit.ly/zdnet-DefenseEvasion>.
- [49] Punnuluk Phaitoonbuathong, Radmehr Monfared, Thomas Kirkham, Robert Harrison, and Andrew West. 2010. Web services-based automation for the control and monitoring of production systems. *International Journal of Computer Integrated Manufacturing* 23, 2 (2010), 126–145.
- [50] Paul Pols and Jan van den Berg. 2017. The Unified Kill Chain. *CSA Thesis, Hague* (2017), 1–104.
- [51] Ron Ross, Michael McEvilly, and Janet Oren. 2016. *Systems security engineering: Considerations for a multidisciplinary approach in the engineering of trustworthy secure systems*. Technical Report. National Institute of Standards and Technology.
- [52] G. Sabaliauskaite and S. Adepu. 2017. Integrating six-step model with information flow diagrams for comprehensive analysis of cyber-physical system safety and security. *Proceedings of IEEE Int. Symposium on High Assurance Systems Engineering* (2017), 41–48. <https://doi.org/10.1109/HASE.2017.25>
- [53] Tara Seals and Tara Seals. [n.d.]. Researcher: Not Hard for a Hacker to Capsize a Ship at Sea. <https://threatpost.com/hacker-capsize-ship-sea/142077/>
- [54] Barry Sheehan, Finbarr Murphy, Martin Mullins, and Cian Ryan. 2019. Connected and autonomous vehicles: A cyber-risk classification framework. *Transportation research part A: policy and practice* 124 (2019), 523–536.
- [55] A. Shostack. 2014. *Threat Modeling: Designing for Security*. Vol. Wiley Publishing.
- [56] Shelley Smith. 2020 (accessed December 8, 2020). *Teen Hacker in Poland Plays Trains and Derails City Tram System*. https://inhomeandsecurity.com/teen_hacker_in_poland_plays_tr/.
- [57] George Stergiopoulos, Marianthi Theocharidou, Panayiotis Kotzanikolaou, and Dimitris Gritzalis. 2015. Using centrality measures in dependency risk graphs for efficient risk mitigation. In *International Conference on Critical Infrastructure Protection*. Springer, 299–314.
- [58] Blake E Strom, Andy Applebaum, Doug P Miller, Kathryn C Nickels, Adam G Pennington, and Cody B Thomas. 2018. Mitre att&ck: Design and philosophy. *Technical report* (2018).
- [59] Kimberly Tam and Kevin Jones. 2018. Cyber-risk assessment for autonomous ships. In *2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*. IEEE, 1–8.
- [60] Kimberly Tam and Kevin Jones. 2019. MaCRA: A model-based framework for maritime cyber-risk assessment. *WMU Journal of Maritime Affairs* 18, 1 (2019), 129–163.
- [61] Christoph A Thieme, Chuanqi Guo, Ingrid B Utne, and Stein Haugen. 2019. Preliminary hazard analysis of a small harbor passenger ferry—results, challenges and further work. In *Journal of Physics: Conference Series*, Vol. 1357. IOP Publishing, 012024.
- [62] Douglas Brent West et al. 1996. *Introduction to graph theory*. Vol. 2. Prentice hall Upper Saddle River, NJ.
- [63] Davey Winder. 2019. U.S. Coast Guard Issues Alert After Ship Heading Into Port Of New York Hit By Cyberattack. <https://bit.ly/USShipAttack>

ACRONYMS

- AADL** Architecture Analysis and Design Language. 21, 23
- AC** Attack Complexity. 17, 18
- AEMC** Autonomous Engine Monitoring and Control. 4, 6, 27, 28
- AIS** Automatic Identification System. 24
- ANS** Autonomous Navigation System. 4, 6, 22, 24–29
- APS** Autonomous Passenger Ship. 1–4, 6, 7, 20–26, 31
- ASC** Autonomous Ship Controller. 6
- ATT&CK** Adversarial Tactics, Techniques, and Common Knowledge. 1–12, 14–21, 23–26, 28–30
- AV** Attack Vector. 17–19
-
- CCST** Component-Criticality-Scores Table. 12, 16, 20, 22, 27, 30
- CDT** Component Description Table. 17, 18, 20, 23, 30
- CIA** Confidentiality, Integrity, and Availability. 8, 10, 25
- CMT** Component-Mitigations Table. 11, 20, 21, 30
- CPS** Cyber-Physical Systems. 1–4, 8, 9, 11–14, 16, 17, 19, 20, 30
- CVE** Common Vulnerabilities and Exposures. 19
- CVSS** Common Vulnerability Scoring System. 3–5, 8, 17, 18, 23, 30
-
- DC** Data Criticality. 14, 15, 22
- DP** Dynamic Positioning. 4, 22
-
- ECT** Emergency Control Team. 4, 6
-
- FC** Financial Criticality. 12, 13, 15, 16, 22, 27–29
- FF** Financial Factor. 12, 16, 27
- FMCT** Failure-Mode-Consequences Table. 12, 13, 16, 20, 22, 24, 29, 30
- FMECA** Failure Modes Effects and Criticality Analysis. 1, 2, 4, 6, 7, 9, 10, 19–21, 28, 30
- FMCT** Failure-Mode-Metric Table. 16, 20, 22, 30
- FMT** Failure-Mitigation Table. 10, 11, 20, 23, 30
-
- GNSS** Global Navigation Satellite System. 4, 22, 27–29
-
- HAZOP** Hazard and Operability. 4, 12, 29
-
- I2CF** Impact to the control functions. 14, 16, 22, 27, 28
- I2MF** Impact to the monitoring functions. 14, 16, 22, 27, 28
- IC** Information criticality. 14–16, 22, 27–29
- ICF** Information Criticality Factor. 12, 16
- ICS** Industrial Control Systems. 3, 7, 8, 11–13, 16, 17, 19
- IMO** International Maritime Organization. 2, 3
- IPC** Intellectual Property Criticality. 14, 15, 22
- IT** Information Technology. 3, 7, 9, 10, 13, 16–18, 21
-
- LIC** Location Information Criticality. 14–16
-
- MCM** Mobile Communication Module. 6, 26
-
- NVD** National Vulnerability Database. 3, 5
-
- OC** Operational Criticality. 13–16, 22, 27–29

- OCC** Overall Component Criticality. 15, 16, 22, 27–29
- ODC** Outbound Degree Centrality. 7, 15, 16, 22, 27–29
- OF** Operational Factor. 12, 16, 27
- OM-AC** Autonomous Control Operational Mode. 21, 23, 27
- OM-AE** Autonomous Execution Operational Mode. 21, 23, 27
- OM-F2S** Fail-to-Safe Operational Mode. 21
- OM-RC** Remote Control Operational Mode. 21, 23, 27
- OOI** Overall operational impact. 14, 16, 22, 27, 28
- OT** Operational Technology. 3, 7, 9, 10, 17, 18, 21, 25

- PHA** Preliminary Hazard Analysis. 4, 12, 22, 29
- PR** Privileges Required. 17, 18

- RCC** Remote Control Centre. 4, 6
- REMS** Remote Engine Monitoring and Control System. 6
- RNS** Remote Navigation System. 6
- RPN** Risk Priority Number. 19, 20, 23, 31
- RPNMI** RPN Calculation and Mitigation Identification. 20, 23, 30
- RSC** Remote Ship controller. 6

- SC** Safety Criticality. 12, 13, 15, 16, 22, 27–29
- SF** Safety Factor. 12, 16, 27
- SPU** Sensor Processing Units. 4, 24
- StC** Staging Criticality. 15, 16, 22, 27–29
- StF** Staging Factor. 12, 16, 25
- STRIDE** Spoofing, Tampering, Repudiation, Information disclosure, Denial of service and Elevation of privileges. 3–5, 7, 8, 25

- TDT** Techniques Description Table. 17, 19, 20, 23, 30

- UI** User Interaction. 17, 18

Paper V

A. Amro and V. Gkioulos, 'Cyber risk management for autonomous passenger ships using threat-informed defense-in-depth,' *International Journal of Information Security*, Nov. 2022, ISSN: 1615-5270. DOI: 10.1007/s10207-022-00638-y. [Online]. Available: <https://doi.org/10.1007/s10207-022-00638-y>

Cyber Risk Management for Autonomous Passenger Ships Using Threat Informed Defense-in-Depth

Ahmed Amro · Vasileios Gkioulos

June 2022

Abstract Recent innovations in the smart city domain have led to the proposition of a new mode of transportation utilizing Autonomous Passenger Ships (APS) or ferries in inland waterways. The novelty of the APS concept influenced the cyber risk paradigm and led to different considerations regarding attack objectives, techniques as well as risk management approaches. The main factor that has led to this is the auto remote operational mode; which refers to autonomous operations and remote supervision and control in case of emergency. The auto-remote operational mode influences the risk of cyber attacks due to the increased connectivity and reliance on technology for automating navigational functions. On the other hand, the presence of passengers without crew members imposes a safety risk factor in cyber-attacks. In this paper, we propose a new cyber risk management approach for managing the cyber risks against cyber-physical systems in general and autonomous passenger ships in particular. Our proposed approach aims to improve the Defense-in-Depth risk management strategy with additional components from the Threat Informed Defense strategy allowing for more evolved cyber risk management capabilities. Moreover, we have utilized the proposed cyber risk management approach for the proposition of a cybersecurity architecture for managing the cyber risks against an APS use case named milliAmpere2. Additionally, we present our results after conducting a Systematic Literature Review (SLR) in cybersecurity evaluation in the maritime domain. Then, the findings of the SLR were utilized for a suitable evaluation of the proposed risk man-

agement approach. Our findings suggest that our proposed risk management approach named Threat Informed Defense-in-Depth is capable of enriching several risk management activities across different stages in the system development life-cycle. Additionally, a comprehensive evaluation of the cybersecurity posture of milliAmpere2 has been conducted using several approaches including risk evaluation, simulation, checklist, and adversary emulation. Our evaluation has uncovered several limitations in the current cybersecurity posture and proposed actions for improvement.

Keywords Autonomous Passenger Ship ; Cybersecurity Architecture ; *ATT&CK* ; Defense in Depth ; Cyber Risk Management

1 Introduction

In a constantly evolving globe, technological advances improve every aspect of modern life. In the maritime domain, automation and digitalization are constantly evolving leading to drastic changes in business models, processes, as well as technology [50]. The impact of the current pandemic has been observed clearly in the maritime transportation sector in the form of a drastic decrease in passengers in 2020 compared to 2019 [11]. At the same time, to adjust to the post-pandemic normal, the development of innovative technologies and services for the transportation community has been proposed. It is already undergoing in the maritime industry to make it greener, cheaper, and more efficient. The pandemic has even emphasized that role [70]. Also, The US Bureau of Transportation Statistics has argued that the increasing demand for extending the capacity and flexibility of transportation systems has fueled the development of innovative technologies and services [17].

Norwegian University of Science and Technology
Teknologivegen 22, 2815, Gjøvik, Norway
E-mail: ahmed.amro@ntnu.no

E-mail: vasileios.gkioulos@ntnu.no

Recent innovations in maritime logistics when meeting activities related to smart city development have led to the creation of innovation in the field of inland passenger transportation through the proposition of Autonomous Passenger Ships (APS) (i.e. ferries). Domestic water transportation in Norway has witnessed the largest increase in passengers during the period between 2015 and 2019 [5]. In that direction, multiple projects have been recently initiated towards the development of autonomous passenger ferries in three regions in Norway [10]. Among these projects is a project named Autoferry which aims to develop an all-electric APS for inland water transport in the city of Trondheim [84]. The work presented in this paper originated from and is part of the Autoferry project. The targeted APS is planned to be autonomous with remote control and monitoring capabilities leading to an unconventional mode of operation in the maritime domain which is referred to as "autoremove" [44]. Although the new operational mode is expected to improve the provisioning of navigational services, it introduced a range of cyber threats with potential safety impacts. Toward addressing such issues, several system-specific requirements have been established during the authors' earlier work [24]. The established requirements were communicated by the identified APS stakeholders with a prime focus on communication reliability and cybersecurity toward safe operations. Towards addressing these requirements, a communication architecture for the APS has been proposed to satisfy the communication-related requirements [25]. This paper on the other hand addresses the cybersecurity requirements through the development of a cybersecurity architecture complementing the previously proposed communication architecture.

The identified stakeholders' requirements and concerns related to cybersecurity can be addressed by a cybersecurity architecture that provides risk management functions, including risk analysis, treatment, and monitoring. Moreover, autonomous ships are expected to include a group of Industrial Control Systems (ICS) and Cyber-Physical Systems (CPS) participating in the provisioning of autonomous and remote control and monitoring functions. For this, the concept of layered defenses; formally known as Defense-in-depth (DiD) is a proposed security strategy for risk management in critical systems [96] and in autonomous and remotely controlled vessels [44, §11]. Despite that, some concerns have been raised regarding the ability of DiD to withstand sophisticated attacks [49] as well as its lack of a Cyber Threat Intelligence (CTI) component that allows organizations to continuously enhance their defenses to match the ever-changing threat landscape [108]. At the same time, CTI is one of the components of another cy-

bersecurity strategy named Threat Informed Defense [80]. In this paper we investigate the utility of combined implementation of the Threat Informed Defense and DiD as a risk management strategy in a maritime use case that is the APS.

The contributions in this paper can be summarized as follows:

- We propose a new risk management approach named Threat Informed Defense-in-Depth that combines components from two cybersecurity strategies, namely, Threat Informed Defense, and Defense-in-Depth.
- We present a cybersecurity architecture for APS that is an outcome of the Threat Informed Defense-in-Depth approach.
- We present the results of our SLR regarding cybersecurity evaluation in the maritime domain highlighting different aspects and approaches.
- We present the results of the conducted cybersecurity evaluation processes for an operational ferry that is a prototype implementation of APS.
- We discuss the observed challenges in carrying cyber risk management functions in the context of the autoremove operational mode.

2 Background

2.1 Maritime Cyber Risk management

The International Maritime Organization (IMO) has issued resolution MSC. 428(98) [41] regarding the consideration of cyber risk management within the safety management systems of the different entities in the maritime industry. In this direction, IMO issued guidelines for cyber risk management [40]. The guidelines suggest several relevant frameworks and resources including the Framework for Improving Critical Infrastructure Cybersecurity by the National Institute of Standards and Technology (NIST) [30]. Additionally, several entities in the maritime domain have discussed approaches to cyber risk management, BIMCO; a global organization for shipowners, charterers, ship brokers, and agents, and DNV; a member of the maritime classification society. The concept of layered defenses known as the Defense-in-Depth (DiD) is the agreed-upon and encouraged strategy among these institutions.

DiD is defined by NIST as an "Information security strategy integrating people, technology, and operations capabilities to establish variable barriers across multiple layers and dimensions of the organization" [33]. A survey of cyber reference architectures and frameworks conducted by Savold et al [92] highlighted that DiD as a security design pattern that is observed in several security frameworks such as the Cisco SAFE [56], Oracle

Reference Architectures [35], and Northrop Grumman Fan. [77]. Nevertheless, the different DiD implementations focused more on Information Technology (IT) systems with the tendency to overlook Operational Technology (OT) systems. For that sake, the department of homeland security in the united states has issued a document for recommended practice as guidance for developing a DiD security program for environments with Industrial Control Systems (ICS) [47]. The document provides a detailed description of the DiD strategy from several viewpoints referred to as defense layers. Also, BIMCO provided guidelines for cyber risk management on board ships and discussed the strategy of DiD as well as Defense-in-Breadth (DiB); referring to the consideration of different technology domains, namely, IT and OT in the cyber risk management. BIMCO proposed a risk management approach including a group of defense layers [88]. Additionally, DNV has suggested the adoption of a DiD strategy for the cybersecurity of autonomous and remotely controlled vessels and discussed several components of a cyber security management system that can be adapted to support the strategy [44, §11].

After surveying the literature regarding cyber risk management approaches in the maritime domain, the adoption of layered defenses has been observed. To mention a few, Svilic et al [98] proposed and conducted a novel cyber risk assessment on board a vessel. The authors surveyed the vessel cybersecurity management system consisting of several defense layers including, physical access, patching, access control, and others. Grigoriadis et al [53] proposed a group of defenses for improving the cybersecurity of ports including vulnerability assessment, communication authenticity, weak password protection, and binary protection. Kavallieratos et al [64] leveraged the STRIDE and DREAD risk analysis techniques to assess cyber risks in cyber-enabled ships; which include autonomous and remotely controlled ships. The authors then followed the ISO 31000 risk management process [61] to propose baseline controls to mitigate the identified risks. The authors relied on the controls suggested by the Guide to industrial control systems (ICS) security [95].

Rajaram et al [86] have proposed guidelines for cyber risk management for shipboard systems with more focus on operational technology. The authors proposed a checklist approach for determining the cyber hygiene of vessels. The approach introduced the concept of security tiers which are aligned with risk priority levels, specifically, low, medium, and high. The concept of security tiers reflects the necessity for implementing security controls to address certain levels of risk.

However, the observed works lacked a clear implementation of the DiD strategy for ensuring that all layers of defenses are systematically considered. Therefore, in this paper, we utilize the DiD as an architecture framework during cybersecurity architecture development. The defense layers are collected from several sources including, the DiD guidelines for ICS in [47], DNV [43], and BIMCO [88]. Additionally, some works in the literature provide valuable artifacts including candidate non-developmental items (NDIs), architectural elements' properties, features as well as their interconnections.

2.2 The *ATT&CK* Framework

The Adversarial Tactics, Techniques, and Common Knowledge from MITRE, shortly known as the *ATT&CK* framework [97] is a recent, widely adopted framework in both academia and the cybersecurity industry. Currently, it encompasses three technology domains which are referred to as matrices, namely, enterprise, mobile, and Industrial Control Systems (ICS). The enterprise matrix covers Information Technology (IT) systems observed in enterprise networks. The mobile matrix covers handheld or mobile devices with Android or IOS. The ICS matrix covers networks and systems with Operational Technology (OT). This inclusion of several technology domains makes *ATT&CK* suitable in a wide range of use cases including the composition of these technologies. Additionally, the *ATT&CK* terminologies are being utilized for mapping adversarial activities by many organizations such as the European Union Agency for Cybersecurity (ENISA) in their annual threat landscape report [20]. Moreover, the *ATT&CK* terminologies are integrated within several Security Incidents and Event Management (SIEM) systems [16, 19] and cybersecurity testing frameworks such as Atomic Red Team [2] aiding the cybersecurity personnel in monitoring, detecting, and emulating adversarial activities in their network. Our risk management approach aims to integrate the *ATT&CK* framework within the different risk management processes. Starting from the risk assessment process, during the cybersecurity architecture development up until the evaluation of the proposed architecture. We argue that this provides a clearer description and traceability of the risks for the organizations as the identified risks in the risk assessment are mapped with the security controls intended to mitigate them and evaluated during the architecture evaluation.

In this direction, a risk assessment approach for the cyber-physical system has been proposed in our earlier work [23]. The approach is based on a design-level Failure Modes Effects and Criticality Analysis (FMECA)

[39] which utilizes the common knowledge encoded within the *ATT&CK* framework. The *ATT&CK* framework was chosen due to its comprehensive and low-level abstraction of adversarial tactics and techniques compared to other high-level models observed in the literature such as STRIDE [94] and the cyber Kill Chain [79]. Provided with a system description and stakeholders' risk thresholds, the approach begins with identifying the possible failure mechanisms (i.e. cyber threat) for each system component. Then the likelihood of these failure mechanisms is estimated utilizing the Common Vulnerability Scoring System (CVSS) [58]. The likelihood estimation also considers the existing mitigation methods. Afterward, the impact of the possible failure modes is estimated for each system component considering the occurrence of the failure mechanism. The *ATT&CK* framework provides the logical mapping of failure mechanisms and failure modes within its threat model. The estimation of the impact relies on a group of metrics including ones that utilize the concept of centrality from graph theory which aids in reducing the effect of biased estimation [104]. These metrics are calculated using a graph of the system. Then, the detectability is calculated which refers to the degree to which the risk of the identified attacks is reduced by the existing controls. Finally, a risk priority number (RPN) is calculated considering the likelihood of failure mechanisms, the impact of the failure mode, and the existing risk reduction measures. The risk is later characterized according to the stakeholders' risk thresholds. In addition to calculating the risks, this approach utilized *ATT&CK* for suggesting suitable risk mitigation methods for each threat against each system component. These mitigation methods are later forwarded for developing a suitable cybersecurity architecture. The reader may refer to our original work [23] for more information regarding the risk assessment approach.

2.3 Evaluation of Cybersecurity controls in the maritime domain

In this paper, we investigate the state of the art of cybersecurity control evaluation in the maritime industry considering the perspectives of both the academic community and relevant organizations including the classification society. The perspective of the academic community is captured through a Systematic Literature Review (SLR) which is discussed in Section 3.2.3. On the other hand, the perspective of the relevant organizations is captured through the collection and analysis of their publications regarding cybersecurity. The organizations were chosen based on the references in the

literature. This includes, IMO, BIMCO, ENISA, and DNV.

The International Maritime Organization (IMO) guidelines for cyber risk management [41] refers to the need for evaluating a cyber risk management regime using effective feedback mechanisms without further description.

BIMCO guidelines [88] refers to the evaluation of cybersecurity controls within the risk assessment process through the assessment of residual risk when considering the existence of security controls. Also, the document refers to the third-party risk assessment process as means of performing accurate risk assessments by identifying whether the defense level matches the accepted level established in the cybersecurity strategy. The document refers mainly to penetration testing as a common approach but argued that it can be intrusive, risky, largely expensive, and requires an understanding of networks and assets. Therefore, other alternative approaches are proposed including asset discovery and inventory, auditing network architecture and design as well as vulnerability assessments.

DNV, another classification society in the maritime industry, refers in their class guidance for cyber-secure systems to the evaluation and assessment of security controls during the acceptance stage for newly built and alteration projects [43]. They highlighted the roles of different stakeholders involved in the cyber security system testing and assessment of controls during the different system development stages.

Another documentation by DNV discusses resilience management of systems onboard ships and mobile offshore units [18]. The document discusses three approaches for assessing the cyber security of a system, namely, high-level assessment, focused assessment, and comprehensive, in-depth assessment. The document refers to cyber security controls as barriers or safeguards. Comparing the current safeguards with the target is conducted through detailed checklists used in interviews with experts and relevant staff and users. After the assessment stage, a verification and validation process is needed to clear any discrepancy between the expected state and the actual state. The document suggests monitoring and testing the barriers at the level of the individual components as well as the system level. The discussed approaches include vulnerability assessment, technical verification such as load testing, network storm simulation (i.e. flooding), fuzz testing, actively provoking failures, and passive measurements. Penetration testing is discussed as a possible approach to systematically employ different methods. Moreover, the document discusses the verification of the information security man-

agement system by accredited third parties through audits and suggests a direction toward certification.

ENISA has published a report regarding cyber risk management for ports [45]. The report refers to assessing the maturity of cybersecurity posture following the maturity levels approach. Each maturity level is described and the organizations are left to self-assess their position within the different levels according to their own risk assessment.

In summary, the evidence collected from the published material by the different relevant organizations suggests the existence of well-established and flexible methods and approaches for the evaluations of cybersecurity controls. Penetration testing is a commonly suggested approach, yet, its discussed challenges pave the way for other possible approaches. However, the increased reliance on the human element within the evaluation process is observed, either through interviews, surveys, or relying on human evaluators. We argue that in autonomous vessels, human involvement is going to be reduced. This motivates the development and integration of automated processes for the evaluation of cybersecurity controls within the different cyber assets involved in the autonomous vessels. In this work, we will investigate the suitability of the identified methods in the literature for the evaluation of cybersecurity architecture for an APS. Moreover, the increased reliance on sensor data supporting systems employing machine learning and artificial intelligence algorithms in the navigation functions exposes the autonomous vessels to new threats such as adversarial machine learning. None of the studied literature or documents from the different organizations have tackled this issue. This suggests the need for future efforts to investigate it.

3 Methodology

3.1 Cyber Risk Management Strategy

The first question in this paper is, "What is a suitable strategy for managing the cyber risks for an autonomous passenger ship?". For this, a comprehensive literature survey was conducted to capture the state of the art in cyber risk management in the maritime domain. The perspectives of both the classification society and academia were considered. For this, academic databases, namely, Scopus and Google Scholar were queried for academic resources while the websites of relevant stakeholders were utilized for extracting documents relevant to maritime cyber risk management. As discussed in Section 2.1, the concept of layer defenses formalized as the DiD is the observed approach in maritime risk management. However, its effectiveness

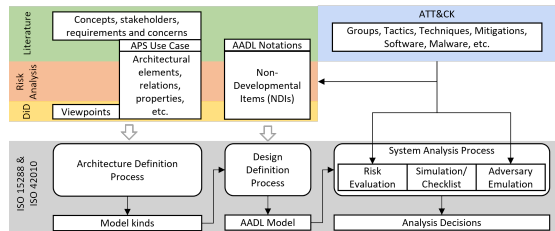


Fig. 1: Overview of the Cybersecurity architecture development methodology

against sophisticated attacks has been questioned [49]. Based on that, we are proposing a new cyber risk management approach in this paper. Our proposed approach is described in Section 4.

3.2 Cybersecurity Architecture Development

The second question is "How a cybersecurity architecture can be developed to support the cyber risk management strategy?". There is a lack of discussion in the literature regarding this topic in the maritime domain. Therefore, we followed a standard system engineering process for the development of the cybersecurity architecture. It is based on a pre-specified set of requirements and concerns. It addresses the analyzed risks and includes components that allow updated risk monitoring and treatment capabilities. To realize this architecture, the system development followed the ISO 15288:2015 technical processes for defining an architecture and its design. Later, the developed design is subject to different system analysis processes to evaluate it. Figure 1 depicts an overview of the methodology followed for the development of the cybersecurity architecture as a system of systems (SoS). Moreover, guidelines from ISO 42010:2011[62] are utilized for the description of the architecture. The figure also reflects the input artifacts as well as the output for each process. Further details are discussed in the following sections.

3.2.1 Architecture Definition Process

The DiD strategy (section 2.1) is utilized as an architecture framework guiding the development of required defense layers also known as viewpoints. The defense layers were defined after studying the documents issued by department of homeland security [47], DNV [43], and BIMCO [88] regarding implementing a DiD risk management approach. Later, and following a top-down approach, the system context, interfaces, and interactions with external entities are defined (Section 6.1). Then,

the architectural entities and their relationships toward the satisfaction of stakeholders' requirements are defined. Afterward, each architectural entity is allocated relevant properties, concepts, etc. For this sake, a use case of the APS is presented to facilitate the description of the aforementioned concepts (Section 5). Then, a detailed description of architectural entities including any required system decomposition is conducted, depicting the interfaces and interactions between the different system elements. The aforementioned resources for DiD guidelines, the literature, and our conducted risk analysis [23] were consulted for useful artifacts during this stage. The outcome of these activities is discussed in detail in section 6. The modeling techniques utilized during these activities are preliminary data flow diagrams and adjacency matrices.

3.2.2 Design Definition Process

Architecture modeling is utilized to allocate system requirements to system elements (Appendix B), establish the structure of system elements (system, process, connection, etc.), defining interfaces among them as well as among external enabling systems. The later activities are achieved through the formalization of a model developed using Architecture Analysis and Design Language (AADL) [48] and OSATE, an open-source tool that supports it [100]. AADL is utilized to facilitate the architecture description and analysis considering that the underlying communication architecture is modeled using the same modeling language [25]. Moreover, AADL which can extend SysML has been utilized for analyzing critical systems due to its ability to describe information related to hardware, operating system, and code, allowing it to be applied at different stages during system development life-cycle [42, 66]. Afterward, the assessment of possible NDIs is performed towards the selection of the preferred solutions. The literature including DiD guidelines and our conducted risk analysis [23] were consulted regarding the possible NDIs to be integrated. Finally, the model is completed, and this paper completes the description of the architecture and its design describing the rationale for the different design decisions (Section 6).

3.2.3 System Analysis Process

The last question is "How the developed architecture can be evaluated?". For this, a Systematic Literature Review (SLR) was conducted following the guidelines for conducting an SLR as proposed by Okoli and Schabram (2010) [85]. The proposed process consists mainly of four phases, planning, selection, extraction, and execution.

During the planning phase, the purpose of the review is defined. In this paper, the aim is to capture the state of the art in evaluating cybersecurity controls in the maritime transport domain, focusing on objectives, approaches, and relevant variables. The study aimed also to identify the relevant safety considerations as well as considerations related to the autonomous mode of operations.

Afterward, the selection phase entails the establishment of the parameters and criteria used to search the literature and filter the results. The following search query was used across three digital libraries, IEEE Xplore, and Scopus (with the appropriate syntax):(ship OR vessel OR maritime) AND (cyber OR "information security") AND (risk OR threat) AND (evaluate OR assess OR validate). The results were filtered to only include works after 2011, English as a language, and only considering documents of types (Conference Paper, article). The choice to only include works that were published in the last 10 years was based on the desire to stay updated. In total 33 articles were identified. A clear criterion has been established for deciding either to include or exclude articles from further steps. The inclusion criterion is to only include works that targeted the evaluation, testing, assessment, or validation of cybersecurity controls in a system that is part of the maritime transport infrastructure.

Later, the extraction phase entails a deeper understanding of the resulted works to perform a quality appraisal and extract relevant data including other relevant articles. The results included a broad range of articles related to the evaluation of cybersecurity in maritime and other relevant domains. The main objective of this work is to identify works that have addressed the evaluation of cybersecurity controls in a maritime transport system. Other works that target systems outside this scope such as marine renewable energy systems were dropped. Additionally, works that targeted the analysis or assessment of the cybersecurity of certain systems without considering the evaluation of security controls were also dropped. The final list of articles proceeded for the next step was 18. Cybersecurity control evaluation is approached in this paper as a system analysis process. Therefore, The data extraction step relies on the ISO 15288 standard to map the observed artifacts in the literature to the relevant aspects in the system analysis process in the standard. For each screened work, the following aspects were captured, the process, approach, method, analysis questions, relevant stakeholders, scope, objectives, enabling systems, assumptions, quality and validity, discussion of corrective actions, and the venues for communication of results.

Finally, the SLR is executed through an overall synthesis of the found literature in addition to discussing and documenting the results and findings. The extracted artifacts from the studies during the data extraction stage are utilized for the identification and classification of evaluation approaches in order to identify those which are suitable for adoption in the evaluation of the cybersecurity controls in the APS. Then, the generation of the final document that is this paper is to be leveraged as a source of knowledge reflecting the current state of the art in the declared scope.

4 Threat Informed Defense-in-Depth

Although DiD is a widely adopted strategy and its usefulness against unsophisticated attacks has been demonstrated, critical discussions have been raised regarding its ineffectiveness against targeted sophisticated attacks [49]. This can be linked to the lack of a Cyber Threat Intelligence (CTI) program, one of the missing elements of DiD [108]. CTI enables defenders to constantly tune their defenses to manage the risks targeting their assets considering the current threat landscape [76]. CTI is one of the three main pillars of the Threat Informed Defense strategy in addition to defensive engagement of the threat and focused sharing and collaboration [14]. The three pillars interact together to provide the *ATT&CK* framework [97] which can be used as an up-to-date resource for encoded common knowledge regarding adversarial behavior. We argue that aligning the *ATT&CK* framework and DiD layers would allow more evolved cyber risk management capabilities. So, in this paper, we propose a cyber risk management approach that integrates elements from the two strategies, namely, the Threat Informed Defense from MITRE [80] (i.e. Threat-Based Defense [14]), and Defense-in-Depth [47]. In the remainder of the paper, we will refer to cyber risk management simply as risk management. The approach is aligned with the risk management process in ISO 15288:2015 [60] as shown in Figure 2 including four stages, planning, managing risk profile, analyzing the risks, and risk treatment and monitoring.

During the planning stage, the scope of the risk management process is defined. This entails the provisioning of a detailed system description including its operational context, stakeholders, requirements, components, their properties, and connections. Then, the stakeholders' risk thresholds are derived and established from their concerns and requirements. Additionally, results of earlier risk analysis and assessments are to be maintained in the risk profile. After that, the risks in the system are analyzed and assessed to identify the required risk controls. For this step, we propose the

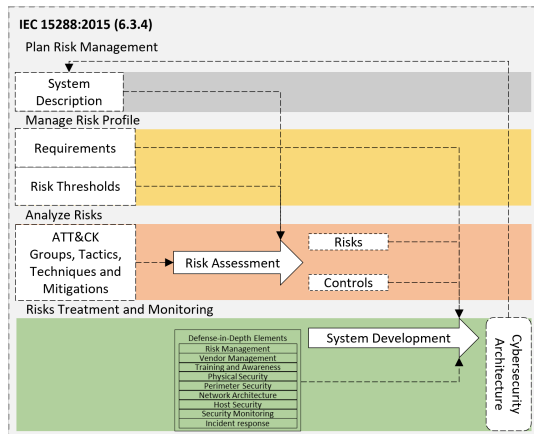


Fig. 2: Threat Informed Defense-in-Depth Risk Management approach

utilization of the *ATT&CK* framework [97] for facilitating the risk assessment process. *ATT&CK* is an integral component in the Threat Informed Defense strategy as it provides a common knowledge repository for adversarial tactics and techniques drawn from several CTI sources. Additionally, *ATT&CK* suggests tailored defensive mechanisms for each technique. Such an approach is demonstrated in our previous work in which we proposed an FMECA-based risk assessment approach utilizing *ATT&CK* [23]. Based on a system description and risk thresholds, the risk assessment process identifies risks and proposes the required controls to mitigate them. Later, a cybersecurity architecture for supporting the risk management approach is developed. The architecture development relies on the proposed controls from the risk assessment process in addition to the stakeholders' requirements.

The DiD elements form architectural viewpoints for guiding a systematic architecture development process. For this, a mapping between the security controls suggested by *ATT&CK* and the DiD viewpoints is needed. The proposed mapping is depicted and discussed in Appendix A. The system development follows the ISO 15288:2015 technical processes for defining an architecture and its design. Later, the developed design is subject to different system analysis processes to evaluate it.

At the design stage, a suitable analysis process is a model-based risk evaluation. Several works have been observed to implement a similar approach [28, 46]. The risk assessment process is conducted in several iterations against the system model considering the different possible defensive strategies. The overall risk reduction

(i.e. residual risk) for each defensive strategy is calculated in order to choose the optimal one. First, the cumulative risk of all the identified risks is aggregated and then the ratio of risks for each defense strategy compared to the base strategy (no controls) is calculated. To facilitate this analysis process we have developed a defense strategy comparison algorithm. The algorithm extends the risk assessment algorithm proposed in our earlier work [23] for comparing the risk reduction of different defensive strategies. The strategy comparison algorithm is shown in Algorithm 1. A defense strategy is modeled using a mapping between the *ATT&CK* controls and the architectural components which are within the scope of the control function.

Algorithm 1 Strategy Comparison Algorithm (SCA)

```

1: procedure SCA(Threat information, Components information,
   Mitigation measures information.)
2:   for each defense strategy do
3:     for each component do
4:       AttackList  $\leftarrow$  IdentifyRelevantAttacks
5:       for each attack in AttackList do
6:         Likelihood  $\leftarrow$  Cal.AttackLikelihood
7:         Impact  $\leftarrow$  Cal.AttackImpact
8:         Detectability  $\leftarrow$  Calc.AttackDetectability
9:         RPN  $\leftarrow$  Likelihood  $\times$  Impact  $\times$  Detectability
10:        MitigationList  $\leftarrow$  GetAttackMitigation
11:       end for
12:     end for
13:     StrategyOverallRisk  $\leftarrow$  Sum.ofAllRPNs
14:   end for
15:   return RiskReduction, AttackLists, RPNs and Mitigation-
   Lists for each defense strategy
16: end procedure

```

In order to demonstrate our approach, we will utilize a use case of an APS during different system development stages, namely, an APS model which has been developed in our earlier work [25] as well as an implemented APS prototype named milliAmpere2.

5 Use Case: Autonomous Passenger Ship

The Autoferry project [84] aims to develop an APS use case named the milliAmpere2; An autonomous ferry capable of carrying 12 passengers across the Trondheim city canal proposed as an alternative to a high-cost bridge [55]. The ferry is designed to operate autonomously with human supervision. Supervision is carried from a Remote Control Center (RCC) encompassing monitoring APS operations and having the ability to intervene at any moment. The operation of the APS relies heavily on its communication architecture. Many stakeholders are involved in the design, development, and expected operations of the APS. The requirements for secure and reliable communication architecture have been collected and adopted from each stakeholder’s per-

spective [24]. The requirements for reliable communication included aspects related to redundancy for high availability, resiliency, network segregation, and others. The communication requirements have been addressed in the design and development of the communication architecture presented in [25]. On the other hand, the requirements related to the cybersecurity of the APS are addressed in this paper with the milliAmpere2 as a use case. A photo of the milliAmpere2 ferry during a test run is shown in Figure 3 including an illustration of its main cyber components.



Fig. 3: The milliAmpere2 ferry with an illustration of its main cyber components

A sufficient level of understanding of the communication architecture is needed to understand the needs and methods for implementing the security practices. An overview of the APS communication architecture in Fig 4 shows the different architectural components and their interconnections. The proposed architecture connects the APS with its operational context through several communication channels. The entities in the operational context include a Remote Control Center (RCC), an Emergency Control Team (ECT), other ships, Vessel Traffic Services (VTS), and others (more details in [24], [25]). The APS communicates externally through several communication modules. A Mobile Communication Module (MCM) connects the APS to the inter-

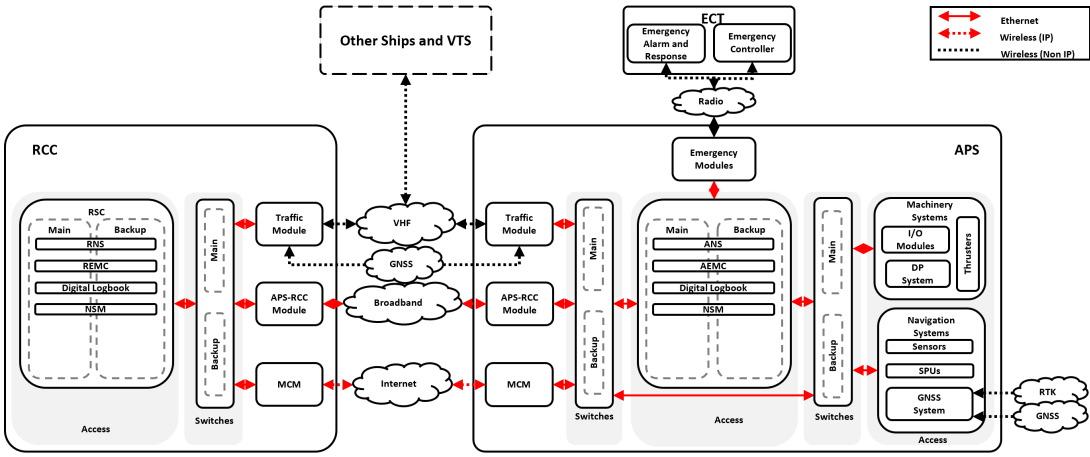


Fig. 4: Overview of the APS Communication Architecture (adapted from[25])

net through a mobile network using suitable technology (e.g.5G). The APS-RCC module provides direct point-to-point communication between the RCC and the APS through a suitable technology such as long-range Wi-Fi or mobile communication through a different service provider. The traffic module is required for ship-to-ship communication by broadcasting and receiving broadcast navigation messages such as ships’ positions, speed, headings, etc, Automatic Identification System (AIS) is a candidate for implementation for the traffic module. Two modules for emergency purposes are integrated into the architecture. One is responsible for providing emergency navigation and control capabilities by the ECT while the other is to transmit emergency signals when passengers press on an emergency push button in case of incidents (e.g. passenger falling from the APS). Finally, the last group of modules is related to positioning and timing. Two Global Navigation Satellite System (GNSS) receivers are implemented one is connected to the traffic module and the other is connected to the GNSS system. Additionally, a single Real-time kinematic (RTK) receiver is connected to the GNSS system providing data for positioning correction.

The internal network of the APS connects different systems needed to carry out the expected functions including navigation, machinery systems, and others. The navigation system is responsible for collecting navigation data from arrays of sensors as well as the GNSS system for determining safe routes through an Autonomous Navigation System (ANS). The Machinery system is responsible for the ship’s movement through active thrusters and a Dynamic Positioning (DP) system which is managed through an Autonomous

Engine Monitoring and Control (AEMC). All the aforementioned components in the different systems are interconnected through an Ethernet network consisting mainly of Layer-3 switches. A compatible arrangement is proposed on the RCC to facilitate communication with the APS network. The three communication modules, namely MCM, APS-RCC, and traffic modules are expected to be similar to their pairs on board the APS. Also, a Remote Navigation System (RNS) supports the APS navigation system, in addition to a Remote Engine Monitoring and Control (REMC) for steering the ship.

The architecture model has been input to the developed defense strategy comparison algorithm (Section 4). Table 1 depicts the outcome of the algorithm for calculating the risk reduction percentage for each considered defense strategy. The strategy that provides the highest risk reduction is Strategy 5 which is based on BIMCO guidelines. However, the results suggest that it might be avoided in case of a reduced budget since Strategy 3 addresses a large portion of the identified risks with a lower amount of controls. Another aspect to consider is if satisfying the stakeholders’ requirements is pursued, then Strategy 4 provides the optimal choice. It addresses the requirements as well as the identified risks while achieving a competing risk reduction score compared to other strategies. Therefore, the architecture development in the following section shall address controls based on Strategy 4. It is worth mentioning that, the financial aspect of the strategy comparison is outside the scope of this paper and can be an item of future work.

Table 1: Risk reduction of the different defense strategies

Strategy	Risk Reduction	Description
1	0,84 %	the included controls in the current system model
2	80,22 %	the suggested or mandated controls in the stakeholders' requirements.
3	69,72 %	the controls suggested after the risk assessment process while considering the current system model.
4	81,94 %	the controls suggested after the risk assessment process while considering the stakeholders' requirements as bases for defense.
5	85,35 %	the controls suggested in the BIMCO guidelines [88]
6	69,26 %	the controls suggested in the ICS DiD guidelines [47]
7	65,74 %	the controls suggested in the DNV guidelines [43]

6 Cybersecurity Architecture

In this section, a cybersecurity architecture is presented which is an outcome of our risk management approach. It describes the different cybersecurity functions carried by the different architectural components across the APS operational context, namely, the ferry, the RCC, and the ECT. The architecture is modeled using AADL [48], thus enabling an extended analysis on one hand, and design modifications in the future on the other. The model code can be accessed through an online repository [1]. It presents the architecture through a group of views encompassing the entire System-of-Systems (SoS) layout (i.e. facilities), the logical view (i.e. service), and the structure view (i.e. system elements). The following sections discuss the different views and present the outcome of the architecture development processes mentioned in Section 3 by providing the rationale behind the different architectural and design decisions as well as attempts to provide a sufficient level of traceability among the different viewpoints, system elements, stakeholders, concerns and requirements.

6.1 Context view

The objective of the cybersecurity architecture is to address stakeholders' concerns regarding managing the risks against the APS systems [25]. An overview of the Narrowest system of interest (NSoI) is depicted in Figure 5. We will refer to the NSoI throughout the paper as the APS ecosystem. This view captures the highest level of abstraction concerning the different architectural components. It captures the System of Systems (SoS) in the operational context that collectively addresses the system objectives. Each SoS is hosted in a

dedicated facility, APS SoS is hosted onboard the autonomous ship, RCC SoS is hosted in a Remote Control Center, and ECT SoS is hosted in a nearby boat carrying an emergency control team. Each SoS integrates additional components or utilizes components within the pre-defined communication architecture discussed in section 5. Additionally, the APS and RCC are expected to utilize enabling systems hosted in remote locations accessed through the internet (e.g. updates). This view aids the understanding of the various interacting entities in the context of the cybersecurity architecture; explicit details related to each architecture viewpoint and its relevant system components in the operational context are discussed in the following subsections.

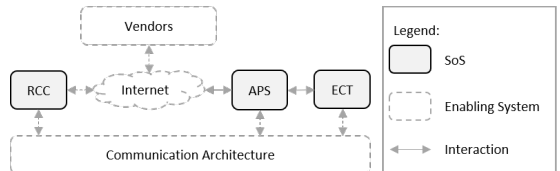


Fig. 5: Overview of the Narrowest system of interest

6.2 Risk Management

After January 1st, 2021, all ship owners must address cyber risk management in their safety management systems for compliance under the ISM code [87]. To aid the efforts toward compliance with these regulations, the need for an Integrated Security, Safety, and Ship Management System (IS3MS) that applies an up-to-date risk management framework has been proposed in our earlier work [24]. Several cyber risk management approaches or frameworks have been cited in the literature, including DNV class guidelines [43], BIMCO guidelines [88], and ICS DiD guidelines in [47]. However, the discussed frameworks are generic and pose no restrictions on the applied methods. So, our proposed risk management approach (Section 4) includes activities that are aligned with all of them as shown in Table 2. Moreover, our proposed approach does not replace organizations' risk management processes. It can be utilized to complement them by identifying existing gaps to rectify them.

The scope of the IS3MS extends the scope of the cybersecurity architecture to include safety-related functions including monitoring and alerting. The development of an IS3MS architecture is a target study for future work. The concept includes the provisioning of the

Table 2: An alignment of our risk management approach with existing relevant approaches proposed by DNV [43], BIMCO [88], and DiD guidelines in [47]

IEC 15288 (6.3.4) [60]	Our Approach	DNV [43]	ICS DiD [47]	BIMCO [88]
Plan Risk Management	Specify Assets		Inventory Assets Categorize Asset Criticality	
Manage Risk Profile	Specify Risk Thresholds			
Analyze Risks	Identify Failure Modes	Identify Risks	Identify Security Risks	Identify Threats
	Identify Controls			
	Identify Effects of Failure Modes	Analyze Risks	Determine Potential Impact	
	Identify Failure Causes			Identify Vulnerabilities
	Estimate Likelihood of Failure Causes			
	Evaluate Risk	Evaluate Risks		Assess Risk Exposure
Risks Treatment and Monitoring	Develop Cybersecurity Architecture	Treat Risks	Implement Security Controls	Develop Protection and Detection Measures
			Monitor and Adjust	Establish Response Plans
				Respond to and Recover from Incidents

different risk management functions by a centralized component. In the scope of this paper, the IS3MS is expected to include the following sub-components each addressing specific requirements or concerns:

6.2.1 Asset and User Inventory

A regularly updated inventory of system users and components is required for the planning stage to define the scope of the risk management process. For the APS, at the design stage, this inventory includes an architecture model. During advanced stages in the development life cycle, this inventory can be conducted through other architecture scanning techniques. Additionally, a User Account Management (UAM) component is included in the architecture to support user inventory activities (more details in section 6.6.4).

6.2.2 Risk Assessment

Periodic risk analysis and assessment activities are required for maritime risk management proposed by the International Maritime Organization (IMO) in Resolution MSC. 428(98) [41] and constitutes an established requirement from the regulators' perspective in the APS system. This component is proposed to facilitate the

conducting of this periodic process. It can be utilized to maintain the risk profile, aid in the identification of threats, assesses their risks, and propose controls. Our developed algorithm that supports risk assessment for cyber-physical systems [23] has been integrated into this module. Our algorithm provides an assessment of the current threat landscape utilizing feeds from active Cyber Threat Intelligence programs delivered through the *ATT&CK* framework. This adds to the architecture the capability to identify weak points as well as directions for improvement. As mentioned in Section 4, we extended the algorithm in this paper to facilitate the comparison of different defensive strategies.

6.2.3 Policies and Procedures

The establishment of policies and procedures is a common practice related to cyber risk management with varying focus areas. DNV's guidelines refer to policies related to personnel security, information classification, change management, and removable media [43]. BIMCO refers to crews' personal devices, use of administrative privileges, and equipment disposal. DiD guidelines in [47] focus on policies and procedures that are related to the human element.

6.3 Physical Security

Controlling physical access to the facilities and components is an agreed-upon defense layer. However, no communicated cybersecurity requirements related to it were identified. In our previous work [24], the requirements were elicited by reviewing stakeholders' publications and documents with a focus on cybersecurity and communication requirements. Physical security is discussed within the realm of safety and security conditions [34, §2.2.2] and access control [44, §4.2.3.2] and [44, §6.4.4]. This suggests that physical security is outside the scope of the cybersecurity architecture of the APS, yet, it is a very important element that is required as an enabling system.

6.4 Training and Awareness

The autoremove operational mode will change the traditional human role in maritime. The need for training regarding cybersecurity policies for APS personnel is a communicated concern and is a common defense layer. This includes personnel who are stationed in the RCC, among the ECT, or any other personnel that may access the APS system including service providers. Moreover, the risk analysis process has identified a group of

threats that can be mitigated with cybersecurity training for both system developers and operators as well as the attack techniques that leverage user actions. Special considerations should be described regarding the implementation of security procedures in ICS to protect mission-critical systems. Training personnel and increasing their awareness regarding IT and OT security threats is an integral aspect to limit opportunities for compromising the systems and enabling the personnel to identify signs of compromise. This component aggregates the management of the aforementioned activities.

6.5 Network Architecture

The segmentation and segregation of the APS network is an established requirement. The network has been designed with segmentation in mind to satisfy a segmentation policy related to communication reliability [24]. Nevertheless, security segmentation considers a different perspective. A network architecture for ICS is proposed in DiD guidelines in [47]. The architecture is described through different zones and levels. The proposed zoning architecture divides the network into six network levels across three zones each connecting a group of components with a specific set of functions.

The zones are the enterprise security zone, the manufacturing security zone, and the Demilitarized Zones (DMZ). The enterprise security zone hosts mostly IT systems that are expected to communicate with external entities. The manufacturing security zone on the other hand hosts mostly OT systems responsible for local or remote control and processing components as well as sensors and field devices. Furthermore, several network security levels reside within each security zone. Table 3 depicts the proposed distribution of components across the different security zones.

Remote access is expected and has been identified as a possible risk, therefore, a secure network architecture should consider the external communications links arriving at the network through insecure networks (e.g. mobile network or wireless medium). For this sake, a DMZ within the APS network is considered to host servers with expected external access (e.g. internet access) including the jump server. Access to the DMZ should be secured using Access Control Lists (ACL). No requirement for a DMZ at the RCC has been identified at this stage. Systems in level 3 are considered among the biggest targets for intruders aiming at affecting a critical infrastructure system according to a “peel-the-onion” analysis due to the ability to control and oversee the control systems residing in lower levels as well as the ability to suppress potential alarms

rising from their malicious actions [47]. A similar conclusion has been drawn from the conducted risk analysis [23]. The remote functions in the APS create additional challenges to the security in level 3 systems. Several systems outside the ship are involved in time-critical processing and control operations on the RCC facility, such as remote navigation and machinery monitoring and control. Such systems are not expected to have external access according to DiD guidelines in [47] nor is this operational mode addressed in the guidelines by DNV or BIMCO. Nevertheless, these systems provide crucial functions for safety and regulatory reasons [24]. Therefore, an additional layer of protection is expected between level 3 systems in different facilities. A proposed solution using VPN tunnels is discussed in Section 6.6.3.

Multiple VLANs are suggested to realize the expected network zones with appropriate Inter-VLAN routing and ACL rules. These configurations can be implemented at the network switches to route traffic between the appropriate zones securely and reliably. The switches act as security domain authorities enforcing the security policies of each security zone.

