Cyber-Physical Tracking of IoT devices: A maritime use case

Ahmed Amro[0000−0002−3390−0772]

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

Abstract. We live in a highly connected world. Many types of devices involved in numerous applications are connected to the internet and their number is increasing day by day. In the maritime domain, maritime entities utilize the Internet to connect geographically dispersed vessels, offshore units, and other sorts of components in the maritime infrastructure. While the locations of some of these components are publicly available through various resources (e.g. MarineTraffic), their cyber-related information is not necessarily intended to be. Obtaining the knowledge of both the physical location as well as the cyber-related information of certain components might provide attackers with opportunities to perform more sophisticated and targeted attacks. With new regulations and guidelines aiming to improve cybersecurity in maritime, investigating possible threats against the maritime infrastructure is required. To this end, this paper investigates the issue of combined cyber and physical tracking of IoT devices with a prime focus on maritime infrastructure. I propose a process for Cyber-Physical tracking of IoT devices including maritime components that are connected to the internet. I employed several IoT scanners (e.g. Shodan) and obtained cyber-related information as well as their physical properties such as location, speed, and others. I have identified 4942 hosts that emit NMEA messages and 331 possible maritime components. Furthermore, I provide discussion regarding the expected risks of such a process while considering both the current state of affairs in maritime as well as futuristic operational modes such as autonomous, unmanned, and remotely connected vessels.

Keywords: IoT · scanning · tracking · cybersecurity · maritime · NMEA

1 Introduction

The amount of connected devices to the internet is growing each year and is expected to double from the year 2021 to 2025 reaching 75.44 billion [28]. This rapid trend of connectivity can pave the way for innovation and an improved way of life, however, it introduces a wide range of cyber attacks if cybersecurity is not considered during the development of such devices and their hosting systems.

Several sectors are following the trend of Internet of Things (IoT) and Industrial IoT (IIoT) including maritime [15, 5]. The maritime sector is undergoing a digital transformation era that drastically impacts its technologies, business

models and operations [10]. Vessel tracking services are among these operations as maritime operators must keep track of their vessels and geographically distinct components for improved management. Therefore, they rely on devices that are connected to the internet and emit marine information that is important for their operations such as location, speed, heading, and others. Some of these devices employ a protocol proposed by National Marine Electronics Association (NMEA) for communicating among maritime components. Under normal circumstances, vessel tracking is a very common domain and field of study. Some marine traffic information is a publicly open resource utilized for legitimate ship tracking purposes. However, fingerprinting and cyber tracking of vessels by unauthorized entities is a less-discussed subject. Previously, ships have been fingerprinted and tracked using data from Automatic Identification Systems (AIS) [24]. Such activities can be conducted by attackers during the reconnaissance stage toward the development of more advanced and targeted attacks. Attackers can collect cyber-related information about target ships such as their connected devices and their vulnerabilities and use this information during exploitation (further discussed in section 4).

Recently, the International Maritime Organization (IMO) has passed Resolution MSC. 428(98) [6] for maritime risk management. The resolution makes it mandatory for ship owners and operators to include cybersecurity in their safety management systems. Among the discussed risk management activities in the resolution is continuous risk analysis considering the threat landscape. My paper supports the efforts in this direction by capturing the current state of a very common maritime protocol that is NMEA observed on the internet. I follow a state-of-the-are process for IoT vulnerability scanning proposed in my earlier work [1] and utilize known IoT scanners (e.g. Shodan) for detecting NMEA emitting devices. Moreover, I develop upon the approach of detecting vessels using AIS data and utilize NEMA messages for fingerprinting maritime components using them. My work aims to shed the light on a possible threat against organizations and systems employing NMEA. My contributions in this paper are summarized as follows:

- I propose a process for Cyber-Physical tracking of IoT devices with a prime focus on maritime components. This process emulates an adversarial behavior against systems using NMEA as an early stage of cyber attacks.
- I present the current status of NMEA service considering the type of messages, devices, ports, and countries. I believe that this information is valuable for the cybersecurity community in maritime and other sectors employing NMEA.
- I provide discussion regarding the risks of my proposed Cyber-Physical tracking process considering both the current state of affairs in the maritime domain as well as considering futuristic operational modes.

The remainder of this paper is organized as follows. Section 2 discusses relevant concepts and artifacts that are utilized in this paper. Then, section 3 discusses in detail my proposed Cyber-Physical tracking process which also resulted in capturing the status of NMEA messages on the internet. Afterward,

section 4 provides a discussion of the risks associated with my proposed process, provides suggestions for mitigation, and discusses limitations. Finally, section 5 concludes the work in this paper.

2 Background and Related Work

Shodan [16]; a known IoT search engine has previously presented a ship tracking capability utilizing AIS data communicated over the Internet which includes position information. This has been argued to be a wake-up call for maritime cybersecurity [24]. Since then, very limited works have discussed this issue as the number of AIS-connected devices visible to the internet are very limited according to my latest search for AIS messages (e.g AIVDM, AIVDO, ABVDM, etc) on Shodan. That work highlighted the ability for unauthorized entities to gain both physical and cyber information regarding vessels by relying on protocols that are accessible through cyber means and disclose physical properties. That work had led me to consider NMEA protocol as another approach. I believe that NMEA provides a suitable link between both the cyber and physical realms.

There are several NMEA protocols including NMEA0183 [2] and NMEA2000 [14]. NMEA messages abiding by the NMEA0183 protocol are textual messages containing structured information intended originally for navigation purposes. The format of NMEA messages includes static information and dynamic information. The static information includes a TalkerID and a MessageID. The dynamic information includes several fields each containing specific information such as time, longitude, latitude, and others (refer to [2, 26] for more details). This information is utilized in legitimate vessel tracking services as well as legitimate navigational functions. The messages are not encrypted or encoded, they are communicated in plain text. Therefore, they can be used to fingerprint devices emitting them.

Originally, NMEA0183 are mostly transmitted over serial links restricting access to them to specific systems and locations [2]. However, adaptations have been proposed to transmit NMEA0183 messages over TCP and UDP protocols making them accessible through IP networks. This transformation introduced a wide range of cyber attacks. The security; or in better terms, the lack of security in NMEA has been discussed by several works. Tran et al [31] have discussed the security of several marine protocols including NMEA0183. The authors referred to the lack of authentication, encryption, and validation of NMEA messages. The authors argued that the messages are susceptible to many attacks if attackers can identify the network device that uses the standard. Other works have argued that NMEA security currently depends on the network and host security [27, 9].

My research targets maritime risk management with a current focus on the risks related to NMEA messages. I employ the $ATT\&CK$ framework [29] for threat modeling to identify threats against maritime systems and components across the different adversarial tactics (i.e. kill chain phases) of cyber attacks. This paper considers attack techniques and mitigation related to NMEA messages during the reconnaissance stage of cyber attacks. I investigate and demon-

strate the ability of attackers to fingerprint devices emitting NMEA messages over the internet. Future work will focus on subsequent kill chain phases.

In this paper, I rely on my previous work [1] in which I presented the state-of-the-art in IoT scanning and vulnerability scanning. I highlighted the increased interest in the field, discussed some challenges, and proposed a systematic process for IoT vulnerability scanning. I also proposed a scanning space in which all scanning processes occur. The space consists of three dimensions, namely, IP addresses, ports, and vulnerabilities. The IP addresses specify the range of hosts to scan for, the ports specify the range of services to connect to, while the vulnerabilities specify which type of vulnerabilities the scan process is looking for. I referred to the Open Web Application Security Project (OWASP) which published the top 10 IoT vulnerability categories [17]. Among the discussed vulnerabilities is insecure data transmission and storage which is relevant to NMEA as the messages are transmitted in plain text and are susceptible to a wide range of attacks. In this paper, I propose a new attack technique by exploiting the insecure manner in which NMEA messages are transmitted and using them in fingerprinting specific targets and gathering victim information for targeted attacks.