6.6 Network Perimeter Security

Additional measures should be put in place to secure communications between the different network security zones in the different facilities, namely, the APS, the RCC, and the ECT. Achieving this can be accomplished by including both physical and logical controls. The physical controls are related to physical security which is outside of the scope of this paper. On the other hand, logical controls can be considered concerning the communication boundaries which are represented by the network gateways. Each gateway should be monitored by a firewall or another security barrier enforcing a security policy for securing the perimeter. The discussed focus areas for perimeter security are discussed in the subsequent sections.

6.6.1 Firewalls

A group of firewalls is placed at the edge of each network security zone in each facility to establish domain separation. Two dedicated network firewall devices are proposed, at the APS and the RCC respectively to handle external connections passing through two IP-based gateways (MCM and APS-RCC Module) such as connections with vendors over the Internet. Additional firewall capabilities for internal networks can be achieved through ACL implemented in the Layer 3 switches. Moreover, utilizing host-based firewalls can extend the

among the threats against the APS operations and the existing mitigation methods are classified by the conducted risk analysis as insufficient. Jamming attacks are proposed to be mitigated in the APS use case through redundant functional components such as other sensors (e.g. lidars, video, etc.) and backup GNSS systems. This has been considered and discussed in the conducted risk assessment [23]. Several solutions have been observed in the literature relying on anomaly detection using machine learning [32], and specification-based detection [26, 72]. We have addressed this issue in our other work in which we have conducted an anomaly analysis and proposed detection rules for detecting simple and sophisticated attacks against GNSS systems communicating using the NMEA protocol (National Marine Electronics Association) [26]. The development of a dedicated anomaly detection solution is an item for future work.

6.6.3 VPN

The risk assessment process identified an important mitigation method through the encryption of sensitive information in transit. Moreover, Virtual Private Network (VPN) tunnels are suggested by DiD strategy to be implemented between systems in the same security level in different facilities for maintaining perimeter security [47]. For instance, the autonomous navigation component in L3 in the APS is expected to communicate navigational information to the remote navigation system in L3 in the RCC, this communication passes insecure mobile and/or broadband networks and may contain passenger-related data (e.g. video stream). Therefore, VPN is suggested to be implemented to secure these communication flows. As shown in Figure 7, VPN tunnels are proposed to be integrated into the APS and RCC dedicated firewalls using router-based IPSec protocol [102] to reduce firewall management. Otherwise, client-based firewalls using PPTP (Point-to-Point Tunneling Protocol) [54] provide another implementation option.

6.6.4 Access Management

The conducted risk analysis process has proposed considerations to be integrated into the access management components including a password policy, multi-factor authentication, and software and device authentication techniques. An overview of the access management services in the APS architecture is depicted in Figure 8. At the higher level of abstraction, the facilities hosting the different SoSs are expected to be

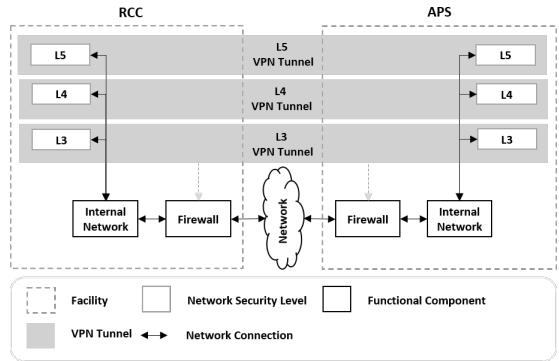


Fig. 7: Overview of the proposed VPN tunnels

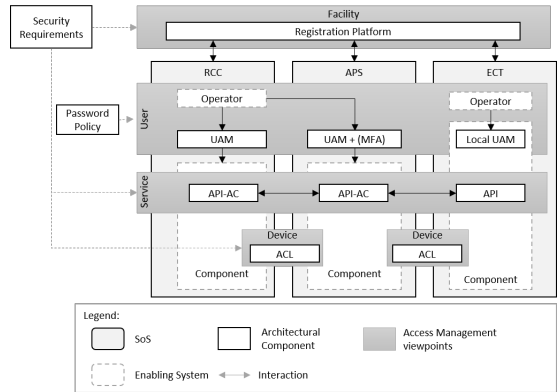


Fig. 8: Overview of the Access Management services

registered on a common platform. This is a communicated requirement for having a ship registry component within the system operational context. Moreover, within the same facility (e.g. RCC), the operator can access components through a User Access Management (UAM) component that implements a password policy. User access to components in another facility requires a multi-factor authentication process (MFA) integrated with the UAM component. Hardware component-to-component access is controlled by ACL while software component-to-component access is controlled by a functionality integrated into the different Application Programming Interfaces (API) which we refer to as API Access Control (API-AC). A group of security requirements is expected to be communicated to the providers of the enabling systems regarding the access management solutions such as the implementation of secure protocols related to Authentication, Authorization, and Accounting (AAA).

DiD guidelines [47] suggests implementation options for the UAM and APIs. Regarding UAM, centralized access management for each facility is favorable over a distributed approach. Lightweight Active Directory Protocol (LDAP) is a possible protocol for implementation as it can provide Role-Based Access Control (RBAC) which is the recommended approach proposed according to the conducted risk analysis. Remote Authentication Dial-In User Service (RADIUS) or Terminal Access Controller Access-Control System (TACACS) are both valid implementation options for the UAM [47]. However, the centralized approach introduces a risk if the authentication server gets compromised therefore, strict security controls should be applied to protect this server as well as establish redundancy. Also, remote connectivity is required for maintenance due to unmanned operations and jump servers hosted in the DMZ are proposed for this sake. Remote access to the jump server should be secure and MFA is proposed for that. Regarding APIs, they are widely popular in ICS and suffer from a wide range of security issues [47]. Therefore, great attention should be spent on the development of the API-AC component. This observation will be forwarded to other project members responsible for the development of APIs.

Further complications regarding access control and authentication are expected in ICS networks due to the provisioning of systems by different manufacturers not necessarily implementing the same authentication mechanisms. Therefore, local authentication and authorization policies and procedures should exist in these components.

6.7 Host Security

Considering the viewpoint of host security, several aspects of interest have been identified through implementing the DiD strategy as well as the learned from the conducted risk analysis. These aspects are detailed in subsequent sections.

6.7.1 Patch and Vulnerability Management

Keeping up-to-date software with security patches is a strong countermeasure to many cyber threats. Integration of components for patch and vulnerability management (PVM) is an agreed-upon mitigation method according to the conducted risk analysis process and the different DiD strategies. Moreover, a PVM component supports the satisfaction of established requirements regarding updates and security analysis. At the same time, the impact of a system patch should be evaluated before implementing the patch on the operating

APS to ensure ongoing operations, especially regarding the operation of critical components.

6.7.2 Malware Protection

The integration of the endpoint malware protection component within the APS architecture is a communicated requirement. Additionally, DiD guidelines suggest malware protection tools for supporting host security. Moreover, the risk assessment process has identified anti-malware among the required risk mitigation measure. Therefore, malware protection software is to be integrated into the relevant architectural components.

6.7.3 Application Isolation and Sandboxing

Identified as a risk mitigation method through the conducted risk assessment to mitigate against high risk imposed by scripting threat. The utilization of virtual machines, docker containers, and other forms of application and component separation is encouraged. Nevertheless, each implementation imposes different security threats and therefore, relevant security controls should be integrated. DiD strategies suggest considerations for the application of virtualized host components. This risk mitigation method is of particular relevance to centralized components hosting autonomous and remote control and navigation functions as well as other components for system and network management and security. The application of virtualization has been proposed in the architecture design earlier [25] and is also adopted in the scope of this cybersecurity architecture.

6.7.4 Backup

Data backup has been identified as the most important risk mitigation method during the conducted risk analysis to mitigate several attack techniques such as defacement attacks and loss of availability [23]. Also, a specific requirement exists concerning the availability of backup facilities for protection and recovery functions following a backup policy. Moreover, remote backup facilities have also been suggested during the risk analysis and the RCC is proposed to host such facilities. For this purpose, two backup servers are proposed, a server on board the APS and another hosted in the RCC. Regarding the APS backup server, the requirement dictates that an early alert indicating storage capacity exhaustion should be implemented and the ability to transfer the data to shore should be made available [24].

6.7.5 System Hardening

Referred to by BIMCO as “Secure configuration of hardware and software” [88]. This component is proposed to address a group of concerns identified through the risk analysis process. It is a required activity to perform several tasks as risk mitigation measures against the high, medium, and low risks including scripting, system timer attacks, and block reporting messages attacks. This component is expected to manage such operations including static network configuration, restrict file and directory permissions, and others.

6.8 Security Monitoring

Intruders are expected to gain access somehow as observed in many attacks against highly secured industrial control systems [36, 73]. A specific requirement exists to integrate monitoring capabilities to detect unusual activities within the network and hosts. Proposed solutions according to the different DiD guidelines and the risk analysis are the application of Intrusion Detection and Prevention Systems, security information and event management (SIEM) systems as well as security audit logging.

6.8.1 Intrusion Detection and Prevention Systems

Intrusion Detection Systems (IDS) and or Intrusion Prevention Systems (IPS) are vital elements for maintaining the security of the APS network. Similar to a typical ICS network, the network traffic within the APS network is predictable, the communicating hosts, IP and MAC addresses, ports, protocols, etc should be known. Strong rules to detect unusual traffic should be feasible, but care must be paid when utilizing IPS since they may stop ongoing vital time-critical operations. Therefore, passive IDS are more favorable than IPS. At the same time, IPS can be utilized to take action against events with high confidence malicious ratings, especially during autonomous operations to reduce human involvement.

IDS are commonly utilized in vehicular systems including maritime vessels as indicated in a survey conducted by Loukas et al. [74]. However, the focus of such systems is mainly on GNSS spoofing and anomalies related to CAN bus protocol. The placement of IDS/IPS according to the DiD strategy is advised to be located in high traffic locations (i.e. connected to network switches) or between security boundaries (i.e. connected to a firewall) [91]. Figure 9 depicts a proposed arrangement of IDS in the APS ecosystem. In the APS network, all traffic between the cell security

zone (i.e. L2, L1, and L0) and L3 should be sent to the internal IDS/IPS with the main focus on inter-system traffic that should be predictable to some extent. Additionally, the conducted risk analysis has identified the need for special host-based IDS/IPS units hosted on the backup ANS and the backup AEMC to monitor traffic from the emergency controller onboard the ECT. Moreover, a main IDS/IPS on each of the APS and the RCC is expected to monitor and possibly control the traffic received from the gateways connected to the enterprise security zone. The IDS/IPS is expected to include capabilities for mitigating a group of identified threats. Such capabilities include Data Loss Prevention (DLP), Endpoint Denial of Service, Restrict Web-Based Content, and others.

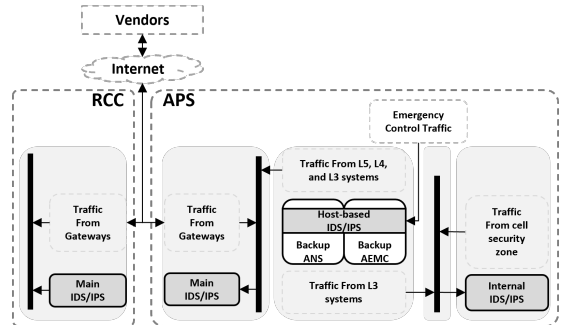


Fig. 9: proposed IDS/IPS architecture in the APS network

6.8.2 Security Incident and Event Monitoring

Logging and monitoring of security related-events is an integral aspect to detect and identify malicious attacks and is among the communicated requirements. Therefore, it is important to enable the logging feature on all the devices within the APS network and facilitate the collection of this information for processing through host-based agents. This feature can be used as one of the data sources for centralized Security Incident and Event Monitoring (SIEM) components. The role of each SIEM component includes but is not limited to monitoring and logging, it can even include detection and post-incident preparation [21]. The centralized SIEM can also receive IDS/IPS data to correlate with other data sources for the detection of possible security events. A possible implementation option is through the utilization of open source Elastic stack instruments [7].

Elastic stack has been proposed in the literature for providing SIEM functionality and more [67, 82].

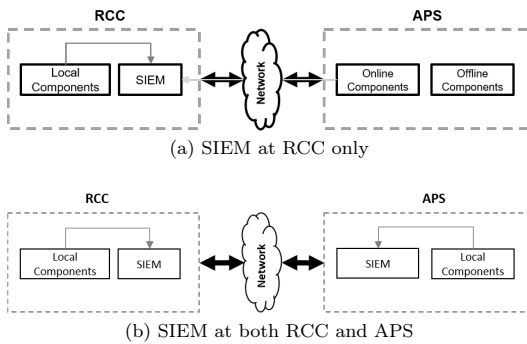


Fig. 10: Options for SIEM placement

During the system analysis process, we have evaluated different possible placements of the SIEM component within the APS ecosystem. The first considered option as shown in figure 10a, is one node at the RCC overseeing the APS network. This option; although it provides a single management location and reduced SIEM cost, has several shortcomings. First, all data collected from the host-based agents at the APS will need to be transferred over the network to the RCC which will consume valuable bandwidth required for critical functions. Second, some components within the APS are not expected to have external connectivity and therefore will not be managed by the SIEM at the RCC. Finally, the APS can be managed by different RCCs at different times, this setup will require re-configuration whenever the RCC in command is changed. The other option is shown in figure 10b, a dedicated SIEM at each of the APS and the RCC. Both nodes are configured to collect all the required information from the other components within their respective network through the host-based agents. This option limits the occupancy of the communication link for SIEM data requires no re-configuration at changing the RCC, and offline components are still managed by the local SIEM. At the same time, the shortcomings of this option are the increased management locations as well as increased SIEM cost.

Moreover, the APS is expected to host several components that are maritime-specific such as the AIS, NMEA speaking components, etc. The inclusion of such components within the coverage of the SIEM component would require various adaptations and domain-specific rules and alerts. This is proposed as an item for future work.

6.9 Vendor Management

Considerations regarding vendor management have been communicated in our previous work [24]. In the same direction, several policies and procedures are proposed by the different DiD guidelines including security in supply chain management, outsourcing, and leveraging cloud services. Establishing security requirements at early stages during the procurement process is proposed to be conducted to control the provisioning of services by third parties [47]. For example, a set of requirements could be specified when purchasing gateways to provide VPN capabilities, firewall capabilities as well as anti-spoofing capabilities. Additionally, several procedures are proposed through the conducted risk analysis to manage the risk in this direction. Such procedures include application vetting, updating software, audit, code signing, vulnerability scanning, and boot integrity. This component is proposed to aggregate the management of these activities.

6.10 Incident Response

The formulation of Incident response plans is a communicated requirement for the APS. Additionally, several DiD guidelines suggests activities related to incident response such as establishing contingency plans [43, 88], and data recovery [88]. Moreover, the conducted risk analysis has identified the utility of out-of-band communication channels as backup channels during incident response or in case of communication failure. This component is proposed to manage the different activities and aspects related to incident response such as the provisioning of appropriate incident response plans.

7 Cybersecurity Evaluation

The conducted SLR has captured the state of the art for the evaluation of cybersecurity controls in the maritime domain. We have identified existing approaches, aspects, scopes, and objectives. In this section, we will summarize our findings and utilize the observed approaches and aspects for evaluating the proposed risk management approach and its produced cybersecurity architecture.

7.1 Evaluation approaches

Several approaches have been observed with a distinct scope and varying levels of rigor. Table 4 depicts the observed approaches and evaluation environments. Com-

monly, approaches are combined for improving the evaluation process. Surveys through checklists and questionnaires are the most common method. Surveys focus on aspects including market research; quality and cost of controls [28], stakeholder engagement [53], usability and quality of risk assessment frameworks [22, 99], existence, revision and awareness of controls [98]. Risk evaluation including risk and residual risk estimation and assessment is usually combined with another approach such as assessing the risks in scenarios using a game-based approach [69] or through questionnaires [22, 28]. Some works evaluate the performance of target systems to assess their security or the impact of security controls on their operation. This is observed to be achieved through emulating the behaviour of adversaries (i.e. adversary emulation) [29, 57] or testing the functionality of a specific functional unit (i.e. unit testing) [53, 69, 71, 106]. Other works target the existing vulnerabilities in the system as an indicator of the efficiency of its security posture [98, 103]. Kuhn et al [68] carried out an exercise for assessing the risk perception of participants for evaluating their decision-making capabilities in cyber incident response. McCreedy et al [78] researched relevant standards and regulations for evaluating the utility, feasibility, and aspects of a maritime compliance regime as an organizational control for improving the cybersecurity posture of organizations in the maritime domain. The authors also indicated the importance of record-keeping for the evaluation of the cybersecurity posture for audit purposes.

Moreover, the approaches vary based on the system development life cycle of the target system of evaluation. This is reflected in the environment used for evaluation. Some works addressing high-level controls such as compliance regimes [78], and risk assessment frameworks [99, 103] utilize abstract descriptions of the target systems including relevant organizations, facilities, and utilized technologies for qualitative evaluation. Other model-based approaches tend to be theoretical with the capacity for simulation [28, 46]. On the other hand, in more advanced system development phases, approaches tend to consider more realistic settings reaching the ability to evaluate the real target system [53, 98] and on some occasions simulating certain elements in the environment [63, 93]. Additionally, some works address the software source code, middleware and hosting operating system for evaluating its security [69, 71, 106].

All the observed approaches rely on a specific threat model encompassing several elements such as target system, threat or attack, mitigation methods, and others. Some approaches are more defensive as they mostly address the established risk mitigation measures or the existence of vulnerabilities in the target system [98], in

Table 4: Evaluation approaches, and environments

Method	References	Environment	References
Risk Evaluation	[28, 46, 93] [22]	Simulation/System model	[28, 57, 68] [46]
Survey	[28, 98, 99] [22, 53]	Real System	[53, 98]
Vulnerability scanning and assessment	[98, 103]	SW	[69, 71, 106]
Performance evaluation	[53, 63, 106] [69, 71]	Simulation/Real Systems	[29, 63, 93]
Unit testing	[53, 71, 106] [69]	Abstract	[22, 78, 99] [103]
Exersize	[68]		
Gamefication	[68, 93]		
Research	[78]		
Record keeping	[78]		
Adversary Emulation	[29, 57]		

other words, blue team activities. Other approaches are more offensive as they aim to evaluate the behavior of the target system against specific adversarial activities [29, 57], in other words, red team activities. Other approaches consider both perspectives [28], similar to the concept of purple teaming, aiming to establish a wider view of the cybersecurity posture of the target system and its risk management capabilities.

7.2 Aspects, scopes and test objectives

A wide range of aspects has been observed encompassing security, privacy, functional, cost, and governance aspects. Table 5 depicts a summary of the observed aspects, their scope, and the test objectives of the system analysis. Regarding the scope, some works have addressed the entire security posture of the organization including the organization, stakeholders, systems, and services. Other works target a certain system such as a ship or the navigation system with or without knowledge of the included security controls. Others target specific security controls such as incident response, policies, and procedures. Regarding the targeted aspects of evaluation, several security aspects are observed including effectiveness, existence, awareness, and revision of controls, the existence of vulnerabilities, requirement satisfaction, and recommendations. Testing the effectiveness of controls is the main goal of evaluation with varying objectives. Some works targeted quantitative metrics such as detection quality, number of vulnerabilities over time, and successful logins. Others targeted qualitative aspects such as the effect of experience on incident response, or how the security control would respond to attacks. Also, evaluating the satisfaction of the stakeholders' requirements related to security and privacy is a common objective. Additionally, the functional aspect is addressed by several works focusing on

the behavior of the system under attack, the integration of controls within the system, and the usability, feasibility, and applicability of the security controls assessed by the system stakeholders. Moreover, the financial and operational cost of controls during different system development life cycles is also addressed. Finally, aspects related to governance such as roles and responsibilities, the obligation of application, penalties of non-compliance, and assessment frequency have also been investigated.

7.3 Cybersecurity Evaluation of the APS use case

Considering the current technology readiness level of the APS use case, there are several aspects and test objectives that are relevant for evaluation. The scope of the evaluation extends to the proposed risk management strategy (Section 3), and the security architecture produced after its application (Section 6).

Proper evaluation of the risk management strategy can be conducted over time by observing the efficiency of the developed cybersecurity architecture. However, its feasibility and usability have been evaluated. The risk management process was initially conducted against a system model of the APS, this has led to the development of the security architecture in section 6. Moreover, another iteration of the process was conducted against the implemented prototype which is the milliAmpere2. The risk assessment process identified a group of risks and required controls that were later integrated into the ferry and the RCC. This reflects the suitability of the process for applications in different system life cycle stages.

The next subject of evaluation is the proposed cybersecurity architecture. The evaluation was conducted again for the developed model as well as the implemented ferry. The evaluation was conducted using several methods, namely, risk evaluation, simulation and checklist, and adversary emulation.

7.3.1 Risk Evaluation

We have implemented our proposed risk assessment approach [23] for the estimation of residual risk before and after the integration of the security controls. The risk evaluation was conducted for two system definitions. The first one is the model of the APS communication architecture discussed in Section 5. The other one is for a model of the implemented milliAmpere2 ferry. The cybersecurity architecture can improve the risk reduction from 0,84% to 81,94% for the APS communication architecture model, and from 58,37% to 85,72% for the milliAmpere2 ferry. The deficiency in the risk

reduction value is mostly related to risks with no existing or limited controls such as resource hijacking and radio jamming; this is inferred from the fact that the *ATT&CK* framework designated such risks to have no or limited existing controls. The risk reduction in the milliAmpere2 before the cybersecurity architecture is due to the existence of controls such as network segmentation, physical security, firewalls, and several others. However, the existing controls weren't sufficient for addressing critical to medium risks.

7.3.2 Simulation and Checklist

An observed method for evaluating the cybersecurity architecture is to check its satisfaction with the cybersecurity requirements. Yi and Kim [106] discussed the evaluation of naval ship combat software against a set of specified technical requirements related to accuracy and adequacy. Additionally, Grigoriadis et al [53] reached out to stakeholders for evaluating their risk assessment process against security, privacy, operational, and usability requirements. As mentioned before, we have identified a group of cybersecurity requirements for the APS [24]. These requirements are then utilized for evaluating the security architecture based on the verification criteria defined during the architecture design (Section 3.2.1). The evaluation was conducted against two system definitions, namely, a simulated cybersecurity architecture, and the implemented milliAmpere2 ferry.

In this paper, simulation is utilized to verify the feasibility of integrating the cybersecurity architecture within the underlying communication architecture and to facilitate later security analysis. A prototype implementation of the IP-based components is provided using the GNS3 simulator [83]. GNS3 (Graphical Network Simulator-3) is a platform for emulating appliances (network, endpoints, etc.) using virtualized images. It enables the configuration, testing, and development of networks with flexibility and lower cost [83]. Later, cybersecurity controls proposed in the cybersecurity architecture in Section 6 were integrated using a variety of open-source and off-the-shelf controls to evaluate their feasibility and suitability for the auto-remote operational mode. Later, the simulated architecture is evaluated for its satisfaction with the requirements. A summary of the verification process including the design-level and implementation-level verification criterion as well as details regarding the supporting components and conducted processes are shown in Table 7 in Appendix B. Detailed description of the simulated network, integrated controls, and attack trees used for evaluation are presented in Appendix D. Ac-

Table 5: Aspects, scopes and test objectives

Category	Aspect	scope	Test objective
Security	Effectiveness	Cybersecurity posture [78] Control [28, 46, 53, 63, 68, 71, 93, 98] Control and Host system [29, 46, 57, 69]	Number of vulnerabilities over time [78]
			Flatness, successful logins [53]
			Defense strategy and risk level [46, 69, 93]
			Will the controls work [53, 63]
			How does the controls work [68]
			How experience affects the control [68]
			Level of Confidentiality, Integrity, and Availability [63]
			Detection quality [63]
			Accuracy, precision, recall, F1-score [71]
			Deterrence [28]
Can attacks be mitigated [57]			
Real time defense [28]			
Restoration [28]			
Existence of controls	Controls [98]	Incident handling and reporting [98] Policies and procedures [98]	
Controls awareness			
Controls revision			
Vulnerabilities	Host system [57, 98] Host system and Controls [103] Application SW [106]	What vulnerabilities exist [57, 98, 106] What errors and defects exist [103]	
Requirements Satisfaction	Controls [53] Application SW, Middleware, OS [106] Host system and Controls [103]	Are the security requirements satisfied [53, 103, 106]	
Recommendations	Controls [53]	e.g. cryptographic strength	
Privacy	Requirements Satisfaction	Controls [53]	Are the privacy requirements satisfied [53]
Functional	Behavior	Host system [57]	How would the system behave under attack [57]
	Integration	Controls [53] Application SW, Middleware, OS [106]	Are the control properly integrated [53, 106]
	Usability	Controls [53, 99]	User acceptance testing [53, 99]
	Feasibility and Applicability	Control [78, 99]	Compliance regime [78] risk assessment framework [99]
Cost	Financial	Controls [22, 28]	Cost of implementation [22, 28] Cost saved by reducing risk [22]
	Operational	Against host system (i.e. safety) [28, 53, 63] Controls [53, 63, 71]	Execution Time [53, 71] Packet loss, delay [28] Overhead [53] Harmlessness and Data lost [63]
	Lifecycle	Control [28]	Future developmnet costs [28]
Governance	Responsibility	Compliance regime [78] Assessment Organization [103]	Enforcement, auditing and reporting [78] Assessment [103]
	Obligation	Compliance regime [78]	Mandatory or voluntary [78]
	Penalties		What are the panalties for non-compliance [78]
	Assessment Frequency		What is the period between audits [78]

cess to our simulated network can be provided upon request.

The design-level verification is of low fidelity and only intended to verify the feasibility of the cybersecurity architecture in the APS design model and simulated implementation. It demonstrated the feasibility of the model for implementation and shed some light on the considerations regarding the provisioning of risk management functions within the auto-remote operational mode. More details are discussed in Section 8.

On the other hand, the implementation-level verification has shown some limitations in the cybersecurity posture of the milliAmpere2 ferry. Due to the involvement of several technology and service vendors, some cybersecurity controls have implementation gaps and limited information regarding their details. The most critical issue observed is related to regular software updates. A high-priority requirement exists to enforce regular software updates for components in the APS network. However, our evaluation uncovered that

some components have outdated software versions and no existing process for updating them. Another issue that has been identified is related to the lack of training exercises related to cybersecurity. Efforts are planned in this regard and are expected to be items for future work. Moreover, the inclusion of some security controls has been found to be not feasible in the current implementation. Some components are protected from manipulation through agreements with vendors which limited the ability to install agent software for a dedicated SIEM and HIDS software. Therefore, reliance on NIDS is considered an alternative. Furthermore, a requirement exists related to the network topology to avoid including components used for navigation and control in the same network. However, it was found that this requirement is not satisfied.

In summary, the simulation was useful in reducing the cost of implementation at the real ferry as well as for trying out several implementing options at a lower cost. However, contextual information such as access limita-

tions to some components was not considered during the simulation. The checklist approach uncovered several limitations in the implemented cybersecurity architecture of the ferry allowing for future improvements.

7.3.3 Adversary Emulation

Having access to the implemented ferry allows for conducting hands-on adversary emulation; a security assessment process applying realistic attack scenarios which emulate the capabilities of real threat actors [97]. Several works in the literature have applied it in demonstrating and evaluating the security of maritime systems. Balduzzi [29] conducted various attacks against AIS protocol and some of its implementation to evaluate its security. Also, Hemminghaus [57] proposed a tool for automating several attacks against integrated bridge systems to evaluate the established security controls. Adversary emulation is another instrument of the threat-informed defense strategy that utilizes the *ATT&CK* of the APS. The DiD has been challenged previously for its ineffectiveness against sophisticated attacks [49]. These findings are confirmed in this work when discussing protocol-specific controls related to Non-IP communication in Section 6.6.2. We have identified the need for a dedicated anomaly detection solution for the NMEA protocol since traditional IDS systems are not tuned to detect anomalies for this specific protocol. This supports the argument that a DiD approach that relies on stacking up controls without proper evaluation of the threat landscape would risk the protected system against sophisticated attacks. Additionally, using simulation we have evaluated and demonstrated the utility of the proposed cybersecurity architecture in withstanding several cyber attacks. We have observed limited discussion regarding system hardening activities in several DiD guidelines while several suggestions in this direction are provided by the *ATT&CK*-based risk assessment. Additionally, the threat-informed defense strategy provided several instruments that enriched the risk management process and aided the development of the cybersecurity architecture. These instruments include the *ATT&CK* framework, and the defensive engagement of threats using adversary emulation. Both instruments are constantly updated from CTI feeds which allows the architecture to constantly evolve in order to match the latest threat landscape. However, some limitations are observed in the defensive functions proposed in *ATT&CK* such as the lack of clear interfaces between threats and incident response functions as well as the lack of high-level risk management elements such as roles and responsibilities. These findings indicate that our proposed threat-informed defense-in-depth risk management approach does provide improved

The tests are intended to be comprehensive, covering all the attackers' objectives (i.e. tactics) proposed in the *ATT&CK* framework and a wide range of techniques and software to implement them. For each tactic, at least one test was planned and developed. The tests were prioritized based on those identified as critical to medium risks and those which are technically feasible for testing. A summary of the planned and conducted tests is depicted in Table 8 in Appendix C. Several tests were not allowed to be carried out due to access limitations by the network vendors. The conducted tests have yielded useful information that will be used for improving the cybersecurity posture of the ferry ecosystem. This includes the discovery of critical vulnerabilities and a large number of open network services. Utilizing the *ATT&CK* framework enriched the adversary emulation process by enabling the development of atomic tests in a systematic manner. Still, a wide range of tests is needed to evaluate the entire architecture.

The utility of the adversary emulation process has been demonstrated. Conducting it at several iterations

is needed to maintain accurate risk awareness. However, it comes with a high cost in time and resources. Therefore, efforts to automate some of the tests and expand their scope are items for future work.

8 Discussion

In this section, we present our reflections after conducting the different research activities presented in this paper. We discuss the challenges and limitations observed in the different applied approaches. Also, we present our observations regarding the provisioning of risk management functions within the autoremove operational mode.

Starting with the risk management strategies. The DiD and the threat-informed defense as risk management strategies are evaluated in this work through the integration of the different elements in the strategies toward the development of the cybersecurity architecture for its ineffectiveness against sophisticated attacks [49]. These findings are confirmed in this work when discussing protocol-specific controls related to Non-IP communication in Section 6.6.2. We have identified the need for a dedicated anomaly detection solution for the NMEA protocol since traditional IDS systems are not tuned to detect anomalies for this specific protocol. This supports the argument that a DiD approach that relies on stacking up controls without proper evaluation of the threat landscape would risk the protected system against sophisticated attacks. Additionally, using simulation we have evaluated and demonstrated the utility of the proposed cybersecurity architecture in withstanding several cyber attacks. We have observed limited discussion regarding system hardening activities in several DiD guidelines while several suggestions in this direction are provided by the *ATT&CK*-based risk assessment. Additionally, the threat-informed defense strategy provided several instruments that enriched the risk management process and aided the development of the cybersecurity architecture. These instruments include the *ATT&CK* framework, and the defensive engagement of threats using adversary emulation. Both instruments are constantly updated from CTI feeds which allows the architecture to constantly evolve in order to match the latest threat landscape. However, some limitations are observed in the defensive functions proposed in *ATT&CK* such as the lack of clear interfaces between threats and incident response functions as well as the lack of high-level risk management elements such as roles and responsibilities. These findings indicate that our proposed threat-informed defense-in-depth risk management approach does provide improved

risk management capabilities by combining both strategies.

Regarding the cybersecurity evaluation. The difference in the evaluation results between the different evaluation methods highlights the importance of diversifying evaluation processes. Some approaches are less costly to implement (e.g. risk-based evaluation), however, their fidelity and accuracy have been questioned. An instance of this has been observed in this work. The risk evaluation assumes that if a mitigation method exists, it is sufficient to reduce the risk. However, the adversary emulation process has uncovered several discrepancies. For instance, a password policy exists regarding default credentials, during the risk assessment this information has rendered all relevant threats to be of negligible risks. During the adversary emulation process, 2 devices with default credentials were found. Hence, inaccurate risk assessment. This issue showcases a usability issue of the risk assessment process, it requires a lot of information that is not easily and readily available, such as the correct status of cybersecurity control coverage for all components in the network. Without active adversary emulation, knowing this with high confidence is not possible for all threats.

The simulated cybersecurity architecture implemented during the system analysis stage has highlighted several challenges in conducting the security functions in the context of the autoremove operational mode. The systems onboard the APS will need to be managed remotely due to the crew-less nature of the APS. This dictates the need to enable a remote management solution. One example is implemented through the installation of a remote desktop service for all components to facilitate their maintenance. This has been observed to be the case in the implemented milliAmper2 ferry. Another solution may include the utilization of Secure Shell Protocol (SSH). However, these particular solutions make the APS susceptible to remote attacks if the proposed cybersecurity architecture is not adopted. From the perspective of the security solutions, it has been observed that solutions that only function through a graphical user interface (GUI) are challenging to manage when integrated within the APS network. So, solutions that provide Command Line Interface (CLI) are more suitable to facilitate their automation and remote management. Additionally, the proposed network architecture by DiD in the context of the autoremove operational mode (refer to Section 6.5). Therefore, we proposed the utilization of VPN tunnels to extend the perimeter of network security levels to span across different facilities. Further analysis of the impact of this proposition on the control and monitoring functions in the APS is within the scope of future work. More-

over, the GNS3 has provided very useful capabilities to demonstrate the feasibility of implementing and integrating the different components. However, we have observed drastic network latency and packet loss which is linked to the nested virtualization capabilities. This drawback has led us to consider migrating the implementation to another platform for future work.

Finally, we acknowledge the following limitations in our proposed approaches and their application:

- Defense strategy comparison algorithm: The comparison is only based on the risk reduction without considering the cost of implementation. Also, the calculation relies on the controls in the *ATT&CK* framework. Some controls in DiD do not map to a clear control in the *ATT&CK* framework. Therefore, some controls do not account for a risk reduction such as "Establish contingency plans" (Other examples can be found in Table 6).
- Cybersecurity evaluation: The simulation of the cybersecurity architecture relied on a group of commonly used open-source or free tools. Some tools are referenced in the literature such as the elastic stack for the SIEM component while others were chosen only for practical and compatibility reasons. On the other hand, the adversary emulation processes were restricted to allowed and feasible tests against the ferry which is only a small subset of the required tests for effective and comprehensive evaluation. Future works can investigate solutions to such limitations.

9 Conclusion

The ongoing digital transformation in the maritime domain has produced novel technologies and modes of operation. An instance of this is the proposition of Autonomous Passenger Ships (APS) for inland transportation. The APS technology is projected to operate under autoremove operational mode; autonomous when possible, remotely controlled when needed. With the recent calls for introducing cyber risk management capabilities in the maritime domain, investigating suitable means for the provisioning of cyber risk management functions for APS is a raising need. This paper investigates cyber risk management for the APS technology. Our research methodology follows a system engineering approach for the application of risk management functions during different system development phases. The approach relies on both the perspective of the classification society in the maritime domain as well as academia. The Defense-in-Depth (DiD) is observed to be an agreed-upon strategy for providing risk management capabilities.

ties. However, some limitations regarding its implementation for defending critical systems have been communicated. Therefore, we proposed a new risk management approach combining DiD with another risk management strategy named threat-informed defense. Our proposed approach has been demonstrated to expand the provisioning of risk management functions using those proposed in the *ATT&CK* framework. Our approach is demonstrated through the development of a cybersecurity architecture for an APS use case. Afterward, a Systematic Literature Review (SLR) for cybersecurity evaluation in the maritime domain has been conducted and its results are presented. Observed approaches and artifacts from the SLR are then utilized for the evaluation of the proposed cybersecurity architecture for the APS. The evaluation has been conducted utilizing a system model as well as a real implemented prototype of an APS named milliAmper2. Several evaluation approaches have been deemed relevant to the current technology readiness level of the APS technology, namely, risk-based evaluation, simulation and checklist, and adversary emulation. Risk evaluation reflects that the risk reduction of the proposed architecture is close to the optimal score considering it addresses the requirements as well as the identified risks, rendering other controls suggested in certain guidelines as non-critical. The adversary emulation and checklist evaluation approaches have identified vulnerabilities and unaddressed risks which have been observed to be a metric of successful risk management strategy. The simulation has uncovered challenges and considerations regarding the provisioning of risk management in the auto-remote operational mode. This includes network segmentation, the reliance on non-IP communication, as well as the placement of SIEM components.

Moreover, the risk assessment process; which is an integral component of the proposed risk management approach, have identified new threats with varying level of sophistication. The proposed cybersecurity architecture has been tuned toward addressing such threats. For instance, a technology-specific intrusion detection system has been proposed and investigated in another work based on the work in this paper. This suggests that the communicated concerns regarding the limitations of DiD in defending against sophisticated attacks are addressed in the proposed architecture.

Several items have been identified suggesting the need for future work. Considering that autonomous vessels rely on machine learning and artificial intelligence, there is a lack of discussion related to the threat of adversarial machine learning in the maritime domain. Also, including other aspects in the strategy comparison algorithm such as the financial aspect for improved

strategy analysis. Additionally, the inter-relations between safety and cybersecurity functions within the context of the APS require additional attention. Finally, additional work is needed relating to the adoption and automation of cybersecurity evaluation methods toward reducing the involvement of the human element.

Compliance with Ethical Standards

- Funding: Not applicable.
- Conflict of Interest: Ahmed Amro declares that he has no conflict of interest. Vasileios Gkioulos declares that he has no conflict of interest.
- Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors

Research Data Policy and Data Availability

An implementation of the strategy comparison algorithm (Section 4) with the data used to generate the results in this paper can be found at the author's online repository [12].

References

1. Aps communication architecture aadl model. <https://github.com/ahmed-amro/APS-CommunicationArchitecture.git>. Accessed : 2022 – 06 – 10.
2. Atomic red team. <https://github.com/redcanaryco/atomic-red-team>.
3. Borgbackup, deduplicating archiver with compression and encryption. <https://www.borgbackup.org/>. Accessed: 11.10.2021.
4. Clamav an open-source antivirus engine. <https://www.clamav.net/>. Accessed: 11.10.2021.
5. Domestic transport. <https://www.ssb.no/en/transport-og-reiseliv/statistikker/transpinn>. Accessed 11.10.2021.
6. Duo security - two factor authentication. <https://duo.com/>. Accessed 11.10.2021.
7. Elk stack: Elasticsearch, logstash, kibana. <https://www.elastic.co/what-is/elk-stack>. Accessed 11.10.2021.
8. FusionInventory - the opensource it inventory solution. <https://fusioninventory.org/>. Accessed 11.10.2021.

9. Gestionnaire libre de parc informatique (glpi). <https://glpi-project.org/>. Accessed 11.10.2021.
10. Nfas - norwegian projects. <https://nfas.autonomous-ship.org/resources/page/projects - page/>.
11. Sea passenger statistics 2020: Short sea routes. <http://bit.ly/PassengerStatistics2020>. Accessed 11.10.2021.
12. Strategy comparison algorithm. https://github.com/ahmed-amro/APS-Communication_Architecture/tree/master/RPNMI/Strategy_Comparison_Algorithm.
13. *Systems and Software Engineering - System Life Cycle Processes*. Geneva, Switzerland: International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC)/Institute of Electrical and Electronics Engineers. ISO/IEC 15288:2015. .
14. Threat-based defense. <https://www.mitre.org/capabilities/cybersecurity/threat-based-defense>.
15. Wazuh - the open source security platform. <https://wazuh.com/>. Accessed 11.10.2021.
16. How mitre attck alignment supercharges your siem. <https://www.securonix.com/how-mitre-attack-alignment-supercharges-your-siem/>, 2019.
17. Transportation statistics annual report 2020. <https://www.bts.gov/tsar>, December 2020.
18. Dnvgl-rp-0496 recommended practice: Cyber security resilience management for ships and mobile offshore units in operation. <https://www.dnv.com/maritime/dnvgl-rp-0496-recommended-practice-cyber-security-download.html>, 2021. Accessed on 16.02.2022.
19. Enhancing with mitre. <https://documentation.wazuh.com/current/user-manual/ruleset/mitre.html>, 2021.
20. Enisa threat landscape 2021. <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021>, 2021.
21. Nurul Hidayah Ab Rahman and Kim-Kwang Raymond Choo. A survey of information security incident handling in the cloud. *computers & security*, 49:45–69, 2015.
22. MD Abkowitz and JS Camp. An application of enterprise risk management in the marine transportation industry. *WIT Transactions on The Built Environment*, 119:221–232, 2011.
23. A. Amro, V. Gkioulos, and S. Katsikas. Assessing cyber risk in cyber-physical systems using the *att&ck* framework. Preprint available online at <http://dx.doi.org/10.13140/RG.2.2.16531.40484>. Submitted for review to ACM Transactions on Privacy and Security (TOPS), 2021.
24. Ahmed Amro, Vasileios Gkioulos, and Sokratis Katsikas. Connect and protect: Requirements for maritime autonomous surface ship in urban passenger transportation. In *Computer Security*, pages 69–85. Springer, 2019.
25. Ahmed Amro, Vasileios Gkioulos, and Sokratis Katsikas. Communication architecture for autonomous passenger ship. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, page 1748006X211002546, 2021.
26. Ahmed Amro, Aybars Oruc, Vasileios Gkioulos, and Sokratis Katsikas. Navigation data anomaly analysis and detection. *Information*, 13(3), 2022.
27. Ahmed Aziz, Pietro Tedeschi, Savio Sciancalepore, and Roberto Di Pietro. Secureais-securing pairwise vessels communications. In *2020 IEEE Conference on Communications and Network Security (CNS)*, pages 1–9. IEEE, 2020.
28. Guy L Babineau, Rick A Jones, and Barry Horowitz. A system-aware cyber security method for shipboard control systems with a method described to evaluate cyber security solutions. In *2012 IEEE Conference on Technologies for Homeland Security (HST)*, pages 99–104. IEEE, 2012.
29. Marco Balduzzi, Alessandro Pasta, and Kyle Wilhoit. A security evaluation of ais automated identification system. In *Proceedings of the 30th annual computer security applications conference*, pages 436–445, 2014.
30. Matthew P Barrett. Framework for improving critical infrastructure cybersecurity. *National Institute of Standards and Technology, Gaithersburg, MD, USA, Tech. Rep*, 2018.
31. Daniel Blauwkamp, Thuy D Nguyen, and Geoffrey G Xie. Toward a deep learning approach to behavior-based ais traffic anomaly detection. In *Dynamic and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS) Workshop, San Juan, PR*. Retrieved from http://faculty.nps.edu/Xie/papers/ais_analysis_18.pdf, 2018.
32. Clet Boudehenn, Olivier Jacq, Maxence Lannuzel, Jean-Christophe Cexus, and Abdel Boudraa. Navigation anomaly detection: An added value for maritime cyber situational awareness. In *2021 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, pages 1–4. IEEE, 2021.
33. Jon Boyens, C Paulsen, Rama Moorthy, Nadya Bartol, and S Shankles. Nist special publication 800-161: Supply chain risk management prac-

- tics for federal information systems and organizations. *NIST*. April, 2015.
34. Bureau Veritas. Ni641 guidelines for autonomous shipping. 2019.
 35. D Chappelle. Security in depth reference architecture release 3.0. *White paper, Oracle Corporation, Redwood Shores*, 2013.
 36. Anton Cherepanov. Win32/industroyer: A new threat for industrial control systems. *White paper, ESET (June 2017)*, 2017.
 37. Howard Chu. LDAP. Washington, D.C., December 2006. USENIX Association.
 38. Cisco. *RV0xx Series Routers, ADMINISTRATION GUIDE*, 2021 (accessed May 13, 2021). <http://bit.ly/RV042>.
 39. IEC 60812 Technical Committee et al. Analysis techniques for system reliability-procedure for failure mode and effects analysis (fmea). 2018.
 40. The Maritime Safety Committee. Interim guidelines on maritime cyber risk management (msc-fal.1/circ.3/rev.1). <https://cutt.ly/6R8wqjN>.
 41. The Maritime Safety Committee. International maritime organization (imo) (2017) guidelines on maritime cyber risk management. <http://bit.ly/MSC428-98>.
 42. Pierre de Saqui-Sannes, Jérôme Hugues, et al. Combining sysml and aadl for the design, validation and implementation of critical systems. *ERTS 2012*, 2012.
 43. DNV. Ddnvgl-cg-0325: Cyber secure class notation. <https://rules.dnvgl.com/docs/pdf/DNVGL/CG/2020-10/DNVGL-CG-0325.pdf>, 2020.
 44. DNV GL. Dnvgl-cg-0264: Autonomous and remotely operated ships. 2018.
 45. Athanasios Drougkas, Anna Sarri, Pinelopi Kyranoudi, and EU Agency for Cybersecurity. Guidelines - cyber risk management for ports. <https://www.enisa.europa.eu/publications/guidelines-cyber-risk-management-for-ports>, 12 2020.
 46. Simon Yusuf Enoch, Jang Se Lee, and Dong Seong Kim. Novel security models, metrics and security assessment for maritime vessel networks. *Computer Networks*, 189:107934, 2021.
 47. M Fabro, E Gorski, N Spiers, J Diedrich, and D Kuipers. Recommended practice: improving industrial control system cybersecurity with defense-in-depth strategies. *DHS Industrial Control Systems Cyber Emergency Response Team*, 2016.
 48. Peter H Feiler, David P Gluch, and John J Hudak. The architecture analysis & design language (aadl): An introduction. Technical report, Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst, 2006.
 49. Andrew Fielder, Tingting Li, and Chris Hankin. Defense-in-depth vs. critical component defense for industrial control systems. In *4th International Symposium for ICS & SCADA Cyber Security Research 2016 4*, pages 1–10, 2016.
 50. Markus Fruth and Frank Teuteberg. Digitization in maritime logistics—what is there and what is missing. *Cogent Business Management*, 4(1):1411066, 2017.
 51. A Goudosis and SK Katsikas. Secure ais with identity-based authentication and encryption. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 14(2), 2020.
 52. Athanassios Goudossis and Sokratis K Katsikas. Towards a secure automatic identification system (ais). *Journal of Marine Science and Technology*, 24(2):410–423, 2019.
 53. Christos Grigoriadis, Spyridon Papastergiou, Panayiotis Kotzanikolaou, Christos Douligeris, Antreas Dionysiou, Athanasopoulos Elias, Karin Bernsmed, Per Håkon Meland, and Liina Kamm. Integrating and validating maritime transport security services: Initial results from the cs4eu demonstrator. In *2021 Thirteenth International Conference on Contemporary Computing (IC3-2021)*, pages 371–377, 2021.
 54. Kory Hamzeh, Grueep Pall, William Verthein, Jeff Taarud, W Little, and Glen Zorn. Point-to-point tunneling protocol (pptp), 1999.
 55. Gina Havdal, Christina Torjussen Heggelund, and Charlotte Hjelmseth Larssen. Design of a small autonomous passenger ferry. Master's thesis, NTNU, 2017.
 56. Americas Headquarters. Cisco safe reference guide. 2009.
 57. C Hemminghaus, J Bauer, and E Padilla. Brat: A bridge attack tool for cyber security assessments of maritime systems. 2021.
 58. Siv Hilde Houmb, Virginia NL Franqueira, and Erleend A Engum. Quantifying security risk level from cvss estimates of frequency and impact. *Journal of Systems and Software*, 83(9):1622–1634, 2010.
 59. Clément Iphar, Cyril Ray, and Aldo Napoli. Data integrity assessment for maritime anomaly detection. *Expert Systems with Applications*, 147:113219, 2020.
 60. ISO. Iec/ieee 15288: 2015. *Systems and software engineering-Content of systems and software life cycle process information products (Documenta-*

- tion), *International Organization for Standardization/International Electrotechnical Commission: Geneva, Switzerland*, 2015.
61. ISO. Iso 31000:2018 risk management — guidelines, 2018.
 62. IEC ISO. Ieee: Iso/iec/ieee 42010: 2011-systems and software engineering—architecture description. *Proceedings of Technical Report*, 2011.
 63. Olivier Jacq, Xavier Boudvin, David Brosset, Yvon Kermarrec, and Jacques Simonin. Detecting and hunting cyberthreats in a maritime environment: Specification and experimentation of a maritime cybersecurity operations centre. In *2018 2nd Cyber Security in Networking Conference (CSNet)*, pages 1–8. IEEE, 2018.
 64. Georgios Kavallieratos and Sokratis Katsikas. Managing cyber security risks of the cyber-enabled ship. *Journal of Marine Science and Engineering*, 8(10):768, 2020.
 65. GC Kessler. Protected ais: a demonstration of capability scheme to provide authentication and message integrity. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 14(2), 2020.
 66. Fabrice Kordon, Jérôme Hugues, Agusti Canals, and Alain Dohet. *Embedded systems: analysis and modeling with SysML, UML and AADL*. John Wiley & Sons, 2013.
 67. Igor Kotenko, Artem Kuleshov, and Igor Ushakov. Aggregation of elastic stack instruments for collecting, storing and processing of security information and events. In *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, pages 1–8. IEEE, 2017.
 68. Kristen Kuhn, Salih Bicakci, and Siraj Ahmed Shaikh. Covid-19 digitization in maritime: understanding cyber risks. *WMU Journal of Maritime Affairs*, pages 1–22, 2021.
 69. Tazim Ridwan Billah Kushal, Kexing Lai, and Mahesh S Illindala. Risk-based mitigation of load curtailment cyber attack using intelligent agents in a shipboard power system. *IEEE Transactions on Smart Grid*, 10(5):4741–4750, 2018.
 70. Yin Lam. Technology will help maritime transport navigate through the pandemic—and beyond. <https://blogs.worldbank.org/transport/technology-will-help-maritime-transport-navigate-through-pandemic-and-beyond>, November 2020. Accessed 05.01.2022.
 71. Ha V Le, Tu N Nguyen, Hoa N Nguyen, and Linh Le. An efficient hybrid webshell detection method for webserver of marine transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 2021.
 72. Dong-Kyeonz Lee, Damian Miralles, Dennis Akos, Andriy Konovaltsev, Lothar Kurz, Sherman Lo, and Filip Nedelkov. Detection of gnss spoofing using nmea messages. In *2020 European Navigation Conference (ENC)*, pages 1–10. IEEE, 2020.
 73. RM Lee and MJ Assante. Analysis of the cyber attack on the ukraine power grid. In *E-ISAC and SANS*. White, 2016.
 74. George Loukas, Eirini Karapistoli, Emmanouil Panaousis, Panagiotis Sarigiannidis, Anatolij Bezemskij, and Tuan Vuong. A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles. *Ad Hoc Networks*, 84:124–147, 2019.
 75. Gordon Fyodor Lyon. *Nmap network scanning: The official Nmap project guide to network discovery and security scanning*. Insecure. Com LLC (US), 2008.
 76. Vasileios Mavroeidis and Siri Bromander. Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In *2017 European Intelligence and Security Informatics Conference (EISIC)*, pages 91–98. IEEE, 2017.
 77. D McCallam. An analysis of cyber reference architectures. In *Presented at NATO 2012 Workshop with Industry on Cybersecurity Capabilities*, 2012.
 78. John W McCready, Winnie Callahan, David Mayhew, and Mark Heckman. Toward a maritime cyber security compliance regime. In *SNAME Maritime Convention*. OnePetro, 2018.
 79. Ioan-Cosmin Mihai, Stefan Pruna, and Ionut-Daniel Barbu. Cyber kill chain analysis. *Int'l J. Info. Sec. & Cybercrime*, 3:37, 2014.
 80. MITRE. Threat-informed defense. <https://www.mitre.org/news/focal-points/threat-informed-defense>. Accessed 05.01.2022.
 81. MITRE. *Chimera, Group G0114*, 2021 (accessed May 11, 2021). <https://attack.mitre.org/groups/G0114/>.
 82. Moukafih Nabil, Sabir Soukainat, Abdelmajid Lakbabi, and Orhanou Ghizlane. Siem selection criteria for an efficient contextual security. In *2017 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–6. IEEE, 2017.