3 Cyber-Physical Tracking of Maritime Components

In this section, I describe my proposed methodology for cyber-physical tracking of maritime components. An overview of my approach is depicted in Figure 1.

My hypothesis is that some maritime components such as vessels have both cyber and physical properties. Cyber properties, include; among others, IP address, services, data, and vulnerabilities. These properties can be recorded by Internet-wide scanners such as Shodan if they are publicly communicated through the Internet. The data might include maritime-specific protocols such as NMEA which I propose to be used to fingerprint maritime components. On the other hand, the physical properties include; among others, navigation information such as location, speed, heading, and time of fix. Some marine tracking services such as MarineTraffic receive such information from various sources, record them and make them available for the public. My approach for correlating these two resources towards the identification and tracking of maritime components is guided by the state-of-the-art process of IoT scanning presented in my earlier work [1]. The process starts with selecting suitable scanner tools and configuring them with the suitable parameters to achieve the objective of the scanning process. Then, the scanning process is initiated and the results are collected. Afterward, the results are validated and analyzed. A detailed description of each step is discussed hereafter.

3.1 Scanner Selection

There are many networks and IoT scanners discussed in the literature. However, Shodan and Censys are the most referenced as stated in my earlier work [1]. The

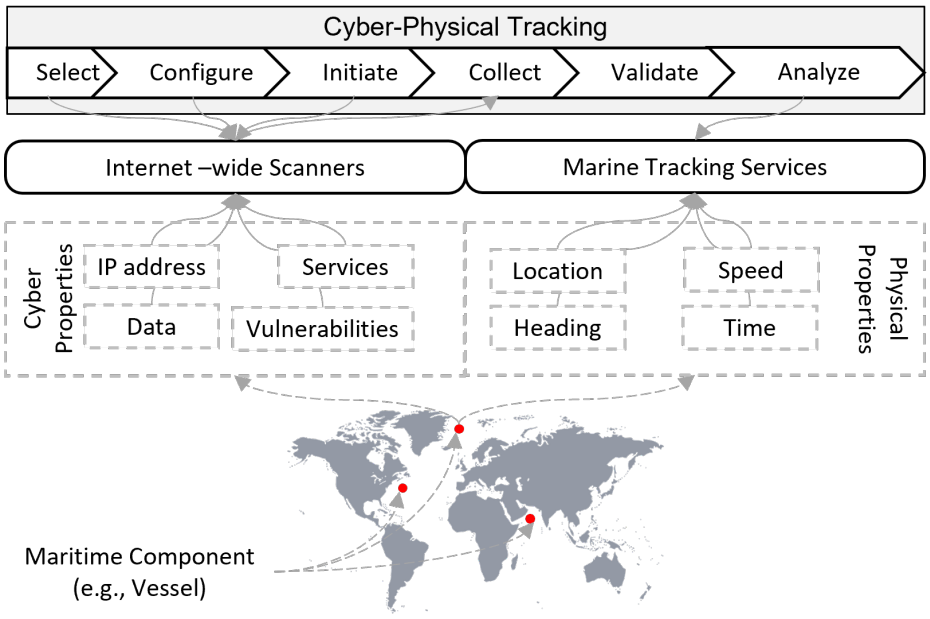

Fig. 1. Overview of my methodology for Cyber-Physical Tracking of maritime components

same notion is observed in several works in the literature. Li et al[13] have presented a survey of Internet-wide scanners including Shodan, Censys, ZoomEye, BinaryEdge, and Fofa. Another review of several public network vulnerability scanners is presented by Tundis et al [32]. The authors have discussed and evaluated Shodan, Censys, ZoomEye, Thingful, and PunkSpider. Moreover, Fofa and BinaryEdge scanners have been utilized in state-supported cyber activities as referred to in the report regarding Irans secret cyber files recently this year [11].

I have previously discussed two types of scanning approaches implemented in the different scanner tools, namely, passive scanning and active scanning [1]. Active scanning is the act of actively attempting to initiate connections with devices in a certain scope that can include the entire internet and recording their responses. On the other hand, passive scanning is the act of querying an indexed database that stores results of previous active scanning activities [1]. In this paper, I have followed the passive scanning approach to avoid any access violations. Therefore, I considered the most common scanners, namely, Shodan [16] and Censys [8], all of which allow for passive scanning while they are conducting active worldwide scanning activities. Other tools such as BinaryEdge, ZoomEye, Fofa, and others are considered for future work.

3.2 Scanner Configuration

Scanner tools utilize specific configurations to be able to query their databases. The configuration plays an important role in the outcome of the scanning process and could lead to the success or failure to meet its objectives. Considering the objective of this scanning process is to fingerprint maritime components worldwide, the configuration should include maritime-specific elements that lead to the desired outcome. Therefore, I propose the utilization of NMEA messages to fingerprint maritime components. The NMEA messages are employed within the queries leading to the identification of possible components. I propose the utilization of the static information in NMEA messages in the fingerprinting process (refer to section 2). Other configurations such as IP range, and ports are not considered relevant in this scanning process.

There are more than a hundred standard types of NMEA messages (i.e. +100 MessageIDs) that can be emitted by more than a hundred types of devices (i.e. +100 TalkerIDs). There is some commonality in the types of messages emitted by certain devices, and some messages are not expected to be emitted by a specific set of devices. However, there are no guidelines that can provide this information. Therefore, I have followed a comprehensive approach for scanning all TalkerID and MessageID pairs in an attempt to scan for all possible NMEA messages. This results in +10000 search queries required to cover all possibilities. A limitation of this approach has been observed during implementation. Some scanner tools limit the number of queries for each user under certain subscription plans. For instance, BinaryEdge and Censys allow for only 250 queries a month for a free subscription. Therefore, for such scanners, I have followed a rather limited yet focused approach for bypassing this issue. My alternative approach is only to query the most common TalkerID and MessageID pairs. Still, I was able to implement the comprehensive approach using Shodan. Additionally, Raymond [26] has compiled comprehensive documentation of NMEA messages which I have relied upon in this paper. Raymond listed a group of uncommon NMEA messages as well as vendor-specific messages. I have included such messages in my scanning scope aiming to achieve comprehensive coverage of NMEA messages. Nevertheless, the standard refers to other vendor-specific messages with a structure that is hard to predict such as starting with the letter "P" or starting with the letter "U" followed by a group of numbers. This means that my coverage of NMEA messages, although comprehensive, it is yet not complete.

The output of both approaches is a group of queries that are used in the next step. To this end, I have developed a group of scripts that can generate all these queries and make them ready to be sent to the Shodan and Censys APIs. The queries are configured to look for banners that include NMEA messages. Table 1 present examples of such queries.

3.3 Scanner Initiation

For this step, I have developed a group of scripts to run all the generated queries against the Shodan and Censys APIs and record the results for analysis. I highlighted the issue of passive scanning with regards to the freshness of results in

Table 1. Examples of query strings

TalkerID	MessageID	Description	Query string
GP	RMC	The static data of a Recommended Minimum Navigation Information (RMC) message emitted by a GPS Device (GP)	"$GPRMC,"
	GGA	The static data of a Global Positioning System Fix Data (GGA) message emitted by a GPS Device (GP)	"$GPGGA,"
GL	HDT	The static data of a Heading - True (HDT) message emitted by a GLONASS Device (GL)	"$GLHDT,"
PGRMZ		A vendor-specific message emitted by Garmin devices containing altitude information.	"$PGRMZ,"

my earlier work [1]. Some queries might return hosts that have been recorded emitting NMEA messages at the time the active scanning was conducted. However, this might not always reflect the correct status of that host. Nevertheless, it has been highlighted by Bennett et al [3] that both Shodan and Censys can reflect updates within 24 hours. Additionally, the IP addresses of the devices emitting NMEA might change overtime, therefore, repeating the scanning process periodically is needed to maintain the most accurate and up-to-date results.