83. Jason C Neumann. *The book of GNS3: build virtual network labs using Cisco, Juniper, and more*. No Starch Press, 2015.
84. NTNU Autoferry. Autoferry - Autonomous all-electric passenger ferries for urban water transport. <https://www.ntnu.edu/autoferry>, 2018.
85. Chitu Okoli and Kira Schabram. A guide to conducting a systematic literature review of information systems research. 2010.
86. Priyanga Rajaram, Mark Goh, and Jianying Zhou. Guidelines for cyber risk management in ship-board operational technology systems. *arXiv preprint arXiv:2203.04072*, 2022.
87. CORE Ramboll. Advokatfirma: Analysis of regulatory barriers to the use of autonomous ships: Final report. *Danish Maritime Authority, Copenhagen*, pages 1374–1403, 2017.
88. DK Rasmus Nord Jorgensen in Copenhagen. Bimco: The guidelines on cyber security onboard ships. <https://iumi.com/news/blog/bimco-the-guidelines-on-cyber-security-onboard-ships>.
89. Ørnulf Rødseth. Munin deliverable 4.3: Evaluation of ship to shore communication links. <http://www.unmanned-ship.org/munin/wp-content/uploads/2014/02/d4-3-eval-ship-shore-v11.pdf>, 2012.
90. Martin Roesch et al. Snort: Lightweight intrusion detection for networks. In *Lisa*, volume 99, pages 229–238, 1999.
91. Ron Ross, Patrick Viscuso, Gary Guissanie, Kelley Dempsey, and Mark Riddle. Protecting controlled unclassified information in nonfederal information systems and organizations. Technical report, National Institute of Standards and Technology, 2016.
92. Risa Savold, Natalie Dagher, Preston Frazier, and Dennis McCallam. Architecting cyber defense: A survey of the leading cyber reference architectures and frameworks. In *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*, pages 127–138. IEEE, 2017.
93. Stefan Schauer, Nineta Polemi, and Haralambos Mouratidis. Mitigate: a dynamic supply chain cyber risk assessment methodology. *Journal of Transportation Security*, 12(1):1–35, 2019.
94. A. Shostack. *Threat Modeling: Designing for Security*, volume Wiley Publishing. 2014.
95. Keith Stouffer, Joe Falco, Karen Scarfone, et al. Guide to industrial control systems (ics) security. *NIST special publication*, 800(82):16–16, 2011.
96. Keith Stouffer, S Lightman, V Pillitteri, Marshall Abrams, and Adam Hahn. Nist special publication 800-82, revision 2: Guide to industrial control systems (ics) security. *National Institute of Standards and Technology*, 2014.
97. Blake E Strom, Andy Applebaum, Doug P Miller, Kathryn C Nickels, Adam G Pennington, and Cody B Thomas. Mitre att&ck: Design and philosophy. *Technical report*, 2018.
98. Boris Svilicic, Junzo Kamahara, Matthew Rooks, and Yoshiji Yano. Maritime cyber risk management: An experimental ship assessment. *The Journal of Navigation*, 72(5):1108–1120, 2019.
99. Kimberly Tam and Kevin Jones. Factors affecting cyber risk in maritime. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–8. IEEE, 2019.
100. SEI AADL Team et al. An extensible open source aadl tool environment (osate). *Software Engineering Institute*, 2006.
101. Ajay Tirumala. Iperf: The tcp/udp bandwidth measurement tool. <http://dast.nlanr.net/Projects/Iperf/>, 1999.
102. Joe Touch, L Eggert, and Y Wang. Use of ipsec transport mode for dynamic routing. *Request for Comments (RFC)*, 3884, 2004.
103. Daniel Trimble, Jonathon Monken, and Alexander FL Sand. A framework for cybersecurity assessments of critical port infrastructure. In *2017 International Conference on Cyber Conflict (CyCon US)*, pages 1–7. IEEE, 2017.
104. Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, NJ, 1996.
105. Wengyik Yeong, Tim Howes, and Steve Kille. Lightweight directory access protocol. 1995.
106. Cheol-Gyu Yi and Young-Gab Kim. Security testing for naval ship combat system software. *IEEE Access*, 9:66839–66851, 2021.
107. Nathan Yocom. pgina administration and users documentation. <http://pgina.org/>. Accessed 11.10.2021.
108. zvelo. Fight ransomware with defense in depth. =<https://zvelo.com/fight-ransomware-with-defense-in-depth/>. Accessed 11.10.2021.

A Mapping between ATT&CK and DiD elements

In this section, we discuss our developed mapping between the ATT&CK controls and DiD elements. This mapping is intended to facilitate the structured architecture development process in order to allocate the controls which are proposed from the risk assessment process and consider them during the architecture development. The mapping is done based on

the description of the controls in the *ATT&CK* framework and the DiD element description. Table 6 depicts the outcome of our mapping efforts. The table reflects that the controls in *ATT&CK* do not support all the DiD elements, and others do not map to any DiD element. This indicates that aligning both would broaden the defensive capabilities of the target system.

Additionally, Table 6 reflects the sources discussing the different DiD elements. The stakeholders' requirements that are related to cybersecurity span across all defense layers except physical security. Additionally, the variance of controls and DiD elements discussed in the different sources demonstrates the need for continuous improvements of the risk management capabilities as no single source has considered all possible controls.

B Verification of Requirements

A detailed analysis of the cybersecurity requirements verification has been conducted and is depicted in Table 7 to demonstrate the satisfaction of the communicated requirements by the proposed architecture. The table details the addressed requirements, their priority, the required verification criteria, the relevant architectural components, and efforts made to verify as well as evaluate the satisfaction of the requirements. The requirements are labeled using a three-level coding scheme (a-b-c). The first level (a) refers to the domain (S for Cybersecurity). The second level (b) refers to the sub-domain which are i) identification (I), ii) protection (P), iii) detection (D), iv) response, and v) recovery (R). Finally, the third level (C) refers to the number of the serial number of the requirement within its sub-domain.

Regarding the requirements priority, this property conveys the exact necessity level of each requirement as communicated by the stakeholders. Using the metrics in the MoSCoW requirement prioritization technique, most of the communicated requirements are "should" except two that are "must". The requirements with a priority "should" as communicated by one of the stakeholders suggests guidelines describing recommended processes to maintain equivalence with conventional designs [44]. In our previous work [24] we adopted these requirements with their indicated priority to propose a feasible architecture that is needed at the current project phase.

C Adversary Emulation Process

A summary of the adversary emulation process conducted against the milliAmpere2 ferry is shown in Table 8. Due to space limitations, the table presents only selected tests of the complete required set of tests. The shown tests cover all the attackers' objectives (i.e. tactics) proposed in the *ATT&CK* framework and a wide range of techniques and software to implement them. For each tactic, the relevant techniques, the test method, results, and proposed corrective action are presented.

D IP Network Simulation

D.1 Development of Simulation Network

Simulation is a proposed system analysis method by the ISO 15288 standard [13] and it is an observed approach in the

maritime domain. [28] utilized simulation for the evaluation of a specific set of security controls against a specific set of attacks against maritime systems. [29] and [57] utilized simulated environments for the evaluation of several maritime systems and protocols. In this paper, simulation is utilized to verify the feasibility of integrating the cybersecurity architecture within the underlying communication architecture and to facilitate later security analysis. A prototype implementation of the IP-based components is provided using the GNS3 simulator [83].

As shown in Figure 11, two physically distinct facilities were implemented at two different locations, emulating both the APS and the RCC. At each location, a dedicated workstation is interfaced with a physical router (i.e. gateway). The rationale behind this is two folds. The first is intended to create a physical division for emulating the remote control and monitoring of the RCC over the APS towards identifying possible challenges in the provisioning of security functions under the autoremove operational mode. The second is related to performance management. The implementation consists of many virtualized components that collectively consume plenty of resources. Therefore, logically distributing these resources over two workstations aids toward an improved testing environment.

The gateways are both interfaced with the same network outside our control. Upon our request, two static IP addresses were reserved to the gateways using Dynamic Host Configuration Protocol (DHCP) to emulate a connection over the Internet from an Internet Service Provider (ISP). The gateways are implemented using Cisco RV042 routers. This model provides several of the required capabilities, namely, firewall, VPN, and DMZ. The firewall capability implements the dedicated firewall component discussed in section 6.6.1. The VPN capability implements the required VPN tunnels discussed in section 6.6.3. Finally, the DMZ capability implements the required DMZ discussed in Section 6.5. On the other hand, each workstation hosts a group of virtualized components emulating either the APS or the RCC networks using the GNS3 simulator. The Core/ Distribution (C/D) switches (C/D at the RCC, C/D A, and C/D B at the APS) are simulated using a Cisco IOS image of a Layer-3 switch, while other components are simulated using different appliances including Windows, Ubuntu Desktop, Ubuntu Server, Kali, and Docker containers. GNS3 provides the capability to interface a simulated component with a physical network using a component called "cloud", this allows the C/D switches to be interfaced with the RV042 routers emulating realistic networking. Several of the required components discussed in section 6 were implemented towards satisfying the established requirements. In the following subsections, detailed discussions are provided for each implemented component grouped according to the DiD architecture viewpoints. We would like to highlight that our choices of tools for the implementation are based on commonly used open-source or free tools. These tools were only used to serve the purpose of the analysis which is feasibility and identification of possible challenges in the management of the cybersecurity risks in autonomous and remotely controlled systems.

D.1.1 Risk Management

In this section, we discuss our implementation of the components that supports the risk management functions discussed in section 6.2. A server with the GLPI software [9] is utilized to support the required asset and user inventory functions discussed in section 6.2.1. Considering that GLPI is managed

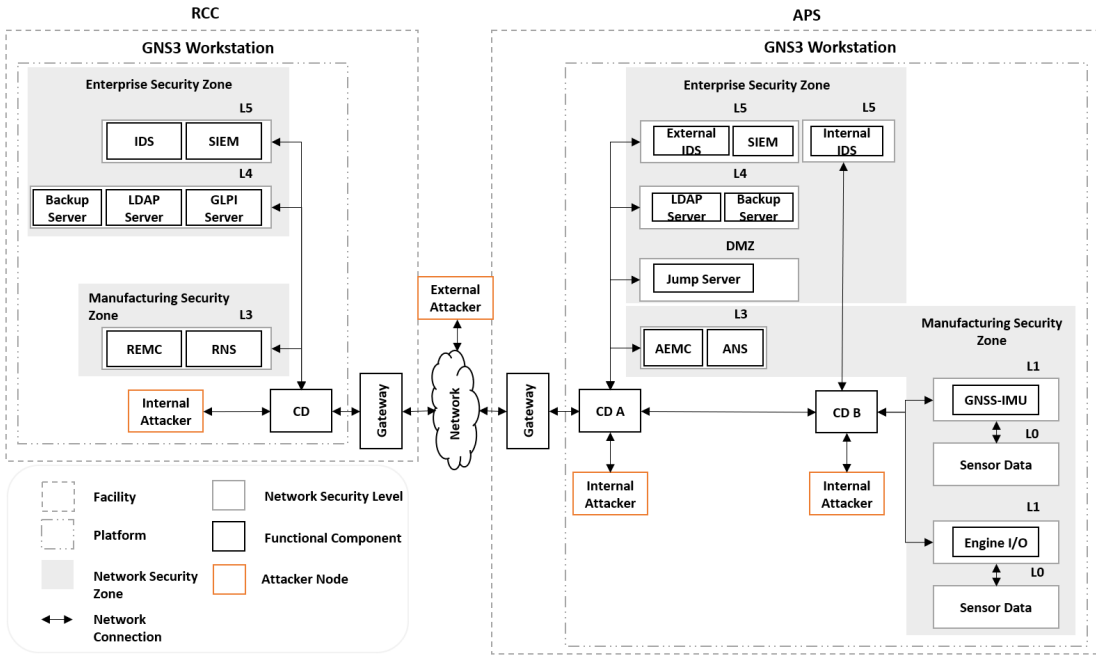


Fig. 11: Overview of the implemented architecture

through a web interface, the server is implemented within the RCC network due to the crewless nature of the APS. Nevertheless, a group of agents using the FusionInventory software [8] are installed at each endpoint and server at both facilities to send inventory information to the GLPI server. This setup is similar to the SIEM setup in figure 10a and has the same shortcomings discussed in section 6.8.2.

D.1.2 Network Architecture

In this section, we discuss our implementations and the analysis for the different components in the network architecture discussed in section 6.5, particularly, the C/D switches, the DMZ, and the jump server. The C/D switches are already proposed in the communication architecture. They are utilized in the cybersecurity architecture to realize the different network security zones and levels discussed in section 6.5, the firewall capabilities discussed in section 6.6.1 as well as the component-to-component access management discussed in section 6.6.4. The security zones and levels are implemented using VLANs mapped to the different network levels. The firewall capabilities, as well as the component-to-component access management, are implemented using ACL. For instance, the GNSS-IMU component is programmed to transmit sensor data to the ANS only. An ACL rule is created to allow this communication and deny any other component to be reached from the GNSS-IMU.

Regarding the DMZ, an IP address has been assigned to the jump server from the network that is outside our control which emulates the internet. The public IP address has been mapped to the local IP address using the one-to-one NAT technique configured at the APS gateway. Firewall rules were

created to restrict inbound access from the server at the DMZ into the internal network to only RDP connections.

As discussed previously in section 6.6.4, a jump server is required for remote maintenance. It is advised to be placed at the DMZ and access to it should be secure using 2FA. This has been implemented using a windows workstation placed at the DMZ. The 2FA is implemented using the free software from Cisco called "Duo" [6]. Duo has been installed at the jump server and is linked to a mobile phone and enforces a policy to approve any remote access using Remote Desktop Protocol (RDP) using the assigned mobile phone through a dedicated app.

D.1.3 Network Perimeter Security

In this section, we discuss our implementations and the analysis for the different components in network perimeter security including the firewalls (section 6.6.1), the VPN tunnels (section 6.6.3) and Access Management (section 6.6.4).

The main firewalls are implemented on the gateways. Additional firewall capabilities are implemented at the C/D switches for achieving the network architecture. Moreover, host-based firewalls are enabled with customized rules to allow the traffic for the required services such as the SIEM agents.

Regarding the implemented VPN tunnels. The implemented protocol is IPSec with strong encryption and policy-enforced complex shared key. We have utilized the iperf tool [101] for latency testing with the activation of VPN tunnels. The average latency was observed to be 1,94 ms for the link between the two workstations which exists outside the GNS3 implementation. This is well within the acceptable latency for the ship to shore communication suggested by the MUNIN

project which is one second [89]. However, we have observed very high latency in the GNS3 implementation which is related to the virtualization technology. For instance, the average latency at the bridged interface between the gateway and the C/D switch is 12,184 ms and it reaches 133,477 ms between the ANS and the RNS with 19% packet loss. This indicates that GNS3 is not suitable for the performance evaluation of the security controls.

An LDAP server running OpenLDAP [37], an open-source software implementing the Lightweight Directory Access Protocol [105] is implemented at each facility to realize the required User Access Management discussed in section 6.6.4. A network domain was created as well as user accounts for the different endpoints. Then the endpoints are joined to the created domain. The open-source pGina plugin [107] is implemented to integrate the endpoints with the Windows operating system with the Linux LDAP server.

D.1.4 Host Security

In this section, we discuss our implementations and the analysis for the different components related to host security, particularly, malware protection, backup, as well as application isolation, and sandboxing.

The malware protection component has been implemented using the open-source ClamAV antivirus engine [4]. This capability realizes the required malware protection function discussed in section 6.7.2.

Backup facilities have been implemented at both the APS and the RCC utilizing the open-source BorgBackup software [3]. A script was written to perform a daily backup of the APS and the RCC endpoints with encryption and compression of the archives.

The development of the APS navigational applications is outside the scope of this paper. Nevertheless, we have developed a group of scripts to receive or generate simulated sensor data and transmit them to the processing and control components in the ASC and the RSC. These scripts were developed as docker containers to realize the application isolation requirement.

D.1.5 Security Monitoring

In this section, we discuss our implementations and the analysis for the different components related to Security Monitoring, particularly, the IDS components and the SIEM components.

Two types of IDS are implemented. Host-based IDS (HIDS) and Network-based IDS (NIDS). HIDSs are implemented using the Wazuh software. This component is installed on all devices including the backup ANS and backup AEMC as discussed in section 6.8.1. This allows for customized rules for these endpoints particularly to monitor the out-of-band emergency communication channel with ECT. On the other hand, the NIDSs are implemented using Snort [90]. Three Snort appliances are implemented and connected to the networks at each facility; two at the APS and one at the RCC, to realize the required IDS capability discussed in section 6.8.1 and shown in Figure 9. This allows for customized rules at each Snort node focusing on a specific set of expected communication flows.

The SIEM components have been implemented using the open-source security platform Wazuh [15]. It is built on top

of the elastic stack and provides a wide range of cybersecurity capabilities. With regards to the APS cybersecurity requirements, Wazuh provides several capabilities; among other things, monitoring and logging (section 6.8), incident response (section 6.10), system hardening (section 6.7.5), and vulnerability management (section 6.7.1). We implemented two SIEM nodes, one at the APS and another at the RCC as suggested in section 6.8.2.

D.2 Adversary Emulation

Having the simulated implementation in section D allows for conducting hands-on adversary emulation. Conducting adversary emulation against the simulated implementation aims to help in understanding the impact of the autoremove operational mode on cybersecurity functions and how would they withstand realistic cyber-attack techniques. Based on that, we define a group of attack scenarios consisting of different *ATT&CK* tactics (i.e. kill chain phase) and techniques against different components in order to showcase the concept of layered defenses. The different attacks techniques applied in this section are a result of the conducted risk assessment process for the APS proposed in our earlier work [23] and discussed briefly in section 2.2.

D.2.1 External attacker aiming to impact the APS operations

In this section, we will describe a fictional attack scenario to showcase the role of the different mitigation methods at each defense layer. Although it is fictional, we have utilized the prototype implementation discussed in section D to emulate some of the attack techniques in order to obtain a realistic attack flow.

In this scenario, an external attacker aims to gain a foothold into the network towards impacting the APS operations in any way possible. The attack exploits the remote desktop feature in the APS which is enabled to allow remote maintenance due to the autoremove operational mode. Figure 12 depicts an attack tree for achieving the attacker objective. The figure also depicts the different stages of the attack, the attacked component as well as the mitigation methods that need to fail for this attack to succeed. A detailed description of the attack scenario is described below:

- **Reconnaissance:** no reconnaissance techniques have been considered during the risk evaluation considering that reconnaissance techniques occur outside the control of the cybersecurity architecture. Nevertheless, the attacker can perform “Scanning IP Blocks” or “Gather Victim Network Information” to learn more about the network and the services it provides. Assuming the attacker has previous knowledge of the public IP addresses of the target network, now the attacker can learn that the jump server at the DMZ is publicly open.
- **Initial Access:** to achieve a foothold into the network, the external attacker utilizes “Internet Accessible Device”. In the simulated architecture, several devices are connected to the internet at both the RCC and the APS. One target can be the jump server using a remote desktop client. The attacker is initially faced with an authentication request to provide credentials to access the jump server (access management). Assuming the attacker can guess the credentials (i.e. default credentials), then the

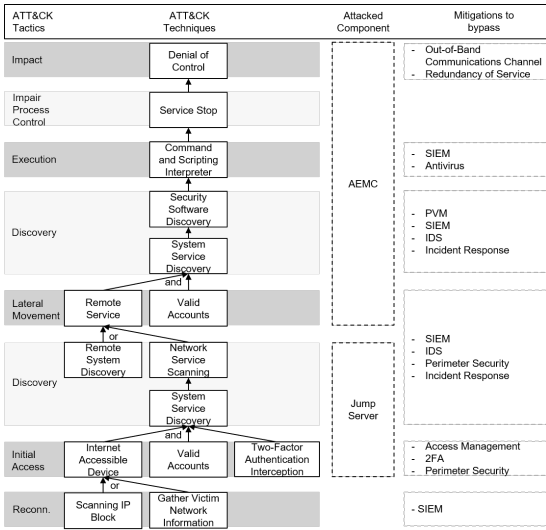


Fig. 12: an attack tree of a possible scenario

attacker gains initial access through "Valid Accounts". Then, the attacker faces the next security control which is the 2FA. Considering that this 2FA is configured to approve access to RDP connections to the jump server by a specific mobile phone outside the attacker's reach, the attack fails. Assuming the attacker can conduct "Two-Factor Authentication Interception" to bypass this security control; similar to the Chimera group [81]. Then the attacker can access the jump server and gain the initial foothold.

- **Discovery:** Assuming the attacker has access to the jump server. The next step is to understand the environment. Initially, the attacker performs "System Service Discovery" to understand the running services on the accessed machine. Considering it's a jump sever and no other important services are hosted on it. The attacker decides to locate a more critical target. So, the attacker performs "Network Service Scanning" or "Remote System Discovery" to discover the running services in the network in an attempt to discover vulnerabilities. Considering that the traffic to and from the DMZ is heavily filtered using ACLs, the attacker is unable to discover services other than the RDP to other hosts in the network. Therefore, the attacker needs to perform a technique to achieve "Lateral Movement" to move within the network to another location with more observability to the network. At the same time, the attacker is performing the activities at the discovery stage, the IDS and the SIEM have raised indicators that some discovery activities are being conducted. If security personnel at the RCC is monitoring these logs and deem suspicious behavior is happening, incident response activities could be initiated. For the sake of this scenario, let's assume that the attacker is still undiscovered.
- **Lateral Movement:** The attacker performs lateral movement from the jump server through the available "Remote Services" (i.e. RDP) on the AEMC and pivots to that lo-

cation using the same valid credentials used at the jump server.

- **Execution:** Assuming the attacker has access to the AEMC. The discovery phase is conducted again to understand the environment. Considering the critical role of AEMC in the control functions, the attacker is able to recognize an industrial control software using "System Service Discovery". Due to the PVM component, this software is patched, and the attacker is unable to discover a vulnerability to exploit. Assuming this software has a non-patched vulnerability that allows for a "Denial of Service" (DoS) attack and there is an available malware that is able to launch the exploit. To avoid detection, before running the malware, the attacker performs a "Security Software Discovery" and discovers the availability of the antivirus software which the attacker discovers is able to recognize the malware binary and prevents it from running as well as alert the RCC of a running attack. So, the attacker utilizes "Command and Scripting Interpreter" to build a custom exploit and execute it.
- **Impair Process Control:** Assuming the DoS attack is in the form of a "Service Stop" that impairs the AEMC from performing control functions of the thrusters.
- **Impact:** Assuming the attack succeeded, the passengers on the APS should feel something out of the ordinary. They can use the previously proposed emergency push button to inform the RCC and the ECT of a problem so they can initiate a suitable response plan. An existing response plane is to intervene using the Out-of-Band communication channels and utilize the redundant control system to navigate the APS to safety (Section 6.10).

D.2.2 Ship to Shore Communication Eavesdropping

In this section, we will describe two attack scenarios in which the attackers aim to eavesdrop on the communication flows between the APS and the RCC. We utilized the prototype implementation to perform a defensive engagement of the threat to evaluate the proposed mitigation methods.

In the first attack, a passenger aims to gain initial access to the network by performing a "Hardware Additions" technique. The passenger inserts a computer into the "C/D tier A" switch and attempts to sniff the traffic. Assuming that no physical barriers are in place, the first security control that the attacker faces is the static network configuration which is a result of the system hardening discussed in section 6.7.5. Assuming that the attacker has previous knowledge of the different sub-nets and can assign a static IP address. Then, the attacker can perform the "Man-in-the-Middle: ARP Cache Poisoning" technique to be able to perform "Network Sniffing" of the traffic. Security events generated from these activities can be observed at the SIEM, if an operator on the RCC is monitoring these events, an incident response plan can be initiated.

In the second attack, an external attacker aims to eavesdrop on the ship to shore communication. Assuming the attacker has previous knowledge of the public IP addresses within the APS ecosystem. Then, the attacker employs "Network Service Scanning" using tools such as *NMAP* [75] to discover the services from each host with public IP. Utilizing *NMAP* OS identification capability, the attacker can identify that two of the public IP addresses are assigned for the gateways (Cisco RV042 routers). Moreover, both routers have port 60443 open which indicates the possible existence of VPN tunnels. Scanning the third IP address discloses the RDP service on the jump server. The attacker then attempts

to perform the "Man-in-the-Middle: ARP Cache Poisoning" technique to be able to perform "Network Sniffing" of the traffic, but the attack failed. The reason this attack failed is due to a requirement in the vendor management process (section 6.9) to have a countermeasure for spoofing attacks, and indeed, the RV042 has such capability [38]. Assuming no such requirement exists, and the gateway doesn't have such capability. The traffic between the two gateways passes an external network that is outside the control of the cybersecurity architecture. Therefore, the implemented VPN tunnels implemented using the IPSec protocol shield the ship-to-shore communication from "Eavesdrop on Insecure Network Communication" techniques.

Table 6: A mapping between DiD layers, elements and their relevant ATT&CK controls. In addition to indication of the sources that discussed their need.

Defense Layers	Elements	Requirement	DNV	ICS DID	BIMCO	ATT&CK
Cyber security and risk management	Policies and procedures		✓	✓	✓	Account Use Policies, Password Policies
	Standards/ Recommendations and System documentation		✓	✓		N/A
	Threat Intelligence		✓	✓	✓	Threat Intelligence Program
	Maintain Asset Inventory	✓	✓	✓		N/A
Physical Security	Management support		✓	✓	✓	N/A
	Roles and responsibilities		✓	✓	✓	Limit Hardware Installation, Limit Access to Resource Over Network
	Training, awareness and competence		✓	✓	✓	User Guidance, User Training, Application Developer Guidance
	Network segmentation		✓	✓	✓	Network Segmentation, Limit Access to Resource Over Network
Perimeter Security	Firewalls		✓	✓	✓	Filter Network Traffic, Limit Access to Resource Over Network, Network Allowlists, SSL/TLS Inspection
	Access Management	✓	✓	✓	✓	Access Management, Account Use Policies, Attestation, Authorization Enforcement, Caution with Device Administrator Access, Communication Authenticity, Multi-Factor Authentication, Operational Information Confidentiality, Password Policies, Privileged Account Management, Software Process and Device Authentication, User Account Control, User Account Management, Minimize Wireless Signal Propagation
	VPN		✓	✓	✓	Communication Authenticity, Encrypt Network Traffic, Encrypt Sensitive Information, Operational Information Confidentiality
	Satellite and radio communication	✓	✓	✓	✓	Communication Authenticity, Encrypt Network Traffic, Encrypt Sensitive Information, Operational Information Confidentiality
Host Security	Patch and Vulnerability Management	✓	✓	✓	✓	Deploy Compromised Device Detection Method, Security Updates, Application Isolation and Sandboxing
	Virtual Machines		✓	✓	✓	Application Isolation and Sandboxing, Execution Prevention, Exploit Protection, Privileged Process Integrity
	Malware protection	✓	✓	✓	✓	Active Directory Configuration, Credential Access Protection, Disable or Remove Feature or Program, Environment Variable Permissions, Execution Prevention, Exploit Protection, Limit Hardware Installation, Limit Software Installation, Lock Bootloader, Operating System Hardening, Restricted Library Loading, Restrict Registry Permissions, Restrict Web-Based Content, Software Configuration, Static Network Configuration, System Partition Integrity, User Account Control
	Secure configuration of hardware and software	✓	✓	✓	✓	N/A
Security Monitoring	Email and web browser protection				✓	N/A
	Intrusion Detection Systems	✓	✓	✓	✓	Audit, Behavior Prevention on Endpoint, Communication Authenticity, Data Loss Prevention, Deploy Compromised Device Detection Method, Exploit Protection, Network Intrusion Prevention, Privileged Process Integrity, SSL/TLS Inspection
	Security Audit Logging	✓	✓	✓	✓	Audit, Behavior Prevention on Endpoint, Data Loss Prevention, Deploy Compromised Device Detection Method, Exploit Protection, Privileged Process Integrity, SSL/TLS Inspection, Vulnerability Scanning
	Security Incident and Event Monitoring	✓	✓	✓	✓	Audit, Behavior Prevention on Endpoint, Data Loss Prevention, Deploy Compromised Device Detection Method, Exploit Protection, Privileged Process Integrity, SSL/TLS Inspection, Vulnerability Scanning
Vendor Management	Establish contingency plans	✓	✓	✓	✓	Application Vetting, Boot Integrity, Code Signing, Supply Chain Management
	Data Recovery	✓	✓	✓	✓	N/A
	Investigating cyber incidents and Effective response	✓	✓	✓	✓	N/A
	Losses arising from a cyber Incident		✓	✓	✓	Communication Authenticity, Encrypt Network Traffic, Encrypt Operational Information, Operational Information Confidentiality, Watchdog Timeout, Redundant Backups, Safety Instrumented Systems, Mechanical Protection Layers
	N/A					

Table 7: Cybersecurity requirements satisfaction checklist

Reference as in [24]	Requirement Description	Priority*	Verification Criteria		Verification Check	Existing and Simulated Components	Verification Criteria		Verification Check	Implemented and Existing Components.
			Design level	Design level			Implementation Level	Implementation Level		
Regulator Perspective	Risk Management	M	Components supporting risk management activities exist	YES	YES	IS3MS 6.2 Risk assessment process	Components supporting risk management activities are implemented	YES	IS3MS 6.2 Risk Assessment Process	Implemented based on the 5G routers at each site. A registration platform is provided by the supplier. The ferry and the RCC are registered and monitored.
Service Provider Perspective	Ship Registry	S	Components supporting secure ship registration exist	YES	YES	Access Management 6.6.4	Components supporting secure ship registration are implemented	YES	YES	Agreements with network providers exist regarding cybersecurity controls. VPN and 2FA for 5G routers. Non-disclosure agreements with network providers, and others.
Service Provider Perspective	Secure service provisioning	S	Security controls for service providers should be planned	YES	YES	Vendor Management 6.9	Security controls for service providers are implemented	YES	YES	IS3MS 6.2 Risk Assessment Process
S-I-1	Cybersecurity Maturity Framework	S	Components supporting activities for the framework are proposed	YES	YES	IS3MS 6.2 Risk assessment process	Components supporting activities for the framework are implemented	YES	YES	Detailed and updated system and network diagrams exist.
S-I-2	Map of IT installation and network architecture	S	Asset inventory capabilities exist	YES	YES	Asset and User Inventory 6.2.1 GLPI software [9]	Detailed inventory of components and network architecture	YES	YES	Distributed: some components have a local user management capability, while others don't.
S-I-3	Inventory of user accounts and privileges	S	Account inventory capability exists	YES	YES	Access Management 6.6.4 OpenLDAP [37]	Network user management capability is implemented	Partially	Partially	Some components have been observed to implement strong password requirements. On the other hand, some components were found to retain default credentials.
S-P-1	User management	S	User management capability exists	YES	YES	Access Management 6.6.4 OpenLDAP [37]	User management capability is with best practices in secure authentication	Partially	Partially	Some components have updating procedures while others don't.
S-P-2	Regular Updates	M	Software updating capability exists	YES	YES	Patch and Vulnerability Management 6.7.1 Wazuh [15]	Software updating according to an update policy is implemented	Partially	Partially	Some components have updating procedures while others don't.
S-P-3	Secure protocols	S	Secure protocols are proposed when applicable	YES	YES	Network Perimeter Security 6.6 Cisco RV042 Router	Secure protocols are implemented when applicable	YES	YES	Site-to-Site VPN was deemed an important control and was implemented using OpenVPN cloud solution.
S-P-4	Malware protection	S	Malware protection capability exists	YES	YES	Malware Protection 6.7.2 ClamAV antivirus [4]	Malware protection capability is implemented	Unknown	Unknown	Access to some components was restricted and no information was provided regarding the existence of malware protection.
S-P-5	Cybersecurity training	S	Training areas for training are proposed	YES	YES	Training and Awareness 6.4	Specific training areas are proposed	NO	NO	No training exercises have yet been proposed.
S-P-6	Regular software security analysis	S	Supporting security analysis are proposed	YES	YES	Patch and Vulnerability Management 6.7.1 Wazuh [15]	Software security analysis is conducted based on a suitable policy	Partially	Partially	Network vulnerability scanning was performed to inform regarding software security analysis.
S-D-1	Monitoring capabilities	S	Monitoring capabilities should exist	YES	YES	Security Monitoring 6.8 Splunk [30] and Snort [90]	Monitoring capabilities are implemented	Partially	Partially	A host-based intrusion detection exists for the routers and provide security alerting features.
S-R-1	Incident response plan	S	Components incident response plans are proposed	YES	YES	Incident Response 6.10	Incident response plans are formulated and implemented when applicable	Partially	Partially	No plans specifically exist for cybersecurity incidents. Incident response plans exist in case of safety-related incidents. Emergency Control Team shall intervene.
S-R-2	Backup facilities	S	Backup facilities exist	YES	YES	Backup 6.7.4 BorgBackup software [3]	Backup facilities are implemented with a suitable backup policy	Partially	Partially	A data backup service is implemented at RCC. No backup facilities exist at the ferry.

* M=SCoW rule (S: Should, M: Must)

Table 8: A summary of planned and executed tests in the adversary emulation process

ATT&CK Tactic	Test objective	ATT&CK techniques	Test method	Results	Corrective action
Reconnaissance	Identify remotely open services	Gather Victim Host Information (T1592), Search Open Websites/Domains (T1593), Network Service Scanning (T1018), Active System Discovery (T0844), Active Scanning (T1595)	Searching network scanners (Shodan, Censys, and BinaryEdge) using the ferry's 5G router's public IP address	Only shodan identified an open port for an IP camera at some time in the past. The port was not open at the time of the test.	None
	Learn ferry network topology	Remote System Discovery (T0846), Active Scanning (T1595)	Scanning the ferry's 5G router's public IP address using Nmap	2 open ports for the router remote authentication and signaling services	None
Initial Access	Discover vulnerabilities	Search Open Technical Databases (T1596)	Using netdiscover to identify hosts and networks. This was only possible after gaining access to the network.	2 local networks were identified. One by the 5G router and another by a network switch. In total, 13 hosts were discovered.	NIDS tuning to detect network scanning.
	gain access to the ferry network	Transient System Asset Additions (T1200)	Using the default Windows policy (NVD) to identify vulnerabilities in the network components.	Several vulnerabilities were found for one critical component in the network. One vulnerability has a 9.5 CVESS rating.	Updating the component to latest version.
Collection	remotely access the 5G network	Valid Accounts (T0859), Default Credentials (T0812)	Insert a Raspberry Pi into the network	Sufficient controls exist providing physical security to mitigate this threat. However, permission to insert the Pi was granted to allow further tests.	NIDS tuning to detect newly installed devices in the network
	Shift network traffic	Network Sniffing (T0842 or T1040)	Using Wirehark to sniff network traffic. Identify and collect navigation messages for planning further attacks.	The platform implements a 2FA service associated with an authenticator mobile application.	None
Execution	Run a script with several commands to collect host information	Adversary-in-the-Middle: ARP Cache Poisoning (T1557,002)	Using a USB stick with customized autorun function	Network traffic is captured at several intervals including NMEA messages emitted from the GPS and broadcasted within the network.	Limit access to resource over network
	Identify hosts and networks	Hardware Additions (T1200), System Information Discovery (T1082) and Exfiltration (T1020), Exfiltration Over Web Service (T1567)	Using a USB stick with customized autorun function	Only ICMP messages between hosts were collected.	NIDS tuning to detect ARP spoofing
Exfiltration	Identify open services at the network	Remote System Discovery (T0846), Active Scanning (T1595)	Using Nmap to identify network services.	No permission was granted from the network vendor to run this test.	None*
	Identify remote services that can allow lateral movement	Remote System Discovery (T0846), Active Scanning (T1595)	Using Nmap to identify remote desktop services.	Was conducted on several occasions without any obstruction.	Data Loss Prevention Solution
Lateral Movement	Remotely access to access other components	Remote Services Remote Desktop Protocol (T1021,001)	Using the identified RDP software, spoofing the IP address and using default credentials found online for accessing network devices.	2 local networks were identified. One by the 5G router and another by a network switch. In total, 13 hosts were discovered.	NIDS tuning to detect network scanning.
	Establish C&C channel between a victim and a remote	Encrypted Channel (T1573)	Using a covert channel software (e.g. Reverb) run a client on a host in the network and run the server on network services.	A lot of services were discovered ranging from HTTP, HTTPS, FTP, SSH, RDP, telnet, and NFS.	NIDS tuning to detect network scanning.
Command & Control	Identify default or weak passwords	Brute Force (T1110)	Using Metasploit to brute force open network services.	The majority of components have open ports for well-known remote desktop software and network sharing service.	None*
	Sniff credentials	Network Sniffing (T0842 or T1040)	Using Wirehark to sniff network traffic. Identify and collect credentials.	No permission was granted from the network vendor to run this test.	Enforcing a password policy
Persistence	Access other hosts in the network to maintain a foothold	Remote Services Remote Desktop Protocol (T1021,001)	Using the identified RDP software, attempt to access other devices.	No permission was granted from the network vendor to run this test.	None*
	Will the network scanning be detected	Network Service Scanning (T1018), Remote System Discovery (T0846), Active Scanning (T1595)	Using Nmap to scan the network with different configurations ranging from aggressive to polite scans.	Aggressive scans were detected and stopped. Polite is unknown. The status of the detection is unknown.	NIDS tuning to detect network scanning.
Defense Evasion	Changing the IP address	Fallback Channels (T1008)	Manually configuring the IP address of the Pi.	When the scans were stopped, the Pi lost access to the network. However, access was regained after manually changing the IP address.	NIDS tuning to detect IP changes in the network
	Gain administrative privileges using operating system vulnerabilities	Abuse Elevation Control Mechanism (T1548)	Run a pre-built malware.	No permission was granted from the network vendor to run this test.	None*
Impact Process Control	Manipulate network messages	Manipulation of Control (T0831)	Using Ettercap, manipulate control commands in the network	Attack did not succeed.	None**
	Drop navigation traffic	Denial of View (T0815)	Using Ettercap Filters to drop some navigation messages (NMEA)	Attack did not succeed.	None**
Network Effect	Jamming GPS data	Denial of View (T0815), Denial of Service (T1464)	Using GPS jammer to impact positioning data collection.	Future work	None
	Obtain device backups stored remotely	Obtain Device Cloud Backups (T1470)	Obtain online stored configurations of hosts.	The configuration files of the 5G routers are secure with 2FA authentication.	None**
Impact	Manipulate network traffic	Manipulation of View (T0832)	Using Ettercap, manipulate some navigation messages (NMEA)	Attack did not succeed.	Limit access to resource over network
	Streamline operational	Operational Security (T0882)	Identifying proprietary information from the network traffic	Proprietary information was identified in the network traffic related to the vendor's navigation system.	Limit access to resource over network




* none at the moment due to lack of information as a result of insufficient testing. Additional testing in the future is needed.
 ** none at the moment due to lack of information as a result of insufficient testing. Additional testing in the future is needed.

Paper VI

A. Amro, A. Oruc, V. Gkioulos and S. Katsikas, 'Navigation data anomaly analysis and detection,' *Information*, vol. 13, no. 3, 2022, ISSN: 2078-2489. DOI: 10.3390/info13030104. [Online]. Available: <https://www.mdpi.com/2078-2489/13/3/104>

Article

Navigation Data Anomaly Analysis and Detection

Ahmed Amro ^{*}, Aybars Oruc , Vasileios Gkioulos and Sokratis Katsikas 

Department of Information Security and Communication Technology, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, 2815 Gjøvik, Norway; aybars.oruc@ntnu.no (A.O.); vasileios.gkioulos@ntnu.no (V.G.); sokratis.katsikas@ntnu.no (S.K.)

* Correspondence: ahmed.amro@ntnu.no

Abstract: Several disruptive attacks against companies in the maritime industry have led experts to consider the increased risk imposed by cyber threats as a major obstacle to undergoing digitization. The industry is heading toward increased automation and connectivity, leading to reduced human involvement in the different navigational functions and increased reliance on sensor data and software for more autonomous modes of operations. To meet the objectives of increased automation under the threat of cyber attacks, the different software modules that are expected to be involved in different navigational functions need to be prepared to detect such attacks utilizing suitable detection techniques. Therefore, we propose a systematic approach for analyzing the navigational NMEA messages carrying the data of the different sensors, their possible anomalies, malicious causes of such anomalies as well as the appropriate detection algorithms. The proposed approach is evaluated through two use cases, traditional Integrated Navigation System (INS) and Autonomous Passenger Ship (APS). The results reflect the utility of specification and frequency-based detection in detecting the identified anomalies with high confidence. Furthermore, the analysis is found to facilitate the communication of threats through indicating the possible impact of the identified anomalies against the navigational operations. Moreover, we have developed a testing environment that facilitates conducting the analysis. The environment includes a developed tool, NMEA-Manipulator that enables the invocation of the identified anomalies through a group of cyber attacks on sensor data. Our work paves the way for future work in the analysis of NMEA anomalies toward the development of an NMEA intrusion detection system.

Keywords: NMEA; cybersecurity; anomaly analysis and detection; maritime



Citation: Amro, A.; Oruc, A.; Gkioulos, V.; Katsikas, S. Navigation Data Anomaly Analysis and Detection. *Information* **2022**, *13*, 104. <https://doi.org/10.3390/info13030104>

Academic Editor: Gianluca Valentino

Received: 19 January 2022

Accepted: 20 February 2022

Published: 23 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The maritime domain is undergoing a major digital transformation, leading to substantial changes in the business models, processes, and technology [1]. The Integrated Navigation System (INS) on conventional vessels of today is deployed to support safe navigation as a result of such digital transformation. However, technological advancements would change the characteristic of vessels dramatically in the near future. Recently, new projects have been proposed to increase autonomy in maritime. This includes automating maritime systems and services until such systems can reach sea-going autonomous operation by the year 2035 [2]. These projects have led to the proposition of a new ship class named Maritime Autonomous Surface Ship (MASS) as defined by the International Maritime Organization (IMO) [3]. Among these new projects is the Autonomous Ferry (Autoferry) project [4]. The project aims to develop an Autonomous Passenger Ship (APS) or ferry for carrying passengers across the Trondheim city canal in Norway. The APS is expected to be remotely monitored and controlled when necessary from a remote center.

The novelty of the Autoferry project influenced the cyber risk paradigm and led to unique attack objectives and techniques. The main factors that have led to this are the auto-remote operational mode as well as the fact that passengers will be on board without a crew. The auto-remote operational mode has led to novel and broader cyber attack vectors due to

remote connectivity, dependency on automated services, reduced human defenses, and the dependency on digital technologies for the remote operator for intervention. Furthermore, the presence of passengers imposes a safety risk factor motivating different types of threat actors to cause them harm through cyber attacks. Among the identified attack vectors in Autoferry is the navigational information which is communicated among the different marine components.

The National Marine Electronics Association (NMEA) defined a group of electronic and data specifications for the communication between different marine electronic systems. These specifications have manifested into a series of standards. The latest versions are NMEA0183 [5] and NMEA2000 [6]. These standards govern the structure and the manner in which messages are communicated among the different devices. NMEA messages are mainly utilized in the maritime domain. However, the positioning information provided by them has found their application in other domains such as those with requirements for location tracking for personal security [7,8], and car theft detection [9]. While these messages provide an abundance of information utilized in different navigational tasks and functions, their security has been investigated and found to be lacking any controls such as authentication, encryption, and validation [10]. This makes them susceptible to a wide range of cyber-attacks.

This paper aims to improve the security of NMEA messages by identifying and proposing relevant approaches for the treatment and monitoring of the risks associated with them. The NMEA0183 standard is considered in this paper, with future plans to extend the work to include the NMEA2000 standard. Therefore, we propose a systematic approach for analyzing NMEA messages, their anomalies, malicious causes of such anomalies (i.e., attacks) as well as the appropriate detection algorithms.

We utilize two maritime use cases throughout this paper to facilitate the description of our approach. The use cases are the APS and conventional vessels equipped with an INS. In this way, we caught an opportunity to prove the importance of our study not only for today's vessels but also for potential vessels of the future. We argue that our approach can aid in the development of resilient navigation systems that are developed and operated under the consideration of adversarial behavior.

The contribution of the paper is as follows:

- We propose a novel systematic approach for anomaly detection in NMEA messages.
- We present an analysis of possible anomalies in NMEA messages and their cause-and-effect relationship with a range of cyber-attacks.
- We propose a method for creating synthetic datasets with both normal and maliciously tampered with NMEA messages, and we implement and use a software package to create such experimental datasets.
- We use the datasets within the context of two use cases to evaluate the performance of anomaly detection approaches specifically designed for the purpose.

The remainder of the paper is structured as follows: In Section 2, we provide the necessary background and we review the relevant literature. In Section 3, we present our proposed method for systematic, multidimensional analysis of anomalies in NMEA messages. In Section 4, we discuss how our proposed method applies to the INS and the APS use cases. In Section 5, we present results of our experimentation with the proposed approaches, towards assessing its usefulness in developing an intrusion detection system for NMEA messages. In Section 6, we evaluate and discuss the results and findings of the experimentation, and in Section 7 we present deployment options of an NMEA IDS. Finally, Section 8 summarizes our conclusions and proposes directions for future research.

2. Background and Related Work

Several publications which point out cyber risks of autonomous ships are available in the literature. Kavallieratos et al. [11] presented the results of cyber risks assessment of remotely controlled and autonomous ships. The risk assessment was performed using the STRIDE threat modeling methodology. According to the results, the Automatic Identifica-

tion System (AIS), Electronic Chart Display and Information System (ECDIS), and Global Maritime Distress Safety System (GMDSS), in particular, include various high risks. Vinem and Utne [12] discussed the possibility of using autonomous ships for damaging the offshore industry. A cyber attack may cause the collision of an autonomous ship and an offshore platform at sea, intentionally or unintentionally. The paper also suggests several mitigation measures. Keeping a small number of crew onboard is argued to be the most effective preventive measure against cyber risk according to the authors.

Not only autonomous ships but also conventional vessels sailing at sea today could be exposed to cyber attacks. Svilicic et al. [13] unveiled the cyber vulnerabilities of an INS onboard ship. The authors acquired a total of 27 pieces of information, and four vulnerabilities using a vulnerability scanner. One of the detected vulnerabilities in the INS was reported as “Critical”. Moreover, a survey of intrusion detection in vehicles, including maritime vessels, is presented by Loukas et al. [14]. The authors discussed several works targeting Global Positioning System (GPS) and AIS spoofing and manipulation. However, no reference is made to the NMEA protocol.

Motivated by the identified threats in the maritime industry and focusing on the NMEA protocol as a possible threat vector, we surveyed the current state-of-the-art of NMEA security. Krile et al. [15] explained the network of an INS onboard ship, including a detailed description of the NMEA 0183 and 2000 standards. The authors focus on NMEA 2000 in particular in different aspects, such as components of an NMEA 2000 network, comparison of ethernet and Controller Area Network (CAN), and the functions of CAN. Moreover, the authors discuss NMEA software, including Sail Soft NMEA Studio, Maretron N2K Analyzer, and N2K Meter. Additionally, several applications of NMEA messages have been observed in digital forensics [16,17], personal security [7,8], car theft detection [9], as well as utilizing NMEA messages in detection Global Navigation Satellite System (GNSS) Spoofing [18]. Nevertheless, these works did not discuss the security of NMEA messages themselves. Some works have argued that NMEA security currently depends on the network and host security [19,20]. However, Seong and Kim [21] addressed the cybersecurity of NMEA messages by utilizing secure hash functions when storing NMEA messages in the voyage data recorder onboard vessels. This is argued to improve the authenticity of stored NMEA messages.

Additionally, Boudehenn et al. [22] proposed a machine learning approach to detect GPS attacks. The GPS device broadcasts NMEA 0183 messages to the ship network. Machine learning software developed by the authors in a Raspberry Pi 3B+ could detect GPS jamming and spoofing attacks successfully. In this way, the officer of the watch on the bridge may be notified about a potential GPS attack. Machine learning could be also used to detect malicious activities in the ship network [23]. Moreover, Hemminghaus et al. [24] have presented a bridge attack tool named BRidge Attack Tool (BRAT) that targets NMEA messages with a wide range of attacks in order to assess the security of maritime systems. The authors discussed the lack of security in marine systems, particularly the ones utilizing the NMEA protocol. Then, they presented the architecture of the tools and evaluated it against the open-source OpenCPN chart plotter.

Another application of the NMEA protocol is found in AIS. The vessels are equipped with an AIS to improve the safety and efficiency of navigation and to protect the marine environment [25]. It is a compulsory component for vessels under specific conditions described in the Safety of Life at Sea (SOLAS) Convention [26]. An AIS transceiver transmits static, dynamic, and voyage related information as well as safety related messages using the format of NMEA messages [25,27]. Several works have addressed anomaly detection in AIS. Iphar et al. [28] proposed an integrity assessment of AIS messages from a data quality perspective. The authors targeted AIS messages for data quality assessment and conducted several manipulation functions on AIS messages to invoke anomalies in the data. Then they proposed a rule-based detection approach. In another work, Blauwkamp et al. [29] utilized machine learning for detecting anomalies in traffic inferred from AIS messages. Although the authors did not target cybersecurity, cyber attacks are among the main

motivations of their research. Although NMEA and AIS messages have a relatively similar format, AIS messages include encoded binary payload instead of a textual payload in the NMEA-0183. Furthermore, AIS messages carry different information than NMEA messages, such as traffic messages from other ships. These differences motivated the work in this paper to investigate suitable anomaly analysis and detection approaches.

The origin of NMEA comes from the CAN protocol or CAN bus, a message-based protocol that enables communication among devices in automobiles [30]. Several works in the literature have addressed anomaly detection in CAN bus. We aim to infer relevant artifacts from the domain of CAN bus anomaly detection and utilize them for NMEA anomaly detection. In this paper, we rely on the state-of-the-art Intrusion Detection Systems (IDS) for CAN bus in the automotive domain which was captured by Lokman et al. [31]. The authors discussed several aspects, namely deployment strategies, detection approaches, attacking techniques, and technical challenges related to the field. Due to the similarities between NMEA and CAN bus, several artifacts were found relevant to our work and they will be discussed throughout this paper.

The systematic anomaly analysis of NMEA messages proposed in this paper is influenced by the Six-Step Model proposed by Sabaliauskaite et al. [32]. The authors proposed six steps for conducting joint safety and security risk analysis process utilizing six dimensions, namely, functions, structure, failures, attacks, safety countermeasures, and security countermeasures. Accordingly, the anomaly analysis in this paper consists of six steps; each step analyzes a different dimension related to NMEA anomaly detection, namely, navigational functions, messages, fields, anomalies, attacks, and detection methods. The systematic and multidimensional nature of the analysis in the Six-Step model has influenced our proposition. Additionally, the analysis of NMEA messages and their anomalies is influenced by the AIS analysis process conducted by Iphar et al. [28] (more details in Section 4.4).

The attack procedures in our work are influenced by the domain-specific information provided in the work of Hareide et al. [33]. The authors in [33] have discussed a cyber kill chain in maritime toward increasing the navigators’ preparedness against cyber attacks. Specifically, they have conducted a contextual attack targeting navigational information received at the ECDIS. Moreover, we relied on the *ATT&CK* framework [34] to describe the conducted attack techniques in the attack scenarios. The *ATT&CK* framework was chosen due to its comprehensive threat model in describing adversarial behavior.

3. Methodology

This paper focuses on the detection of anomalies in NMEA messages that can be caused by malicious actors. Figure 1 depicts the proposed meta-model of the NMEA anomaly detection system. Several NMEA message types support several navigational functions. Each message consists of several fields, each holding specific information. Attackers conduct attack procedures to impact navigational functions by targeting message types or fields. Defenders implement detection algorithms to protect the navigational functions by monitoring message types and fields to detect attack procedures. The meta-model is general in nature; as such, it is relevant to any use case that utilizes sensor data communicated in NMEA messages for navigational functions.

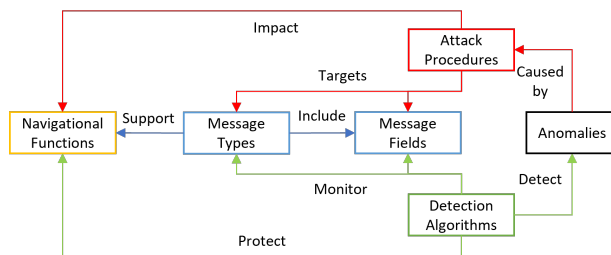


Figure 1. NMEA anomaly detection meta model.

We propose a method for systematic and multidimensional anomaly analysis of NMEA messages toward the development of an NMEA-focused anomaly detection solution. Anomaly analysis, as defined in the data quality domain, is a process for analyzing the values in a data set empirically, looking for unexpected behavior [35]. In this process, NMEA messages are analyzed for identifying possible anomalies, their impact, and ways in which they can be invoked and detected. A detailed description of the proposed method is provided hereafter:

3.1. Step 1—*Navigational Functions (i.e., Tasks)*

The identification of the navigational functions which rely on NMEA messages. These functions are defined for an INS by the IMO and for autonomous ships by classification societies describing the different tasks to be carried by marine systems or personnel such as route monitoring, collision avoidance, engine monitoring and control, and others. This information can later be utilized in risk analysis and, specifically, impact assessment.

3.2. Step 2—*Message Types*

The identification of targeted message types, and categorizing them according to relevant attributes such as their navigational functions (e.g., engine monitoring) and source (e.g., engine). This step specifies the scope of the analyzed messages and is expected to be system-dependent, since each system supports a specific list of messages.

3.3. Step 3—*Message Fields*

The identification of relevant message fields, the identification of the type of information they hold (e.g., speed, heading, etc.), and the format of each field. The type of information is useful for the identification of related message fields within the same message and across different message types. The information is useful for both attack and detection activities. A sophisticated attacker is expected to reflect a completely modified view, while the defender can detect anomalies by observing relevant fields for inconsistencies. On the other hand, the format and other aspects such as the range of each field are useful for the development of anomaly detection methods.

Steps 2 and 3 rely heavily on the format of the analyzed messages. The format of an NMEA-0183 message is depicted in Figure 2. After the starting delimiter, (\$), a 2-letter NMEA talker ID (e.g., GP for GPS) is attached to a 3-letter message ID specifying the message type (e.g., DTM, RMC, etc.). Then, each NMEA message has several fields, each corresponding to a certain piece of information, such as time, longitude, Speed Over Ground (SOG) etc. Then, a 2-digit in hexadecimal format that represents the calculated sentence checksum is separated from the last field value using the checksum delimiter (*). Finally, a carriage return and a line field specify the end of each message.



Figure 2. The format of an NMEA message.

3.4. Step 4—*Anomalies*

The identification of anomalous patterns (e.g., unusual values and events) that may appear during operations. During this step, all messages within the scope and their fields are analyzed to identify anomalous patterns based on some categorization of anomalies.

3.5. Step 5—Attack Techniques

The identification of attack techniques that can be carried out to invoke anomalous patterns in the selected message types and their message fields.

3.6. Step 6—Detection Algorithms

The identification of suitable detection algorithms for detecting anomalous patterns caused by attack techniques carried against NMEA messages. This step is tightly coupled with the previous step as the detection algorithm is continuously challenged and enhanced with improved attack techniques until a sufficient efficiency level is achieved.

4. Systematic NMEA Analysis Considering APS and INS Use Cases

In this section, we discuss the activities, artifacts, and results of our proposed NMEA analysis approach presented in Section 3, considering both the INS and the APS use cases. This is aimed to demonstrate the utility of the proposed anomaly analysis process in addition to the development of a suitable anomaly detection solution.

4.1. Step 1—Navigational Tasks and Functions

The tasks and functions for the INS and APS use cases were identified. The tasks of the INS were defined in the Resolution MSC.252(83) “Adoption of the revised performance standards for Integrated Navigation System (INS)” by the IMO [36]. On the other hand, the functions for the APS are defined by Amro et al. [37].

4.1.1. Navigational Tasks of the INS

The concept of the INS was developed to enhance the safe navigation of vessels with integrated and augmented functions. The INS consists of six navigational tasks [36], as follows:

- Route Monitoring (INS-RM): continuous monitoring of the own vessel as per the planned route [38].
- Route planning (INS-RP): capability of route planning (e.g., store and load, import, export, documentation), route checking based on minimum under keel clearance, drafting and refining the route plan against meteorological information [36].
- Collision Avoidance (INS-CA): detecting and plotting other ships and objects in the vicinity in order to prevent collisions [38].
- Navigation Control Data (INS-NCD): providing data to the task station for the manual and automatic control of the ship [38].
- Navigational Status and Data Display (INS-NSDD): displaying several information (e.g., AIS data, Maritime Safety Information (MSI) messages, INS configuration), and providing management functions [36].
- Alert management (INS-AM): centralized alert management on the bridge for the monitoring, handling, distribution, and presentation [38].

The INS facilitates the performing of the aforementioned navigational tasks. The tasks of “route monitoring” and “collision avoidance” are mandatory as per the IMO’s regulations [38]. Moreover, the requirements of “presentation of navigation control data for manual control” of the navigation control data task and “Module C” of the alert management task should be fulfilled [36]. Given that the lack of some navigational tasks and requirements in the INS may increase risks in the safe navigation of the vessel, they are classified as mandatory by the IMO. For instance, while “collision avoidance” and “route monitoring” are mandatory navigational tasks for an INS, the “route planning” and “navigational status and data display” are left optional by the IMO [36]. In the next step, the relevant NMEA messages to each navigational task is identified. Such a matching enables us to understand the risk level of potential NMEA anomalies by considering mandatory and optional navigational tasks defined by the IMO.

4.1.2. The Functions of the APS

There exists no regulatory framework or globally accepted guidelines that define the functions of an APS. Nevertheless, in our previous work [37], we have compiled a group of expected APS functions based on a group of relevant works including the work of Rødseth et al. [39] in the “Maritime Unmanned Navigation through Intelligence in Networks (MUNIN)” project as well as class guidelines for autonomous and remotely operated vessels by DNV [40]. A brief summary of the expected APS functions is discussed below (refer to [37] for more details):

- Engine Monitoring and Control functions: the monitoring and control of APS engine. They can be conducted by the APS itself (APS-AEMC), a Remote Control Center (RCC) (APS-REMC), or an Emergency Control Team (ECT) (APS-EEMC).
- Navigation Functions: establishing situational awareness. They can be conducted by the APS itself based on the sensor data (APS-AN), at the RCC based on the sensor data transmitted from the APS (APS-RN), or by the ECT based on the sensor data transmitted from the APS (APS-EN).

Impacting these functions through cyber attacks could cause safety, financial, and operational consequences according to a previously conducted risk assessment [41].

4.2. Step 2—Message Types

There are many NMEA messages (i.e., sentences) defined in the IEC 61162-1 standard [42]. In this paper, we will restrict our analysis on the NMEA messages broadcasted by the Bridge Command simulator (<https://www.bridgecommand.co.uk/> (accessed on 16 February 2022)) in order to evaluate the proposed analysis process. The messages were investigated using the guideline of the IEC 61162-1 standard [42] which is compatible with NMEA 0183. The messages in addition to their descriptions are shown in Table 1.

Table 1. NMEA messages within the analysis scope.

Msg.	Description	Msg.	Description
DTM	Datum reference	GGA	Global positioning system (GPS) fix data
GLL	Geographic position—Latitude/longitude	HDT	Heading true
RMC	Recommended minimum specific GNSS data	ROT	Rate of turn
RPM	Revolutions per minute	RSA	Rudder sensor angle
TTM	Tracked target message	ZDA	Time and date

After identifying the targeted NMEA messages for analysis, their involvement in the navigational functions is analyzed. This analysis can reflect the impacted navigational function for each NMEA message that is a subject of an attack. Table 2 depicts the identified relevance between messages and the APS and INS functions. A message is considered relevant to a function if it provides a piece of information that influences performing the function. For instance, the APS-AN function relies on the location information (i.e., coordinates) communicated in either one of the GGA, GLL, or RMC messages for route planning. Regarding the INS use case, the targeted NMEA messages are involved in all the INS functions except “Alert Management” using Resolution MSC.252(83) [36]. On the other hand, as the APS use case is in its the early development stages, we considered the possible involvement of each message in the APS navigational functions based on the designs and concepts communicated in the literature. Our analysis suggests that all considered messages are expected to be involved in the navigation functions except for the RPM messages, which are expected to be involved in the engine monitoring and control functions.

Table 2. Mapping of the messages and their supported navigational function.

Message \ Function	Function											
	INS-RM	INS-RP	INS-CA	INS-NCD	INS-NSDD	INS-AM	APS-AN	APS-RN	APS-EN	APS-AEMC	APS-REMC	APS-EEMC
DTM	✓				✓		✓	✓	✓			
GGA	✓				✓		✓	✓	✓			
GLL	✓				✓		✓	✓	✓			
HDT	✓			✓	✓		✓	✓	✓			
RMC	✓			✓	✓		✓	✓	✓			
ROT	✓	✓		✓	✓		✓	✓	✓			
RPM				✓	✓					✓	✓	✓
RSA				✓	✓		✓	✓	✓			
TTM	✓		✓				✓	✓	✓			
ZDA		✓		✓	✓		✓	✓	✓			

4.3. Step 3—Message Fields

During this step of our analysis, we have analyzed the fields of all the messages identified during step 2 (Section 4.2). The goal of this analysis is to understand the utility, and format of the piece of information depicted in each field. This understanding facilitates the activities to be conducted in the upcoming steps. Furthermore, the related NMEA messages are identified and depicted in Table A1 in Appendix B. Two messages are considered to be correlated if a change in information contained in one message that occurs under normal circumstances, will (direct effect) or might (indirect effect) change information in the other message. An example of an indirect effect can be observed in the relation between the RMC and RPM messages: changes in the Speed over Ground (SOG) in the RMC message might reflect different engine speed which is captured in the revolutions per minute field within the RPM message. On the other hand, an example of a direct effect has been observed in the location information (i.e., longitude and latitude) that is communicated in three messages, namely, GGA, GLL, and RMC. If any value changes in any message, it should be reflected in the other messages. Such information is valuable for both attack and detection activities.

4.4. Step 4—Anomalous Patterns

In this step, the possible anomalies that can be observed in NMEA messages are identified. Iphar et al. [28] proposed 13 possible anomalous patterns in AIS messages, that follow the same standard as NMEA with some differences in the message format as well as in content. However, they have the same abstraction of message types, each consisting of several message fields. In this paper, we adopt the relevant anomalies and neglect those that are not relevant to NMEA messages. Seven main anomalous patterns have been identified; these, along with brief descriptions are shown in Table 3.

Table 3. The anomalies within scope.

Anomaly	Description
Sudden unexpected change	an abnormal change in a field value in a certain period of time.
Nonexistent value	a value that does not match the system specification.
Unexpected value	a field value that is outside the usual norm.
Incorrect value	a value that does not match a reference value (e.g., time).
Data field evolution	an abnormal pattern in a field value over time.
Conformity issues	values that are not within the protocol specifications.
Unusual reporting	reduced or increased reporting rate over a period of time

We have analyzed all the aforementioned anomalies against all message types and their corresponding fields and recorded our results for the next steps. Some examples of the identified anomalies are presented in Table 4 while a list of all the identified anomalies is provided in Table A2 in Appendix C.

Table 4. Examples of possible NMEA anomalies.

Anomaly	Message	Field	Description
Sudden unexpected change	RMC	SOG	The Speed over Ground (SOG) has abnormally changed.
Nonexistent value	TTM	Target Distance	Target distance larger than radar range
Unexpected value	ROT	Rate Of Turn	Abnormal rate of turn value
Incorrect value	RMC	UTC	The UTC timestamp is not correct compared to a reference time value
Data field evolution	TTM	Target Status	Abnormal patten in the target status over time
Conformity issue	RPM	Source	The source field contains values that are not either E (Engine) or S (Shaft).
Under Reporting	RMC	-	The rate of receiving RMC messages is less than usual
Over Reporting	DTM	-	The rate of receiving DTM messages is more than usual

4.5. Step 5—Attack Techniques

In this section, we discuss the activities performed during the fifth step in the analysis concerning attack techniques that are expected to invoke one or several of the anomalies identified during step 4. In this direction, we propose the application of the *ATT&CK* framework [34] for threat modeling due to its comprehensive nature and suitable level of abstraction [41]. However, other threat modeling methods can still be applied if they propose attack techniques that can be technically achieved. The threat modeling approach considers both simple attacks as well as sophisticated attacks. Lokman et al. [31] discussed several attack types against the CAN bus, namely, packet insertion, erasure, reply, and payload modification. In the *ATT&CK* framework, insertion, reply, and payload modification may fall under the Manipulation of View (MoV) attack technique [43] while packet erasure may fall under Denial of View (DoV) attack technique [44]. A brief description of each technique is provided below:

- DoV attacks entail denying the seafarers or the depending systems the ability to render a live perception of the physical environment. This is achieved by dropping one or several NMEA messages to hinder the relevant navigational functions.
- MoV attacks entail the modification of the live perception of the physical environment. This can be done in several ways:
 - Fixed: the attacker modifies the values in original NMEA messages to specific fixed values. For example, no matter what is the real speed reflects another fixed speed value. This emulates a simple threat actor using simple Man-in-the-Middle (MitM) attack rules (i.e., filters).
 - Context attacks: the attacker manipulates the messages based on the values observed in the original messages to create a gradual change. This emulates a more advanced threat actor using more sophisticated MitM attack rules. Avoiding detection is among the attacker’s objectives.
 - Confusion attacks: the attacker sends crafted or repeated messages in addition to the original messages.
 - Replay attacks: the attacker replays a fixed set of messages instead of the original stream of messages.

We have analyzed the anomalies and the possible attacks that can invoke them. A mapping between an anomaly and an attack is identified if the attack, based on its definition, may result in invoking the anomaly. The identified relations are depicted in Table 5. The table can be read as follows: a DoV attack is expected to only invoke an “Under Reporting”

anomaly, while a confusion MoV attack is expected to invoke all possible anomalies except “Under Reporting”.

Table 5. Mapping of the anomalies and the attacks that are expected to invoke them.

Attack/Anomaly	DoV	MoV			
		Fixed	Context	Confusion	Replay
Sudden unexpected change		✓		✓	✓
Nonexistent value		✓	✓	✓	✓ ¹
Unexpected value		✓	✓	✓	✓ ¹
Incorrect value		✓	✓	✓	✓ ¹
Data field evolution		✓	✓	✓	✓ ¹
Conformity issue		✓	✓	✓	✓ ¹
Under Reporting	✓	✗ ²	✗ ²	✗ ²	✗ ²
Over Reporting		✗ ²	✗ ²	✓	✗ ²

¹ If the attacker replays a fixed set of messages that are not generated by normal means (i.e., forged). ² If the attacker maintained the normal message transmission rate.

To realize such attack techniques, we have developed a system called NMEA-Manipulator to facilitate the process of invoking the identified NMEA anomalies. NMEA-Manipulator intercepts and controls the flow of NMEA messages following a set of rules (detailed description in Section 5.1).

4.6. Step 6—Detection Algorithms

Lokman et al. [31] discussed several detection algorithms for detecting attacks against CAN bus messages. Three main detection approaches have been observed in the literature, signature-based, anomaly-based detection, and specification-based. A brief description of each approach is provided below in addition to our rationale for its utility in our analysis:

- Signature-based detection refers to the utilization of a specific signature or event for the detection of a specific malicious activity [45]. This would require documented attacks against NMEA messages to generate suitable signatures.
- Anomaly-based detection refers to the observing of real-time activities in a system and comparing them to normal behavior and raising an alarm when a deviation of normal behavior is observed [46]. This approach includes machine learning, frequency, statistical, and hybrid-based approaches. We argue that the machine learning and statistical approaches require a large set of data to effectively train robust models and, consequently, they are currently not viable options in our case. We have reached this conclusion after experimenting with a one-class support vector machine, and decision trees for detecting anomalies. The model evaluation has reflected poor performance mostly associated with the limited size of the data set. Since there exists no publicly available data set for the NMEA messages in the scope of our analysis, we have not pursued machine learning and statistical based approaches any further. On the other hand, frequency-based detection, considering message arrival frequency, was found relevant and is further considered for evaluation.
- Specification-based detection refers to the application of suitable thresholds and rules for describing the well-known behavior of a component [47]. We argue that this approach is the most suitable in the scope of our analysis because it does not require a large amount of data for learning. Moreover, considering the dynamic, yet predictable nature of NMEA messages, their behavior might be confined within a set of rules and thresholds (i.e., specifications). We have identified several categories of specifications, namely, physical, system, protocol, and environment specifications. A brief description of each category is provided below:
 - Physical specifications restrict the manner in which the values change over time among consecutive messages (e.g., maximum change in distance). This is related to the physical environment the NMEA messages are intended to reflect.

- System specifications restrict the values in the NMEA fields and their evolution over time for each system (e.g., maximum engine rpm, SOG acceleration, etc.). This needs to be defined for each target system.
- Protocol specifications restrict the format of the NMEA messages and their fields (e.g., UTC format in the UTC field of GGA, GGL, and RMC). This needs to be defined for each target protocol; in our analysis, NMEA0183 is utilized.
- Environment specifications restrict a range of values that are related to the operational environment. This includes time, date, longitude, latitude, datum code, and others.

We have analyzed the identified anomalies by considering the expected useful detection methods. A mapping between an anomaly and a detection method is identified if the anomaly can violate a certain specification or threshold in the corresponding detection method, based on their definitions. For instance, a sudden unexpected change in any field value, within the predefined format and range, does not violate the protocol specification of that field. Furthermore, changing the filed values alone is not expected to change the frequency of message arrival. Therefore, protocol specifications and frequency-based detection are not expected to be useful for detecting this type of anomaly. However, a sudden change in certain fields related to system, physical or environmental parameters such as speed, time, and distance, would violate the corresponding specifications. Table 6 depicts the results of our analysis. Our analysis suggests that frequency-based anomaly detection might only be suitable for under and over reporting anomalies. On the other hand, the specification-based approach can be used for the remaining anomaly types, using different specification categories.

Table 6. Mapping of the Anomalies and proposed detection methods.

Anomalies	Specification-Based				Anomaly-Based
	Physical	System	Protocol	Environment	Frequency
Sudden unexpected change	✓	✓		✓	
Nonexistent value		✓			
Unexpected value	✓	✓		✓	
Incorrect value				✓	
Data field evolution	✓	✓			
Conformity issue			✓		
Under Reporting					✓
Over Reporting					✓

In the sequel, we analyze the different detection methods against the messages and their fields to identify the required specification rules for detection.

5. Data Generation and Preparation

In this section, we present the results of our experiments throughout our analysis in order to evaluate its usefulness in the development of an intrusion detection solution. Initially we generated data, prepared it, and enriched it to facilitate the analysis. Afterwards we utilized the generated data for the identification of candidate specifications and rules for detecting anomalies. The activities in this process have been conducted using our developed maritime-themed cybersecurity testbed, which we proposed and presented in our earlier work [48]. The testbed includes several components that support the development of the anomaly detection solution.

5.1. Data Generation

The data generation process is facilitated by the availability of a simulator software or a device that generates NMEA messages. There are several NMEA simulator software, such as BridgeCommand simulator (<https://www.bridgecommand.co.uk/> (accessed on 18 February 2022)) and NMEA simulator (<https://github.com/panaaj/nmeasimulator>

(accessed on 18 February 2022)). The BridgeCommand simulator software was utilized in this paper since it allows for customized scenarios. The customization includes the navigational area, ship class, weather, time, and options for the included technologies on board.

Additionally, the data generation process is supported by capabilities for conducting various attack scenarios in order to invoke the different analyzed anomalies. For this reason, we have developed a system called NMEA-Manipulator. Similar to the previously proposed BRAT assessment tool [24] NMEA-Manipulator enables conducting a wide range of attacks against NMEA traffic. However, the main design goal of it is not to assess the security of marine systems, rather to facilitate the analysis of NMEA anomalies toward the development of intrusion detection systems. NMEA-Manipulator is utilized in two steps, namely, step 5 of the NMEA anomaly analysis process to observe the impact of the suggested attack techniques, and it is utilized as well in the data generation step for evaluating the different detection algorithms. An overview of NMEA-Manipulator is depicted in Figure 3.

The NMEA-Manipulator must be hosted in a device that is connected to the same LAN that connects the original NMEA speaker and listener. Furthermore, the NMEA-Manipulator requires two additional Non-Developmental Items (NDI), namely, a MitM tool such as *Ettercap* and a network sniffer tool such as *tshark*. The MitM tool provides the ability to access and control the LAN traffic, while the sniffer tool allows the recording of the original NMEA messages into an NMEA Messages File (NMF). Additionally, NMEA-Manipulator requires attack rules to govern its behavior. The attack rules are inserted into an Attack Rules File (ARF). The ARF contains multiple lines, each corresponding to a certain attack scenario (i.e., invoked anomaly). The attack rules command the NMEA-Manipulator to perform one or more of the attack techniques discussed in Section 4.5.

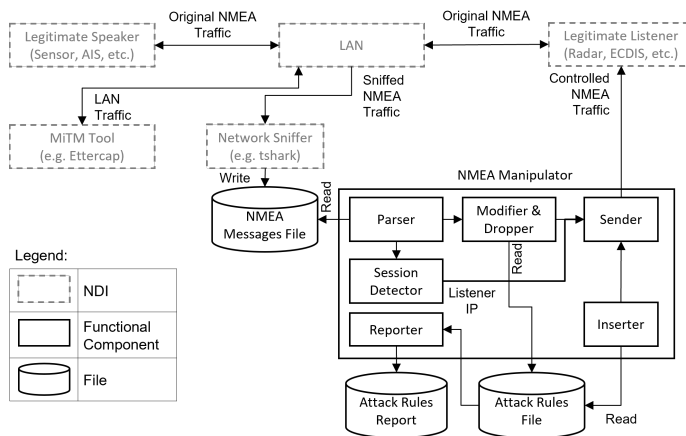


Figure 3. Overview of NMEA-Manipulator.

The NMEA-Manipulator includes six main subcomponents (i.e., modules), a parser, a session detector, a modifier and dropper, an inserter, a sender, and a reporter. The parser reads the NMEA messages from the NMF and invokes the modifier and dropper in case a new message is observed. The modifier and dropper then apply the activated attack rules specified in the ARF while the inserter applies the activated insertion attack rules as well as the replay and confusion attacks. The session detector identifies the IP address of the NMEA listener from the sniffed traffic and forwards it to the sender which sends the modified and inserted messages to the listener. Finally, the reporter generates a log file regarding the activated, deactivated, and modified attack rules to facilitate the later steps of the analysis.

In this direction, we utilized NMEA-Manipulator during the data generation process by conducting three experiments. Each experiment consists of several attack scenarios running in conjunction with a normal navigational scenario. The attack scenarios were chosen to be comprehensive so that they invoke a wide range of the identified anomalies. All the exercises included a normal navigational scenario which is following a predefined path in an area near the UK using a large ship equipped with a RADAR Detection Furthermore, Ranging (RADAR) and a GPS. A brief summary of the conducted experiments is presented below:

1. A combination of different MoV attacks, namely, fixed and context attacks was conducted in an attempt to invoke five anomalies, namely sudden unexpected change, nonexistent value, unexpected value, incorrect value, and data field evolution. The attack scenarios targeted several messages and message fields such as going back in time 1 day by changing UTC fields in GGA and GLL messages. Another example is increasing the distance of RADAR targets as well as other fields in the TTM message, to create a collision scenario.
2. Several MoV context attacks and DoV attacks were conducted to invoke data field evolution and under reporting anomalies, respectively.
3. A combination of different MoV attacks was conducted, namely fixed, confusion and replay attacks. The goal is to invoke several anomalies, including conformity issues and over reporting.

More details regarding the conducted experiments can be found in Table A3 in Appendix D.

The testing environment used to realize the different scenarios is depicted in Figure 4. The ship view is produced using the bridge command simulator, which is utilized as the NMEA sender. The simulator includes a simulated GPS device sending NMEA messages over UDP to the listener. On the other hand, the chart plotter view is produced by the OpenCPN chart plotter software, which is utilized as the NMEA listener. Moreover, the *Wireshark* software [49] is utilized for capturing the network traffic at the receiving node. The attacker node operates the Kali Linux operating system with the NMEA-Manipulator system.

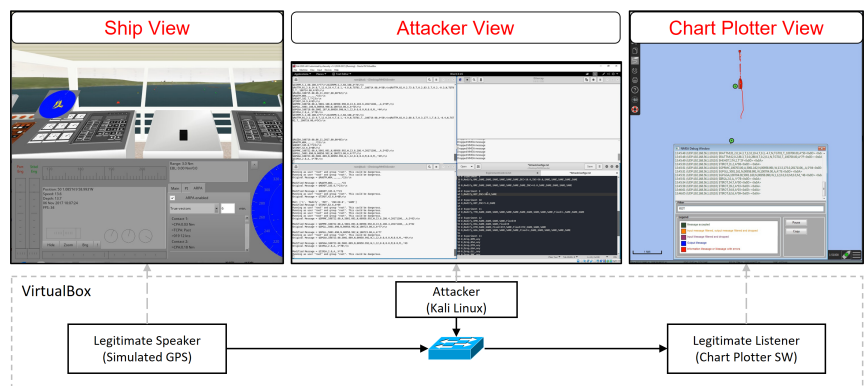


Figure 4. The testing environment.

Two artifacts are generated from each experiment, an experiment log and a packet capture of the traffic arriving at the NMEA receiver. The log is utilized to facilitate the labeling of NMEA messages (e.g., attacked or normal) and to support traceability of the events occurred during the experiments. The packet capture on the other hand is utilized for later steps in evaluating the different anomaly detection methods. Finally, additional experiments were conducted to capture NMEA messages in normal operations to aid the efforts in the identification of suitable specifications and rules for detecting anomalies.

In these experiments, only the NMEA sender and receiver were operational while the attacker node was kept idle.

5.2. Preparation and Enrichment of Data

The captured network traffic from the conducted experiments was utilized as input in this step. *Tshark*, the command-line interface of *Wireshark*, was utilized for extracting the NMEA messages with the associated time of observation of each message. The time information is needed since some NMEA messages do not contain such information. Then, data cleaning is conducted to fix some issues in the messages such as splitting concatenated messages and removing carriage return characters. Afterwards, the messages with the timing information are transformed into feature vectors. Moreover, the messages are labeled as “Normal” or “Attack” by referring to the experiment documentation. For instance, during the different experiments, some NMEA message types were not targets of attacks, which leads to a “Normal” classification of such messages and an “Attack” classification of the targeted message types during the period of activity of the attack rules. Additional fields are added for the enrichment of the captured data, such as distance from the last position, rate of message arrival etc. This is intended to facilitate the analysis of the specifications and anomalies. Finally, the output feature vectors with the headers and labels are exported into Comma-Separated Value (CSV) format. Due to different features for each NMEA message type, it has been deemed necessary to evaluate the specifications for each type separately. For this reason, the output from the previous step is a group of CSV files, one for each message type. Each file holds all the messages of a single message type that were generated in a single experiment. All activities performed during this stage except for the manual labeling of some feature vectors are automated using python scripts.

6. Evaluation and Discussion

We have analyzed the generated and processed data to evaluate the proposed detection approaches. A selected group of results is highlighted hereafter.

6.1. Specification-Based Detection

The results of the evaluation of the different proposed specification-based detection categories are presented in this section, namely, environment, physical, system, and protocol specifications.

6.1.1. Protocol Specifications

The approach is implemented through a validation process for each message to check the compliance of its fields with the protocol specifications. Considering the known structure and format for each message type, this approach can detect conformity issues with high confidence. For instance, an attack scenario was carried during experiment 3 (Table A3) to manipulate the view regarding the number of satellites that may be utilized for position fixing. The attack is carried by modifying the original expected integer value with the character “x”. The impact of the attack was observed on the chart plotter software as the indicator of the number of satellites was changed to red, reflecting a bad signal, however, the true value was 13. A detection rule for this attack is implemented to check if the field format is in compliance or not and the results reflect correct detection.

6.1.2. Environment Specifications

The evaluation of this approach is not implemented, rather it is conceptually evaluated in the section. This is due to the rationale that such category of specification relies on true sources for reference information, such as correct time and datum code for coordinate ranges. The utilized simulator relies on static scenario definition files. The time of each simulation experiment is manually defined. Relying on the true system time to detect inconsistency generates false positives. Therefore, we argue that in a live implementation,

if the ship systems receive time information from a source other than GPS signals, detecting any modifications of the time information is straightforward.

6.1.3. System Specifications

This category of specifications is encoded in a file for each ship to host the anomaly detection solution. This file dictates to the detection software the specifications for the different systems on board to identify anomalies due to the existence of values that are not supposed to be observed in the host system.

Figure 5 depicts a visualization of the anomaly type “Non existent value” that is invoked by an MoV attack during experiment 1 (Table A3). The figure reflects two observations, the left part reflects detection of the RSA message of a sensor reading receiving from a port rudder sensor. However, the specification file of the simulated system dictates that it only has a starboard rudder sensor, which means a single rudder sensor. On the right part of the figure, a similar anomaly is detected in the RPM message when observing a new shaft (i.e., source number 4) that is not within the system specifications which is only having two shafts (i.e., source numbers 1 and 2). These two anomalies and similar ones are detected with high confidence due to the static nature of systems in ships. Furthermore, anomalies due to sudden changes can be detected using system specifications. For instance, the possible changes in the revolutions per minute (rpm) in the RPM message in one of the ship models in the simulator is in the range (25 to 2000 rpm) while the range for another model is (1 to 100). Detection rules for each one is formulated to detect any rpm changes outside the appropriate range. These rules will return anomalies with high confidence. However, if the attackers maintained a change in the rpm values that is within the system specifications, their actions will not be detected using this approach based on the rpm value as a feature. Other anomalies that can be detected using this approach are unexpected values and data field evolution. An example of an unexpected value can be observed in the rate of turn (ROT) value in the ROT message. Each ship has a certain expected limit within which the ship can turn in a minute. For instance, the difference in ROT value between two consecutive messages in one ship model in the simulator can change within the range (−100 to 100). An attack during experiment 2 demonstrates this effect by increasing the ROT value by a hundred while the true difference was 11.2, leading to a total difference of 111.2 which allowed the identification of the anomaly. However, similarly to the previous anomaly, an attacker can stay undetected if the resulted change is within the system specifications.

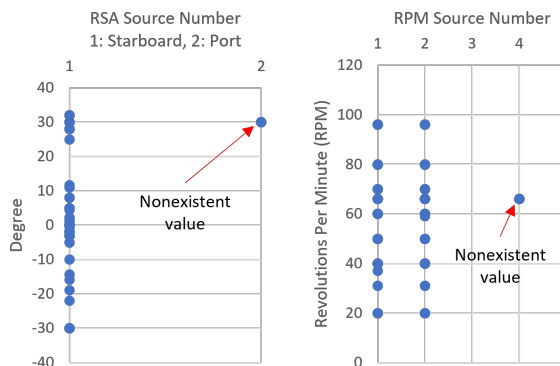


Figure 5. A plot visualizing the detection of a nonexistent value in a specific system.

6.1.4. Physical Specifications

We will demonstrate the utility and identified challenges in this approach in a selected group of messages, namely, RMC, RPM, and HDT, for detecting sudden unexpected change and data field evolution anomalies. An example of a sudden unexpected change

can be invoked through an MoV attack to reflect a position change in a strange manner. A demonstration of this was carried out during experiment 1 (Table A3). The impact of this attack can be observed clearly when visualizing the difference between each two consecutive latitude and longitude values, as depicted in Figure 6.

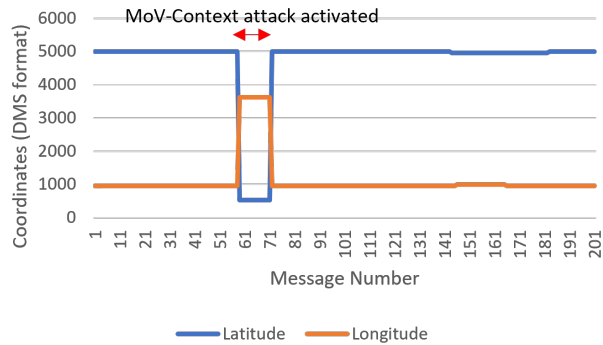


Figure 6. A plot visualizing the detection of a sudden change in position information in the RMC message.

A data field evolution type of anomaly can be invoked through gradually changing the SOG and Course Over Ground (COG) fields in the RMC message to reflect a different speed and heading details. An example of this attack is demonstrated during experiment 2 (Table A3). The SOG and COG values under normal conditions change in a specific manner. The SOG value between two consecutive RMC messages cannot physically change outside a certain range. Still, a visualization of a data field anomaly is depicted in Figure 7a. This anomaly can be detected by defining a rule specifying the range in which the SOG can change over a certain period of time. Then, any change outside this rule would suggest an anomaly with high confidence. Figure 7b depicts the manner in which the difference between consecutive SOG values behaves under normal and attack conditions. The impact of the attack is clearly visible and can be deduced from the figure. The behavior of the attacker can be described as reducing the SOG value by 1 several times in some time frame then returning the value to its normal value, which leads to a sudden increase. Similarly, the COG is not expected under normal conditions to have large differences between consecutive messages except if the value is reaching the limits in the range (−360 to 360); only then a large change is normal. In this paper, the threshold for a normal difference was determined based on the largest calculated difference in value between the same field in two consecutive messages in the normal recorded traffic. The determination of a more accurate threshold would require a larger data set of NMEA messages from real systems.

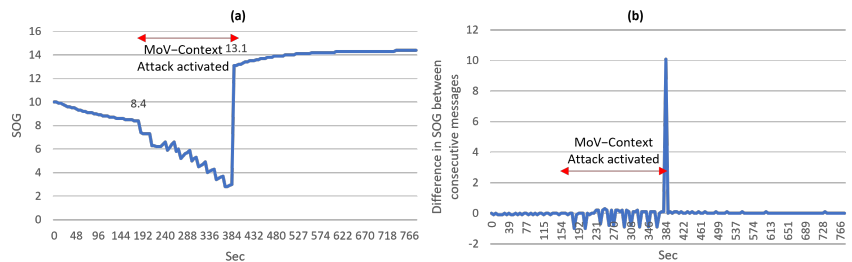


Figure 7. (a) A plot visualizing a data field anomaly in the SOG field in the RMC message (b) A plot visualizing the detection of a data field anomaly in the SOG field using the difference between consecutive messages.

We have mentioned previously that we are considering simple and advanced threat actors. For this reason, we have proposed the process of identifying the relevant NMEA messages and fields during step 3 (Section 3.3). If attackers crafted messages with differences within the normal range, correlating the relevant fields in other messages can be utilized to increase the confidence in the detection. A correlation between the COG field in the RMC message and the heading in the degrees field in the HDT message can provide evidence of anomalies in both messages under the condition that only one of them is subject to attack at a certain time. Figure 8a,b look very similar; this depicts the direct correlation between the two fields in two different messages. To evaluate the ability of detection based on the correlation of message fields, two attack scenarios were conducted at different times during experiment 2 (Table A3). One attack targeted the COG field by gradually increasing the COG value several times over a period of time and later returning it to normal value. The effect of this attack can be observed in Figure 8c. During the same time frame, the differences in the heading field do not reflect similar behavior; this is strong indication of an anomaly. Similarly, the other attack targeted the heading field in the HDT, causing a similar effect which is observed in Figure 8d. This also can provide strong indication of an anomaly. Notably, minor delay is sometimes observed before the COG and heading values match each other. Accommodating this constitutes a challenge for the detection algorithm and will be considered in future work.

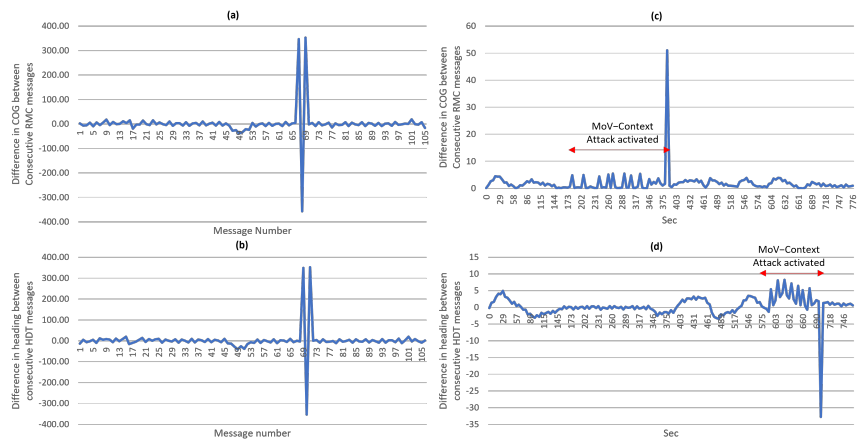


Figure 8. (a) A plot visualizing the normal manner the COG field changes between consecutive messages (b) A plot visualizing the normal manner in which the heading field in degrees changes between consecutive messages (c) A plot visualizing the detection of a data field anomaly in the COG field when correlating with the heading field in the HDT (d) A plot visualizing the detection of a data field anomaly in the heading field when correlating with the COG field.

6.2. Frequency-Based Detection

This category of specifications is solely suggested for detecting over and under reporting anomalies. The rationale behind it is that under normal conditions NMEA messages of each type have a specific amount of messages or packets arriving at the target system in a specific period. We refer to this metric as arrival rate. DoV attacks will drop some or all messages; this will lead to a decrease in the arrival rate which suggests under reporting. Furthermore, replay MoV attacks during which attackers replay messages at a reduced Transmission Rate (TR) will also cause an under reporting anomaly. On the other hand, if attackers increased the transmission rate to go beyond normal, the arrival rate will increase, which suggests over reporting. Similarly, confusion MoV can cause over or under reporting based on the transmission rate controlled by the attacker. To evaluate

the proposed detection for these anomalies, we utilize *Wireshark* to graph the number of packets per second in the recorded traffic of the relevant experiments.

The traffic contains one NMEA message per packet except for RPM messages; each pair is joined in one packet. Figure 9a,b depict the arrival rate for the recorded traffic in experiments 2 and 3 (Table A3), respectively. The impact of the DoV attacks is clearly visible in Figure 9a. During this attack, all messages were dropped, resumed, then dropped again. The third visible drop in the graph is due to the switching from the traffic generated through the NMEA-Manipulator and normal traffic. Figure 9b depicts the arrival rate during different attacks with different employed TRs. After starting with normal traffic, NMEA-Manipulator controlled the traffic and applied several fixed MoV attacks, which had a limited impact on the arrival rate. Then, a confusion MoV attack started by sending pre-recorded normal messages at TR of 1 message every 0.5 s in addition to the normal traffic. After that, the transmission rate was increased to send 1 message every 0.1 s using alternately recorded normal messages and recorded forged messages. Finally, the transmission rate was increased to 1 message every 0.05 s, similar to the last iteration, by using normal and forged recorded messages. The impact of confusion and replay attacks using different TR is clearly visible in Figure 9b as their arrival rate deviates in a clear manner from the normal arrival rate. However, if attackers targeted only a few message types, considering the arrival rate for the entire traffic might miss the targeted attack. Therefore, we propose that a dedicated arrival rate for each message type be monitored separately. This is particularly suitable in real ship systems in which different NMEA messages have different arrival rates. In order to avoid detection using frequency-based detection, attackers need to appropriately adjust the TR during their attacks, which complicates their task.

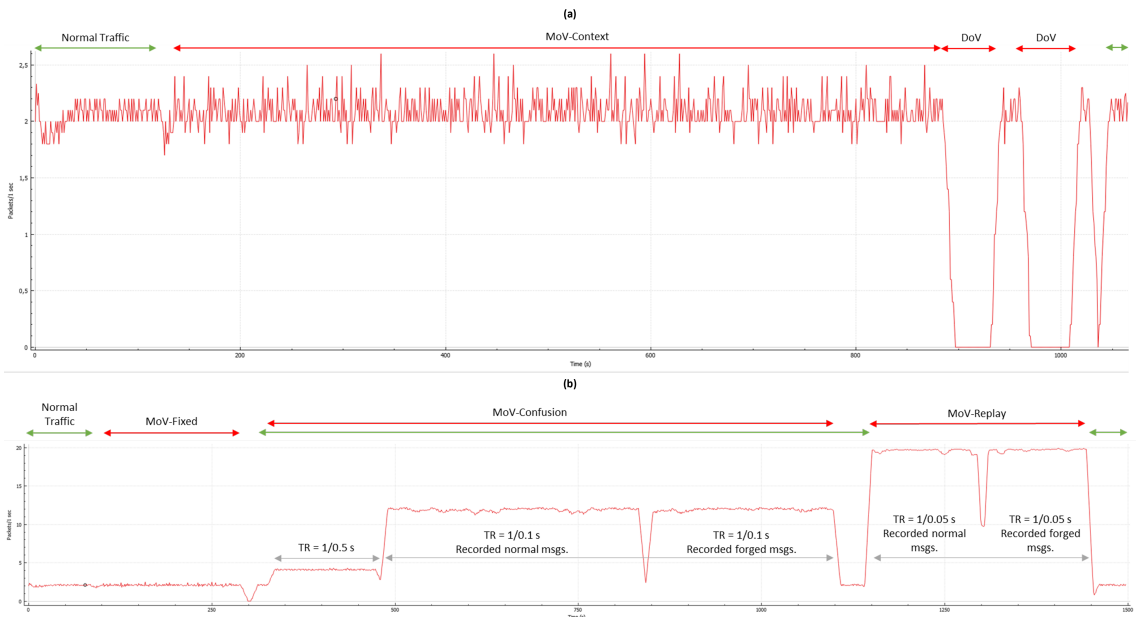


Figure 9. (a) A plot visualizing the number of packets per second during normal traffic, traffic with context MoV attacks and two DoV attacks. (b) A plot visualizing the number of packets per second during normal traffic, traffic with fixed MoV attacks, confusion MoV attacks, and replay MoV attacks.

6.3. Communication of Risk Associated with Detection

Our proposed analysis aims to develop an anomaly detection solution that is enriched with the operational context of the NMEA messages that are associated with the different functions. When an anomaly is detected using the different detection mechanisms for a

specific message or a group of messages, the impacted functions are known. Consequently, the multi-tier communication of the risk from the technical level to the operational level is facilitated. An example of this is as follows: if the anomaly depicted in Figure 8d against the HDT message is detected, the following is known:

- An inconsistency is identified by a physical-based specification concerning the heading information in HDT messages; this is due to steps 6 and 3. In step 3 the heading field was designated for identifying relevant anomalies, attacks, and detection methods. In step 6 the detection methods for the heading field were proposed.
- The anomalous messages are arriving from the gyro compass; this is known through the TalkerID identified in step 2.
- This might indicate a data field evolution anomaly; identified in step 4.
- This might be a result of a MoV attack; the relationship between the anomaly and the attack that possibly causes it is identified in step 5 (see Table 5).
- This could impact the Route Monitoring, Navigation control data and Navigational status and data display functions in the INS which might cause safety, financial and reputation loss and environmental damage; the relationships between the targeted message and the relevant functions are identified in step 1 (see Table 2).

6.4. Identified Limitations

The experiments have highlighted some possible limitations in the different anomaly detection approaches. A summary of the identified limitations is provided below:

- False positives: weakly configured specifications can generate false positive alerts. For instance, a system specification for one simulated ship is a range for RPM change of 100 between two consecutive RPM messages. Using the same specification in another ship with a different RPM change range would generate a false positive. Therefore, it is highly encouraged to fine-tune the system specifications for each target system. Another aspect to consider is the sensor error or noise. A noise in the sensor might invoke an anomaly and a false indication of malicious behavior. An instance of this issue has been observed in one of the experiments. A glitch in the simulator caused the speed of the vessel to abnormally change due to low water depth. An anomaly in the data is observed which is not caused by malicious behavior. These issues need to be considered during the development of the anomaly detection system.
- False negatives: malicious behavior operating within the thresholds defined in the specifications will not be detected but still can cause an impact. For instance, reducing the RPM value by 50 while the change threshold is 100 will not generate an alert, but, the speed value will appear less than what is expected, this can cause a speed increase command which in turn can cause an issue in safe navigation.

Moreover, other limitations in our analysis are mentioned hereafter:

- The number of analyzed messages is limited to those supported by the available simulator. Still, four message types, namely, GLL, RMC, GGA, ZDA are among the top 10 NMEA messages observed on the internet during a scan in Shodan [50]. Future work should focus on more message types.
- Our analysis included only NMEA0183. Other protocols are being integrated into the maritime systems including NMEA2000 [6] and OneNet [51]. Yet, NMEA0183 is still utilized in the maritime industry as observed in the literature and the internet-wide Shodan scan [50].
- Our analysis only considers attacks with objectives to impact navigational tasks by denying and manipulating the view that is rendered using the NMEA messages. Other attack techniques that can cause anomalies can be investigated in future work.
- We utilized the ATT&CK framework for the threat modeling. Using other threat modeling techniques might identify other attacks. Still, our considered attacks are in line with the attacks discussed in the literature.

- We utilized certain categories of anomalies during our analysis based on the observed anomalies in the literature. Other anomalies can exist. If new anomalies are identified in the future, this would require another iteration of the analysis process to consider relevant messages, fields, attacks, and detection methods.

7. Deployment Options

Lokman et al. [31] discussed the observed deployment strategies of IDS for CAN bus. It has been observed that CAN bus IDSs are placed either in the central or end nodes or within the CAN network. NMEA IDSs are expected to have the same options. However, the choice of placement is sensitive to the use case. Still, we discuss here the possible placement of NMEA IDSs in marine systems. Two main categories of IDS are observed in the literature, namely Host-based IDS (HIDS) and Network-based IDS (NIDS). Jacq et al. [52] have discussed the concept of situational awareness in naval systems and indicated the challenge in the application of HIDS due to possible warranty disruptions. On the other hand, NIDS is a more suitable deployment option, as it can be added to the networks for monitoring NMEA traffic and detecting anomalies. However, we argue that if attackers are able to target the ship network and successfully carry the attacks presented in this paper, NIDS might also be targeted using similar attack techniques, to avoid detection. Therefore, we argue that the optimal implementation can be achieved through the integration of the anomaly detection solution within the NMEA receiver node, as part of the software that consumes NMEA messages. Still, considering that this solution would allow real-time anomaly detection, a performance evaluation is crucial to validate that the solution does not hinder the navigation functions. Future work can target a proof-of-concept implementation through the development of an NMEA IDS and integrate it with the OpenCPN software.

8. Conclusions

The ongoing digitization in maritime is leading to new operational modes. The shipping operations are being gradually transferred to remote shore locations, relying on sensor data transmitted from the vessels. This mode of operation makes the shipping operations susceptible to a wide range of cyber-attacks including manipulation and denial of view, which subsequently hinders safe navigation. This paper targeted the detection of anomalies in NMEA messages caused by malicious actors. NMEA messages carry information that is crucial for several navigational functions, such as collision avoidance. The consequences of targeting them in cyber attacks could cause safety, operational and financial consequences.

Initially, a systematic analysis of NMEA messages was proposed. The analysis aims to identify anomalies in NMEA messages, their cause, and possible detection methods. Afterwards, several of the identified anomalies were invoked using some of the identified attacks in a testing environment against a group of simulated NMEA messages. Then the identified detection methods were evaluated.

Our analysis suggests relevant NMEA messages and fields which have demonstrated utility for detecting inconsistencies. Moreover, the systematic analysis provides a multi-tier overview of the risks associated with the detected anomalies. When an anomaly is detected at the technical tier, technical-level information can be induced, such as source device, candidate attack technique, and relevant messages for investigation. Furthermore, risk information is utilized at the operational or mission tier regarding the possible risk of the detected anomaly against the relevant functions. The employed attack techniques reflect several possible threat actors with varying degrees of complexity; this constitutes a good coverage of possible threats against the NMEA messages within scope. Other attack techniques to achieve other objectives than impacting the navigational functions will be considered in future work.

Specification-based and frequency-based detection has been demonstrated to provide detection capability using different approaches. Protocol specifications can provide a confident indication of conformity issues in the NMEA messages. Furthermore, system specifications can provide a confident indication of system-specific anomalies related to

some values and events that are not expected in a specific system. However, the efficiency of this approach relies on the strength of the defined specifications for each host system. Advanced attacks can still avoid detection. Additionally, physical specifications can provide indicators of anomalies, yet they require careful development of the specifications. We have demonstrated the utility of identifying relationships among some NMEA messages and their fields in the identification of data field evolution anomalies. The relationships can provide strong indications of anomalous patterns. Still, a comprehensive analysis of all the identified relationships is required to generalize the findings. Moreover, frequency-based detection can provide a strong indication of over and under reporting anomalies. However, advanced attackers can avoid detection by maintaining a transmission rate that is consistent with normal traffic.

Future work can utilize the results of our analysis to develop, implement, and evaluate an NMEA anomaly detection solution that is suitable for deployment onboard vessels. Furthermore, another direction could utilize a variation of our testing environment by including other NMEA simulators or a physical NMEA source (e.g., GPS or AIS); this would include other NMEA messages and fields. Additionally, this paper targeted attack techniques that can cause an impact on navigational functions. Future work can explore different attack techniques that do not aim to impact the navigational functions and can still invoke anomalies such as ex-filtration or command and control. Moreover, we have identified a possible application of machine learning techniques in NMEA anomaly detection. Yet, the limited amount of data was a challenge that can be targeted in future work. Finally, another direction for future work can be the extraction of physical and system specifications from a large amount of NMEA messages recorded in several systems, to improve the specification rules.

Author Contributions: Conceptualization, A.A., V.G. and S.K.; Data curation, A.A.; Formal analysis, A.A. and A.O.; Investigation, A.A.; Methodology, A.A.; Software, A.A.; Supervision, V.G. and S.K.; Validation, A.A., A.O. and V.G.; Writing—original draft, A.A., A.O.; Writing—review & editing, A.A., V.G. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by (a) the NTNU Digital transformation project Autoferry; (b) the Research Council of Norway through the Maritime Cyber Resilience (MarCy) project, Project no. 295077; and (c) the Research Council of Norway through the SFI “Norwegian Centre for Cybersecurity in Critical Sectors (NORCICS)”, Project No.310105

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data generated during the data generation process discussed in Section 5.1 can be shared upon request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NMEA	National Marine Electronics Association
APS	Autonomous Passenger Ship
INS	Integrated Bridge System
IMO	International Maritime Organization
AIS	Automatic Identification System
ECDIS	Electronic Chart Display and Information System
GPS	Global Positioning System
CAN	Controller Area Network
MoV	Manipulation of View
DoV	Denial of View

ARG	Attack Rules File
NMF	NMEA Messages File
UTC	Coordinated Universal Time
CSV	Comma Separated Valu
TR	Transmission Rate
SOG	Speed Over Ground
COG	Course Over Ground
DTM	Datum Reference Message
GLL	Geographic position Message
RMC	Recommended Minimum specific GNSS data Message
RPM	Revolutions Per Minute Message
TTM	Tracked Target Message
GGA	Global positioning system (GPS) fix data Message
HDT	Heading true Message
ROT	Rate of Turn Message
RSA	Rudder Sensor Angle Message
ZDA	Time and Date Message

Appendix A. NMEA Messages and Their Fields

The investigated NMEA messages and their fields in steps 2 and 3 are depicted in Figure A1. For more details please refer to the IEC 61162-1 standard.

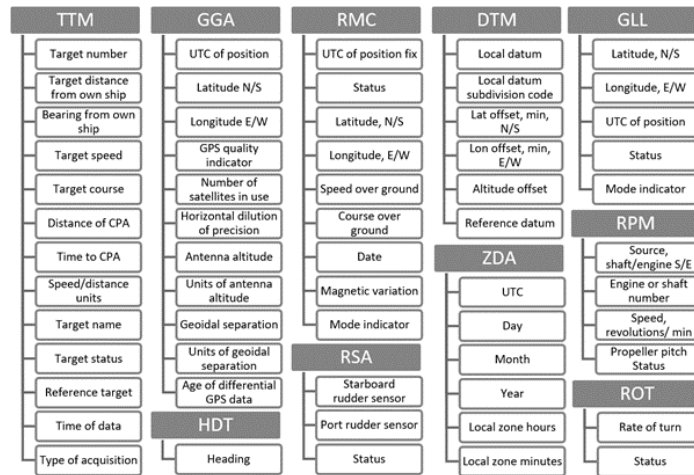


Figure A1. A list of the analyzed messages and their fields.

Appendix B. Interaction of NMEA Messages

In Table A1, the relevance among the analyzed NMEA messages is presented. This is an outcome of step 3.

Table A1. Relevant fields across the selected NMEA messages.

Msg	Fields	Related Msg	Related Fields
DTM	Local datum	GGA, GLL, RMC	Longitude, E/W, Latitude, N/S
		RMC	COG, Longitude, E/W, Latitude, N/S
		ROT	Rate of turn
		HDT	Heading
		RSA	Rudder angle
		DTM	Local datum
GGA	Longitude, E/W, Latitude, N/S	GLL	Longitude, E/W, Latitude, N/S
		GGA	Longitude, E/W, Latitude, N/S
		RSA	Rudder angle
		DTM	Local datum
		RMC	COG, Longitude, E/W, Latitude, N/S
		ROT	Rate of turn
GLL	Longitude, E/W, Latitude, N/S	ROT	Rate of turn
		HDT	Heading
		RSA	Rudder angle
		RMC	COG
		ROT	Rate of turn
		GLL, GGA	Longitude, E/W, Latitude, N/S
HDT	Heading	TTM	Bearing Time to CPA, Distance of CPA
		RSA	Rudder angle
		DTM	Local datum
		ROT	Rate of turn
		HDT	Heading
		RMC	COG
RMC	COG	DTM	Local datum
		ROT	Rate of turn
		HDT	Heading
		GLL, GGA	Longitude, E/W, Latitude, N/S
		RPM	RPM (if FPP) Propeller pitch (if CPP)
		RMC	COG
ROT	Rate of turn	HDT	Heading
		RSA	Rudder angle
		GLL, GGA	Longitude, E/W, Latitude, N/S
		TTM	Bearing Time to CPA, Distance of CPA

FPP: Fixed Pitch Propeller, CPP: Controllable Pitch Propeller, CPA: Closest Point of Approach.

Appendix C. The Identified Anomalies

The identified anomalies in the investigated messages and their fields in step 4 (Section 3.4) are presented in Table A2. These anomalies are relevant to the NMEA0183 protocol and can therefore be utilized in use cases other than the INS and APS.

Table A2. The identified anomalies in the NMEA messages within scope.

Anomaly	Message	Field	Anomaly	Message	Field	
Sudden unexpected change	RPM	RPM in FPP	Unexpected value	TTM	Target Speed	
		Propeller pitch in CPP			Time until CPA	
	GGA	Location (long. & lat.)		DTM	All Fields	
		UTC		ZDA	Local zone	
		GPS Quality Indicator			Local zone minutes	
		Age of differential GPS data		RMC	SOG	
	GLL	Location (long. & lat.)			Magnetic Variation	
		UTC			GPS Quality Indicator	
	ZDA	Status			Number of satellites	
		Time (UTC, day, month, year, zone, and zone minutes)			GGA	Horizontal Dilution
	HDT	Heading				Antenna Altitude
		T = True				Geoidal separation
	ROT	Rate Of Turn				Age of differential GPS data
		Status			ROT	Rate Of Turn
	RSA	Starboard rudder sensor			GGA	UTC
		Port rudder sensor		Incorrect value		UTC
		Location (long. & lat.)			ZDA	Day
		Status				Month
	SOG				Year	
	RMC	COG			GLL	UTC
Date			RMC	UTC		
UTC				Date		
Nav status			TTM	UTC		
TTM	Distance, bearing, speed, course, distance of CPA, time until CPA, units, and target status	Data field evolution	All Messages	All fields		
	UTC	Conformity issue				
	Type	Under Reporting				
	DTM	Over Reporting				
Nonexistent value	TTM	Target Distance				
		Bearing from own ship				
		T or R				
		Distance of CPoA				
	DTM	Speed/ distance units				
		Local datum code				
		Local datum subcode				
		Datum name				
	RPM	Source				
		Source number				
		RPM (i.e., speed)				
	RSA	Propeller pitch				
		Starboard rudder sensor				
	HDT	Port rudder sensor				
		Heading				
			T = True			

Appendix D. Data Generation Experiments

A summary of the conducted experiments during the data generation process discussed in Section 5.1 is presented in Table A3. The experiments resulted in generating normal and attack data for consequent anomaly detection steps. The table depicts the

targeted anomalies for invocation, the conducted attack techniques, the targeted messages, and a brief description of the attack.

Table A3. Summary of conducted experiments during data generation.

#	Attack Type(s)	Anomaly(s)	Message(s)	Description
1	MoV: Fixed	Sudden unexpected change	RPM	Fixed RPM to (60) and propeller pitch to (10) and Fixing True field in HDT to (R)
			HDT	
	MoV: Context	Nonexistent value	GGA, GLL, RMC	Changing position: Decreasing latitude and longitude by 45 degrees (4500) and minutes by (30.000)
			RPM, RSA	Changing source of RPM to Engine and number of source to 4. Streaming two sensor values for RSA starboard and port
	Mov: Fixed	Unexpected value	GGA, ROT	Fixed HDOP to (50.0) and increase ROT by 100
MoV: Context	GGA, GLL		Go back in time 1 day by changing the UTC field	
	TTM, DTM		Increase target distance by app 460 meters or 0.25 nautical miles (0.25), distance of CPA by (0.25), time until CPA by (1.0) and Modify Datum specs for default datum	
2	MoV: Context	Data field evolution	RPM	Increment RPM by (10)
			RMC	Decreasing SOG by (10.0) and Increasing COG by (45.0)
			ROT	Increasing ROT by 10
			HDT	Increasing heading by 5
			TTM	Losing/recovering targets fluctuation
DoV	Under reporting	ALL	Drop messages	
3	Mov: Fixed	Conformity issue	RPM, RSA, DTM, GGA	RPM and RSA status fields to M. DTM offsets to string. GGA number of sat. to x
	MoV: Confusion	Over reporting	All	Confusion attack
	Mov: Replay	Several	All	Replay attack

References

- Fruth, M.; Teuteberg, F. Digitization in maritime logistics—What is there and what is missing? *Cogent Bus. Manag.* **2017**, *4*, 1411066. [CrossRef]
- Levander, O.; Marine, R.R. Ship intelligence—A new era in shipping. In Proceedings of the Royal Institution of Naval Architects, Smart Ship Technology, International Conference Proceedings, London, UK, 26–27 January 2016; pp. 26–27.
- IMO. Autonomous Ships: Regulatory Scoping Exercise Completed. Available online: <https://bit.ly/3gFLigk> (accessed on 18 February 2022).
- Autonomous All-Electric Passenger Ferries for Urban Water Transport. Available online: <https://www.ntnu.edu/autoferry> (accessed on 18 February 2022).
- N.M.E. Association. NMEA0183 Standard. 2002. Available online: https://www.nmea.org/content/STANDARDS/NMEA_0183_Standard (accessed on 18 February 2022).
- Luft, L.A.; Anderson, L.; Cassidy, F. Nmea 2000 a digital interface for the 21st century. In Proceedings of the 2002 National Technical Meeting of The Institute of Navigation, San Diego, CA, USA, 28–30 January 2002; pp. 796–807.
- Jethwa, B.; Panchasara, M.; Zanzarukiya, A.; Parekh, R. Realtime Wireless Embedded Electronics for Soldier Security. In Proceedings of the 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2–4 July 2020; pp. 1–6.
- Singh, A.K.; Balamurugan, S.; Aroul, K.; Marimuthu, R. Design of universal module for personal security. *Indian J. Sci. Technol.* **2016**, *9*, 99031. [CrossRef]
- Aishwarya, K.; Manjesh, R. A Novel Technique for Vehicle Theft Detection System Using MQTT on IoT. In *International Conference on Communication, Computing and Electronics Systems*; Springer: Singapore, 2020; pp. 725–733.
- Tran, K.; Keene, S.; Fretheim, E.; Tsikerdekis, M. Marine Network Protocols and Security Risks. *J. Cybersecur. Priv.* **2021**, *1*, 239–251. [CrossRef]
- Kavallieratos, G.; Katsikas, S.; Gkioulos, V. Cyber-Attacks Against the Autonomous Ship. In *Computer Security; Lecture Notes in Computer Science*; Katsikas, S.K., Cuppens, F., Cuppens, N., Lambrinoudakis, C., Antón, A., Gritzalis, S., Mylopoulos, J., Kalloniatis, C., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Volume 11387, pp. 20–36.
- Vinnem, J.E.; Utne, I.B. Risk from cyberattacks on autonomous ships. In *Safety and Reliability—Safe Societies in a Changing World*; Haugen, S., Barros, A., van Gulijk, C., Kongsvik, T., Vinnem, J.E., Eds.; Taylor & Francis: London, UK, 2018.
- Svilicic, B.; Rudan, I.; Jugović, A.; Zec, D. A Study on Cyber Security Threats in a Shipboard Integrated Navigational System. *J. Mar. Sci. Eng.* **2019**, *7*, 364. [CrossRef]

14. Loukas, G.; Karapistoli, E.; Panaousis, E.; Sarigiannidis, P.; Bezemskij, A.; Vuong, T. A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles. *Ad Hoc Netw.* **2019**, *84*, 124–147. [CrossRef]
15. Krile, S.; Kezić, D.; Dimc, F. NMEA Communication Standard for Shipboard Data Architecture. *Int. J. Marit. Sci. Technol.* **2013**, *60*, 68–81.
16. De Sousa, J.P.C.; Gondim, J.J.C. Extraction and analysis of volatile memory in android systems: An approach focused on trajectory reconstruction based on nmea 0183 standard. In Proceedings of the 2016 11th International Conference on Availability, Reliability and Security (ARES), Salzburg, Austria, 31 August–2 September 2016; pp. 328–337.
17. Cantelli-Forti, A. Forensic Analysis of Industrial Critical Systems: The Costa Concordia’s Voyage Data Recorder Case. In Proceedings of the 2018 IEEE International Conference on Smart Computing (SMARTCOMP), Taormina, Italy, 8–20 June 2018; pp. 458–463.
18. Lee, D.K.; Miralles, D.; Akos, D.; Konovaltsev, A.; Kurz, L.; Lo, S.; Nedelkov, F. Detection of GNSS Spoofing using NMEA Messages. In Proceedings of the 2020 European Navigation Conference (ENC), Dresden, Germany, 23–24 November 2020; pp. 1–10.
19. Sivkov, Y. Transformation of NMEA ship network from sensor-based to information-based model. In Proceedings of the 2018 20th International Symposium on Electrical Apparatus and Technologies (SIELA), Bourgas, Bulgaria, 3–6 June 2018; pp. 1–4.
20. Fiorini, M. Maritime awareness through data sharing in VTS systems. In Proceedings of the 2012 12th International Conference on ITS Telecommunications, Taipei, Taiwan, 5–8 November 2012; pp. 402–407.
21. Seong, K.T.; Kim, G.H. Implementation of voyage data recording device using a digital forensics-based hash algorithm. *Int. J. Electr. Comput. Eng.* **2019**, *9*, 5412. [CrossRef]
22. Boudehenn, C.; Jacq, O.; Lannuzel, M.; Cexus, J.C.; Boudraa, A. Navigation anomaly detection: An added value for Maritime Cyber Situational Awareness. In Proceedings of the 2021 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), Dublin, Ireland, 4–18 June 2021; pp. 1–4. [CrossRef]
23. Furumoto, K.; Kolehmainen, A.; Silverajan, B.; Takahashi, T.; Inoue, D.; Nakao, K. Toward automated smart ships: Designing effective cyber risk management. In Proceedings of the 2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Rhodes, Greece, 2–6 November 2020; pp. 100–105.
24. Hemminghaus, C.; Bauer, J.; Padilla, E. BRAT: A BRidge Attack Tool for Cyber Security Assessments of Maritime Systems. *TransNav* **2021**, *15*, 35–44. [CrossRef]
25. IMO. *Resolution A.1106(29) Revised Guidelines for the Onboard Operational Use of Shipborne Automatic Identification Systems (AIS)*; IMO: London, UK, 2015.
26. IMO. *SOLAS Ch. V Safety of Navigation, Regulation 19 Carriage Requirements for Shipborne Navigational Systems and Equipment*; IMO: London, UK, 2013.
27. ITU. *Recommendation ITU-R M.1371-5 Technical Characteristics for an Automatic Identification System Using Time Division Multiple Access in the VHF Maritime Mobile Frequency Band*; ITU: Switzerland, Geneva, 2014.
28. Iphar, C.; Ray, C.; Napoli, A. Data integrity assessment for maritime anomaly detection. *Expert Syst. Appl.* **2020**, *147*, 113219. [CrossRef]
29. Blauwkamp, D.; Nguyen, T.D.; Xie, G.G. Toward a Deep Learning Approach to Behavior-based AIS Traffic Anomaly Detection. Dynamic and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS) Workshop, San Juan, PR. 2018. Available online: http://faculty.nps.edu/Xie/papers/ais_analysis_18.pdf (accessed on 18 February 2022).
30. Bosch, R. *CAN Specification Version 2.0*; Rober Bousch GmbH: Postfach, Germany, 1991; Volume 300240, p. 72.
31. Lokman, S.F.; Othman, A.T.; Abu-Bakar, M.H. Intrusion detection system for automotive Controller Area Network (CAN) bus system: A review. *EURASIP J. Wirel. Commun. Netw.* **2019**, *2019*, 1–17. [CrossRef]
32. Sabaliauskaite, G.; Adepu, S.; Mathur, A. A six-step model for safety and security analysis of cyber-physical systems. In *International Conference on Critical Information Infrastructures Security*; Springer: Cham, Switzerland, 2016; pp. 189–200.
33. Hareide, O.S.; Jøsok, Ø.; Lund, M.S.; Ostnes, R.; Helkala, K. Enhancing navigator competence by demonstrating maritime cyber security. *J. Navig.* **2018**, *71*, 1025–1039. [CrossRef]
34. 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. Available online: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf (accessed on 16 February 2022).
35. Loshin, D. *The Practitioner’s Guide to Data Quality Improvement*; Morgan Kaufmann Publishers Inc.: Burlington, NJ, USA, 2010.
36. IMO. *Resolution MSC.252(83) Adoption of the Revised Performance Standards for Integrated Navigation Systems (INS) Introduction Contents Module A-B*; IMO: London, UK, 2018.
37. Amro, A.; Gkioulos, V.; Katsikas, S. Communication architecture for autonomous passenger ship. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2021**. [CrossRef]
38. IMO. *Resolution MSC.252(83) Adoption of the Revised Performance Standards for Integrated Navigation Systems (INS)—Appendices*; IMO: London, UK, 2018.
39. Rødseth, Ø.J.; Kvamstad, B.; Porathe, T.; Burmeister, H.C. Communication architecture for an unmanned merchant ship. In Proceedings of the OCEANS-Bergen, 2013 MTS/IEEE, Bergen, Norway, 10–14 June 2013; pp. 1–9.

40. DNV GL. DNVGL-CG-0264: Autonomous and Remotely Operated Ships. 2018. Available online: <https://rules.dnv.com/docs/pdf/DNV/cg/2018-09/dnvgl-cg-0264.pdf> (accessed on 16 February 2022).
41. Amro, A.; Gkioulos, V.; Katsikas, S. Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework. Submitted for Review to ACM Transactions on Privacy and Security (TOPS). Available online: <https://www.researchgate.net/publication/355203975> (accessed on 16 February 2022).
42. Commission I.I.E. IEC 61162-1. 2010. Available online: <https://webstore.iec.ch/publication/25754> (accessed on 16 February 2022).
43. Manipulation of View—ATT&CK ICS. 2021. Available online: <https://cutt.ly/MoV> (accessed on 16 February 2022).
44. Denial of View—ATT&CK ICS. 2021. Available online: <https://cutt.ly/DoV> (accessed on 16 February 2022).
45. Kruegel, C.; Toth, T. Using decision trees to improve signature-based intrusion detection. In *International Workshop on Recent Advances in Intrusion Detection*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 173–191.
46. Garcia-Teodoro, P.; Diaz-Verdejo, J.; Maciá-Fernández, G.; Vázquez, E. Anomaly-based network intrusion detection: Techniques, systems and challenges. *Comput. Secur.* **2009**, *28*, 18–28. [CrossRef]
47. Tseng, C.Y.; Balasubramanyam, P.; Ko, C.; Limprasittiporn, R.; Rowe, J.; Levitt, K. A specification-based intrusion detection system for AODV. In Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks, Fairfax, VA, USA, 30 October 2003; pp. 125–134.
48. Amro, A.; Gkioulos, V. Communication and Cybersecurity Testbed for Autonomous Passenger Ship. In *Computer Security, ESORICS 2021 International Workshops*; ESORICS 2021; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; Volume 13106. [CrossRef]
49. Orebaugh, A.; Ramirez, G.; Beale, J. *Wireshark & Ethereal Network Protocol Analyzer Toolkit*; Elsevier: Amsterdam, The Netherlands, 2006.
50. Amro, A. Cyber-Physical Tracking of IoT Devices: A Maritime Use Case. In *Norsk IKT-Konferanse for Forskning og Utdanning*; Number 3; 2021; Available online: <https://ojs.bibsys.no/index.php/NIK/article/view/961> (accessed on 16 February 2022).
51. OneNet Standard for IP Networking of Marine Electronic Devices. Available online: <https://www.nmea.org/content/STANDARDS/OneNet> (accessed on 7 February 2022).
52. Jacq, O.; Brosset, D.; Kermarrec, Y.; Simonin, J. Cyber attacks real time detection: Towards a cyber situational awareness for naval systems. In Proceedings of the 2019 International Conference on Cyber Situational Awareness, Data Analytics Furthermore, Assessment (Cyber SA), Oxford, UK, 3–4 June 2019; pp. 1–2.

Paper VII

A. Amro and V. Gkioulos, 'From click to sink: Utilizing ais for command and control in maritime cyber attacks,' in *European Symposium on Research in Computer Security*, Springer, 2022, pp. 535–553

From Click To Sink: utilizing AIS for command and control in maritime cyber attacks

Ahmed Amro^[0000-0002-3390-0772] and Vasileios Gkioulos

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

Abstract. The maritime domain is among the critical sectors of our way of life. It is undergoing a major digital transformation introducing changes to its operations and technology. The International Maritime Organization urged the maritime community to introduce cyber risk management into their systems. This includes the continuous identification and analysis of the threat landscape. This paper investigates a novel threat against the maritime infrastructure that utilizes a prominent maritime system that is the Automatic Identification System (AIS) for establishing covert channels. We provide empirical evidence regarding its feasibility and applicability to existing and future maritime systems as well as discuss mitigation measures against it. Additionally, we demonstrate the utility of the covert channels by introducing two realistic cyber attacks against an Autonomous Passenger Ship (APS) emulated in a testing environment. Our findings confirm that AIS can be utilized for establishing covert channels for communicating Command & Control (C&C) messages and transferring small files for updating the cyber arsenal without internet access. Also, the establishment and utilization of the covert channels have been found to be possible using existing attack vectors and technologies related to a wide range of maritime systems. We hope that our findings further motivate the maritime community to increase their efforts for integrating cyber security practices into their systems.

Keywords: maritime · cybersecurity · Automatic Identification System (AIS) · cover channel · *ATT&CK*

1 Introduction

We live in a highly connected world that depends on various means of transportation for the delivery of goods, services, and the transportation of people all around the globe. Thus, the transportation sector is regarded internationally as a critical infrastructure. In the European Union, five modes of transport have been recognized: air, road, rail, maritime, and inland waterways [4]. Among these sectors, this paper targets the maritime domain. The maritime domain is linked to the well-being, prosperity, and security of the citizens of Europe [1]. It is also involved in 90% of the global trade of goods [3] making it a domain worthy of increased attention in the research community.

Maritime systems include a variety of cyber systems including Information Technology (IT) and Operational Technology (OT) which are distributed across port facilities, ships, and other components within the maritime infrastructure. These systems are applied in specific applications in navigation, propulsion and steering, cargo handling, and others. These applications rely on a group of maritime-specific systems such as the Automatic Identification System (AIS), and the Electronic Chart Display and Information System (ECDIS). Additionally, such systems rely on maritime-specific protocols and standards including among others, the National Marine Electronics Association (NMEA) standard, and the AIS protocol. NMEA standard is utilized in the communication between marine systems including the communication of sensor data through message-based protocol [49]. AIS is a special message-based protocol based on the NMEA standard which is utilized in many maritime services including; among others, traffic management, search and rescue, and collision avoidance [41].

Disruptive attacks against the maritime domain can have devastating effects as witnessed in the cyber attack against Mærsk shipping company, which lead to weeks of interrupted operations and losses beyond 300 million US dollars [36]. Also, insufficient security in the maritime systems and protocols has been demonstrated in the literature. To mention a few examples, Balduzzi et al [25] have demonstrated a wide range of attacks against AIS including spoofing, jamming, and other sorts of misuse while Tran et al [60] discussed the limited authentication, encryption, and validation in one of the NMEA protocols. Positively, there are demands for the consideration of cyber threats and cyber risk management in the current state of affairs in the maritime domain. The International Maritime Organization (IMO) has adopted Resolution MSC.428(98) [32] encouraging the maritime industry stakeholders to include cyber risk management into their safety management systems. The resolution provides guidelines and requirements for cyber risk management [31]. The guidelines suggest the continuous analysis and assessment of the threat landscape against the maritime infrastructure.

In this direction, this paper investigates attacks in the maritime industry in order to identify novel attacks that can surface into reality in the future. We have identified a limitation in the literature when discussing Command and Control (C&C) activities. Then, we investigate the utility of the Automatic Identification System (AIS) as a covert channel for conducting C&C activities during the development of cyber attacks against maritime infrastructure. In our investigation, we initially developed a threat model of the covert channel focusing on the threat requirements, scope, objectives, and techniques. Afterward, we developed and evaluated a proof of concept of the covert channel. Moreover, we demonstrated the utility and application of the covert channel in two realistic attack scenarios against a modern maritime use case which is an Autonomous Passenger Ship (APS). We aspire to motivate the maritime community to further adopt cybersecurity into their operations and system development practices.

2 Background and Related Work

2.1 Autonomous Passenger Ship

This paper is part of an ongoing research project titled "Autoferry" [50]. The project targets the development of an APS prototype which is named milliAmpere2; an autonomous ferry with the capacity to carry 12 passengers and their luggage across the Trondheim city canal as an alternative for a high-cost bridge [38]. MilliAmpere2 is designed to be fully autonomous with the ability to be supervised and controlled from a Remote Control Center (RCC). The ferry includes an Autonomous Navigation System (ANS) which utilizes data from various sensors for establishing situational awareness and safe navigation. The sensors include lidar, radar, Automatic Identification System (AIS), Global Positioning System (GPS), and others. The ANS forwards sensor data to a Remote Navigation System (RNS) at the RCC through a ship-shore communication link. More details can be found in our earlier article [22]. In this paper, we utilize this APS as a use case for demonstrating two cyber kill chains (i.e attack scenarios) to showcase the application and utility of the discussed covert channel.

2.2 ATT&CK Framework

Recently, wide adoption is observed for the Adversarial Tactics, Techniques, and Common Knowledge from MITRE, shortly known as the *ATT&CK* framework [57]. *ATT&CK* captures adversarial behavior in enterprise environments, industrial control systems, and other technology domains making it suitable for modeling cyber attacks in a wide range of use cases. The European Union Agency for Cybersecurity (ENISA) utilizes *ATT&CK* terminologies for mapping adversarial activities in their annual threat landscape report [11]. Also, Security Incidents and Event Management (SIEM) systems utilize *ATT&CK* terminologies for detecting adversarial activities [2, 10].

The recent adoption of *ATT&CK* as a threat model is observed for modeling threats against maritime systems. Kovanen et al [45] utilized *ATT&CK* for mapping threat actors' objectives to a remote pilotage system for improved risk assessment and design. Also, Jo et al [43] proposed a cyber attack analysis method based on *ATT&CK*. The authors described four documented cyber attacks in the maritime domain using *ATT&CK* tactics and techniques. Moreover, in our earlier work [23] we utilized *ATT&CK* as a threat model for describing attacks against navigational functions. In this paper, we will also utilize *ATT&CK* for modeling cyber attacks and provide a proof of concept of some of the *ATT&CK* techniques in common maritime systems.

The *ATT&CK* threat model provides useful terminologies for describing the different elements of threats. In this paper, we rely heavily on both, namely tactics and techniques. Tactics describe the adversarial objectives also referred to as stages of cyber attacks. Techniques on the other hand describe the adversarial method for realizing an objective [57].

2.3 Maritime Kill Chains, Threats and Attacks

In this paper we investigate and aim to answer the following question; what are the adversarial tactics (i.e. objectives) and techniques that are discussed in the literature in the maritime domain and do they cover the current threat landscape. In our research, we rely on the *ATT&CK* framework due to its comprehensive threat model and increased adoption as a new standard for adversarial tactics, techniques, and procedures. We have conducted a comprehensive literature review to identify relevant works that have discussed adversarial techniques across the different stages of cyber attacks (i.e. tactics). This allows for a clearer understanding of the current threat landscape in the maritime domain.

Starting with the reconnaissance stage, Enoch et al [33] briefly discussed the utility of OpenVAS and NMAP for conducting reconnaissance-related activities in a vessel system. Also, Standard et al [54] discussed the teaching of network reconnaissance for naval officers during a cybersecurity course for capacity development. Additionally, Lund et al [47] mentioned that activities at the reconnaissance stage were conducted through physical access to the vessel and access to the network, and ECDIS software. Moreover, Amro [20] has demonstrated the utility of AIS and NMEA communicated messages for gaining both cyber and physical attributes of possible maritime targets.

For gaining access to maritime components and networks; also known as attack delivery, Lund et al [47] discussed the utilization of a USB flash drive to deliver a malicious payload into the ECDIS machine and execute it. Also, Papastergiou et al [51] referred to the possibility of gaining access to maritime infrastructure through compromising the supply chain. Additionally, Pavur et al [52] demonstrated the feasibility of VSAT TCP session hijacking for reaching and controlling maritime VSAT communication. Moreover, Tam and Jones[58] argued that users can be tricked into downloading and executing malicious software or guided into malicious websites.

After gaining access to systems and networks attackers aim to achieve a group of objectives including discovery, credential access, and collection. Hemminghaus et al [39] target the network for discovery through sniffing and collection of network traffic including navigation data. Jo et al [43] categorized vulnerability scanning of ship systems, eavesdropping on Voice over Internet Protocol (VoIP), and Wi-Fi communication in the discovery stage of cyber attacks. Pavur et al [52] demonstrated the ability to collect credit card information, visa, passport, ship manifest, and non-encrypted REST API credentials communicated through eavesdropping on VSAT connections.

In certain cases, attackers desire to perform privilege escalation to execute commands and programs with higher privilege. Lund et al [47] mentioned that the operator station utilized as the pivot point of their attack demonstration was running already within administrator privilege and therefore doesn't require escalation. However, they referred to hijacking execution flow through a malicious Windows socket dynamic-link library (Winsoc DLL), this is among the techniques utilized to achieve privilege escalation, persistence in the target system, and evade defensive measures [13].

Many works have discussed attacks that aim to impact maritime operations. Lund et al [47] and Hemminghaus et al [39] discussed the manipulation of sensor messages for impacting the operation of navigation systems. Amro et al [23] formalized manipulation and denial of view based on navigational data as attacks that can impact navigational functions. Moreover, Hemminghaus et al [39] referred to alarm suppression for inhibiting response functions as well as spoof reporting messages to impair process control.

Many stages of cyber attacks in the maritime domain are demonstrated and discussed in the literature in sufficient detail. Still, a limited discussion is observed regarding Command and Control (C&C) activities. Hooper [40] has investigated the potential of covert communications in pulsed or continuous-wave radar and discussed the cyber implications of that in the maritime domain. The authors argued that communication links utilizing spectrum-sharing may pave the way for unintended channels (i.e covert channels); an inclination which we agree with. Hareide et al. [37] bypassed the need for the C&C channel by implementing a specific condition for an attack to be launched when arriving at a certain position. Jo et al [43] described three maritime cyber incidents including C&C stages with a limited description of the implementation. Enoch et al [33] have briefly mentioned C&C in the attack model but without details of the implementation. Leite et al [46] proposed a triggering mechanism for cyber attacks based on radar and AIS messages. The authors proposed and demonstrated a pattern matching technique that can identify false plots depicted on the ECDIS which can be used for triggering cyber attacks. Other than that, to the best of our knowledge, no other work has discussed C&C in the maritime domain in more detail. Therefore, a contribution of this paper is an investigation of the utilization of AIS as a covert channel for C&C attack techniques using real maritime systems. This is intended to raise awareness of yet another possible attack utilizing the AIS protocol and hopefully drive the maritime community to consider cybersecurity more seriously and deeply within their systems.

The concept of a kill chain; a multi-staged cyber attack scenario, is observed in the maritime domain. Hareide et al.[37] have discussed a maritime kill chain for demonstrating the feasibility of cyber attacks in order to increase awareness. The authors relied on a previously developed attack by Lund et al [47] which also discusses the development of the attack through a kill chain. Also, Jo et al [43] utilized consequent tactics from *ATT&CK* for describing cyber attacks against maritime systems. In this paper, we will also utilize the concept of kill chains for discussing complete scenarios for cyber attacks that implement our novel Command and Control (C&C) covert channel.

3 AIS as a Covert Channel

In this section, we discuss our analysis of the utility of the AIS as a covert channel supporting adversarial activities throughout different phases of cyber attacks. The analysis considers both the AIS protocol itself as well as AIS devices. This section also describes the threat model with details from different

viewpoints. Context (i.e. physical and cyber architecture), Objectives (i.e. tactics), and techniques. Additionally, a proof of concept of the attack is developed and demonstrated in this section in addition to a discussion of relevant counter-measures.

3.1 Context view

Following a top-down approach, the context of utilizing AIS as a covert channel is discussed in this section. A physical view of the context is demonstrated in Figure 1a. A threat actor needs to be located in physical proximity to the victim ships either at land or sea. The range is limited by the VHF range of the attacker station and the placement of the antennas on both sides; the range can reach up to 60 nautical miles [19]. The VHF radio frequencies for AIS belong to the licensed portion of the radio spectrum and require a proper license to operate in most countries. Therefore, an attacker without a proper license can be detected and addressed. However, an attacker with a proper license such as an industrial competitor or a maritime entity belonging to a nation-state might operate undetected at this level.

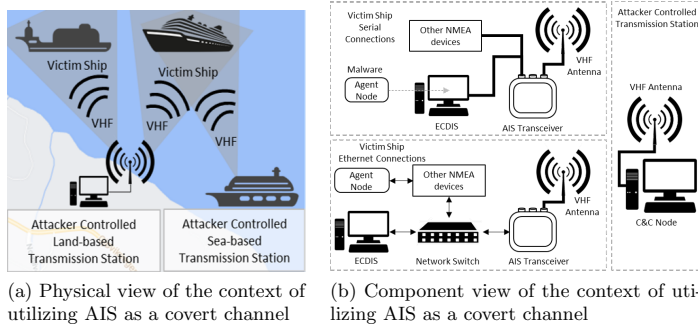


Fig. 1. Physical and component view for utilizing AIS as a covert channel

A component view of the context is depicted in Figure 1b. The attacker station consists of a Command and Control (C&C) node that is able to transmit AIS traffic over VHF. On the other hand, the victim ships network might have either serial [29, 55, 56] or Ethernet connections [30] from the AIS device to internal components. An internal agent node to be controlled by the attacker is needed to receive and execute the (C&C) commands. The agent node is assumed to either be a machine infected with an attacker’s controllable malware or a standalone malicious machine. In a ship network consisting of serial connections, malware is expected to infect an existing machine. On the other hand, in an

Ethernet network, a standalone machine is a possible alternative. Different attack techniques are needed to establish a covert channel in each network (More details in Section 4).

3.2 Tactics and Techniques

The threat model is developed considering variant attacker capabilities and communicated as tactics and techniques using the *ATT&CK* terminology. The objectives (i.e tactics) of the attackers are assumed to be the following:

- Command and Control: send unidirectional C&C messages from an attacker to victims (1 to many). The messages can carry either simple commands or files (e.g. malware). This is assumed to be achievable through properly encoding commands and files into AIS messages. More advanced threat actors are expected to pursue secure C&C communication. They might aim to secure the communication from being revealed, or tampered with. Even if their activities are detected, the executed commands or transferred files are aimed to be kept a secret. This is assumed to be established through hiding command messages into AIS messages with additional obfuscation, steganography, or cryptography. A bi-directional channel is expected to require additional components, tactics, and techniques which are items for future work.
- Defense Evasion: this includes avoiding raising the operators' attention or other detection measures. This means that limited impact on legitimate operations is pursued. This is assumed to be achievable through careful selection of AIS message types and fields.

To achieve the C&C objective the attacker can establish the covert channel using a combination of Alternate Network Medium (i.e. VHF) [5] and Protocol Tunneling [17] command and control attack techniques. This combination entails the utilizing of VHF radio communication as a medium for the C&C communication which is tunneled through the AIS protocol. Based on the attacker capabilities to secure it, attackers can apply Data Encoding [7], Data Obfuscation [8] or Encrypted Channel [9]. According to ATT&CK, data encoding can be achieved using standard or non-standard encoding (e.g. Base32), Data obfuscation can be achieved using stenography, protocol impersonation, or junk data, and Encrypted channel can use asymmetric or symmetric cryptography [15]. On the other hand, to avoid detection, the different types of AIS messages and fields are considered to best serve the objectives. The criteria for choosing the most suitable message type and field is that they should provide the largest capacity of transfer and limited impact on operations. The rationale for choosing the largest capacity is to reduce the amount of AIS messages needed to encode C&C messages.

We have considered all possible 27 AIS message types using the description provided By Rayomon[53]. As shown in Table 1, messages 8 and 14 were found to provide the largest capacity while at the same time having a common appearance, unlike message type 26. Moreover, the messages types; if carefully

configured, do not provide navigational data that will influence the navigational functions and therefore are expected to have no impact on operations. Message 14 can be utilized in managing distress signals and might invoke a response from a nearby rescue unit [48]. Therefore, we will restrict our discussion in this paper in the utility of message type 8 for C&C. Furthermore, the structure of message 8 content itself is controlled. We analyzed the different content categories to identify the category that allows for the largest capacity and flexible field format. We relied on IMO circulation SN.1/Circ.289 [28] in our analysis. We have identified that a text description message is the best candidate as it includes a text string field with a maximum limit of 161 ASCII characters. Although there is a standard format for this field, it is only recommended and not mandatory to follow.

Table 1. The top 5 AIS message types with the largest fields

Message Type	Field	Max Size (bits)	Rational
Type 26: Multiple Slot Binary Message	Data	1004	Extremely rare
Type 14: Safety-Related Broadcast Message	Text	968	Suitable.
Type 8: Binary Broadcast Message	Data	966	Suitable.
Type 12: Addressed Safety-Related Message	Text	936	Addressed to a specific target. Reduced C&C channels
Type 6: Binary Addressed Message	Data	920	Addressed to a specific target. Reduced C&C channels

3.3 Proof of Concept

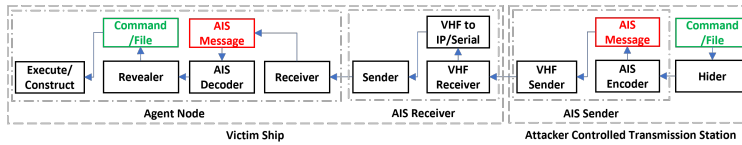


Fig. 2. A logical view of the components of the AIS covert Channel

In this section, we present the development of the proof of concept for utilizing AIS as a covert channel. Figure 2 depicts the required logical components to achieve the attackers' objectives. First, the C&C message or file is input into a hider function to evade detection and the output is then encoded into an AIS message. Then, the message is transmitted over VHF using an AIS transmitter. Should it be received and accepted at the AIS on the victim ship, protocol conversion is expected to forward the AIS messages through a serial link or IP protocol to the ship network; this is traditionally performed by AIS receivers. The agent

node then eavesdrops on the AIS message stream, decodes the messages to identify C&C messages (e.g. based on the MMSI or other signal) reveals the hidden message, and executes it, or reconstructs it if its part of a file. Through this channel, attackers gain the capabilities to remotely and covertly update their cyber attack arsenal and techniques.

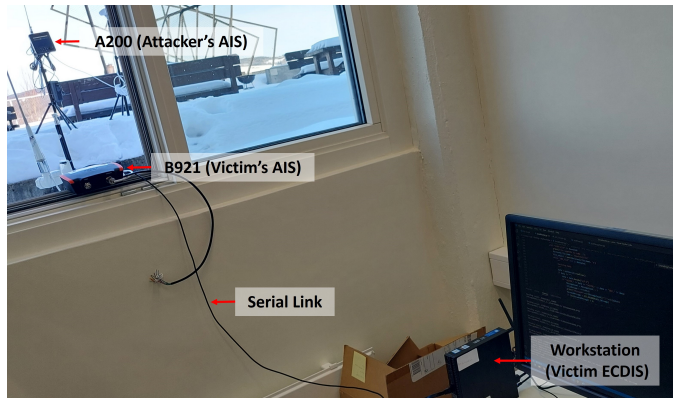


Fig. 3. Setup for the proof of concept of AIS as a covert channel

Figure 3 depicts the setup for the proof of concept. It is implemented using two AIS transceivers, namely, em-trak A200 and em-trak B921. A200 is used as the attacker-controlled transmission station. B921 is used as the AIS receiver and is connected through a serial link to a workstation simulating the victim ECDIS. The workstation is equipped with a script that simulates the agent node or malware that is monitoring the AIS messages over the serial link. The script decodes AIS messages and when a C&C message is identified it executes the encoded command or reconstructs the transmitted file.

We conducted several experiments to test if the implementation works. We attempted to send and execute commands as well as construct files at the victim ECDIS. Due to space restrictions, we will present one of these experiments. First, the ciphertext which includes the hidden C&C message is prepared using a python script. In this example, the attacker will send a directory listing command, the plaintext of the hidden message "CM:dir" is encrypted using Advanced Encryption Standard (AES), the ciphertext is "9C6ED8600E1F" and then encoded into an AIS message "!AIVDM,1,1,,B,83o0F400@00;@uQA0ed;1LAP,0*39". The "CM:" string is used to identify a command execution message at the agent node while the "dir" string is the directory listing command in Windows.

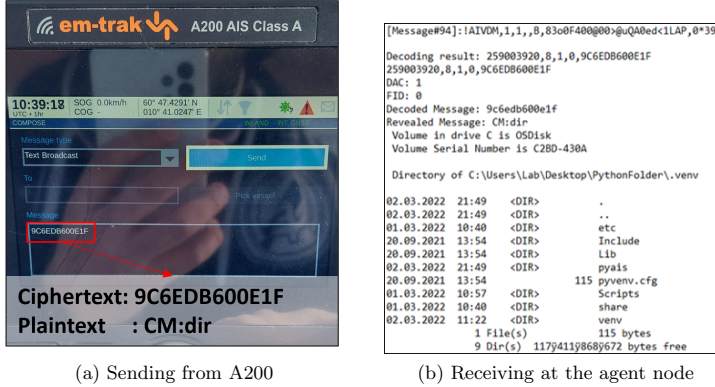


Fig. 4. Demonstration for sending and receiving covert C&C message over AIS

Figure 4(a) depicts a photo of the message composer at the A200 AIS transceiver with the ciphertext as the content of the message. After the message was sent, Figure 4(b) depicts a screenshot of the agent node receiving and executing the command.

3.4 Evaluation of the Covert Channel

In this section, we will evaluate the utility of the covert channel to attackers to better analyze the associated risk. The evaluation is discussed based on their type, throughput, and robustness to detection and countermeasures. Then, suggesting suitable improvement for the detection and prevention is provided.

Our analysis considers two hider functions and two settings for the covert channel. The hider functions are Base32 encoding and AES-CFB encryption; with a 16-byte key and a 16 bytes Initialization Vector (IV). The settings are either based on the protocol specifications or the em-track A200 commercial AIS device. The type of the channel is a unidirectional covert channel. The C&C node can transmit messages that the agent node can receive, however, the agent node; on its own, cannot establish an outbound channel through the AIS device. This limits the attackers' capabilities in managing the agent node in the targeted environments. Regarding the throughput, the maximum capacity for the text string field is 966 or 480 bits in the protocol specifications or the A200, respectively. The implementation of encoding or encryption further restricts the capacity. Table 2 depicts the maximum size of the field that can hold the clear segment of a command or a file as well as the corresponding throughput considering the two hider functions, two settings, and two transmission rates (TR).

From the attacker's perspective, using AES as a hider function is a reasonable option since it provides secure communication with only a relatively less

Table 2. Covert Channel Throughput Evaluation

Hider Function	Based on	Max Field Capacity (bit)	Throughput (bit/sec)	
			2 sec TR	10 min TR
Base32	Protocol specs	600	300	1
	AIS200 em-trak	276	138	0,46
AES	Protocol specs	480	240	0,8
	AIS200 em-trak	240	120	0,4

TR: Transmission Rate

throughput than the Base32. Still, secure key establishment and handling is an additional burden the attacker needs to consider. While the Base32 encoding is simpler to implement and provides slightly better throughput, it doesn't provide secure communication and can expose the content of the covert channel. We have also evaluated the utility of this channel for delivering malware to the victim ship and allowing threat actors to update their adversarial cyber arsenal at sea. With such a transmission rate, transporting a 338 Kb malware; the average malware size in 2010 [14] at a 2-sec transmission rate would take 3 hours considering the protocol specifications. However, transporting the NotPetya malware which is 1,5 Gb [6] would take 29826 hours at the same transmission rate. Therefore, the utility of this covert channel is limited to commands and small malware. Regarding robustness to detection and countermeasures, several works have discussed countermeasures for securing AIS communication using encryption for authentication and integrity [44, 35, 24, 25]. Although a wide adoption of such countermeasures is not observed we argue that encryption doesn't eliminate the threat of covert channels against AIS. In the case of utilizing a public key infrastructure (PKI) for authenticating the different entities participating in the AIS communication, threat actors with legitimate credentials such as boat and ship owners, competitors, and nation-states would still be able to utilize the channel. Moreover, there is a discussion regarding anomaly detection algorithms for AIS such as the work of Iphar et. al [42], Blauwkamp et. al [27] and Balduzzi et. al [25]. However, there is no discussion regarding anomalies associated with AIS message type 8. Additionally, if the attacker maintained a reduced transmission rate, the likelihood of detecting anomalies is expected to be reduced. Real maritime infrastructure is required for formal evaluation of the robustness of this covert channel against detection. Therefore, we argue that such channels constitute a threat to the maritime infrastructure that is utilizing AIS communication and countermeasures should be tuned to detect them. Future efforts are advised for investigating the utility of anomaly detection in detecting the covert channel.

4 Adversary Emulation against an Auto-remote Vessel

To demonstrate the utility of the proposed covert channel for attackers, and its technical application in realistic attack scenarios, we will apply an adversary emulation process; a security assessment process applying realistic attack scenarios which emulate the capabilities of real threat actors [57]. This enables the elicitation and evaluation of relevant security control.

In this section, we present two cyber kill chains emulating two attack scenarios against an Autonomous Passenger Ship (APS) use case which is discussed in Section 2.1. The kill chains are constructed based on the observed adversarial techniques in the maritime industry across the different kill chain phases which are discussed in Section 2.3. Additionally, we improve the kill chains by utilizing the proposed C&C channel discussed in Section 3 to demonstrate its application. We argue that the kill chains are also relevant for other maritime use cases encompassing similar technologies.

We utilize our previously proposed maritime-themed testbed [21] for the development of the adversarial techniques. The utilization of the testbed with regards to this paper is system replication and system analysis. During system replication, we developed a replica of the target system using real and simulated components, and then target the developed replica with a group of attack techniques emulating an adversarial behavior.

4.1 Target Environment

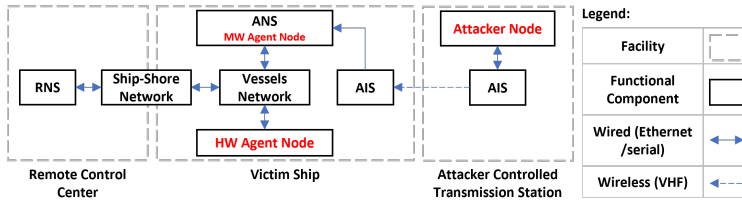


Fig. 5. A model of the target environment for the development of the kill chains

A model of the target environment is depicted in Figure 5. It emulates three facilities, namely, an attacker-controlled transmission station, a victim ship, and a remote control center. The attacker station consists of capabilities to create and transmit command and control traffic encapsulated within AIS messages over VHF. The A200 AIS is utilized at this station. The victim ship consists of an AIS transceiver; in this setup, the B921 is utilized. The receiver AIS receives AIS messages and forwards them over a serial link to the Autonomous Navigation System (ANS) which in turn forwards it to the Remote Navigation System (RNS) over a ship-shore network. The ANS and RNS are emulated using virtual machines while the vessel and ship-shore networks are emulated using virtual networking using Virtualbox. Due to the lack of available ANS and RNS software, both components are simulated as chart plotters using the OpenCPN software. The difference between them is that the ANS is not intended to be monitored by a human operator while the RNS is. The autonomous and remote

navigation functions are simulated only through rendering the AIS and companion NMEA messages in the chart plotter. No control functions are simulated in this environment. Additionally, another virtual machine with Kali Linux is added to simulate a hardware agent node. This environment will be utilized in the demonstration of the later kill chains and is added as part of our testbed for further research.

4.2 Cyber Kill Chains

In this section, we present and discuss two attack scenarios. We will utilize the *ATT&CK* terminologies to facilitate the communication of a threat. In this paper, we utilized the abstract concept of the tactics and techniques and positioned them in a maritime context. We utilized attack trees for the description of the kill chain as it has been observed to be a common approach in the literature [33, 34]. These kill chains can later be used as adversary emulation exercises for the evaluation of cybersecurity controls in maritime systems with technologies similar to the ones in the testing environment.

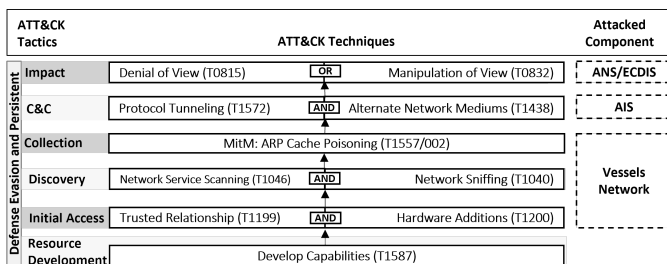


Fig. 6. Remotely and covertly controlling a malicious hardware agent node

Kill Chain 1: Impact through malicious hardware agent node The first kill chain depicted in Figure 6 describes the following scenario. A motivated threat actor invests in the development of attacking capabilities into the attacker agent node to be boarded on the vessel and remotely controlled from a place within range by utilizing the covert channel described in Section 3. The capabilities include a hardware component with Ethernet and software to receive and execute commands from the C&C node. In our environment, this is achieved through the Kali Linux virtual machine which can later be shipped into a Raspberry Pi or small hardware. The node is also equipped with scripts that are needed to conduct the later attack techniques. First, the developed capability needs to be connected to the ship network. Considering the lack of crew on the

autonomous vessels, an attacker may attempt to access the vessel and locate the network and insert the agent node (Hardware Additions [12] or Transient Cyber Asset [18]). The success of this depends on the imposed physical security controls. In the case that physical controls exist, threat actors could exploit trusted relations and gain access to the network for several reasons (e.g. maintenance) and insert the node. This is a communicated concern in the maritime community. BIMCO; a global organization for shipowners, charterers, shipbrokers, and agents, discussed the issue of the lack of control of the onboard systems during ship visits in their latest guidelines [26]. They argued that knowing whether malicious software has been left in the systems onboard vessels is difficult. After the insertion of the node, assuming it received valid network configurations (e.g. through DHCP), the node is developed to conduct network service scanning using a scanning tool (e.g. NMAP) and sniffing using a network sniffer (e.g. tshark) to identify other components in the network. Later, target components with specific criteria are identified; certain operating system versions, or certain network services. The chosen targets are then targeted by a MitM attack in the form of ARP spoofing using a MitM tool (e.g. Ettercap). If that is successful, the node should be capable of eavesdropping on network traffic passing to and from the attacked components in the vessel network including AIS messages. When reaching this vantage point, the node stays dormant and only monitors the AIS messages to identify commands from the C&C node. On the other side, the threat group utilizes an alternate network medium that is the VHF radio used in the AIS to send C&C messages. The attacker node can send either command to be executed by the agent node upon reception or send files including malware. This capability allows attackers to bypass traditional network defenses if the AIS link is not monitored. In traditional vessels, the ECDIS which is usually connected to the AIS is considered air-gapped and not connected to the internet [37]. However, this attack would remove the gap and provide attackers with an offensive capability not possible before. At this stage, the threat group has a tactical advantage of observing the physical operational environment and launching an attack under certain conditions (e.g. difficult weather conditions in which visibility is limited). Their next step is targeting the NEMA messages in a combination of denial of view and manipulation of view attacks. The options for the attackers are a lot, only limited by the number of NMEA messages utilized in the vessels and their criticality to the navigation functions. In our earlier work, we formalized and demonstrated a group of such attacks [23]. One instance could be that the attackers choose to drop radar messages (TTM messages) going to the ANS denying it from establishing accurate rendering of the vessels in the physical environment. Also, attackers can manipulate the actual Speed Over Ground (SoG) estimated from the GPS to impact the speed of the vessel. According to a previously conducted Preliminary Hazard Analysis for an autonomous ferry use case, manipulation of sensor data could lead to collisions or ship sinking [59]. This concludes the first kill chain which can; in the lack of proper defenses, cause few clicks to sink a vessel.

Throughout the kill chain, several evasion and persistence techniques can be employed to challenge the detection and countermeasures and maintain a foothold in the network. This can include the utilization of the hider functions in the covert channel (Section 3), applying slight modification to the sensor data, and others.

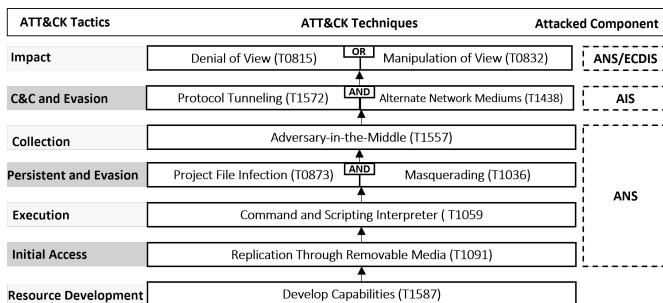


Fig. 7. Remotely and covertly controlling a malware agent node

Kill Chain 2: Impact through malware The second kill chain depicted in Figure 7 describes the following scenario. A motivated threat actor targets the APS through the maintenance personnel boarding the APS. It is assumed that the malware is loaded into the ANS through a USB stick. The malware relies on commands and scripting for executing its tasks. Upon execution, the malware aims to eavesdrop on the AIS messages communicated over the serial link at the ANS. However, serial interfaces allow only a single listener. In principle, there are several options to bypass this constraint. One option is discussed by Lund et al [47] through malicious Winsock DLL (Section 2.3). This direction, however, would require escalating privilege. Another option, which is explored in this paper, is to modify the configuration file of the ANS regarding the sources of AIS messages. A similar technique suggested in *ATT&CK* is called Project File Infection [16]. This option, in principle, doesn't require escalated privileges under the assumption that the permissions to modify the configuration files are granted to normal users. This is the case for the OpenCPN software. Therefore, the malware is programmed to first close the OpenCPN software to release the serial interface and update the data source configuration to receive AIS and NMEA messages from the malware over UDP and then reopen the software quickly. In this manner, the malware masquerades as a legitimate data source. However, during testing, it was observed that this activity can be detected by the local firewall. A message is shown on the monitor requesting acceptance for the creation of a new connection. Assuming that a local firewall is activated at

the ECDIS, the attacker needs to implement techniques to bypass it. Now the malware is actually in the middle between the AIS and the OpenCPN software. It has access to the serial link, can collect the messages, and forwards them to the OpenCPN software to avoid disrupting the operations. At this vantage point, the malware keeps monitoring the messages waiting for a C&C message. When one arrives the malware can distinguish if it's a command to be executed or file segments to be reconstructed. From this point forward, similar to the previous kill chain, the range of possible activities the malware can perform is wide open and relies on the C&C messages sent from the attacker-controlled transmission station. Among the options are also manipulating or denying the view and possibly causing a collision and sink. The malware is developed using python and is compiled as an executable for windows.

This scenario relies on a group of assumptions regarding the knowledge needed by the threat group while developing the malware. First, the name and path of the ANS or ECDIS executable, as well as the name, path, and structure of the configuration file, are all assumed to be known. This is likely possible for commonly deployed software such as OpenCPN. Also, altering the configuration without causing operation disruption is not trivial if there are multiple AIS data sources and destinations. In our proof of concept, the modification is done using a simple rule which is to remove a serial data source and replace it with a UDP data source. These kill chain conditions render it a targeted attack that requires a sufficient level of the domain and system knowledge in addition to a moderate level of complexity.

5 Conclusion

Recent efforts are undergoing to introduce cyber risk management into the maritime community. This includes the continuous identification and analysis of the threat landscape. In this direction, this paper presents an overview of the maritime cyber threat landscape and presents the results of an investigation of a novel cyber attack against maritime systems. The attack is in the form of a covert channel utilizing the prominent Automatic Identification System (AIS) for sending Command & Control messages and delivering malware. We have investigated the feasibility of this attack by developing a threat model utilizing the *ATT&CK* framework, developing a proof of concept of the attack, as well as presenting two cyber attack scenarios (i.e. kill chains) that can utilize this attack. The feasibility of the attack has been demonstrated using existing technology that is relevant to a wide range of traditional and future maritime systems including autonomous vessels. The findings are hoped to urge the maritime community to increase their integration of cybersecurity practices. Future work can be dedicated to the investigation and development of mitigation solutions against the proposed covert channel. Additionally, the proposed kill chains can be utilized as adversary emulation plans for the evaluation of cybersecurity of maritime systems.

References

1. European defence agency, maritime domain. <https://eda.europa.eu/docs/default-source/eda-factsheets/2017-09-27-factsheet-maritime> (2017)
2. How mitre attck alignment supercharges your siem. <https://www.securonix.com/how-mitre-attack-alignment-supercharges-your-siem/> (2019)
3. Ocean shipping and shipbuilding. <https://www.oecd.org/ocean/topics/ocean-shipping/> (2019)
4. Transport modes. https://ec.europa.eu/transport/modes_en (Jan 2019)
5. Alternate network mediums. <https://attack.mitre.org/techniques/T1438/> (2021), accessed on 30.01.2022
6. Backdoor built in to widely used tax app seeded last week's notpetya outbreak. <https://arstechnica.com/information-technology/2017/07/heavily-armed-police-raid-company-that-seeded-last-weeks-notpetya-outbreak/> (2021), accessed on 20.12.2021
7. Data encoding. <https://attack.mitre.org/techniques/T1132/> (2021), accessed on 30.01.2022
8. Data obfuscation. <https://attack.mitre.org/techniques/T1001/> (2021), accessed on 30.01.2022
9. Encrypted channel. <https://attack.mitre.org/techniques/T1573/> (2021), accessed on 30.01.2022
10. Enhancing with mitre. <https://documentation.wazuh.com/current/user-manual/ruleset/mitre.html> (2021)
11. Enisa threat landscape 2021. <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021> (2021)
12. Hardware additions. <https://attack.mitre.org/techniques/T1200/> (2021)
13. Hijack execution flow: Dll search order hijacking. <https://attack.mitre.org/techniques/T1574/001/> (2021), accessed on 14.03.2022
14. How large is a piece of malware? <https://nakedsecurity.sophos.com/2010/07/27/large-piece-malware/> (2021), accessed on 20.12.2021
15. Mitre attck. <https://attack.mitre.org/> (2021), accessed on 14.12.2021
16. Project file infection. <https://collaborate.mitre.org/attackics/index.php/Technique/T0873> (2021)
17. Protocol tunneling. <https://attack.mitre.org/techniques/T1572/> (2021), accessed on 30.01.2022
18. Transient cyber asset. <https://collaborate.mitre.org/attackics/index.php/Technique/T0864> (2021)
19. Two-way radio range, the facts about distance. <https://quality2wayradios.com/store/radio-range-distance> (2021), accessed on 14.12.2021
20. Amro, A.: Cyber-physical tracking of iot devices: A maritime use case. In: Norsk IKT-konferanse for forskning og utdanning. No. 3 (2021)
21. Amro, A., Gkioulos, V.: Communication and cybersecurity testbed for autonomous passenger ship. In: European Symposium on Research in Computer Security. pp. 5–22. Springer (2021)
22. 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)

23. Amro, A., Oruc, A., Gkioulos, V., Katsikas, S.: Navigation data anomaly analysis and detection. *Information* **13**(3) (2022). <https://doi.org/10.3390/info13030104>, <https://www.mdpi.com/2078-2489/13/3/104>
24. Aziz, A., Tedeschi, P., Sciancalepore, S., Di Pietro, R.: Secureais-securing pairwise vessels communications. In: 2020 IEEE Conference on Communications and Network Security (CNS). pp. 1–9. IEEE (2020)
25. Balduzzi, M., Pasta, A., Wilhoit, K.: A security evaluation of ais automated identification system. In: Proceedings of the 30th annual computer security applications conference. pp. 436–445 (2014)
26. BIMCO: The Guidelines on Cyber Security Onboard Ships. BIMCO (2016)
27. Blauwkamp, D., Nguyen, T.D., Xie, G.G.: Toward a deep learning approach to behavior-based ais traffic anomaly detection. In: Dynamic and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS) Workshop, San Juan, PR. Retrieved from http://faculty.nps.edu/Xie/papers/ais_analysis_18.pdf (2018)
28. Circular, I.D.S.: Guidance on the use of ais application-specific messages|. IMO NAV55/21/Add 1
29. Commission, I.I.E., et al.: Iec 61162-1 (2010)
30. Commission, I.I.E., et al.: Iec 61162-450 (2016)
31. Committee, T.M.S.: Interim guidelines on maritime cyber risk management (msc-fal.1/circ.3/rev.1). <https://cutt.ly/6R8wqjN>
32. Committee, T.M.S.: International maritime organization (imo) (2017) guidelines on maritime cyber risk management. <https://www.imo.org/en/OurWork/Security/Pages/Cyber-security.aspx>
33. Enoch, S.Y., Lee, J.S., Kim, D.S.: Novel security models, metrics and security assessment for maritime vessel networks. *Computer Networks* **189**, 107934 (2021)
34. Glomsrud, J., Xie, J.: A structured stpa safety and security co-analysis framework for autonomous ships. In: European Safety and Reliability conference, Germany, Hannover (2019)
35. Goudosis, A., Katsikas, S.: Secure ais with identity-based authentication and encryption. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* **14**(2) (2020)
36. Greenberg, A.: The untold story of notpetya, the most devastating cyberattack in history, <https://bit.ly/MaerskAttack>
37. Hareide, O.S., Jøsok, Ø., Lund, M.S., Ostnes, R., Helkala, K.: Enhancing navigator competence by demonstrating maritime cyber security. *The Journal of Navigation* **71**(5), 1025–1039 (2018)
38. Havdal, G., Heggelund, C.T., Larssen, C.H.: Design of a Small Autonomous Passenger Ferry. Master’s thesis, NTNU (2017)
39. Hemminghaus, C., Bauer, J., Padilla, E.: Brat: A bridge attack tool for cyber security assessments of maritime systems (2021)
40. Hooper, J.L.: Considerations for operationalizing capabilities for embedded communications signals in maritime radar. Tech. rep., NAVAL POSTGRADUATE SCHOOL MONTEREY CA (2018)
41. IMO: Resolution a.1106(29) revised guidelines for the onboard operational use of shipborne automatic identification systems (ais) (2015)
42. Iphar, C., Ray, C., Napoli, A.: Data integrity assessment for maritime anomaly detection. *Expert Systems with Applications* **147**, 113219 (2020)
43. Jo, Y., Choi, O., You, J., Cha, Y., Lee, D.H.: Cyberattack models for ship equipment based on the mitre att&ck framework. *Sensors* **22**(5), 1860 (2022)

44. Kessler, G.: Protected ais: a demonstration of capability scheme to provide authentication and message integrity. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* **14**(2) (2020)
45. Kovanen, T., Pöyhönen, J., Lehto, M.: epilotage system of systems' cyber threat impact evaluation. In: *ICCWS 2021 16th International Conference on Cyber Warfare and Security*. p. 144. Academic Conferences Limited (2021)
46. Leite Junior, W.C., de Moraes, C.C., de Albuquerque, C.E., Machado, R.C.S., de Sá, A.O.: A triggering mechanism for cyber-attacks in naval sensors and systems. *Sensors* **21**(9), 3195 (2021)
47. Lund, M.S., Hareide, O.S., Jøsok, Ø.: An attack on an integrated navigation system (2018)
48. Maritime, N.R.F.N. 46 ais safety-related messaging. <https://puc.overheid.nl/insi/doc/PUC.2045.14/1/>
49. NMEA: National marine electronics association - nmea0183 standard (2002)
50. NTNU Autoferry: Autoferry - Autonomous all-electric passenger ferries for urban water transport. <https://www.ntnu.edu/autoferry> (2018)
51. Papastergiou, S., Kalogeraki, E.M., Polemi, N., Douligeris, C.: Challenges and issues in risk assessment in modern maritime systems. In: *Advances in Core Computer Science-Based Technologies*, pp. 129–156. Springer (2021)
52. Pavur, J., Moser, D., Strohmeier, M., Lenders, V., Martinovic, I.: A tale of sea and sky on the security of maritime vsat communications. In: *2020 IEEE Symposium on Security and Privacy (SP)*. pp. 1384–1400. IEEE (2020)
53. Raymond, E.S.: Aivdm/aivdo protocol decoding, <https://gpsd.gitlab.io/gpsd/AIVDM.html>
54. Standard, S., Greenlaw, R., Phillips, A., Stahl, D., Schultz, J.: Network reconnaissance, attack, and defense laboratories for an introductory cyber-security course. *ACM Inroads* **4**(3), 52–64 (2013)
55. Std, I.: 61162-2. Maritime Navigation and radiocommunication equipment and systems–Digital interfaces–Part2: Single talker and multiple listeners, high-speed transmission (1998)
56. Std, I.: 61162-3. Maritime Navigation and radiocommunication equipment and systems–Digital interfaces–Part3: Serial data instrument network (2008)
57. 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)
58. Tam, K., Jones, K.: Macra: a model-based framework for maritime cyber-risk assessment. *WMU Journal of Maritime Affairs* **18**(1), 129–163 (2019)
59. Thieme, C.A., Guo, C., Utne, I.B., Haugen, S.: Preliminary hazard analysis of a small harbor passenger ferry–results, challenges and further work. In: *Journal of Physics: Conference Series*. vol. 1357, p. 012024. IOP Publishing (2019)
60. Tran, K., Keene, S., Fretheim, E., Tsikerdekis, M.: Marine network protocols and security risks. *Journal of Cybersecurity and Privacy* **1**(2), 239–251 (2021)

Paper VIII

A. Amro and V. Gkioulos, 'Communication and cybersecurity testbed for autonomous passenger ship,' in *European Symposium on Research in Computer Security*, Springer, 2021, pp. 5–22

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

autoremode, this entails that the APS will be mainly autonomous, with human supervision from a remote control center (RCC) [12]. Although this unconventional mode of operation is expected to improve the provisioning of navigational services, it introduces a wide range of cyber threats with possible safety impacts as it relies on a group of interconnected Industrial Control Systems (ICS) as well as several communication technologies.

Communication and cybersecurity are considered among the biggest challenges for the advancement of the autonomous shipping concept [12]. 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 [12]. 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 [9]. On the other hand, the cybersecurity requirements in addition to a group of risk analysis processes for the APS as a cyber physical system [8,10] 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. 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. 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.

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.

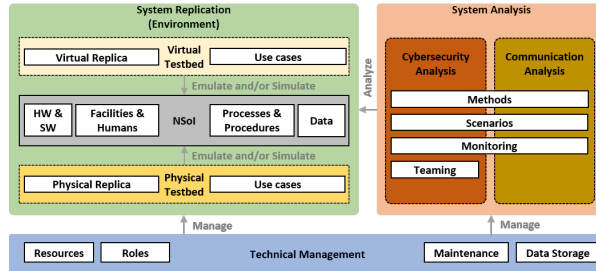


Fig. 1. Process view of the testbed

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.

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

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	<i>OpenCPN</i> 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. <i>TcpReplay</i>)
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

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 data 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.

Additionally, several maritime processes and procedures are addressed including the different communication functions specified in the APS communication architecture [9], 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 [9]. 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, particularly, communication or cybersecurity analysis. Different aspects are relevant for each direction. 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.

Communication and Cybersecurity Testbed for Autonomous Passenger Ship

Table 2. Tools utilized in the virtual testbed

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

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
		Furuno 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)	
	Network Equipment	Maretron IPG100 (2)	NMEA Internet Protocol Gateway
		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
	Portable Power Sources	Cisco RV042G (2)	Load balancer, VPN router, and firewall
Omnicharge Ultimate (7) 9V power bank (3)		Portable power source with 38400 mAh. Providing DC, AC, and USB output. Additional power sources	
System Analysis	Software Defined Radio (SDR)	SDRplay RSPdx (1) ADALM-PLUTO (4)	Wideband SDR Active SDR learning module
	Technical Management	Data Backup	LaCie 2TB (1)

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)

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.	

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 [9] and a cybersecurity architecture [7] proposed in our earlier works based on a group of predefined communication and cybersecurity requirements in [12] 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 [9]. 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 cybersecurity 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 autoremove 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 simu-

lator 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 Management		Each facility at a dedicated workstation
	Role Management	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 [11]. 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 Automatic Identification System (AIS) device. Additionally, the *OpenCPN* chart plotter software³ is used and configured to receive the transmitted NMEA messages. Additional scripts are 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

¹ <https://www.bridgecommand.co.uk> (accessed July 2021)

² <https://cutt.ly/NMEASimulator> (accessed July 2021)

³ <https://opencpn.org> (accessed July 2021)

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, Fururu 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 of 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 by Yamin et al [27], 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.

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

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	✗		Emulation tools	✓
	Domain	Hybrid network applications	✓	Tools	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	✓	Red team	✓		
	Management	✓	Teaming	Blue team	✓	
	Learning	✗		White team	✓	
Monitoring	✓	Green team		✗		
Remote Access	✓	Autonomous Team		✓		
Mobility	✓	Purple teaming		✓		
Scalability	Restricted					

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 un-

dergoing 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. Thus, 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 de-

velopment 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.: Communication architecture for autonomous passenger ship. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability p. 1748006X211002546 (2021)
10. 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)
11. Amro, A., Oruc, A., Yildirim Yayilgan, S.: Nmea anomaly analysis and detection (2020), to be submitted
12. 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)
13. Committee, T.M.S.: Maritime cyber risk management in safety management systems (2017)

Amro A. and Gkioulos V.


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 iee 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)

Paper IX

A. Amro and V. Gkioulos, 'Evaluation of a cyber risk assessment approach for cyber-physical systems: Maritime- and energy-use cases,' *Journal of Marine Science and Engineering*, vol. 11, no. 4, 2023, ISSN: 2077-1312. DOI: 10.3390/jmse11040744. [Online]. Available: <https://www.mdpi.com/2077-1312/11/4/744>

Article

Evaluation of a Cyber Risk Assessment Approach for Cyber–Physical Systems: Maritime- and Energy-Use Cases

Ahmed Amro *  and Vasileios Gkioulos

Department of Information Security and Communication Technology, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, 2815 Gjøvik, Norway; vasielios.gkioulos@ntnu.no

* Correspondence: ahmed.amro@ntnu.no

Abstract: In various domains such as energy, manufacturing, and maritime, cyber–physical systems (CPSs) have seen increased interest. Both academia and industry have focused on the cybersecurity aspects of such systems. The assessment of cyber risks in a CPS is a popular research area with many existing approaches that aim to suggest relevant methods and practices. However, few works have addressed the extensive and objective evaluation of the proposed approaches. In this paper, a standard-aligned evaluation methodology is presented and empirically conducted to evaluate a newly proposed cyber risk assessment approach for CPSs. The approach, which is called FMECA-ATT&CK is based on failure mode, effects and criticality analysis (FMECA) risk assessment process and enriched with the semantics and encoded knowledge in the Adversarial Tactics, Techniques, and Common Knowledge framework (ATT&CK). Several experts were involved in conducting two risk assessment processes, FMECA-ATT&CK and Bow-Tie, against two use cases in different application domains, particularly an autonomous passenger ship (APS) as a maritime-use case and a digital substation as an energy-use case. This allows for the evaluation of the approach based on a group of characteristics, namely, applicability, feasibility, accuracy, comprehensiveness, adaptability, scalability, and usability. The results highlight the positive utility of FMECA-ATT&CK in model-based, design-level, and component-level cyber risk assessment of CPSs with several identified directions for improvements. Moreover, the standard-aligned evaluation method and the evaluation characteristics have been demonstrated as enablers for the thorough evaluation of cyber risk assessment methods.



Citation: Amro, A.; Gkioulos, V. Evaluation of a Cyber Risk Assessment Approach for Cyber–Physical Systems: Maritime- and Energy-Use Cases. *J. Mar. Sci. Eng.* **2023**, *11*, 744. <https://doi.org/10.3390/jmse11040744>

Academic Editors: Kevin Jones and Kimberly Tam

Received: 7 March 2023

Revised: 21 March 2023

Accepted: 27 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: cyber risk assessment; evaluation; cyber–physical systems; ATT&CK; FMECA; maritime; energy; autonomous passenger ship; digital substation

1. Introduction

Interest in cyber–physical systems (CPSs) has increased in recent years across different application domains such as maritime and energy. The maritime industry is undergoing a major transformation leading to changes in operations and technology [1]. As an example of this trend, this work is part of the “Autoferry” project [2] that aims to develop a ferry to transport passengers autonomously across the Trondheim canal. In our previous work [3], we classified the ferry as an autonomous passenger ship (APS).

Cyber attacks targeting the maritime domain are increasing both in number and severity [4]. Attacks of this type target all segments of the maritime infrastructure, including ships, ports, and shipping companies. The denial of service attack against the COSCO shipping company [5], stealing of confidential designs from Austal naval shipbuilders [6], and ransomware attack against Maersk [7] are notable and well-known examples. Arguably, attacks against ships are of relatively low complexity [8]. In real events, ships themselves have also been targets of attacks, and incidents involving their global positioning system (GPS) and communication technologies [9] indicate the feasibility of cyber attacks and potential impact.

Regarding the APS, it includes a wide range of information and communication technology (ICT), and industrial control systems (ICS) allowing it to be conceptualized as a CPS. The APS security risks may directly or indirectly endanger passengers' safety and adversely affect the operational environment. The ship can be steered into a collision with the surrounding environment or other ships if its remote or autonomous control capabilities are hijacked. To increase the trustworthiness, security, and resilience of integrated systems, risk management is considered for implementation in the APS architecture. As discussed in ISO 31010 [10] and ISO 27005 [11], risk management includes several processes, with risk assessment at the core. To ensure the safety of people and the systems themselves, the relationship between safety and security in the risk management of CPSs, such as autonomous ships, requires additional attention.

Moreover, surveyed risk assessment approaches in CPSs have been observed to rely heavily on experts' judgment which increases the required efforts for continuous risk assessment and management as well as having results that are heavily subject to bias [12]. In this direction, the authors of this article have previously proposed an approach for assessing the risks in cyber-physical systems [12]. The approach is based on failure mode, effects and criticality analysis (FMECA) risk assessment process and enriched with the semantics and encoded knowledge in the Adversarial Tactics, Techniques, and Common Knowledge framework (ATT&CK). We refer to this approach throughout this article as FMECA-ATT&CK. The approach reduces the need for expert judgment in several steps of the risk assessment leading to reduced efforts and impact of bias on judgment. In this article, further evaluation of FMECA-ATT&CK is carried and its results are presented. The evaluation relies on the engagement of a group of experts for conducting the risk assessment process using the FMECA-ATT&CK approach and another common approach, the Bow-Tie, against two different use cases in different application domains, one in maritime and another in energy. This allows for the evaluation of the approach based on a group of characteristics, namely, applicability, feasibility, accuracy, comprehensiveness, adaptability, scalability, and usability.

The contribution of this article can be summarized as follows:

- An evaluation of an open-source risk assessment process that is FMECA-ATT&CK supporting its development as a semi-automated cyber risk assessment tool for CPSs.
- Key characteristics for the evaluation of risk assessment methods. These characteristics can be utilized as a basis for comparison among existing and newly proposed methods for risk assessment.
- A standard-aligned methodology for the evaluation of risk assessment methods. The methodology allows for the evaluation according to a group of characteristics while reducing the impact of bias.

The remainder of this paper is structured as follows. In Section 2, background information regarding the relevant standards, methods, and use cases are provided. In Section 3, a group of related works is discussed to highlight the observed relevant methods and characteristics for evaluation. Then, the evaluation methodology is presented in Section 4. The evaluation of FMECA-ATT&CK is detailed in Section 5. The evaluation results are presented in Section 6. Then, reflections from conducting the evaluation including identified limitations and future directions are discussed in Section 7. Finally, concluding remarks are provided in Section 8.

2. Background

FMECA-ATT&CK has been developed as an approach for risk assessment. It has a defined set of inputs, and procedures, and produces an output. This allows it to be conceptualized as a system. Therefore, the evaluation is approached as a system analysis process following the ISO 15288:2015 system development standard [13]. Furthermore, the chosen system analysis method relies on experts' judgment through brainstorming, then techniques for eliciting expert views are utilized from the IEC 31010:2019 standard for risk assessment techniques [10]. Additionally, the applicability of FMECA-ATT&CK

for assessing risks in different use cases and application domains is among the targeted characteristics for evaluation. Therefore, two use cases are utilized to carry out the assessment procedure. One use case is a maritime-use case that is an APS while the other is from the energy domain that is a generic digital substation. In order to establish a reference for comparison, the evaluation includes the utilization of another well-established assessment process, the Bow-Tie, for evaluating the same use cases. In this section, several resources, approaches, and use cases are introduced to facilitate later discussion of the evaluation process of the FMECA-ATT&CK approach.

2.1. Standards, Methods and Approaches

The evaluation process is aimed to be aligned with the relevant standards and common approaches in the industry. The relevant standards are the IEC 31010:2019 [10], ISO 15288:2015 [13], and IEC 60812 [14]. Additionally, a commonly utilized method for risk assessment, the Bow-Tie, is utilized to provide a basis for a comparison with regards to FMECA-ATT&CK. Moreover, details regarding the use cases are presented hereafter.

2.1.1. IEC 31010:2019, ISO 15288:2015 and IEC 60812

The FMECA-ATT&CK approach has been developed based on the IEC 60812 FMECA standard [14]. The standard provides detailed steps for conducting a FMECA process including guiding criteria and suggested methods. The IEC 31010:2019 standard [10] was utilized for the identification of relevant risk analysis and assessment techniques to be adopted during the different steps in the FMECA process, such as the utilization of threat taxonomies as a threat identification method. Additionally, the IEC 31010:2019 standard was consulted during the evaluation of FMECA-ATT&CK regarding guidelines for eliciting expert opinions and judgment. Additionally, the system analysis process in the ISO 15288:2015 [13] standard for system development was consulted for the development of the evaluation methodology, particularly, the system analysis process. This highlights how aligned the FMECA-ATT&CK approach and its evaluation is with the relevant standards.

2.1.2. FMECA-ATT&CK

The Adversarial Tactics, Techniques, and Common Knowledge from MITRE, shortly known as the ATT&CK framework [15] is witnessing widespread adoption in both academia and the cybersecurity industry as a source of knowledge regarding adversarial tactics, techniques, and procedures (TTP). ATT&CK includes several technology domains such as enterprise information technology (IT) and the operational technology (OT) in industrial control systems (ICSs) and mobile technology making ATT&CK suitable in a wide range of use cases hosting a collection of these technologies. As opposed to other high-level models observed in the literature such as STRIDE [16] and the cyber Kill Chain [17], the ATT&CK framework presents a comprehensive and low-level abstraction of adversarial tactics and techniques. Additionally, the witnessed utilization of ATT&CK terminologies in threat reports [18] and cybersecurity testing frameworks, such as Caldera [19], highlights the utility of integrating the ATT&CK framework within different risk management processes starting with risk assessment.

Yet, ATT&CK is not a method for risk assessment. Therefore, a number of approaches have been considered in order to determine what method is most appropriate for risk assessment. We referred to the IEC 31010:2019 [10] standard for risk assessment techniques. The standard describes and compares the most commonly employed techniques in the different steps of risk assessment. We considered scope, time horizon, specialist expertise requirements, and the amount of effort required to apply risk assessment techniques. The scope of our risk assessment in a CPS includes components, equipment, and processes. In order to support continuous risk assessment and management as well as reduce the effect of biased assessment associated with expert judgment, the time horizon should be flexible. In addition, the amount of specialist expertise and effort needed should be at most moderate. We have chosen failure modes, effects, and criticality analysis (FMECA) [14] based on the

mentioned criteria. Further, the standard emphasizes FMECA’s application to all stages of the risk assessment process, which include identifying risk, assessing consequence, estimating likelihood, and evaluating risk.

Based on FMECA, FMECA-ATT&CK makes use of common knowledge encapsulated in the ATT&CK framework as shown in Figure 1. The components within the scope, their functions, and performance standards are defined. Then, the relevant failure modes are identified, and for this, the ATT&CK tactics are considered. Then, the existing detection methods are identified and their efficiency is estimated. Later, the impact of the consequences of failure is estimated based on five elements of impact, namely, operational, safety, financial, information, and staging (stage further attacks). The operational and staging impacts are estimated based on the centrality measures of the component after a graph of the system is modelled. The remaining elements are estimated based on expert judgment. Each failure mode is assigned a weighting of the expected impact elements. For instance, the collection tactic as a failure mode is only expected to cause information and staging consequences. In order to calculate the estimated failure mode for each component, each component is assigned criticality scores covering all five elements of impact. Afterwards, the possible failure mechanisms causing the failure modes are identified. For this, the ATT&CK techniques are utilized and their properties, such as relevant assets and platforms, are used to match them with the relevant components. After this, the likelihood of failure mechanisms is estimated based on the exploitability score in the common vulnerability scoring system (CVSS) which considers attack vector (AV), attack complexity (AC), privilege required (PR), and user interaction (UI). Later, the risk rating criteria are defined (e.g., based on value distribution). Finally, the relevant mitigation measures for each failure mechanism are defined. This is derived from the encoded knowledge in ATT&CK. When all the aforementioned information is collected, a risk priority number calculation and mitigation identification (RPNMI) algorithm is executed to calculate the risk of each failure mechanism and suggest the relevant mitigation measures. A detailed description of the FMECA-ATT&CK steps, tables, data types and sources of knowledge is presented in Appendix A. Additionally, the reader may refer to our original work [12] for more information regarding the risk assessment approach including a detailed comparison with other approaches ([12] §2.1).

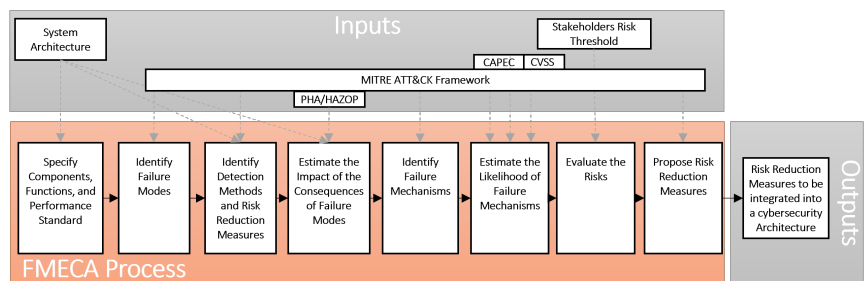


Figure 1. Steps of FMECA-ATT&CK with the knowledge sources (adapted from [12]).

2.1.3. Bow-Tie

The Bow-Tie approach allows for the assessment of cyber risks and the identification of barriers needed to control them without focusing on the likelihood. This aids in quick visualizations of the measures that are needed for implementation [20]. The Bow-Tie is a well-known method in the maritime sector and was found suitable for implementation in this paper to provide a basis for comparing the risk assessment results achieved through FMECA-ATT&CK and evaluating their soundness. The Bow-Tie method, as shown in Figure 2, begins with defining the scope of the target system for evaluation. This can be performed by interviewing system users, operators, and other stakeholders to answer scoping questions regarding the system components, and existing mitigation measures

to identify gaps. Then, threats and consequences are identified using any suitable threat modelling approach. This includes the identification of a top event, threat scenarios leading to it, and possible arising consequences. Then, the incident prevention and consequence reduction barriers are identified. Finally, the robustness and effectiveness of the barriers are considered to identify directions for improvement.



Figure 2. The Bow-Tie method (adapted from [20]).

2.2. Use Cases

Two use cases are utilized to evaluate FMECA-ATT&CK, namely, an autonomous passenger ship (APS) and a generic digital substation (DS). The APS represents a use case from the maritime domain while the DS represents a use case from the energy domain.

2.2.1. Autonomous Passenger Ship (APS)

The first use case is a prototype of an APS named milliAmper2. An overview of the use case description is depicted in Figure 3.

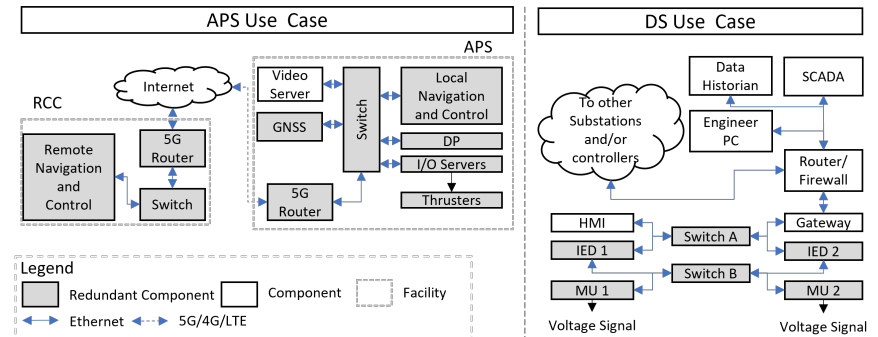


Figure 3. Overview of the APS and DS use cases.

The assessment scope was defined in one network among several networks of the milliAmper2 to reduce the assessment time and required efforts. However, information regarding redundant systems in other networks was utilized during the assessment for accurate risk estimation. In summary, the ferry includes several components such as a global navigation satellite system (GNSS) and video cameras as sensor data. These data are sent to a local navigation and control system to identify safe routes and then instruct the dynamic positioning (DP) system to control the thrusters through a group of I/O servers. The ferry is connected through a 5G network to a remote control centre (RCC) hosting a remote navigation and control system that can intervene in case of an unsafe situation. Further details about the APS architecture can be found in [21].

2.2.2. Digital Substation (DS)

To evaluate FMECA-ATT&CK applicability in different use cases in the different application domains, another use case is needed. For this, we have identified the digital substation from the work of Khodabakhsh et al. [22] as a suitable use case. An overview of the use case description is depicted in Figure 3. The digital substation includes a supervisory control and data acquisition (SCADA) system with an engineering PC for monitoring and control. A data historian is hosted for storage. These components are connected through a router and a gateway to the lower devices, intelligent electronic devices (IED),

human-machine Interface (HMI), and merging units (MU) for monitoring and controlling the voltage signal.

3. Related Work

The area of cyber risk assessment in cyber-physical systems is rich with relevant literature. A wide range of risk assessment methods and approaches exist. However, limited works have been observed regarding a structured and systematic evaluation of such works.

Some works have been observed that targets the evaluation of risk assessment approaches. Tam [23] conducted a qualitative evaluation of the author's risk assessment framework. The author relied on expert judgment to measure the usability and applicability of the risk assessment framework using a survey. Abkowitz and Camp [24] investigated the applicability of the enterprise risk management (ERM) framework in marine transportation. The authors utilized a group of experts to implement the ERM framework against a case study including a marine transportation carrier. Grigoriadis et al. [25] engaged system stakeholders to evaluate a risk assessment tool regarding satisfaction of the stakeholders' security and privacy requirements as well as the feasibility of its application. The evaluation approach entails a demonstration of the tool and asking the participants to answer a questionnaire. Moreover, ref. [26] examined the feasibility of using the system theoretic process analysis (STPA) for risk analysis and quantitative risk modelling of autonomous ships. The author identified and assessed 35 risk analysis methods and found seven methods that can be used to enhance STPA for risk analysis of autonomous ships.

Additionally, several works proposing risk analysis and assessment approaches in CPS have been observed in the literature. The authors' evaluation of their contributions tend to include the utilization of certain use cases to demonstrate the applicability of their proposed approach (e.g., [27,28]). Some works have utilized other approaches to provide a ground for comparison (e.g., [23,26]).

Moreover, several guidelines and standards are available with relevant artefacts to evaluate risk assessment approaches. This includes the NIST assessment guidelines (NIST.SP.800-53Ar5) [29] and the ISO 31010:2019 risk assessment standard [10]. The NIST guidelines discuss several approaches for evaluation, namely, examine, interview, and test with different levels of rigour and scope ranging from basic to comprehensive. The ISO 31010 standard [10] suggests characteristics for comparison among risk assessment and analysis methods including application, scope, specialist expertise, and efforts to apply.

Lastly, our original work [12] proposing FMECA-ATT&CK as a risk assessment approach for CPS discussed the background and rationale. Among the original objectives is to include the applicability in different application domains utilizing information technology (IT) and operational technology (OT). Furthermore, the approach must be comprehensive in its consideration of risk elements. Additionally, the system must also reduce the need for expert judgment through employing the concept of curated knowledge to support the automation of some elements to allow a continuous risk assessment process. Moreover, the approach's adaptability to include additional components of risks has been demonstrated in the original work. Therefore, measuring the applicability of FMECA-ATT&CK, its comprehensiveness, and the accuracy of the results based on curated knowledge, automated elements and adaptability, was deemed necessary to assess the satisfaction of the original objectives and therefore these characteristics are targeted in this work.

4. Evaluation Methodology

An evaluation methodology for evaluating a risk assessment approach is proposed in this section. The approach under evaluation, FMECA-ATT&CK in this paper, is conceptualized as a system. Therefore, the evaluation is approached as a system analysis process following the ISO 15288:2015 [13] system development standard. This includes preparation, conducting, and managing the analysis results, as shown in Figure 4.

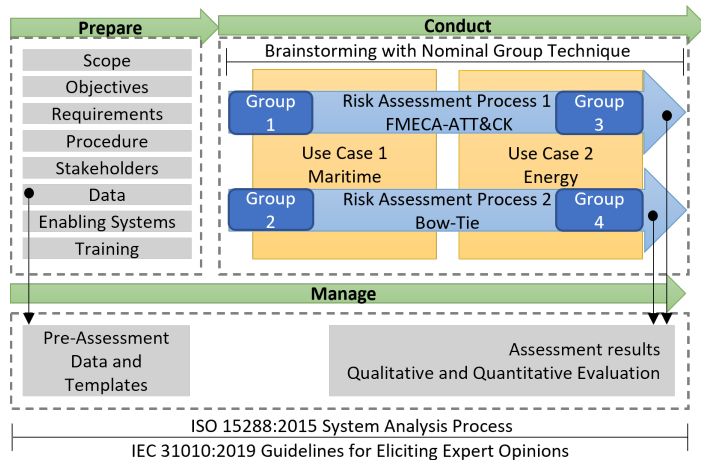


Figure 4. Evaluation methodology.

4.1. Preparing for the Evaluation

The preparation entails specifying the evaluation scope, objectives, and requirements. Then, the evaluation procedure is determined, the stakeholders are identified, and the required data and enabling systems are prepared. The scope is defined within the evaluation of a risk assessment method through application against several CPS use cases in different application domains. The evaluation objective is to evaluate the risk assessment method according to a group of characteristics, namely, applicability, feasibility, comprehensiveness, adaptability, scalability, usability, and accuracy. The characteristics were chosen based on what has been observed in the literature as well as the original motivations that led to the proposition of FMECA-ATT&CK. The evaluation characteristics are summarized in Table 1. Methods for measuring the characteristics are discussed in Section 5.1.2.

Table 1. Evaluation characteristics for the cyber risk assessment approaches.

Characteristics	Objective	Related Work
Applicability	suitability for application in different use cases in different application domains.	[23,24]
Feasibility	the ability to implement the different steps in the approach.	[25,26]
Comprehensiveness	the extent to which different aspects of risks have been considered. Aspects of risks include, threats identification, likelihood and impact estimation, mitigation measures, etc.	[12]
Adaptability	The extent to which the missing aspects can be integrated to improve the method.	
Scalability	The performance of the process in large and complex networks.	
Usability	The ability to follow and conduct the process with limited training/consultation.	[23]
Accuracy	The soundness of the results.	[23,26]

Then, the analysis requirements are identified. The requirements are expected to be different according to each evaluation process. The following requirements are derived based on the method itself:

- Due to the reliance on expert judgment, measures for reducing bias in the assessment must be integrated in order to improve the assessment quality.
- Diversity in the use cases should be pursued to include various application and technology domains in order to measure applicability.
- Another common risk assessment method needs to be chosen that performs a similar function to the method that is subject to evaluation and provides categorically aligned results.

Additionally, an evaluation procedure should be defined. This includes applying the risk assessment that is subject to evaluation as well as another common and similar method against the same set of use cases. This is intended to provide a reference to compare the results. Moreover, the relevant stakeholders for the evaluation should be identified. The identification should consider their expertise in the application domain of the use cases. Finally, the data and enabling systems needed for the evaluation need to be prepared. This includes training the participants for the assessment.

4.2. Executing the Evaluation

The evaluation is proposed to be executed over several sessions spanning the different groups. As shown in Figure 5, the procedure is divided into three stages for each group, the first stage aims to run the assessment process step by step, describe to each participant the individual tasks, and address their questions. The participants should be given a sufficient period of time to provide their individual input. After receiving the participants' input, the results are evaluated to identify conflict areas and generate initial results based on consolidated inputs. Proposed consolidation rules can be found in Appendix B. In the third stage, the results are discussed in a group to reach a conclusion. Then, feedback from the participants applying the method under evaluation should be queried regarding their experience with the evaluated risk assessment approach utilizing a questionnaire.

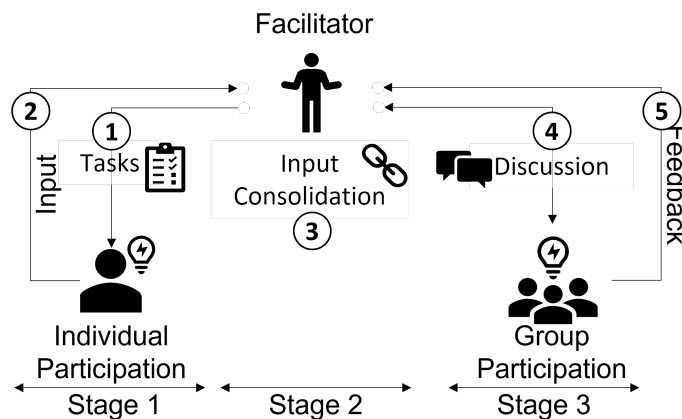


Figure 5. Execution Procedure.

4.3. Managing the Evaluation

All the prepared data for the evaluation, the participants' input, and the results should be maintained for future reference. This includes documents for conducting risk assessment processes by experts, the description of the use cases supporting their assessments, and the tools necessary to conduct the assessment processes, as well as their outputs.

5. FMECA-ATT&CK Evaluation

In this section, the evaluation process of FMECA-ATT&CK is presented. The evaluation is based on the methodology discussed in Section 4.

5.1. Preparing for the Evaluation

Preparation involves defining the scope, objectives, and requirements of the evaluation. Following this, the evaluation procedure is determined, stakeholders are identified, and the data and enabling systems are prepared.

5.1.1. Scope, Objectives and Requirements

The scope is constrained to the evaluation of FMECA-ATT&CK as a design-level cyber risk assessment approach in cyber-physical systems (CPSs). Two use cases of CPSs were chosen in different application domains, namely, an autonomous passenger ship (APS) or ferry representing the maritime domain and a generic digital substation (DS) representing the energy domain. The objectives include evaluating FMECA-ATT&CK according to the defined characteristics, namely, applicability, feasibility, comprehensiveness, adaptability, scalability, usability, and accuracy.

5.1.2. Evaluation Procedure

The evaluation procedure combines both qualitative and quantitative evaluation. All the characteristics are qualitatively evaluated after applying FMECA-ATT&CK in different use cases. The evaluation is based on experts' feedback through a questionnaire. A detailed evaluation criteria can be found in Appendix C. Additionally, to evaluate the applicability of FMECA-ATT&CK, we considered the utilization of two distinct use cases in different application domains as the target system of analysis. Additionally, usability is quantitatively measured by monitoring the experts' progression throughout the execution process. Moreover, the accuracy is qualitatively measured through a categorical comparison of the results obtained after the utilization of another commonly adopted risk analysis and assessment method carried out by different experts. The results of both methods are used as the basis for comparison to evaluate whether FMECA-ATT&CK is able to provide sound results.

Moreover, brainstorming to elicit experts' comments, concerns, and ideas regarding FMECA-ATT&CK was found to be a suitable approach for evaluation. However, since the chosen evaluation approach relies on expert judgment, the techniques for eliciting expert views are utilized from IEC 31010:2019 [10]. Additionally, since such views are subject to bias, the following measures for reducing bias were implemented:

- To reduce bias based on the bandwagon effect, the nominal group technique is implemented [30]. The bandwagon effect refers to the tendency of group ideas to converge rather than diverge. The nominal group technique has been found to generate more ideas than brainstorming alone [10].
- Group communication is hindered to avoid information bias.
- FMECA-ATT&CK itself implements measures to reduce bias through the utilization of metrics based on graph theory and data from the ATT&CK framework.
- Inputs from previous relevant risk assessment processes are avoided as much as possible. However, the utilization of some previous data was unavoidable. More details will be discussed later on when such a case occurred.

Then, the additional risk assessment process to be conducted was chosen to be the Bow-Tie method since it is a common approach in evaluating the risks in CPSs. Therefore, brainstorming with the nominal group technique while implementing FMECA-ATT&CK and Bow-Tie against different use cases was determined as the assessment procedure. Excel sheets were utilized as the medium for guiding the tasks and collecting the input from the experts. Noteworthy, each risk assessment process might require specific system-level information in a specific format. Therefore, a coherent state of the system description of both use cases must be maintained when applying the two processes to ensure a symmetric basis for evaluation. Furthermore, several groups each working on a different use case and applying a specific risk assessment process need to be formulated.

5.1.3. Identifying Stakeholders

The relevant stakeholders including the evaluation participants are identified and approached. In this direction, several subject matter experts (SMEs) were considered based on their experience in the use case application domain. As shown in Table 2, both academic and industrial SMEs were pursued with various experiences and backgrounds to improve the quality of the evaluation process.

Table 2. Experts roles and background.

Group	Assessment Process	Participants	
		Current Roles	Background and Previous Roles
1	FMECA-ATT&CK on APS	PhD candidate in maritime cybersecurity	Working experience on off-shore vessels
		Researcher in cybersecurity	Maritime, energy, and CPS cybersecurity
		PhD candidate in maritime cybersecurity	Seafarer (AB apprentice, AB, Deck Cadet, Junior Officer) and FMEA Auditor of DP systems
2	Bow-Tie on APS	Cybersecurity Consultant	IT/OT cybersecurity
		Cybersecurity Consultant	IT/OT cybersecurity
3	FMECA-ATT&CK on DS	PhD Candidate in CPS cybersecurity	Cybersecurity in the smart grid
		Postdoctoral researcher	Postdoctoral researcher in smart grid communication and security simulation
		Researcher in cybersecurity and privacy	Cybersecurity in the smart grid and IoT privacy
4	Bow-Tie on DS	Industrial PhD/Cybersecurity Engineer	Cybersecurity in the smart grid
		Industrial PhD/Senior adviser information security	Working with SCADA/OT—systems in the electricity sector for over 30 years

5.1.4. Preparing Data, Enabling Systems and Training for the Assessment

The required data and enabling systems for the assessment were identified after a detailed study of the two risk assessment processes, namely, FMECA-ATT&CK and Bow-Tie. The common starting point for both is to define the analysis scope and this includes the targeted system for evaluation (i.e., use case). Descriptions of the two use cases to be targeted by both assessment processes were formulated. Several views of each use case architecture were prepared and made ready for analysis (Sections 2.2.1 and 2.2.2). The following steps are different for each risk assessment process. The steps for conducting FMECA-ATT&CK were followed based on our original article where it was proposed [12], while the steps for conducting Bow-Tie were based on the class guidelines published by De Norsk Veritas (DNV) [20].

When preparing the data required to conduct FMECA-ATT&CK, some data do not rely on expert judgment as it is extracted from the continuously updated ATT&CK framework, others required modelling the use cases in graphs to calculate the centrality metrics, while others were identified to require input from the experts.

The DNV class guidelines [20] were utilized for preparing the data required for conducting Bow-Tie. This entails answering a list of questions to help guide the experts in identifying the scope of the analysis and assessing the risks.

Finally, to facilitate a productive risk assessment process with the limited time the experts were willing to provide, the participants received a briefing and training on the assessment procedure and were provided with the required preliminary data using a combination of meetings and email communications.

5.2. Executing the Evaluation

The evaluation was executed over several sessions spanning the different groups following the procedure depicted in Figure 5. During the first stage, the assessment processes, namely Bow-Tie and FMECA-ATT&CK, were conducted step by step, describing to each participant the individual tasks, and addressing their questions. A sufficient period of time was provided so that participants could provide their individual input. On the basis of consolidated inputs, the results were evaluated to identify conflict areas and generate initial results. To reach a conclusion, the results were discussed in a group stage. Participants applying FMECA-ATT&CK were then surveyed regarding their experience with the risk assessment approach utilizing a questionnaire. An exception to that procedure occurred with the fourth group as we were forced to accommodate the participants' time constraints by running the first stage as a group by facilitating and stimulating discussion among the group participants and receiving their input for the assessment. After consolidating their input, the results were sent to the participants to receive their confirmation on the final assessment results.

5.2.1. Delivering Tasks and Receiving Input from Experts

The tasks for conducting FMECA-ATT&CK were compiled in an Excel worksheet including all the steps, tables to provide guiding notes, and extra room to receive detailed comments. The utilized template for each use case can be found in the authors' public repository (<https://github.com/ahmed-amro/FMECA-ATT-CK-Evaluation>) (accessed on 28 February 2023). On the other hand, the tasks for conducting Bow-Tie were communicated to the relevant groups in meetings. They were provided with the data prepared for the assessment. They were requested to deliver input to draft Bow-Tie diagrams and highlight the top threats, and mitigation measures required. After a sufficient period, the experts provided their answers.

5.2.2. Evaluating the Results

After receiving the experts' input. A consolidation process was executed to produce the risk assessment results (see Appendix B). The inputs for the FMECA-ATT&CK process were fed into a semi-automated tool to generate the results. On the other hand, the inputs for the Bow-Tie process were utilized to draft Bow-Tie diagrams. Finally, the results were presented to the experts, and a discussion was opened to reach a conclusion. Both sets of outputs were then compared to evaluate the soundness of the FMECA-ATT&CK results.

Another input was received from the groups applying the FMECA-ATT&CK process. This includes quantitative evaluation in the form of a rating of the process based on the characteristics specified earlier (Section 5.1.1) as well as qualitative evaluation in the form of comments on the process.

5.3. Managing the Evaluation

All the prepared data for the evaluation, the participants' input, and results are maintained in a public repository for future reference (<https://github.com/ahmed-amro/FMECA-ATT-CK-Evaluation>) (accessed on 28 February 2023). This includes the following:

- The utilized template for receiving experts' input for each use case when conducting FMECA-ATT&CK.
- The scoping questions and the prepared answers for the Bow-Tie process.
- The FMECA-ATT&CK scripts, inputs, and outputs.
- The generated Bow-Tie diagrams.

Moreover, this paper constitutes a report of the executed evaluation with the lessons learned.

6. Evaluation Results

The results of the evaluation are presented in this section. The evaluation relied on three types of input, namely, categorical comparison between the results of the risk

assessment processes, namely, FMECA-ATT&CK and Bow-Tie, experts’ feedback through questionnaires, and experts’ comments.

6.1. Risk Assessment Results

The results from the risk assessment processes conducted by the four groups have been collected, categorized, and compared. The categorization is based on the identified risks and suggested controls. More details in this regard are presented hereafter.

6.1.1. Top Risks

The Bow-Tie and FMECA-ATT&CK methods are categorically compatible. Threats, consequences, and mitigations in Bow-Tie can be mapped to techniques, tactics, and mitigations in FMECA-ATT&CK, respectively. This allows for a comparison of the results of the two methods and consequently provides evidence regarding the soundness of the results obtained through the application of FMECA-ATT&CK. The top risks identified through Bow-Tie and their relevant identified techniques in FMECA-ATT&CK and their corresponding risks are presented in Table 3. The results suggest that FMECA-ATT&CK can identify similar risks to Bow-Tie with more granularity-defined atomic techniques. The highest risks identified through FMECA-ATT&CK are all identified through Bow-Tie while several threats identified through Bow-Tie were rendered low risks. The rationale for these discrepancies is the consideration of existing mitigation measures. FMECA-ATT&CK does consider the existing mitigation measures in the risk calculation while the experts applying Bow-Tie appear to have either dropped them from their considerations or found them inefficient. Still, some threats identified through Bow-Tie (e.g., employees wrongdoing) are not supported by the current version of FMECA-ATT&CK which only considers adversarial threats. Due to the technical nature of some attack techniques, we provided the ATT&CK ID for the reader to refer to them in the ATT&CK framework repository. <https://attack.mitre.org/> (accessed on 28 February 2023).

Table 3. The relations between the top risks identified through Bow-Tie and FMECA-ATT&CK.

Bowtie Threats	FMECA-ATT&CK Techniques (ATT&CK ID)	FMECA-ATT&CK Risk
APS Use Case		
Valid Accounts Stolen from a Student	Valid Accounts (T0859/T1078)	Low risk due to the inclusion of many relevant mitigation methods (e.g., access management)
Remote Desktop Protocol (RDP)	Remote Desktop Protocol (T1021.001)	Low risk due to the inclusion of many relevant mitigation methods (e.g., network segmentation)
Compromises Hosts	High-level threat. Relevant techniques: Drive-by Compromise (T1189), Compromise Client Software (T1554)	Both relevant techniques have a low risk either due to the inclusion of several relevant mitigations methods or low estimated impact and likelihood (e.g., update software)
Internal Spear phishing	Internal Spear phishing (T1534)	Low due to low estimated likelihood
Malicious Software	Malicious File (T1204.002)	Low risk either due to the inclusion of many relevant mitigation methods, and low estimated likelihood (e.g., execution prevention)
Compromised Credentials	High-level threat. Relevant techniques: Valid Accounts (T0859/T1078) Default Credentials (T0812)	Low risk due to the inclusion of many relevant mitigation methods (e.g., access management)
Single 4G/5G link	Outside the scope of FMECA-ATT&CK which only considers adversarial threats.	Although no techniques are identified for this specific threat, FMECA-ATT&CK does consider the existing redundant services to calculate the detectability (risk reduction degree).
Malicious Remote Access Tools	Exploitation of Remote Services (T1210)	Low risk due to the inclusion of many relevant mitigation methods (e.g., update software)
Legitimate Credentials with Native Network and Operating System Tools	Remote Services (T1021)	High risk for some components due to high likelihood, impact, and lack of existing relevant mitigation measures
Remote Services		
Commonly used port (RDP, SMB, SSH, etc.)	Commonly Used Port (T0885)	Lw risk due to the inclusion of many relevant mitigation methods (e.g., network segmentation)
Repetitive Change of the I/O point values at the Control computer	Brute Force I/O (T0806)	Low risk due to the inclusion of many relevant mitigation methods (e.g., network segmentation)

Table 3. Cont.

Bowtie Threats	FMECA-ATT&CK Techniques (ATT&CK ID)	FMECA-ATT&CK Risk
DS Use Case		
Supply Chain Compromise	Supply Chain Compromise (T1195)	High risk for some components due to high likelihood, impact, and lack of existing relevant mitigation measures
Wrongdoing by Employees	Outside the scope of FMECA-ATT&CK which only considers adversarial threats.	N/A
External Environmental Threats		
Gaining Access to the System	20 techniques in the “Initial Access” Tactic (TA0001 and TA0108).	High risk for some components due to high likelihood, impact, and lack of existing relevant mitigation measures
Ransomware	Data Encrypted for Impact (T1486)	Low risk due to low likelihood and existing relevant mitigation measures
Malware Injection	Malicious File (T1204.002)	Low risk due to low impact and existing relevant mitigation measures
Rogue Devices	Rogue Master (T0848)	Low risk due to low likelihood and existing relevant mitigation measures

Another categorical view is the desired attacker objectives or expected consequences in the evaluated systems. The identified consequences through Bow-Tie and their corresponding tactics (i.e., objectives) identified by FMECA-ATT&CK are presented in Table 4. The results suggest an alignment of the identified possible consequences in both use cases. The results of Bow-Tie cover all the high risk objectives identified by FMECA-ATT&CK. Some consequences from Bow-Tie are rendered medium to low risks in FMECA-ATT&CK due to existing risk mitigation measures. Additionally, FMECA-ATT&CK identified additional objectives which Bow-Tie did not. This includes privilege escalation, exfiltration, credential access, and others.

Table 4. The relations between the top consequences/objectives identified through Bow-Tie and FMECA-ATT&CK.

Bow-Tie Consequences	FMECA-ATT&CK Tactics	C	H	M	L
APS Use Case					
Malicious actions with logged in user privileges	Initial Access	0	0	7	313
Attackers with more information about the system	Discovery	0	0	0	398
Loss of view and control of the ferry from RCC	Impact	0	15	78	301
Attackers propagate and move freely within the network	Lateral movement	0	3	18	239
Malicious control over compromised hosts	Command and Control	0	54	206	208
An undesired system state or action is reached	Impair Process Control	0	0	0	51
DS Use Case					
Covert access to the system	Command and Control	0	53	69	239
Gaining physical access to the system	Initial Access	0	4	2	40
Losing trust of the system	Impact *	0	6	21	263
Credibility and societal trust					
Human harm					
Reputation damage					
Loss of revenue	Impair Process Control	0	0	0	35
Render system non-functional					

C: Critical, H: High, M: Medium, L: Low. * Trust is not an element of impact estimation; however, losing trust in the system is perceived by the assessors because of the functional impact.

6.1.2. Suggested Risk Controls

The identification of required risk mitigation measures or controls is a main objective of FMECA-ATT&CK. It was originally proposed as an instrument for the identification of risks and the proposition of the required controls to be considered in a subsequent process which includes the development of an architecture for cyber risk management. Table 5 depicts the controls suggested by Bow-Tie and the corresponding controls suggested by FMECA-ATT&CK. The controls already included in the use case are highlighted. Additionally, the number of identified high and medium risks for which the corresponding controls are suggested are presented as well. This suggests a certain priority of certain controls over others. The results suggest that the controls suggested by Bow-Tie and FMECA-ATT&CK are comparable. Most

of the controls proposed by Bow-Tie are also identified by FMECA-ATT&CK to address high to medium risks in both use cases. Some controls are proposed in FMECA-ATT&CK but not in Bow-Tie such as data backups for the APS use case. On the other hand, some controls suggested through Bow-Tie are not supported by FMECA-ATT&CK due to the scope. FMECA-ATT&CK only addresses controls that are relevant to the system’s components.

Table 5. The relations between the top controls identified through Bow-Tie and FMECA-ATT&CK.

Bow-Tie Mitigations	FMECA-ATT&CK Mitigations	Already Included	H	Suggested for M
APS Use Case				
Audit the Remote Desktop Users group membership regularly.	Audit	Yes	0	0
Remove unnecessary accounts and groups from Remote Desktop Users groups.	Use Account Management	Limited **	3	34
Secure remote access to internal PC’s and PLC’s	Access Management, Account Use Policies, Authorization Enforcement, Human User Authentication, Password Policies, Software Process and Device Authentication, User Account Management, Multi-factor Authentication	Partially *	3	4
Secure portable media	Limit Hardware Installation, Antivirus/ Anti-malware, Behaviour Prevention on Endpoint, Execution Prevention, Exploit Protection	Limited **	3	90
Clean support computers	Antivirus/ Anti-malware	Limited **	0	6
Regular patching, minimal applications, AV scan etc. for the jump server	Security Updates, Update Software, Use Recent OS Version, Vulnerability Scanning	Partially *	0	8
Email Gateways	Not supported			
Redundancy of 4G/5G Service	Redundancy of Service	Yes	0	0
Network Segmentation	Network Segmentation, Limit Access to Resource Over Network	Yes	0	3
Strict Access Control and Management of Change (MoC) with proper Validation	Not supported			
Firewalls	Filter Network Traffic, Limit Access to Resource Over Network, Network Allow lists, SSL/TLS Inspection	Limited **	33	126
Intrusion Detection Systems	Behaviour Prevention on Endpoint, Network Intrusion Prevention	Very Limited	48	195
Not Discussed	Data Backup	Very Limited	3	27
DS Use Case				
Following Standards and Routines	Not supported	No		
Asset Management	Not supported	No		
Security Testing	Deploy Compromised Device Detection Method, Vulnerability Scanning	No	4	10
Redundancy and Resilience	Redundancy of Service	Partially *	0	0
Access Control and Management	Access Management, Account Use Policies, Authorization Enforcement, Human User Authentication, Password Policies, Software Process and Device Authentication, User Account Management, Multi-factor Authentication, User Account Control	Partially *	0	5
Segmentation	Network Segmentation, Limit Access to Resource Over Network	Yes	0	14
Certification	Not supported	No		
Awareness, Competence, and Skills Building	User Guidance, User Training, Application Developer Guidance	Yes	0	0
Business Continuity Plan (BCP)	Not supported	No		
Recovery Capability	Data Backup, Remote Data Storage	Yes	0	1
Isolation Mode	Not supported	No		
Incident Response, Detection, and Logging	Audit, Behaviour Prevention on Endpoint, Deploy Compromised Device Detection Method, Exploit Protection, SSL/TLS Inspection, Network Intrusion Prevention	Very Limited	59	75
Not Discussed	Filter Network Traffic	Partially *	12	23
	Update Software	Limited **	4	14
	Execution Prevention	No	3	3
	Encrypt Sensitive Information	No	1	4

H: High Risks, M: Medium Risks. * Partially: the controls are included only for some components. Furthermore, some controls are not included. ** Limited: the controls are included only for very few components.

6.1.3. Usability Metric

The experts applying FMECA-ATT&CK were given an Excel worksheet with detailed instructions for delivering their input. The experts were instructed to leave a field empty if the task was not clear or they lacked the relevant knowledge needed for delivering a sound judgment. This procedure allows for estimating the usability of the current FMECA-ATT&CK version. In this direction, we define a usability metric to be the ratio of the number of decisions made by an expert to the number of decisions asked to be made by the expert. Table 6 depicts the number of decisions provided to the experts and the number of decisions made; subsequently, used to calculate the usability metric. The experts were given mandatory and optional tasks regarding the risk assessment. The mandatory tasks were system-specific; the expert judgment was expected to be different for different use cases. On the other hand, the optional tasks were non-system-specific, such as the threat checklist, likelihood, and mitigation effectiveness. Furthermore, a decision on an aspect added to the process, not existing in the original proposition is considered optional. Such as the estimation of the environmental and reputation impacts. Offering the experts the option to provide a decision was intended to reduce bias from previous risk assessment processes. Table 6 only depicts the statistics related to the required decisions. The estimated usability of the current FMECA-ATT&CK process was 94.32%. This is an excellent indication of the readiness of the process for application in other use cases. Feedback was received from experts regarding the challenges faced during the execution. The main reason for the lack of ability to provide a judgment was the lack of sufficient background.

Table 6. Calculation of the usability metric.

Use Case Expert	APS			DS			Usability
	1	2	3	4	5	6	
Required decisions	700	700	700	624	624	624	
Required decisions made	677	608	692	608	538	623	
% of required decision made	96.71%	86.86%	98.86%	97.44%	86.22%	99.84%	94.32%

6.2. FMECA-ATT&CK Questionnaire

After the execution of the FMECA-ATT&CK risk assessment process, the experts were asked to anonymously answer a questionnaire to rate the method according to the targeted characteristics (Section 5.1.1). The questionnaire is not specific for each of the use cases. Therefore, the compiled results from all the experts are presented in this section. The questionnaire included nine questions, seven of which were related to the targeted characteristics, one regarding the execution time, and the last to record their comments. Additional details regarding the questions are presented in Appendix C. Regarding the execution time, it ranged from 3 to 4 h per expert. The main reason behind this can be linked to the comprehensive nature of the approach which according to the majority of experts was found to be from comprehensive to very comprehensive. The approach was also perceived to be suitable, feasible and highly adaptable for application in several use cases, but requires certain adaptations for its implementation in real systems. The majority of results in the scalability rating suggest that the approach is suitable for implementation in a system of systems with a moderate number of components. Finally, the majority of experts found some of the results to make sense while others did not. This was expected due to the fact that the input used to generate the risk assessment results were consolidated from all the experts in each use case with various diversion in the experts' inputs. Nevertheless, experts' critical comments were received and are presented in Section 6.3 and will be considered for future improvement of FMECA-ATT&CK. Additionally, the risk assessment results when compared to the Bow-Tie results suggest that FMECA-ATT&CK is capable of producing sound results that are comparable with a more granular risk description, adaptable and comprehensive approach.

6.3. Experts Comments

The experts were asked to provide their critical comments regarding each step of FMECA-ATT&CK. Several comments were received from different experts. They can be summarized as follows:

- Scope definition (Step 1): The classification criteria for certain components is not clear. Some components can be classified in different ways, others were outside the knowledge field of some experts. Furthermore, additional technical and non-technical components should be considered, such as the human operator. Moreover, there exist several performance standards for defining safety-related failure modes.
- Relevant failure modes (Step 2): The criteria for defining the relevant failure mode was characterized as difficult. Some emphasized existing failure modes that are safety-related are easier to consider than security-related failure modes. Furthermore, human errors were proposed for consideration.
- Impact estimation of failure modes (Step 4): The current estimation criteria are generic and require additional methods such as a hazard and operability study (HAZOP) or event tree analysis (ETA). Furthermore, some failure modes were unclear to some experts and therefore were unable to estimate their impact. Additionally, quantifying the safety, financial, environmental, and reputation criticality scores for certain components was found to be challenging.
- Training: Additional training was required for better execution.
- Scope: Experts with more operational than technical expertise found the approach difficult to apply due to the lack of knowledge of component-level failures. Furthermore, the human element is under-represented in the current approach. Humans can be an asset in the system as well as a risk.
- Background: the approach requires several experts with diverse backgrounds, including operational and technical experts. Some components require specific knowledge to provide a more sound judgment.

7. Discussion

This paper presents an empirical study aimed to evaluate the recently proposed FMECA-ATT&CK risk assessment approach for CPSs. The evaluation approach relied on expert judgment. FMECA-ATT&CK was subjected to detailed application and critical comments from a group of experts with various expertises and diverse backgrounds to elicit improvements. In this section, we will summarize the limitations of the FMECA-ATT&CK approach and discuss directions for future work.

Input from experts implementing Bow-Tie referred to the demanding task of conducting a component-level assessment. With the time provided for assessments, only high-level assessment was possible. FMECA-ATT&CK, on the other hand, was originally proposed as a method to reduce the need for expert judgment while at the same time being comprehensive and systematic in its coverage. The expert spent no time identifying threats, estimating their likelihood, or figuring out the required risk controls. Such information was utilized based on the encoded knowledge provided by the ATT&CK framework. The threats were drawn from the list of ATT&CK techniques and their properties. Based on the components' properties in the system model, the relevant ATT&CK techniques are automatically identified as relevant threats. The likelihood values of the ATT&CK techniques are estimated based on the CVSS method and relying on a group of heuristics (more details can be found in [12]). The relevant risk controls for each ATT&CK technique are queried from the ATT&CK repository. Furthermore, the experts were not required to estimate the operational nor staging impact of threats as it was pre-calculated based on a modelled graph of the system. The graph is modelled based on the components' network and application level connections provided as input for the risk assessment. The observed average time for conducting FMECA-ATT&CK was 3 h per expert to provide a comprehensive output. This highlights the utility of FMECA-ATT&CK in achieving its original objective.

However, several aspects were observed when comparing the results obtained through Bow-Tie and FMECA-ATT&CK. One aspect related to the experts' ability to contextualize the system is unmatched in the current form of FMECA-ATT&CK. For instance, in the APS use case, the RCC is expected to be hosted at a university facility. This information is not encoded in the system model. However, it was communicated during the initial session introducing the use case. Although valid accounts are an identified risk by FMECA-ATT&CK, the contextual information that these accounts can be stolen from students is not yet encoded in FMECA-ATT&CK. This affects the communication of the identified risks. Additionally, when discussing possible mitigation methods during the execution of Bow-Tie in the DS use case, the experts suggested future directions that are relevant but not yet implemented, such as zero trust and resilient design. Such strategic directions cannot be made by the FMECA-ATT&CK approach. This sheds additional light on the component level in which FMECA-ATT&CK operates.

7.1. Limitations in the Evaluation

We acknowledge the following limitations in the evaluation process:

- The results received from the fourth group might include bias due to the bandwagon effect. This was an unavoidable effect in order to accommodate the participants' time limitations. Efforts to reduce the bias were taken in the form of seeking individual confirmation of the results.
- The FMECA-ATT&CK approach for calculating threat likelihood is based on the calculated CVSS metrics for the techniques in the different ATT&CK matrices which are system-independent and pre-estimated and discussed in previous work [12]. The experts were offered a chance to provide their own estimation but due to time limitations, they were unable to do so. Therefore, we resorted to utilizing the pre-estimated data which is subject to bias.
- We are not claiming that FMECA-ATT&CK is straightforwardly applicable in application domains of CPSs other than maritime and energy. This would require extending the evaluation to include additional and diverse use cases.

7.2. Future Work

In summary, based on the results from the evaluation process, the identified future work to improve FMECA-ATT&CK are listed below:

- The scope of considered failure modes focuses on adversarial threats. Considerations of non-adversarial threats, such as human errors, could be useful as a future direction.
- Additional guidelines and supporting methods are needed to estimate the impact of certain failures. Particularly, the estimation of safety and financial impacts.
- The current asset categorization does limit the scope of relevant use cases. Categorizing some components according to the existing asset categorization criteria was found to be challenging. This suggests the proposition of domain-specific categorization. Consequently, the approach requires additional adaptations to accommodate the change of scope. This can include domain-specific threats, failure modes, and risk controls.
- FMECA-ATT&CK is suitable for tier 3 activities according to NIST risk management tiers which address risk from the perspectives of system components [31]. The conducted risk assessment process using Bow-Tie yielded some risk mitigation measures that are at higher tiers, such as a business continuity plan (BCP). The consideration of such mitigation measures requires additional tasks to be conducted after FMECA-ATT&CK which focus on multi-tier risk management rather than tier 3 risk assessment. In this direction, the expansion of the list of supporting controls will be considered in the future.
- The utilization of additional use cases and different application domains for the application of FMECA-ATT&CK will expand its applicability.
- Investigating the efficiency of integrating FMECA-ATT&CK for cyber risk management in real decision-making units (DMUs) would be an interesting direction. For that,

Wang et al. [32] proposed the utilization of data envelopment analysis (DEA) to measure the efficiency of cybersecurity DMUs. This approach would provide quantitative measurements for the reduced cost which FMECA-ATT&CK is hypothesized to achieve as a consequence to reduce the need for expert judgment.

8. Conclusions

There is increased interest in cyber-physical systems (CPSs) as their application has been observed in various domains such as energy, manufacturing, and maritime. The cybersecurity aspect of such systems has been the focus of many in academia and industry. In order to improve the risk management capabilities, a number of approaches and methods have been proposed to assess the cyber risks of CPSs. However, there is a lack of dedicated work in the literature that addresses the evaluation of proposed risk assessment approaches. Our evaluation approach in this paper can be useful to evaluate other risk assessment processes. We proposed a set of characteristics to evaluate risk assessment processes and multi-staged execution procedures to measure the process according to a group of characteristics: applicability, feasibility, usability, adaptability, scalability, accuracy, and comprehensiveness. At the same time, reducing the effect of bias introduced by the reliance on experts' judgment was pursued. Recently, a new FMECA-ATT&CK approach has been proposed to address the issue of increased reliance on expert judgment. The approach is based on the failure mode, effects and criticality analysis (FMECA) risk assessment process, enriched with the semantics and encoded knowledge in the ATT&CK framework. FMECA-ATT&CK was subjected to empirical evaluation by applying it to different use cases from different application domains by several groups of experts with various expertise and backgrounds. To provide a comparison basis, Bow-Tie was used as an additional common risk assessment process.

When comparing FMECA-ATT&CK with Bow-Tie for risk assessment, it was found that FMECA-ATT&CK is capable of identifying similar risks, consequences, and risk controls for the same use cases although the assessment was conducted by different groups of experts without any communication between them. This finding highlights the accuracy of the results obtained through the application of FMECA-ATT&CK. Additionally, the comprehensiveness, adaptability, feasibility, and usability of the approach were measured by experts through a questionnaire and were found to be excellent. On the other hand, the scalability was restricted to systems with a moderate number of components. Furthermore, the applicability of the approach was demonstrated through its application in assessing the risks for two CPS use cases in two different application domains, providing logically sound results. In summary, the overall results are positive and suggest that FMECA-ATT&CK is a viable option for design- and component-level cyber risk assessment for CPSs.

However, several areas for improvement have been identified based on experts' input. This includes asset categorization, identification of relevant failure modes, impact estimation, lack of human element, and the scope of the suggested controls. All of these have been discussed in the paper and rendered as suggested directions for future work.

Author Contributions: Conceptualization, A.A. and V.G.; methodology, A.A. and V.G.; software, A.A.; validation, A.A. and V.G.; formal analysis, A.A.; investigation, A.A.; resources, A.A. and V.G.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A. and V.G.; visualization, A.A.; supervision, V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the NTNU Digital transformation project Autoferry.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data prepared for the evaluation, results, and utilized templates are maintained in a public repository <https://github.com/ahmed-amro/FMECA-ATT-CK-Evaluation> (accessed on 28 February 2023).

Acknowledgments: The authors would like to express their gratitude to the participating experts, namely, Aida Akbarzadeh, Andre Jung Waltoft-Olsen, Arne Roar Nygård, Erlend Erstad, Filip Holik, Georgios Kavallieratos, Marie Haugli-Sandvik, Mohamed Abomhara, and Vijayan Manogara. The experts’ time and efforts invested in this work as well as their valuable comments are highly appreciated and will contribute to advancing the research in the field.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Detailed FMECA-ATT&CK Description

FMECA-ATT&CK approach relies on a group of tables to collect and process the different aspects of risks, namely, threat identification, likelihood estimation, impact estimation, and detectability estimation. Table A1 summarizes the steps of FMECA-ATT&CK, the relevant tables to be filled, the data description, expected values, and the data sources. This highlights areas where experts’ judgments were spared. After conducting the steps in the table are completed, a risk priority number (RPN) calculation and mitigation identification (RPNMI) algorithm is executed to generate the results. The algorithm is described in Algorithm A1.

Table A1. Detailed description of the FMECA-ATT&CK steps, tables, data types, and knowledge sources.

Step	Table	Column	Data Description	Data Values	Data Source
Step 1: Specify Components	Component Description Table (CDT)	Class	Relevant ATT&CK Matrices	Enterprise, ICS, Mobile, Combination	Experts
		Comp Name	Component Name		Architecture Model
		Type	Component ICS Categorization	Control Server, Data Historian, Engineering Workstation, Field Controller/RTU/PLC/IED, HMI, I/O Server, SIS/Protection Relay, Sensor	Experts choice based on ATT&CK categorization
		Platform	Component IT Platform	Windows, Linux, Network, macOS, Cloud, Containers	Architecture Model
		Technology	Component Technology	App-Based or Other	
		Additions	Component Additions	Radio, GPS, Cell, Wi-Fi, Video, etc.	
Step 2: Identify Failure Modes	-	-	Relevant Failure Modes	All ATT&CK Tactics (16)	Experts choice based on ATT&CK Tactics
Step 3: Identify Controls	Failure-Mitigation Table (FMT)	Matrix	ATT&CK Matrix	Enterprise, ICS, Mobile	
		Technique	ATT&CK Technique	All ATT&CK Techniques (>700)	ATT&CK
		Mitigation	ATT&CK Mitigation	All ATT&CK Mitigations (>70)	
	Component-Mitigation Table (CMT)	Efficiency	Mitigation Efficiency	(0.0–1.0)	Experts
		Mitigation	ATT&CK Mitigation	All ATT&CK Mitigations (>70)	ATT&CK
		Component 1			Architecture Model
	Component 2				
	Component Name	(0: not covered or 1: covered)	Architecture Model	
	Component N				
Step 4: Estimate the Impact of the Consequences of Failure Modes	Failure-Mode-Consequences Table (FMCT)	Matrix	ATT&CK Matrix	Enterprise, ICS, Mobile	
		Tactic	ATT&CK Tactics and Impact Techniques	All Tactics and Impact Techniques (>90)	ATT&CK
		Operational	Wight of Operational Consequence		
		Safety	Wight of Safety Consequence		
		Information	Wight of Information Consequence	(0.00–infinity)	Experts
		Financial	Wight of Financial Consequence		
		Staging	Wight of Staging Consequence		

Table A1. Cont.

Step	Table	Column	Data Description	Data Values	Data Source			
	The Failure-Mode-Metric Table (FMMT)	Matrix	ATT&CK Matrix	Enterprise, ICS, Mobile	ATT&CK			
		Tactic	ATT&CK Tactics and Impact Techniques	All Tactics and Impact Techniques (>90)				
		Operational	Operational Metric to be used	Overall Operational Impact (OOI), Impact to Control Functions (I2CF), Impact to Monitoring Functions (I2MF)				
		Safety	Safety Metric to be used	Safety Criticality (SC)				
		Information	Information Metric to be used	Location Information Criticality (LIC), Information Criticality (IC), Intellectual Property Criticality (IPC)		Process Defined		
		Financial	Financial Metric to be used	Financial Criticality (FC), Occurring Financial Criticality (FC2)				
		Staging	Staging Metric to be used	Out-Degree Centrality (ODC), Overall Component Criticality (OCC)				
	Component-Criticality-Scoring Table (CCST)	Comp Name	Component Name	(0.0–1.0)	Experts			
		OOI	OOI score of component					
		I2CF	I2CF score of component					
		I2MF	I2MF score of component					
		SC	SC score of component					
		LIC	LIC score of component					
		IC	IC score of component					
IPC		IPC score of component						
FC		FC score of component						
FC2		FC2 score of component						
	Techniques-Description Table (TDT)	Matrix	ATT&CK Matrix	Enterprise, ICS, Mobile	ATT&CK			
		Technique	ATT&CK Technique	All ATT&CK Techniques (>700)				
		Tactic	ATT&CK Tactics	All ATT&CK Tactics (16)				
		Platform	Technique IT Platform	Windows, Linux, Network, macOS, Cloud, Containers				
		Type	Technique ICS Assets	Control Server, Data Historian, Engineering Workstation, Field Controller/RTU/PLC/IED, HMI, I/O Server, SIS/Protection Relay, Sensor				
		Technology	Technique Technology	App-Based or Other		Experts		
		Additions	Technique Additions	Radio, GPS, Cell, Wi-Fi, Video, etc.				
		Step 6: Estimate the Likelihood of Failure Mechanisms	Techniques-Description Table (TDT)	CVSS		Technique Exploitability Score based on CVSS	(0.00–3.89)	ATT&CK-based heuristics and Experts
		Step 7: Evaluate the Risks	-	-		Risk Rating Criteria such as thresholds	e.g., Risk <3 = Low	Experts
Step 8: Propose Risk Reduction Measures	-	-	Suggested mitigation methods for each technique	All ATT&CK Mitigations (>70)	ATT&CK			

Algorithm A1 Risk Priority Number (RPN) Calculation and mitigation identification (RPNMI) (adapted from [12]). Check Table A1 for acronyms

```

1: procedure RPNMI(TDT, CDT, FMCT, CCST, FMMT, FMT, CMT)
2:   for each component in CDT do
3:     AttackList IdentifyRelevantAttacksByMatchingAttributes(CDT, TDT)
4:     for each attack in AttackList do
5:       Likelihood CalculateAttackLikelihood(CVSSinTDT)
6:       Impact CalculateAttackImpact(RelevantConsequencesinFMCTandmetricsinFMMT
andcomponentscoresinCCST)
7:       Detectability CalculateAttackDetectability(MaxEfficiencyamongrelevantMitigationsinFMTandCMT)
8:       RPN Likelihood × Impact × Detectability
9:       MitigationList GetAttackMitigation(FMT)
10:     end for
11:   end for
12:   return AttackLists, RPNs and MitigationLists
13: end procedure

```

Appendix B. Consolidation Process

The consolidation was utilized as a means to implement voting on conflicting decisions in the assessment. Voting applies brainstorming with the nominal group technique. The following protocol was followed for consolidating the results. If a majority is identified for a decision point, the majority decision will be directly used as the input for the assess-

ment. Otherwise, if only a single response is found for a decision point, the response is directly used as the input for the assessment. Conversely, if a response agrees with other non-matching responses, that response is considered inclusive and is used as the input for the assessment. For instance, if the component classification is IT, OT, or IT/OT, then IT/OT is considered inclusive of the other responses. Otherwise, the average is calculated for decisions including numerical values. The conflicting decision points were moved for discussion in stage 3 in the groups. Moreover, an additional step is conducted to rectify any implementation errors. For instance, expert input was considered incorrect under the scope and semantics of the conducted process. For instance, some experts categorized certain components based on their own definition rather than the definition proposed in the process. Additionally, the ratio of consensus is tracked to measure the assessment quality; under the assumption that when a consensus is reached, the input quality for the assessment is higher than in the case of no consensus.

Appendix C. Questionnaire Details

The experts’ feedback was queried through a questionnaire to evaluate FMECA-ATT&CK based on the chosen characteristics. Table A2 depicts the questions sent to the experts and the answer guide.

Table A2. Questions and choices used for expert feedback.

	Question	Choice	Meaning
1	How applicable is the approach for application in different CPS use cases?	1	Very limited applicable use cases
		2	Only few number of applicable use cases
		3	Several applicable use cases
		4	Many applicable use cases
		5	So many applicable use cases
2	How feasible was the implementation of the different steps? Note: This is related to the a pproach itself and not the current mode of execution as delivered through the excel sheet	1	The entire process is not feasible for implementation
		2	Some steps are not feasible for implementation
		3	The process is feasible but require some adaptation for implementation
		4	The process is feasible and can be implemented in its current form
3	How reasonable were the results?	1	The results did not make sense at all
		2	Some of the results did not make sense while others did
		3	The results do make sense
4	How difficult it is to integrate additional aspects? (asset categories, threats, mitigation measures, impact elements, etc.)	1	It would be extremely difficult to integrate additional aspects
		2	It would require a lot of modifications to integrate additional aspects
		3	Integrating additional aspects is possible with minor modifications
5	How comprehensive is the approach in its inclusion of elements required for sufficient cyber risk assessment processes?	1	The approach scope is very limited
		2	The approach scope is limited
		3	The approach scope is sufficient, but many elements should be added
		4	The approach scope is comprehensive; but some elements can be added
		5	The approach scope is very comprehensive
6	How would it perform in large and complex networks or Systems of Systems (SoS)?	1	Suitable and efficient only for small SoS
		2	Suitable and efficient for moderate SoS
		3	Suitable and efficient for large SoS
		4	Suitable and efficient for very large SoS

Table A2. Cont.

	Question	Choice	Meaning
7	How easy was it to follow with limited training/ Consultation? Note: this is related to the current mode of execution as delivered through the excel sheet	1	I could not execute the assessment with the amount of training I received.
		2	I could only execute some steps of the assessment due to ambiguous tasks.
		3	I executed all the required steps but could not finish some of the tasks due to ambiguity
		4	I executed all the required steps and finished all the tasks
8	Would you like to elaborate on the applicability, feasibility, accuracy, adaptability, scalability, and required training to apply the approach?	Open Ended	
9	How many hours in total did the process took to be completed, approximately. (Filling the Excel sheet)	1	an hour or less
		2	around 2 h
		3	around 3 h
		4	around 4 h
		5	5 h or more

References

- Duru, O. The Future Shipping Company: Autonomous Shipping Fleet Operators. Available online: <https://www.maritime-executive.com/editorials/the-future-shipping-company-autonomous-shipping-fleet-operators> (accessed on 28 February 2023).
- NTNU Autoferry. *Autoferry—Autonomous All-Electric Passenger Ferries for Urban Water Transport*; Norwegian University of Science and Technology: Trondheim, Norway, 2018.
- Amro, A.; Gkioulos, V.; Katsikas, S. Connect and Protect: Requirements for Maritime Autonomous Surface Ship in Urban Passenger Transportation. In *Computer Security*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 69–85.
- Johnson, B. Maritime Cyber Incidents Increased at Least 68 Percent in 2021, Coast Guard Reports. Available online: <https://www.hstoday.us/featured/maritime-cyber-incidents-increased-at-least-68-percent-in-2021-coast-guard-reports/> (accessed on 28 February 2023).
- offshore energy.biz. COSCO Shipping Lines Falls Victim to Cyber Attack, 2018. Available online: <https://www.offshore-energy.biz/cosco-shipping-lines-falls-victim-to-cyber-attack/> (accessed on 28 February 2023).
- Norman, J. Iranian Hackers Suspected in Cyber Breach and Extortion Attempt on Navy Shipbuilder Austal, 2018. Available online: <https://www.abc.net.au/news/2018-11-13/iranian-hackers-suspected-in-austal-cyber-breach/10489310> (accessed on 28 February 2023).
- Greenberg, A. *The Untold Story of NotPetya, the Most Devastating Cyberattack in History*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2018.
- Seals, T. Researcher: Not Hard for a Hacker to Capsize a Ship at Sea. Available online: <https://threatpost.com/hacker-capsize-ship-sea/142077/> (accessed on 28 February 2023).
- Hambling, D. Ships Fooled in GPS Spoofing Attack Suggest Russian Cyberweapon, 2017. Available online: <https://www.newscientist.com/article/2143499-ships-fooled-in-gps-spoofing-attack-suggest-russian-cyberweapon/> (accessed on 28 February 2023).
- IEC 31010; Risk Management—Risk Assessment Techniques. ISO: Geneva, Switzerland, 2019.
- ISO/IEC 27005:2018; Information Technology. Security Techniques. Information Security Risk Management. ISO: Geneva, Switzerland, 2018.
- Amro, A.; Gkioulos, V.; Katsikas, S. Assessing Cyber Risk in Cyber-Physical Systems Using the ATT&CK Framework. *ACM Trans. Priv. Secur.* **2022**, *26*, 1–33. [CrossRef]
- 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.
- IEC. *Analysis Techniques for System Reliability-Procedure for Failure Mode and Effects Analysis (FMEA)*; IEC: Geneva, Switzerland, 2018.
- 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; MITRE: Bedford, MA, USA, 2018.
- Shostack, A. *Threat Modeling: Designing for Security*; Wiley Publishing: Hoboken, NJ, USA, 2014.
- Mihai, I.C.; Pruna, S.; Barbu, I.D. Cyber kill chain analysis. *Int. J. Info. Sec. Cybercrime* **2014**, *3*, 37. [CrossRef]
- ENISA. ENISA Threat Landscape 2021. Available online: <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021> (accessed on 28 February 2023).
- Alford, R.; Lawrence, D.; Kouremetis, M. *CALDERA: A Red-Blue Cyber Operations Automation Platform*; MITRE: Bedford, MA, USA, 2022.
- DNV GL. *Cyber Security Resilience Management for Ships and Mobile Offshore Units in Operation*; Technical Report, DNVGL-RP-0496; DNV GL: Oslo, Norway, 2016.

21. Amro, A.; Gkioulos, V.; Katsikas, S. Communication architecture for autonomous passenger ship. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2021**. [[CrossRef](#)]
22. Khodabakhsh, A.; Yayilgan, S.Y.; Abomhara, M.; Istad, M.; Hurzuk, N. Cyber-risk identification for a digital substation. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Virtual, 25–28 August 2020; pp. 1–7.
23. Tam, K.; Jones, K. Factors affecting cyber risk in maritime. In Proceedings of the 2019 International Conference on Cyber Situational Awareness, Data Analytics Furthermore, Assessment (Cyber SA), Oxford, UK, 3–4 June 2019; pp. 1–8.
24. Abkowitz, M.; Camp, J. An application of enterprise risk management in the marine transportation industry. *WIT Trans. Built Environ.* **2011**, *119*, 221–232.
25. Grigoriadis, C.; Papastergiou, S.; Kotzanikolaou, P.; Douligeris, C.; Dionysiou, A.; Elias, A.; Bernsmed, K.; Meland, P.H.; Kamm, L. Integrating and Validating Maritime Transport Security Services: Initial results from the CS4EU demonstrator. In Proceedings of the 2021 Thirteenth International Conference on Contemporary Computing (IC3-2021), Noida, India, 5–7 August 2021; pp. 371–377.
26. Johansen, T.; Utne, I.B. Risk Analysis of Autonomous Ships. In Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference (ESREL2020 PSAM15), Venice, Italy, 1–5 November 2020.
27. Tam, K.; Jones, K. Cyber-risk assessment for autonomous ships. In Proceedings of the 2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), Scotland, UK, 11–12 June 2018; pp. 1–8.
28. Kavallieratos, G.; Katsikas, S.; Gkioulos, V. Cyber-attacks against the autonomous ship. In *Computer Security*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 20–36.
29. Pillitteri, V.Y.; Pierre, J.; Stine, K.; Scholl, M.; Stine, K. *Assessing Security and Privacy Controls in Information Systems and Organizations*; NIST: Gaithersburg, MD, USA, 2022.
30. Boddy, C. The nominal group technique: An aid to brainstorming ideas in research. *Qual. Mark. Res. Int. J.* **2012**, *15*, 6–18. [[CrossRef](#)]
31. Boyens, J.; Paulsen, C.; Moorthy, R.; Bartol, N.; Shankles, S. *NIST Special Publication 800-161: Supply Chain Risk Management Practices for Federal In-Formation Systems and Organizations*; NIST: Gaithersburg, MD, USA, 2015.
32. Wang, C.N.; Yang, F.C.; Vo, N.T.; Nguyen, V.T.T. Wireless communications for data security: Efficiency assessment of cybersecurity industry—A promising application for UAVs. *Drones* **2022**, *6*, 363. [[CrossRef](#)]



Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Paper X

A. Oruc, A. Amro and V. Gkioulos, 'Assessing cyber risks of an ins using the mitre att&ck framework,' *Sensors*, vol. 22, no. 22, p. 8745, 2022

Article

Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework

Aybars Oruc , Ahmed Amro  and Vasileios Gkioulos

Department of Information Security and Communication Technology, Norwegian University of Science and Technology, 2815 Gjøvik, Norway

* Correspondence: aybars.oruc@ntnu.no

Abstract: Shipping performed by contemporary vessels is the backbone of global trade. Modern vessels are equipped with many computerized systems to enhance safety and operational efficiency. One such system developed is the integrated navigation system (INS), which combines information and functions for the bridge team onboard. An INS comprises many marine components involving cyber threats and vulnerabilities. This study aims to assess the cyber risks of such components. To this end, a methodology considering the MITRE ATT&CK framework, which provides adversarial tactics, techniques, and mitigation measures, was applied by modifying for cyber risks at sea. We assessed cyber risks of 25 components on the bridge by implementing the extended methodology in this study. As a result of the assessment, we found 1850 risks. We classified our results as 1805 low, 32 medium, 9 high, and 4 critical levels for 22 components. Three components did not include any cyber risks. Scientists, ship operators, and product developers could use the findings to protect navigation systems onboard from potential cyber threats and vulnerabilities.

Keywords: maritime cyber security; risk assessment; INS; integrated navigation system; MITRE ATT&CK framework



Citation: Oruc, A.; Amro, A.;

Gkioulos, V. Assessing Cyber Risks of an INS Using the MITRE ATT&CK Framework. *Sensors* **2022**, *22*, 8745. <https://doi.org/10.3390/s22228745>

Academic Editor: Keshav Dahal

Received: 17 September 2022

Accepted: 7 November 2022

Published: 12 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over 80% of goods in international trade are carried by ships [1]. One of the most essential elements of maritime transportation is explicitly ships. In 2020, the worldwide merchant fleet grew by 3% and reached 99,800 ships of 100 gross tons and above [1]. Contemporary ships are equipped with computerized systems for different purposes, such as navigation, communication, propulsion, and cargo handling. The safety and operational efficiency of vessels are improved because of such systems. However, these systems are accompanied by growing cyber security concerns in the maritime industry because of experiencing cyber incidents and revealing research results.

The International Maritime Organization (IMO) is the responsible agency in the United Nations for the safety and security of shipping and the prevention of environmental pollution by ships [2]. Maritime cyber risk is defined by the IMO as “a measure of the extent to which a technology asset is threatened by a potential circumstance or event, which may result in shipping-related operational, safety or security failures as a consequence of information or systems being corrupted, lost or compromised” [3]. In 2017, the IMO issued a resolution to prevent maritime cyber risks [4]. As per the resolution in force, cyber risks must be assessed by ship operators and addressed in their approved Safety Management Systems (SMS). Moreover, they should make reference to the Ship Security Plan (SSP) as per the International Ship and Port Facility Security (ISPS) Code [5,6]. This requirement has been verified in the Document of Compliance (DOC) audits of ship operators since 2 January 2021.

This paper reveals the significance of cyber risks onboard vessels. We contributed to the literature by extending a methodology using the MITRE ATT&CK framework to assess the cyber risks of systems onboard ships. Moreover, the method was implemented to specifically assess the cyber risks of an INS in this study. A total of 1850 risks were classified

as 1805 low, 32 medium, 9 high, and 4 critical levels. Given that no marine casualty (e.g., collision, explosion, injury, and oil spill) caused by cyber attacks was found in the literature, safety and environmental impacts of cyber risks are outside of the scope of this study.

We organised the remainder of the paper as follows. Section 2 gives a review of the related literature. In Section 3, the methodology is discussed and implemented for the cyber risks of an INS. Section 4 offers a summary and suggests additional research topics for further investigation. Consequently, in Appendix A, cyber risks of medium, high, and critical levels are listed.

2. Background

2.1. INS Concept

The IMO defines an INS as “A system in which the information from two or more navigation aids is combined in a symbiotic manner to provide an output that is superior to any one of the component aids” [7]. The INS aims to improve safe navigation by combining and integrating information and functions for the Officer of the Watch (OOW) in planning, monitoring, and controlling ship navigation [8]. An INS constitutes six navigational tasks as mandatory and optional, as follows:

- Route Monitoring: “The navigational task of continuous surveillance of own ships position in relation to the pre-planned route and the waters” [9].
- Route Planning: The task that provides procedures for voyage planning, route planning functions and data for the Electronic Chart Display and Information System (ECDIS), administering the route plan, checking route plan against hazards, manoeuvring limitation (e.g., rate of turn (ROT)), drafting and refining the route plan against meteorological information [8].
- Collision Avoidance: “The navigational task of detecting and plotting other ships and objects to avoid collisions” [9].
- Navigation Control Data: “Task that provides information for the manual and automatic control of the ship’s movement on a task station” [9].
- Navigational Status and Data Display: The task that displays data for the manual and automatic control of the ship’s primary movement [8].
- Alert Management: “Concept for the harmonized regulation of the monitoring, handling, distribution and presentation of alerts on the bridge” [9].

2.2. MITRE ATT&CK Framework

The ATT&CK framework (which stands for Adversarial Tactics, Techniques, and Common Knowledge) has been developed by MITRE since 2013 [10]. It is a globally accessible database including attack tactics, techniques, and mitigation measures for the matrices of enterprise, mobile, and industrial control systems (ICS). The *Enterprise Matrix* covers offensive information (i.e., tactics and techniques) for information technology (IT) networks and cloud services, such as operating systems (i.e., Windows, Linux, and macOS), network components, Office 365, and Google Workspace [11,12]. The *Mobile Matrix* includes offensive knowledge for iOS and Android platforms [13]. The *ICS Matrix* provides offensive information for the ICS [14]. The *Tactics* represents the attack objective, such as initial access, credential access, and lateral movement [15]. *Techniques* expresses methods to achieve an attack objective [16]. The ATT&CK framework also provides mitigation measures to avoid a technique from being successfully executed [17]. Moreover, malware and tools which can be used for malicious purposes are described under the name of *Software* [18]. Another important dimension of ATT&CK is to offer cyber-threat intelligence. *Groups* refers to adversary actor and give techniques implemented and software used by them for an attack in the past [19]. *Data Sources* provides information about various subjects and notions [20].

2.3. Literature Review

In the literature, papers implementing various methods have assessed the cyber risks of autonomous ships and conventional ships. Kavallieratos and Katsikas [21] implemented

STRIDE and DREAD methods for the cyber risk assessment of several systems on the autonomous ship, such as a collision avoidance system, Radio Detecting Additionally Ranging (RADAR), closed-circuit television (CCTV), Voyage Data Recorder (VDR), cargo management system, and autopilot. Kavallieratos et al. in [22] also implemented STRIDE for an Automatic Identification System (AIS), engine automation system, bridge automation system, shore control center, engine efficiency system, navigation systems, autonomous ship controller, and so on. Tusher et al. [23] have a cyber risk assessment work for autonomous ships, as well. In their study, the Bayesian best–worst method was implemented, and the authors revealed navigation systems as the most vulnerable element in the context of future autonomous shipping operations. Shang et al. [24] implemented the combination of fuzzy set theory and the Attack Tree method to assess cyber risks of the control system for a gas turbine onboard ship. Oruc [25] also combined fuzzy set theory with another risk assessment method, Fine–Kinney. In the study, 31 cyber risks in the bridge, engine room, and cargo control room onboard a tanker were assessed. Moreover, the efficiency of proposed mitigation measures is shown by implementing the method a second time after taking precautions. Kessler et al. [26] focused on 16 different cyber risks of an AIS. Their study reveals that the disruption of an individual AIS message is more crucial than being unusable of an entire AIS. Svilicic et al. [27] also performed a risk assessment for a specific component. The authors made a cyber risk assessment for the ECDIS on a training vessel by using a vulnerability scanner, named Nessus Professional, and interviewing the ship crew. Several cyber threats were determined regarding the operating system, procedures, awareness, and so on. iTrust published a guideline [28] to uncover cyber risks of operational technology (OT) systems on conventional vessels, including navigation, machinery, communication, and cargo management systems. The traditional risk calculation formula ($\text{risk} = \text{severity} \times \text{likelihood}$) was implemented to assess cyber risks. The study also proposes actionable mitigation measures. You et al. [29] focused on risk assessment methods in other fields and discussed their adaptation to the maritime industry. According to the study, Attack Tree, simulations, and models can be implemented for the cyber risk assessment of marine systems.

Novel methods other than well-established methods are also available in the literature for cyber risks onboard ships. Tam and Jones [30] developed a model-based framework for maritime cyber-risk assessment, entitled Maritime Cyber-Risk Assessment (MaCRA). The authors also implemented the method to assess the cyber risks of three autonomous ship projects in a separate paper [31]. Bolbot et al. [32] proposed a novel method, named Cyber-Risk Assessment for Marine Systems (CYRA-MS), by considering the Preliminary Hazard Analysis (CPHA) method to assess cyber risks of ship systems. The authors implemented the method on navigation and propulsion control systems of a fully autonomous inland ship. Meland et al. [33] offered an alternative method for cyber risk assessment. The likelihood of a threat in new design systems is a challenge. The authors propose the threat likelihood approach to support security decision-making for new design systems in particular. Their method is the combination of current concepts, techniques, expert judgements, and domain-specific information.

The ISO 31000 is the root standard and comprises principles, a framework, and a process for risk management [34]. The standard offers a common approach for any size of organization to manage any kind of risk, including the decision-making process [34]. The ISO/TR 31004 explains the effective implementation of ISO 31000 in detail [35]. The IEC 31010 clarifies the selection and application of risk assessment techniques in different situations [36]. The ISO 27000 is another root standard and gives a general approach to information security management systems [37]. The IEC 63154 identifies requirements, test methods, and required test results against cyber incidents for shipborne navigational aids, radio, and navigational equipment [38]. The Formal Safety Assessment (FSA) [39] published by the IMO is a systematic methodology to enhance safety in the maritime industry, including the protection of human life, health, the marine environment and property by using risk analysis. The circular describes the notions, methods, and control

measures for a risk assessment. The FSA gives an overall knowledge for a risk assessment in the maritime industry but is not designed specifically for cyber risk assessment.

As mentioned before, IMO issued a regulation for the assessment of cyber risks [4]. After this regulation particularly, several guidelines were published by class societies and other IMO-recognized organizations to support the maritime industry against cyber risks [40–42]. The Guidelines on Maritime Cyber Risk Management [42] jointly developed by several industry associations are officially recommended by the IMO [3,43]. The guidelines provide detailed explanations in different dimensions of cyber security, such as cyber threats, risk management, technical and procedural protection measures, and contingency plans, including response and recovery procedures for the maritime industry.

Various comparisons among high-level models, such as the ATT&CK framework, Cyber Kill Chain, OWASP top 10, STRIDE, and the Diamond Model exist [44–47]. Even though such models are effective in understanding processes and adversary goals, models other than the ATT&CK framework are not useful for explaining the impact of an action to another [48]. Furthermore, the ATT&CK framework depicts correlations of actions with data sources, defenses, configurations, and other countermeasures used for the security of a platform [48].

Even though ATT&CK framework is not a risk assessment method, papers using ATT&CK framework are available for different purposes in other domains, such as risk assessment and risk identification [49,50]. In our study, we reveal that the ATT&CK framework can be used for cyber risk assessment of ship systems as well. Moreover, in the literature, any risk assessment focusing on an INS was not found. Papers in the literature typically assessed the cyber risks of a few components. In our study, we assessed cyber risks for 25 marine components.

3. The Extended Methodology and Implementation

Our methodology was derived from the [51] to specialize cyber risks of vessels. The method is based on a Failure Mode Effects and Criticality Analysis (FMECA) and the MITRE ATT&CK framework. The core advantage of the original method is to reduce the need for expert judgement. Thus, the impact of bias in a risk assessment reduces. Moreover, the method is comprehensive and semi-automated. Mitigation measures for cyber risks are included. Our adapted methodology for marine systems is performed as follows:

1. Components are specified and classified.
2. Functions of components and data flow among components are identified.
3. The failure modes for components are determined.
4. Failure modes are mapped with consequences and impacts.
5. Estimation criteria for criticalities are identified.
6. Detection methods and existing controls are identified.
7. The impact scores of components are identified.
8. Risk scores are calculated and risk levels are identified.

3.1. Component Specification and Classification

Our methodology starts with the specification and classification of marine components. We implemented our risk assessment methodology on an INS in this study. An INS consists of various marine components. We found 25 components for an INS in our previous study [52]. Such components were classified by IMO and method definitions, respectively. The method definitions for the classification of components are given in Table 1 (e.g., IT, OT, Wireless). Classification by the method definitions is required for the risk assessment process. However, the classification by the IMO definitions is given to provide an additional contribution and to understand the differences between classifications in Table 2.

Table 1. Component classification by method [51].

Classification	Description
IT	Components that are hosted on a traditional IT system such as multipurpose computers or network devices.
OT	Components that are involved in monitoring and controlling functions.
Wireless	Components that are connected to a mobile network or communicate with an external infrastructure, such as Aids to Navigation, to acquire location-related information in the maritime domain.
IT/OT	Dual-homed components that are hosted on a traditional IT system and are involved in monitoring and controlling functions.
IT/OT/Wireless	Components that are classified as IT/OT and are connected to a mobile network or communicate with an external infrastructure.

According to the IMO, components are divided into two groups, such as information technology (IT) and operational technology (OT), and the difference between IT and OT systems is defined as “*Information technology systems may be thought of as focusing on the use of data as information*”, and “*Operational technology systems may be thought of as focusing on the use of data to control or monitor physical processes*” [3]. Moreover, the IMO-recommended document, Guidelines on Maritime Cyber Risk Management, expresses that “*IT covers the spectrum of technologies for data storing and processing, including software, hardware, and communication technologies*”, and “*OT includes hardware and software that directly monitors/controls physical devices and processes, typically on board.*” [42]. Various maritime cyber security-related guidelines were reviewed to find a reliable classification list for marine components by such definitions. However, some marine components, such as ECDIS, RADAR, gyro compass, AIS, global positioning system (GPS), and Bridge Navigational Watch Alarm System (BN-WAS) are classified as OT by several organizations [40–42]. A full list for INS components has not been found. We classified INS components considering IMO definitions as shown in Table 2. The table also includes columns for *Type*, *Platform*, and *Technology*. The *Type* of the components, such as sensors, Human–Machine Interface (HMI), control server, and engineering workstation was determined. For switches (e.g., the Rudder pump selector switch), we ignored the *Type*. If a component needs an operating system to run, it was stated in the *Platform*. The *Technology* refers to attached technologies such as Wi-Fi, cellular, and Bluetooth.

Table 2. Components and classification of components.

Component	Classification		Type	Platform	Technology
	IMO	Method			
AIS	OT	IT, OT, Wireless	Sensor		radio, GPS
Anemometer	OT	IT, OT	Sensor		
BNWAS	OT	IT, OT	Sensor		
Central Alert Management HMI	OT	IT, OT	HMI		
Controls for main engine	OT	OT	Control Server		
Controls for main rudder	OT	OT	Control Server		
Controls for thruster	OT	OT	Control Server		
ECDIS	OT	IT, OT	Engineering workstation	OS	
Echo Sounder	OT	IT, OT	Sensor		
GPS	OT	IT, OT, Wireless	Sensor		GPS
Gyro-Compass	OT	IT, OT	Sensor		
Heading Control System (HCS)	OT	IT, OT	Control Server		
Indicators	OT, IT	IT	HMI		
Magnetic Compass	OT	IT, OT	Sensor		
Multi Function Display (MFD)	OT	IT, OT	Engineering workstation	OS	
Navigational Telex (NAVTEX)	OT	IT, OT, Wireless	Sensor		radio
RADAR	OT	IT, OT	Sensor	OS	radio
ROTI	OT	IT, OT	Sensor		
Rudder pump selector switch	OT	OT	N/A		
Sound reception system	OT	IT, OT	Sensor		
Speed and Distance Measuring Equipment (SDME)	OT	IT, OT	Sensor		
Steering mode selector switch	OT	OT	N/A		
Steering position selector switch	OT	OT	N/A		
Track Control System (TCS)	OT	IT, OT	Control Server		
Transmitting Heading Device (THD)	OT	IT, OT	Sensor		

3.2. Functions of Components and Data Flow among Components

In the second step of the method, the functions of the components and data flow among the components are investigated. Such knowledge for an INS was taken from our previous article, as shown in Table 3 [52]. Data flow in the table was identified as per the minimum requirements of the IMO. However, additional connections among the components are allowed.

Table 3. Functions of components and data flow [52].

Component	Function	Data Flow
AIS	identifying ships, assisting in target tracking, assisting in search and rescue operation, information exchange, providing additional information to assist situation awareness	Sends to: RADAR
Anemometer	detecting and indicating wind speed and direction	N/A
BNWAS	monitoring bridge activity, detecting operator disability and then alerting automatically	N/A
Central Alert Management HMI	reporting abnormal situation which requires an attention	Receives from: sensors connected
Controls for main engine	Control buttons or levers of the main engine for different purposes such as rpm, load, emergency stop button, sailing mode selection button, and so on	N/A
Controls for main rudder	commanding the rudder angel, activating the override mode	N/A
Controls for thruster	commanding the thrusters such as starting, stopping, load /stage, etc.	N/A
ECDIS	offering the functions of route planning, route monitoring and positioning for officers in ECDIS instead of paper charts	Receives from: GPS, gyro compass, SDME. If the ships are not equipped with gyro compass, ECDIS receives data from the transmitting heading device
Echo Sounder	measuring the depth of water under the ship, and presenting graphically	N/A
GPS	providing space-based positioning, velocity and time system	Sends to: AIS, RADAR, ECDIS, HCS, TCS, Gyro compass
Gyro-Compass	determining the direction of the ship's head in relation to geographic (true) north	Sends to: AIS, RADAR, ECDIS, HCS, TCS Receives from: GPS
HCS	keeping the vessel in preset heading by using heading information	Receives from: Gyro compass or Transmitting Heading Device. Moreover, GPS or SDME
Indicators	shows data or status information received from sensor	Receives from: Sensors connected.
Magnetic Compass	determining and displaying the ship's heading without any power supply	Sends to: THD
MFD	A display unit presents information from more than a single function of the INS	depends on connected equipment
NAVTEX	receiving and automatically printing or displaying Maritime Safety Information (MSI)	N/A
RADAR	indication, in relation to own ship, of the position of other surface craft, obstructions and hazards, navigation objects and shorelines	Receives from: AIS, GPS, SDME Moreover, Gyro compass or Transmitting Heading Device
ROTI	indicating rates of turn to starboard and to port of the ship to which it is fitted	Sends to: AIS

Table 3. Cont.

Component	Function	Data Flow
Rudder pump selector switch	selection of primary and secondary (emergency) hydraulic or electrohydraulic pumps for rudder direction	N/A
Sound reception system	offers the OOW who can hear and determine the direction of the sound signals of the vessels nearby	N/A
SDME	measuring and indicating speed and distance of the vessel	Sends to: HCS, RADAR, ECDIS, TCS
Steering mode selector switch	selection of steering modes, such as "Auto", "Non-Follow Up", or "Follow Up".	N/A
Steering position selector switch	determining the active steering workstation (i.e., port wing, starboard wing or center)	N/A
TCS	Track control system keeps the vessel on a pre-planned track over ground by using position, heading and speed information of the vessel	Receives from: GPS, SDME, Gyro compass
Transmitting Heading Device	indicating ship's true heading by means of magnetic compass	Receives from: magnetic compass Sends to: AIS, HCS, TCS, ECDIS, RADAR

The ORA is a network tool to analyze, visualize, fuse, and forecast behaviour given network data [53]. Vulnerabilities, model network changes over time, and key players can be identified and formatted reports can be received [54]. Moreover, it consists of tools for optimizing a network's design structure [54]. In our study, the ORA was employed to calculate various centrality metrics, such as authority, betweenness, and in-degree. Then, the dependency graph was drawn, based on Table 3. The dependency graph among the components is illustrated in Figure 1. In this graph, the nodes represent the investigated component in the INS while the edges represent the identified data flow between components. For instance, as stated in Table 3, the GPS component sends positioning information to the AIS component. This dictates the definition of an edge originating from the GPS component to the AIS component. Additionally, the node size highlights the importance of the node in the network, which is inferred from the nodes' centrality measurements.

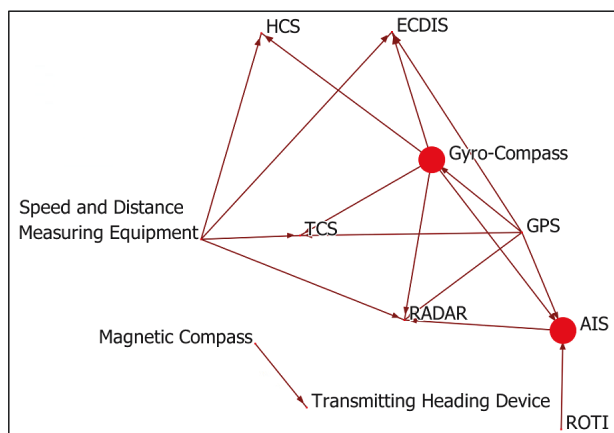


Figure 1. Graph on dependency of INS components.

3.3. Identifying Failure Modes

The literature was reviewed to understand occurred cyber incidents onboard ships and threats and vulnerabilities of the marine components found in research activities. Moreover, the guidelines of products were reviewed to understand potential failures of components. Component damages and installation mistakes were ignored. In this way, potential failures caused by a cyber attack were determined. Then, failure modes were determined. In this study, failure mode refers to *Tactics* [15] in the ATT&CK framework and is given in three categories, such as Mobile, Enterprise, and ICS. Samples of findings are represented in Table A2.

Then, the possible causes of failure modes or attack techniques were identified and their likelihood was estimated. The identification was performed component-by-component by detecting relationships between components and techniques based on matching attributes. The ATT&CK framework provides attributes of relevant asset types and platforms for each technique. This allows for the identification of the relevant techniques for each component in the system based on the system category. For instance, “Alarm Suppression” is an attack technique against several categories of ICS components such as “RTU”; therefore, “Alarm Suppression” technique would be assigned among the threats identified for any system component that can be categorized as an “RTU”. Afterwards, the likelihood of each technique was calculated based on the exploitability score in the Common Vulnerability Scoring System (CVSS). This entails the estimation of the techniques likelihood based on a Bayesian network of four elements, namely, Attack Complexity (*AC*), Privilege Required (*PR*), Attack Vector (*AV*) and User Interaction (*UI*) using Equation (1):

$$\text{Likelihood}_T = 8.22 \times AV \times AC \times PR \times UI \quad (1)$$

(*T* : Technique)

Equation (1) is adapted from the CVSS for calculating the exploitability score to maintain alignment with a widely recognized approach for calculating likelihood [55]. The *AC*, *PR*, *AV*, and *UI* information was system-independent and encoded in a Threat Description Table (TDT), and was adopted for all the list of techniques from the original methodology [51].

3.4. Mapping Failure Modes with Consequences and Impacts

The consequence is an outcome of an accident [39]. In the original method, consequences are identified as operational, safety, information, financial, and staging. The IMO recommends assessing environmental risks in the FSA [39]. Moreover, we investigated several risk assessment matrices used in the maritime industry and noticed that reputation consequence is also assessed by tanker operators, in particular. Because of such reasons, we extended the method with reputation and environmental consequences.

Safety Consequence depicts the potential to cause harm to persons (e.g., crew and passengers). *Operational Consequence* describes potential disruptions, such as errors in the systems during cargo handling. *Financial Consequence* refers to economic losses such as component damages, or commercial losses (e.g., charter party violations). *Information Consequence* explains possible privacy or/and confidentiality violations, such as hosted and processed data in a component. *Staging Consequence* describes the effect of a failure mode which facilitates the staging of future attacks. *Environmental Consequence* describes the potential to cause harm to the environment (e.g., air and water pollution). *Reputation Consequence* describes harm to company prestige and business life.

Operational, Information, and Staging consequences were broken into impacts. Three metrics are available for estimating the impact on operational consequence, namely the Overall Operational Impact (OOI), Impact to the Control Functions (I2CF), and Impact to the Monitoring Functions (I2MF). If a failure mode (e.g., manipulation of control) impacts the control, it is estimated using the I2CF. If a failure mode (e.g., loss of view) impacts

monitoring, it is estimated using the I2MF. Others are estimated using the OOI metric. Staging was estimated using Overall Component Criticality (OCC) and Outbound Degree Centrality (ODC). The failure modes of persistence, defense evasion, and privilege were estimated using the OCC. Others are estimated using the ODC metric. Three types of metrics exist for the information consequence. These are Data Criticality (DC), Intellectual Property Criticality (IPC), and Location Information Criticality (LIC). DC relates to hosted and processed data in a component (e.g., crew information). IPC relates to the hosting of processes with intellectual value. LIC relates to the location information of a component (e.g., position information of a vessel).

Any components in the context of an INS do not process or host personal and confidential data. One feature of an AIS is to transmit location information frequently. When an AIS is equipped mandatorily, it must be always active at anchor and underway unless the master decides to switch it off due to safety and security concerns [56]. However, this decision should be recorded in the logbook with reasons and reported to authorities [56]. Moreover, Long-Range Identification and Tracking (LRIT) onboard also transmit position information [57]. Because of such regulations, the position information of a vessel can not be confidential. Components of an INS are easily found in the market. Furthermore, component standards are identified by the IMO. This is why intellectual property does not exist for an INS. Because of such reasons, an INS is not subject to information consequences. Failure modes were mapped with other consequences and potential impacts for an INS, as illustrated in Table 4.

Table 4. Mapping failure modes, consequences, and impacts.

Matrices	Failure Modes	Consequences						
		Operational	Reputation	Environmental	Safety	Information	Financial	Staging
Mobile	Network Denial of Service	I2MF		EC	SC			
	impact	I2MF		EC	SC			
IT	collection							ODC
	credential access		RC					ODC
	data encrypted for impact	OOI	RC	EC	SC		FC	
	data manipulation	OOI	RC	EC	SC		FC	
	discovery							ODC
	execution	OOI	RC	EC	SC		FC	ODC
	exfiltration							ODC
	firmware corruption	OOI		EC	SC		FC	
	initial access							ODC
	lateral movement							ODC
	system shutdown/reboot	OOI		EC	SC		FC	

Table 4. Cont.

Matrices	Failure Modes	Consequences						
		Operational	Reputation	Environmental	Safety	Information	Financial	Staging
ICS	collection							ODC
	discovery							ODC
	execution	OOI	RC	EC	SC		FC	ODC
	initial access							ODC
	lateral movement							ODC
	loss of availability	OOI	RC	EC	SC		FC	ODC
	loss of control	I2CF	RC	EC	SC		FC	
	loss of safety	OOI	RC	EC	SC		FC	
	loss of view	I2MF	RC	EC	SC		FC	ODC
	manipulation of control	I2CF	RC	EC	SC		FC	
manipulation of view	I2MF	RC	EC	SC		FC	ODC	

SC: Safety criticality, FC: financial criticality, EC: environmental criticality, RC: reputational criticality.

3.5. Identified Estimation Criteria for Criticalities

The estimation criteria were identified for safety, financial, environmental, and reputational criticalities. We proposed estimation criteria for such criticalities. The scores in the estimation criteria tables were identified between 0 and 1 using their impact degrees. Table 5 was used to estimate the impact of a failure mode on the safety consequence. Table 6 was used to forecast financial criticality. The estimation criteria for environmental criticality are depicted in Table 7. Tables 5 and 7 were derived from the *Appendix 4—Initial Ranking of Accident Scenarios* in the FSA published by the IMO [39].

Table 5. Estimation criteria for safety criticality.

Safety Criticality	Description	Score
None	No injury or insufficient data	0
Minor	Single or minor injuries	0.25
Significant	Multiple or severe injuries	0.50
Severe	Single fatality or multiple severe injuries	0.75
Catastrophic	Multiple fatalities	1

Table 6. Estimation criteria for financial criticality.

Financial Criticality	Description (USD)	Score
None	No financial loss or insufficient data	0
Minor	1–10,000	0.25
Significant	10,001–100,000	0.50
Severe	100,001–1,000,000	0.75
Catastrophic	Financial loss > 1,000,000	1

Table 7. Estimation criteria for environmental criticality.

Environ. Criticality	Description	Score
None	No environmental damage or insufficient data	0.00
Minor	Oil spill size < 1 tonne	0.20
Significant	Oil spill size between 1–10 tonnes	0.40
Severe	Oil spill size between 11–100 tonnes	0.60
Catastrophic	Oil spill size between 101–1000 tonnes	0.80
Extreme	Oil spill size > 1000 tonnes	1

Because of cyber incidents, the seaworthiness and cargo worthiness of a ship may be lost or the ship might be delayed to its destination port. In such cases, the master may need to inform charterers or maritime regulators, such as the port state, flag state, and class society. This would explicitly damage the reputation of the ship operator. This is why we identified two criteria for reputation criticality, as shown in Table 8.

Table 8. Estimation criteria for reputational criticality.

Reputation Critical.	Description	Score
None	None	0
Significant	Notification requirement to third parties	1

3.6. Identifying Detection Methods and Existing Controls

Technical and procedural mitigation measures for enterprise [17], mobile [58], and ICS [59] matrices are given in the ATT&CK framework. Over 70 mitigation measures were assessed for each component in the context of an INS. In Table 9, samples of mitigation measures for components are illustrated. The number “1” in the table refers to that the mitigation measure can be implemented for the component. On the other hand, “0” in the table denotes that the mitigation measure cannot be implemented for the component.

This table assists in calculating the detectability of techniques that can be addressed by certain mitigation measures. Detectability is a term utilized in the original methodology [51] that refers to the degree of risk reduction due to the availability of risk mitigation measures. The detectability of a technique when targeting a specific component is calculated based on Equation (2):

$$Detectability_{T,C,M} = Coverage_{M,C} \times Efficiency_{T,M} \quad (2)$$

(T : Technique, C : Component, M : Mitigation measure)

The coverage of a mitigation measure (*M*) for a component (*C*) is referred to in Table 9 while the efficiency of a mitigation measure (*M*) in reducing the risk of a technique (*T*) is estimated for each mitigation measure. In this paper, for simplicity, the efficiency was assumed as 0.5 for all mitigation measures due to the lack of such estimation.

Table 9. Samples for risk-mitigation measures.

Component	Samples for Mitigation Measures														
	Account Use Policies	Active Directory Configuration	Antivirus/Antimalware	Application Developer Guidance	Application Isolation and Sandboxing	Audit	Behavior Prevention on Endpoint	Boot Integrity	Code Signing	Credential Access Protection	Data Backup	Data Loss Prevention	Disable or Remove Feature or Program	Do Not Mitigate	Encrypt Sensitive Information
AIS	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Anemometer	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BNWAS	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Central Alert Management HMI	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Controls for M/E	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Controls for main rudder	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Controls for thruster	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
ECDIS	0	1	1	0	0	1	1	1	0	1	1	0	1	0	1
Echo Sounder	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
GPS	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Gyro-Compass	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
HCS	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Indicators	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Magnetic Compass	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
MFD	0	1	1	0	0	1	1	1	0	1	1	0	1	0	1
NAVTEX	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
RADAR	0	1	1	0	0	1	1	1	0	1	1	0	1	0	1
ROTI	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Rudder pump selector switch	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Sound reception system	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
SDME	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Steering mode selector switch	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Steering position selector switch	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
TCS	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Transmitting Heading Device	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0

3.7. Identifying Impact Scores of Components

Information impacts (i.e., IPC, DC, LIC) were not available for an INS as mentioned in Section 3.4. During the literature review, no incidents harming humans or the environment were found to be caused by cyber attacks against a vessel. This is why safety criticality and environmental criticality were assumed to be in the category *None—No injury or insufficient data*. Various aspects affect financial losses, including violation of the charter party agreement, daily operational expenses, repair costs, and so on. It is difficult to estimate a potential loss; however, it is highly possible for this to be over \$10,000. This is why financial criticality was assumed as *Significant—\$10,001—\$100,000*. The loss of various components may cause the delay of a vessel or the need to inform maritime regulators, such as AIS, GPS, or RADAR. Such components are assumed as *Significant* for reputational criticality. The OOI is the normalized average of all centrality metrics of a component calculated using ORA. ODC denotes the out-degree centrality of a component calculated using ORA. OCC is the overall component criticality, which is calculated using an equation

in [51]. It is basically the average of all impacts (e.g., safety, financial, and information). All such assumptions and calculations are represented in Table 10.

Table 10. Component criticality score table.

Component	Information			SC	EC	FC	RC	OOI	Staging	
	IPC	DC	LIC						ODC	OCC
AIS	0	0	0	0	0	0.5	1	0.872174439	0.042	0.402362407
Anemometer	0	0	0	0	0	0.5	0	0	0	0.083333333
BNWAS	0	0	0	0	0	0.5	0	0	0	0.083333333
Central Alert	0	0	0	0	0	0.5	0	0	0	0.083333333
Manageme. HMI	0	0	0	0	0	0.5	1	0	0	0.25
Controls for M/E	0	0	0	0	0	0.5	1	0	0	0.25
Controls for main rudder	0	0	0	0	0	0.5	1	0	0	0.25
Controls for thruster	0	0	0	0	0	0.5	0	0	0	0.083333333
ECDIS	0	0	0	0	0	0.5	1	0.438221675	0	0.323036946
Echo Sounder	0	0	0	0	0	0.5	0	0	0	0.083333333
GPS	0	0	0	0	0	0.5	1	0.7350904	0.208	0.407181733
Gyro-Compass	0	0	0	0	0	0.5	1	1	0.208	0.284666667
HCS	0	0	0	0	0	0.5	0	0.301782611	0	0.133630435
Indicators	0	0	0	0	0	0.5	0	0	0	0.083333333
Magnetic Compass	0	0	0	0	0	0.5	0	0.149697807	0.042	0.115282968
MFD	0	0	0	0	0	0.5	1	0	0	0.25
NAVTEX	0	0	0	0	0	0.5	1	0	0	0.25
RADAR	0	0	0	0	0	0.5	1	0.735171456	0	0.372528576
ROTI	0	0	0	0	0	0.5	0	0.177510045	0.042	0.119918341
Rudder pump selector switch	0	0	0	0	0	0.5	0	0	0	0.083333333
Sound reception system	0	0	0	0	0	0.5	0	0	0	0.083333333
SDME	0	0	0	0	0	0.5	0	0.552742648	0.167	0.203290441
Steering mode selector switch	0	0	0	0	0	0.5	0	0	0	0.083333333
Steering position selector switch	0	0	0	0	0	0.5	0	0	0	0.083333333
TCS	0	0	0	0	0	0.5	0	0.438221675	0	0.156370279
Transmitting Heading Device	0	0	0	0	0	0.5	0	0.156940387	0	0.109490065

3.8. Calculating Risk Scores and Identifying Risk Levels

The last element that is required for calculating the risk is the impact of techniques targeting components. This is achieved by utilizing the information in Tables 4, 10 and A2. Table A2 specifies the relevant failure modes for a component. Table 4 specifies the metric to be utilized for estimating the impact of failure mode, and Table 10 specifies the quantification of the impact for each impact element. The final value of the impact of failure mode (F) for component (C) was calculated using Equation (3):

$$Impact_{F,C} = (SF_F \times SC_C) + (FF_F \times FC_C) + (ICF_F \times IC_C) + (OF_F \times OC_C) + (StF_F \times StC_C) \quad (3)$$

where SF_F , FF_F , ICF_F , OF_F , and StF_F are the weighting factors for safety, financial, information criticality, operational, and staging impact elements. These factors are expected to be driven from the risk management strategy to prioritize certain impact elements (e.g., safety). In this paper, all impact elements are treated equally, rendering all the factors to be (=1). Additionally, SC_C , FC_C , IC_C , OC_C , and StC_C are the quantification of the impact element for the component (C) based on which metric specified for the failure (in Table 4)

and the value of that metric (in Table 10). Afterwards, a risk priority number (RPN) can be calculated for each identified technique, leading to a failure mode for each component based on Equation (4):

$$RPN_{T,C} = Likelihood_T \times Impact_{F,C} \times Detectability_{T,M} \quad (4)$$

T: Technique, *C*: Component, *F*: Failure, *M*: Mitigation measure

The likelihood quantification is derived from Equation (1), the impact is derived from Equation (3), and the detectability is derived from Equation (2).

Our findings were prepared in Excel tables as described in [51]. Then, risk scores were calculated by the script, which was specifically coded for the methodology. In the original method, the risks are classified for levels of low risk rating (0–4.86), medium risk rating (4.87–9.72), high risk rating (9.73–14.58), and critical risk rating (14.59–19.44). However, in this study, we ignored several consequences, as described in Section 3.7. This is why we re-defined the risk levels by scores. According to our findings, risks are in the range of 0.041624847 and 8.68820705893103. The range was divided into four classes to prioritize the risks, as shown in Table 11.

Table 11. New risk scores with levels.

Range	Level
0.00–2.18	Low
2.19–4.36	Medium
4.37–6.54	High
6.55–8.72	Critical

In this study, cyber risks for 25 components in an INS were investigated. Three components, such as rudder pump selector switch, steering mode selector switch, and steering position selector switch do not include any cyber risks. A total of 1850 risks belonging to the rest of 22 components were found. Our results classified 1805 risks as low, 32 as medium, 9 as high, and 4 as critical. Risk numbers for each component and risk levels by the original method and our study definitions are represented in Table 12. Medium, high, and critical risks are listed in Appendix A.

Table 12. Results of risk assessment.

Component	Total Risk	Risk Level (Original)	Risk Level (Study)
AIS	5	5 low	3 low 1 medium 1 high
Anemometer	5	5 low	5 low
BNWAS	5	5 low	5 low
Central Alert Management HMI	41	41 low	41 low
Controls for M/E	40	40 low	35 low 5 medium
Controls for main rudder	40	40 low	35 low 5 medium
Controls for thruster	40	40 low	40 low

Table 12. Cont.

Component	Total Risk	Risk Level (Original)	Risk Level (Study)
ECDIS	499	496 low 3 medium	489 low 7 medium 1 high 2 critical
Echo Sounder	5	5 low	5 low
GPS	5	5 low	4 low 1 medium
Gyro-Compass	5	5 low	5 low
HCS	40	40 low	39 low 1 medium
Indicators	41	41 low	41 low
Magnetic Compass	5	5 low	5 low
MFD	499	497 low 2 medium	492 low 3 medium 2 high
NAVTEX	11	10 low 1 medium	9 low 1 medium 1 high
RADAR	504	501 low 3 medium	492 6 medium 4 high 2 critical
ROTI	5	5 low	5 low
Rudder pump selector switch	0		
Sound reception system	5	5 low	5 low
Speed and Distance Measuring Equipment	5	5 low	5 low
Steering mode selector switch	0		
Steering position selector switch	0		
TCS	40	40 low	38 low 2 medium
Transmitting Heading Device	5	5 low	5 low
Total	1850	1841 low 9 medium	1805 low 32 medium 9 high 4 critical

Nine high risks were related to AIS, ECDIS, MFD, NAVTEX, and RADAR. RADAR solitarily included four of nine high risks. In total, 1502 risks of 1850 total were related to ECDIS (499 risks), MFD (499 risks), and RADAR (504 risks). The remaining risks related to 19 components. Moreover, four critical risks related to ECDIS and RADAR. A total of 1497 risks for enterprise, 342 risks for ICS, and 11 risks related to the mobile matrix; in total, 443 different techniques led to 1850 risks, 13 of which might compromise over 9 risks as represented in Table 13.

Table 13. Techniques compromising over 10 risks.

Matrix	MITRE ID	Techniques	Risk Number
ICS	T0858	Change Operating Mode	24
ICS	T0829	Loss of View	14
ICS	T0832	Manipulation of View	14
ICS	T0849	Masquerading	14
ICS	T0859	Valid Accounts	14
ICS	T0886	Remote Services	14
ICS	T0815	Denial of View	12
Enterprise	T1078	Valid Accounts	12
Enterprise	T1078.001	Valid Accounts: Default Accounts	12
Enterprise	T1078.002	Valid Accounts: Domain Accounts	12
Enterprise	T1078.003	Valid Accounts: Local Accounts	12
ICS	T0822	External Remote Services	10
ICS	T0856	Spoof Reporting Message	10

4. Conclusions

We proposed a derived method to assess the cyber risks of ships. The original method was developed to assess cyber risks of cyber-physical systems by following the FMECA and MITRE ATT&CK framework. We adapted the method for marine systems in particular. Then, we implemented the method to assess the cyber risks of an INS, and 1850 risks related to 22 components were found. Any risks for three components (i.e., switches) were not available. The risks were classified as 1805 low, 32 medium, 9 high, and 4 critical.

The high and critical risks reflect adversarial objectives to cause an impact on the INS functions. This includes a wide range of threats, such as several variations of denial of service attacks, denial of the processing of sensor data, jamming attacks, and hijacking the resources of sensitive components.

The ECDIS, MFD, and RADAR are the only components that need an operating system to run. According to our results, the operating system increases the cyber threats to and vulnerabilities of a component dramatically. Other components underlying the operating system onboard, such as the ballast water management system and any transfer systems (e.g., bunker), would involve many cyber risks similar to the ECDIS, MFD, and RADAR.

In the original method, consequences are identified as operational, safety, information, financial, and staging. Because of the industry's necessities, we also took into environmental and reputational consequences. The impact estimation criteria for each consequence were adapted by considering FSA. Information consequence was not available for an INS. Safety and environmental consequences could be possible; however, any marine casualty (e.g., collision, injury, and explosion) caused by cyber incidents does not exist in the literature to date. This is why safety and environmental criticalities could be assumed or ignored. We decided to ignore both. For this reason, we also re-classified risk levels by risk scores. If we had not re-classified the risk levels, the risks would have been underestimated. Once the literature is enriched, other consequences must be considered as well.

The IMO only defines the minimum standards for marine components. Each manufacturer is usually free in various aspects, such as product design, working principle, software, hardware, and operating system. Features, more than requirements, may be attached to products by makers to create added value. This is why failure modes and mitigation measures could be changeable by products. In this study, an implementation of our proposed method is represented and the risk assessment was performed for a typical INS. However, the method is convenient to be implemented in the cyber risk assessment of marine systems other than INS. In further studies, cyber risks of other systems in the bridge, such as safety, security, and communication systems, can be assessed. Moreover, cyber risks of equipment in other locations, such as the engine room and cargo control room, may be assessed.

Our study is based on several assumptions, as many risk assessments were conducted. A few records of cyber incidents and experimental studies against marine systems are

available in the literature. This is why we also investigated troubleshooting sections of product brochures to assume the impact of a potential attack. The mapping of failure modes and their consequences are subjective and might change under expert judgement. Financial criticality was considered as significant (USD 10,001–100,000). However, commercial losses (e.g., cargo claims, charter party violations, and loss of potential charterer) and costs for components, service, mooring and so on could directly affect the financial losses of a cyber incident. This is why financial impact is based on assumptions, as well. Despite several assumptions, the method is comprehensive and detailed. It can be perfectly implemented to assess the cyber risks of well-defined marine systems under a specific scenario.

The study offers two classifications for components of an INS. The IMO classifies the components as IT and OT. However, our method can classify IT, OT, wireless, and combinations of these. Our method and IMO differently define IT and OT notions. For the risk assessment method, IMO definitions are not required. Given that any complete list could not be found in the literature, component classification for an INS by the IMO definition was also given in our study as an additional contribution.

Author Contributions: Conceptualization, A.O. and V.G.; methodology, A.O. and A.A.; software, A.A.; formal analysis, A.O.; investigation, A.O.; validation, A.O. and A.A.; writing—original draft preparation, A.O. and A.A.; writing—review and editing, A.O., A.A. and V.G.; visualization, A.O. and A.A.; data curation, A.O. and A.A.; supervision, V.G.; project administration, V.G.; funding acquisition, V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by (a) the Research Council of Norway through the Maritime Cyber Resilience (MarCy) project, Project no. 295077; and (b) the NTNU Digital transformation project Autoferry.

Institutional Review Board Statement: There is no institutional review board statement.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available upon request via corresponding author email.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Medium, high, and critical risks of an INS are given in Table A1.

Table A1. Medium, high, and critical risks of an INS.

No.	Component	MITRE ID	Techniques	Risk
1	AIS	T0815	Denial of View	High
2	AIS	T0829	Loss of View	Medium
3	Controls for M/E	T0879	Damage to Property	Medium
4	Controls for M/E	T0809	Data Destruction	Medium
5	Controls for M/E	T0826	Loss of Availability	Medium
6	Controls for M/E	T0828	Loss of Productivity and Revenue	Medium
7	Controls for M/E	T0856	Spoof Reporting Message	Medium
8	Controls for main rudder	T0879	Damage to Property	Medium
9	Controls for main rudder	T0809	Data Destruction	Medium
10	Controls for main rudder	T0826	Loss of Availability	Medium
11	Controls for main rudder	T0828	Loss of Productivity and Revenue	Medium
12	Controls for main rudder	T0856	Spoof Reporting Message	Medium

Table A1. Cont.

No.	Component	MITRE ID	Techniques	Risk
13	ECDIS	T1498.002	Reflection Amplification	Medium
14	ECDIS	T1499.004	Application or System Exploitation	Medium
15	ECDIS	T1499.003	Application Exhaustion Flood	Medium
16	ECDIS	T1499.002	Service Exhaustion Flood	Medium
17	ECDIS	T1499.001	OS Exhaustion Flood	Medium
18	ECDIS	T1531	Account Access Removal	Medium
19	ECDIS	T1529	System Shutdown/Reboot	Medium
20	ECDIS	T1499	Endpoint Denial of Service	Critical
21	ECDIS	T1498	Network Denial of Service	Critical
22	ECDIS	T1496	Resource Hijacking	High
23	GPS	T0815	Denial of View	Medium
24	HCS	T0826	Loss of Availability	Medium
25	MFD	T1531	Account Access Removal	Medium
26	MFD	T1529	System Shutdown/Reboot	Medium
27	MFD	T1499	Endpoint Denial of Service	High
28	MFD	T1498	Network Denial of Service	High
29	MFD	T1496	Resource Hijacking	Medium
30	NAVTEX	T1464	Network Denial of Service	High
31	NAVTEX	T1463	Manipulate Device Communication	Medium
32	RADAR	T1498.002	Reflection Amplification	High
33	RADAR	T1499.004	Application or System Exploitation	Medium
34	RADAR	T1499.003	Application Exhaustion Flood	Medium
35	RADAR	T1499.002	Service Exhaustion Flood	High
36	RADAR	T1499.001	OS Exhaustion Flood	High
37	RADAR	T1491.001	Internal Defacement	Medium
38	RADAR	T1531	Account Access Removal	Medium
39	RADAR	T1529	System Shutdown/Reboot	Medium
40	RADAR	T1499	Endpoint Denial of Service	Critical
41	RADAR	T1498	Network Denial of Service	Critical
42	RADAR	T1496	Resource Hijacking	High
43	RADAR	T1491	Defacement	Medium
44	TCS	T0809	Data Destruction	Medium
45	TCS	T0826	Loss of Availability	Medium

Appendix B

Table A2 represents samples of failures, cyber incidents, vulnerabilities and failure modes.

Table A2. Samples of failures, incidents and vulnerabilities, and failure modes.

Component	Failure	Occurred Incidents & Discovered Vulnerabilities	Failure Modes		
			Mobile	Enterprise	ICS
AIS	<ul style="list-style-type: none"> not receiving AIS messages; not transmitting AIS messages; transmitting the wrong AIS messages; displaying invalid AIS information; difference between internal and external GPS data; mismatching heading data. 	<ul style="list-style-type: none"> spoofing; hijacking; availability; tampering. 	<ul style="list-style-type: none"> network denial of service; impact. 	<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
Anemometer	<ul style="list-style-type: none"> inaccurate wind speed; missing wind speed; inaccurate wind direction; missing wind direction. 	N/A		<ul style="list-style-type: none"> data manipulation. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of view; manipulation of control; manipulation of view.
BNWAS	<ul style="list-style-type: none"> not activating/deactivating it in automatic mode; not rising alarm; rising alarm constantly; not working motion detectors if equipped. 	N/A		<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
Central Alert Management HMI	<ul style="list-style-type: none"> not stopping alert; not rising alert; not keeping alert history; displaying alerts with wrong date/time stamp. 	N/A		<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.

Table A2. Cont.

Component	Failure	Occurred Incidents & Discovered Vulnerabilities	Failure Modes		
			Mobile	Enterprise	ICS
Controls for M/E	<ul style="list-style-type: none"> not changing or RPM; missing or wrong information; not working of command. 	N/A			<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
ECDIS	<ul style="list-style-type: none"> collapsing the operating system; wrong position of own vessel; not updating ENC/RNC; not receiving/displaying information from connected components; not allowing route planning or monitoring; data manipulation in functions such as past track or planned course. 	<ul style="list-style-type: none"> operating system vulnerabilities; middleware vulnerabilities; manipulation of the ship position. 		<ul style="list-style-type: none"> collection; discovery; execution; exfiltration; initial access; data encrypted for impact; credential access; data manipulation; lateral movement; system shutdown/reboot; defense evasion. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
Echo Sounder	<ul style="list-style-type: none"> inaccurate depth value; no depth value. 	N/A		<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
GPS	<ul style="list-style-type: none"> not fixing the position; wrong position; not transmitting the data to other components. 	<ul style="list-style-type: none"> jamming; spoofing. 	<ul style="list-style-type: none"> network denial of service; impact. 	<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.

Table A2. Cont.

Component	Failure	Occurred Incidents & Discovered Vulnerabilities	Failure Modes		
			Mobile	Enterprise	ICS
Gyro-Compass	<ul style="list-style-type: none"> displaying wrong heading information; not receiving GPS messages; not transmitting information to other components. 	N/A		<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
HCS	<ul style="list-style-type: none"> not receiving NMEA messages from connected components. 	N/A		<ul style="list-style-type: none"> data manipulation; firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
Indicators	<ul style="list-style-type: none"> not receiving NMEA messages from connected components. 	N/A		<ul style="list-style-type: none"> data manipulation. 	<ul style="list-style-type: none"> loss of availability; loss of safety; loss of view; manipulation of view.
MFD	<ul style="list-style-type: none"> not receiving NMEA messages from connected components; collapsing operating system. 	<ul style="list-style-type: none"> operating system vulnerabilities; middleware vulnerabilities. 		<ul style="list-style-type: none"> collection; defense evasion; discovery; execution; exfiltration; initial access; data encrypted for impact; credential Access; data manipulation; lateral movement; system shut-down/reboot. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.
NAVTEX	<ul style="list-style-type: none"> not receiving MSI 	N/A	<ul style="list-style-type: none"> network denial of service; impact. 	<ul style="list-style-type: none"> firmware corruption; initial access. 	<ul style="list-style-type: none"> loss of availability; loss of control; loss of safety; loss of view; manipulation of control; manipulation of view.

References

1. UNCTAD. *Review of Maritime Transport 2021*; United Nations Publications: New York, NY, USA, 2021; Available online: <https://unctad.org/webflyer/review-maritime-transport-2021> (accessed on 20 November 2021).
2. IMO. Introduction to IMO. Available online: <https://www.imo.org/en/About/Pages/Default.aspx> (accessed on 29 September 2022).
3. IMO MSC-FAL.1-Circ.3-Rev.1; Guidelines on Maritime Cyber Risk Management. IMO: London, UK, 2021.
4. IMO Resolution MSC.428(98); Maritime Cyber Risk Management in Safety Management Systems. IMO: London, UK, 2017.
5. IMO. *Guide to Maritime Security and the ISPS Code: Section 4 Security Responsibilities of Ship Operators—4.13 Cyber Security on Board Ships*; IMO: London, UK, 2021; Available online: <https://shop.witherbys.com/guide-to-maritime-security-and-the-isp-code-2021-edition/> (accessed on 10 July 2022).
6. IMO. *ISPS Code: Part A Mandatory Requirements—9 Ship Security Plan*; IMO: London, UK, 2002.
7. Resolution A.915(22); Revised Maritime Policy and Requirements for a Future Global Navigation Satellite System (GNSS). IMO: London, UK, 2001.
8. IMO MSC.252(83); Adoption of the Revised Performance Standards for Integrated Navigation Systems (INS): Introduction, Contents, Module A-B. IMO: London, UK, 2018.
9. IMO MSC.252(83); Adoption of the Revised Performance Standards for Integrated Navigation Systems (INS): Appendices. IMO: London, UK, 2018.
10. Strom, B. ATT&CK 101. Available online: <https://medium.com/mitre-attack/att-ck-101-17074d3bc62> (accessed on 6 November 2022).
11. MITRE. Enterprise Matrix. Available online: <https://attack.mitre.org/matrices/enterprise/> (accessed on 10 July 2022).
12. Georgiadou, A.; Mouzakitis, S.; Askounis, D. Assessing MITRE ATT&CK Risk Using a Cyber-Security Culture Framework. *Sensors* **2021**, *21*, 3267. [CrossRef] [PubMed]
13. MITRE. Mobile Matrix. Available online: <https://attack.mitre.org/matrices/mobile/> (accessed on 10 July 2022).
14. MITRE. ICS Matrix. Available online: <https://attack.mitre.org/matrices/ics/> (accessed on 10 July 2022).
15. MITRE. Enterprise Tactics. Available online: <https://attack.mitre.org/tactics/enterprise/> (accessed on 10 July 2022).
16. MITRE. Enterprise Techniques. Available online: <https://attack.mitre.org/techniques/enterprise/> (accessed on 10 July 2022).
17. MITRE. Enterprise Mitigations. Available online: <https://attack.mitre.org/mitigations/enterprise/> (accessed on 10 July 2022).
18. MITRE. Software. Available online: <https://attack.mitre.org/software/> (accessed on 10 July 2022).
19. MITRE. Groups. Available online: <https://attack.mitre.org/groups/> (accessed on 10 July 2022).
20. MITRE. Data Sources. Available online: <https://attack.mitre.org/datasources/> (accessed on 10 July 2022).
21. Kavallieratos, G.; Katsikas, S. Managing cyber security risks of the cyber-enabled Ship. *J. Mar. Sci. Eng.* **2020**, *8*, 768. [CrossRef]
22. Kavallieratos, G.; Katsikas, S.; Gkioulos, V. Cyber-attacks against the autonomous ship. In *Computer Security*; Katsikas, S.K., Cuppens, F., Cuppens, N., Lambrinouidakis, C., Antón, A., Gritzalis, S., Mylopoulos, J., Kalloniatis, C., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 20–36. ISBN 978-3-030-12785-5.
23. Tusher, H.M.; Munim, Z.H.; Notteboom, T.E.; Kim, T.-E.; Nazir, S. Cyber security risk assessment in autonomous shipping. *Marit. Econ. Logist.* **2022**, *24*, 208–227. [CrossRef]
24. Shang, W.; Gong, T.; Chen, C.; Hou, J.; Zeng, P. Information security risk assessment method for ship control system based on Fuzzy Sets and Attack Trees. *Secur. Commun. Netw.* **2019**, *2019*, 3574675. [CrossRef]
25. Oruc, A. Cybersecurity Risk Assessment for Tankers and Defence Methods. Master’s Thesis, Piri Reis University, Istanbul, Turkey, 2020.
26. Kessler, G.C.; Craiger, P.; Haass, J.C. A taxonomy framework for maritime cybersecurity: A demonstration using the Automatic Identification System. *TransNav Int. J. Mar. Navig. Saf. Sea Transp.* **2018**, *12*, 429–437. [CrossRef]
27. Svilicic, B.; Kamahara, J.; Rooks, M.; Yano, Y. Maritime cyber risk management: An experimental ship assessment. *J. Navig.* **2019**, *72*, 1108–1120. [CrossRef]
28. iTrust. Guidelines for Cyber Risk Management in Shipboard Operational Technology Systems. 2022. Available online: <https://itrust.sutd.edu.sg/news-events/news/guidelines-for-cyber-risk-management-in-shipboard-ot-systems/> (accessed on 6 April 2022).
29. You, B.; Zhang, Y.; Cheng, L.-C. Review on cybersecurity risk assessment and evaluation and their approaches on maritime transportation. In Proceedings of the 30th Annual Conference of International Chinese Transportation Professionals Association, Houston, TX, USA, 19–21 May 2017.
30. Tam, K.; Jones, K. MaCRA: A model-based framework for maritime cyber-risk assessment. *WMU J. Marit. Aff.* **2019**, *18*, 129–163. [CrossRef]
31. Tam, K.; Jones, K. Cyber-risk assessment for autonomous ships. In Proceedings of the 2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), Glasgow, UK, 11–12 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–8, ISBN 978-1-5386-4683-0.
32. Bolbot, V.; Theotokatos, G.; Boulougouris, E.; Vassalos, D. A novel cyber-risk assessment method for ship systems. *Saf. Sci.* **2020**, *131*, 104908. [CrossRef]
33. Meland, P.H.; Nesheim, D.A.; Bernsmed, K.; Sindre, G. Assessing cyber threats for storyless systems. *J. Inf. Secur. Appl.* **2022**, *64*, 103050. [CrossRef]

34. ISO 31000; Risk Management Guidelines: Guidelines. ISO: Geneva, Switzerland, 2018. Available online: <https://www.iso.org/iso-31000-risk-management.html> (accessed on 12 July 2022).
35. ISO/TR 31004; Guidance for the Implementation of ISO 31000. ISO: Geneva, Switzerland, 2013. Available online: <https://www.iso.org/standard/56610.html> (accessed on 12 July 2022).
36. IEC 31010; Risk Management: Risk Assessment Techniques. IEC: Geneva, Switzerland, 2019. Available online: <https://www.iso.org/standard/72140.html> (accessed on 12 July 2022).
37. ISO/IEC 27000; Information Technology—Security Techniques—Information Security Management Systems—Overview and Vocabulary. ISO/IEC: Geneva, Switzerland, 2018. Available online: <https://www.iso.org/standard/73906.html> (accessed on 12 July 2022).
38. IEC 63154; Maritime Navigation and Radiocommunication Equipment and Systems: Cybersecurity—General Requirements, Methods of Testing and Required Test Results. IEC: Geneva, Switzerland, 2021. Available online: <https://webstore.iec.ch/publication/61003> (accessed on 12 July 2022).
39. IMO MSC-MEPC.2 Circ.12/Rev.2; Revised Guidelines for Formal Safety Assessment (FSA) for Use in the IMO Rule-Making Process. IMO: London, UK, 2018.
40. Witherbys; BIMCO; ICS. *Cyber Security Workbook for on Board Ship Use*; Witherby Publishing Group: Scotland, UK, 2022.
41. DNV-RP-0496; Cyber Security Resilience Management for Ships and Mobile Offshore Units in Operation. DNV: Oslo, Norway, 2021. Available online: <https://www.dnv.com/maritime/dnv-rp-0496-recommended-practice-cyber-security-download.html> (accessed on 27 June 2022).
42. BIMCO; CSA; DCSA; ICS; INTERCARGO; InterManager; INTERTANKO; IUMI; OCIMF; WSC; et al. The Guidelines on Cyber Security Onboard Ships. 2020. Available online: <https://www.ics-shipping.org/wp-content/uploads/2021/02/2021-Cyber-Security-Guidelines.pdf> (accessed on 21 March 2022).
43. MSC.1-Circ.1639; The Guidelines on Cyber Security Onboard Ships. IMO: London, UK, 2021.
44. Sheraz, M. Cyber Kill Chain vs. MITRE ATT&CK. Available online: <https://www.linkedin.com/pulse/cyber-kill-chain-vs-mitre-attck-muhammad-sheraz/> (accessed on 1 October 2022).
45. Poston, H. Top threat modeling frameworks: STRIDE, OWASP Top 10, MITRE ATT&CK Framework and More. Available online: <https://resources.infosecinstitute.com/topic/top-threat-modeling-frameworks-stride-owasp-top-10-mitre-attck-framework/> (accessed on 1 October 2022).
46. Stack Exchange. Difference between STRIDE and Mitre ATTACK. Available online: <https://security.stackexchange.com/questions/184083/difference-between-stride-and-mitre-attack> (accessed on 1 October 2022).
47. CyCraft Technology Corp. CyCraft Classroom: MITRE ATT&CK vs. Cyber Kill Chain vs. Diamond Model. Available online: <https://medium.com/cycraft/cycraft-classroom-mitre-att-ck-vs-cyber-kill-chain-vs-diamond-model-1cc8fa49a20f> (accessed on 1 October 2022).
48. Strom, B.E.; Applebaum, A.; Miller, D.P.; Nickels, K.C.; Pennington, A.G.; Thomas, C.B. *MITRE ATT&CK: Design and Philosophy*; MITRE Corporation: McLean, VA, USA, 2020; Available online: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf (accessed on 10 January 2022).
49. Khodabakhsh, A.; Yayilgan, S.Y.; Abomhara, M.; Istad, M.; Hurzuk, N. Cyber-risk identification for a digital substation. In Proceedings of the 15th International Conference on Availability, Reliability and Security, ARES 2020, Virtual Event Ireland, 25–28 August 2020; Volkamer, M., Wressnegger, C., Eds.; ACM: New York, NY, USA, 2020; pp. 1–7, ISBN 978-1-4503-8833-7.
50. He, T.; Li, Z. A model and method of information system security risk assessment based on MITRE ATT&CK. In Proceedings of the 2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT), Sanya, China, 27–29 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 81–86, ISBN 978-1-6654-3757-8.
51. Amro, A.; Gkioulos, V.; Katsikas, S. Assessing cyber risk in cyber-physical systems using the ATT&CK framework. *ACM Trans. Priv. Secur.* **2021**. [CrossRef]
52. Oruc, A.; Gkioulos, V.; Katsikas, S. Towards a Cyber-Physical Range for the Integrated Navigation System (INS). *J. Mar. Sci. Eng.* **2022**, *10*, 107. [CrossRef]
53. Carley, K.M. ORA: A Toolkit for Dynamic Network Analysis and Visualization. In *Encyclopedia of Social Network Analysis and Mining*; Alhajj, R., Rokne, J., Eds.; Springer New York: New York, NY, USA, 2014; pp. 1219–1228. ISBN 978-1-4614-6169-2.
54. Altman, N.; Carley, K.M. *ORA User's Guide 2022*; Carnegie Mellon University: Pittsburgh, PA, USA, 2022; Available online: <http://www.casos.cs.cmu.edu/publications/papers/CMU-ISR-22-107.pdf> (accessed on 1 October 2022).
55. FIRST. Common Vulnerability Scoring System v3.1: Specification Document. Available online: <https://www.first.org/cvss/v3.1/specification-document> (accessed on 21 October 2022).
56. IMO. A.1106(29) Revised Guidelines for the Onboard Operational Use of Shipborne Automatic Identification Systems (AIS); IMO: London, UK, 2015.
57. IMO. SOLAS Chapter V Safety of Navigation: Regulation 19-1 Long-Range Identification and Tracking of Ships; IMO: London, UK, 2006.
58. MITRE. Mobile Mitigations. Available online: <https://attack.mitre.org/mitigations/mobile/> (accessed on 30 June 2022).
59. MITRE. ICS Mitigations. Available online: <https://attack.mitre.org/mitigations/ics/> (accessed on 30 June 2022).

ISBN 978-82-326-6160-2 (printed ver.)
ISBN 978-82-326-5701-8 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)



NTNU

Norwegian University of
Science and Technology