I repeated the search process several times against the Shodan API following the comprehensive approach discussed in section 3.2. However, I followed the focused approach against Censys without repetition due to the limited subscription plan.

3.4 Collection

The query results are stored in files with different formats corresponding to the different scanning tools. I collected records with information including:

- The number of hosts observed to emit each NMEA message. This would shed a light on the most common messages.
- For each observed occurrence of NMEA message by a host, record the host IP, port number or service, country, and banner data containing the message. This information can be utilized for vulnerability analysis and the identification of maritime components. The port numbers as well as the banner data are expected to provide information regarding the device or software that is used for this service. Such information is valuable to attackers during the reconnaissance stage of cyber attacks.
- Record the results of all queries for validation.

3.5 Validation

The validation at this step refers to ensuring that the scanning results are correct and are useful to achieve the scanning objectives. Otherwise, the process is re-initiated with different scanner tools, configuration, or collection approaches. For this use case, it is necessary to validate that the identified hosts are actually emitting NMEA messages.

It is crucial to understand the different employed scanners as each scanner employs a unique query functionality that determines the quality of returned results concerning the scanning objective. For instance, Shodan does not have an "exact match" feature for queries. Instead, queries return results that approximately contain the query string. This has lead to getting false positive matches. The reason behind this is that the string of certain NMEA messages may appear in banners grabbed by hosts but that banner is not relevant to an NMEA service. An observed example of this issue is the query string "$TRACK," which is employed to scan for NMEA message "ACK" (Alarm Acknowledgement) emitted by Talker ID "TR" (TRANSIT Navigation System). Shodan removes the special characters from the query string and the remaining phrase "TRACK" appears in the banner data of many hosts but not as NMEA messages. Therefore, a validation process is required to ensure that only hosts emitting NMEA messages are identified and their data are collected. For this, I have developed scripts that will read all the returned results and only return results that contain correct NMEA messages.

3.6 Analysis

During this step, I analyzed the collected search results discussed in section 3.4. The analysis is different for each scanner tool as each one returns different results with different information. I will highlight the analysis process for the search results obtained from Shodan since it returned the largest amount of results. After removing the duplicate results, I have observed 4992 unique NMEA sessions emitted by 4942 hosts. The session information includes the host IP, port number, country code, banner data as well as a summary of NMEA messages in the banner data. The latter led me to the identification of additional NMEA messages that were outside the scope of the search (refer to section 3.2) but appeared to accompany the messages within the scope. The analysis process included four activities, namely, device identification, general statistics about the NMEA service, maritime component identification, and comparison between the different scanner tools.

Device Identification The identification of IoT devices and their operating system (OS) is among the challenges highlighted in my earlier work [1]. Banner data, port numbers, certificates, and other information have been employed in the literature to identify device types and OSs. This information is afterward employed in the identification and analysis of the vulnerability of such devices and OSs.

I have attempted to identify devices following several approaches. First, a generic classification is possible using the NMEA format. The type of device from which the NMEA message is coming is encoded in the TalkerID. Although the type of IoT device that might be forwarding the messages cannot be identified through this approach, nevertheless, it can shed a light on the type of devices connected to the host. Such information is useful for attackers at the reconnaissance stage. Table 2 depicts the number of detected hosts for each NMEA talker.

The table reflects that the majority of devices are receivers of the major positing systems, namely, Global Positioning System (GPS), GLONASS, and a combination of many systems. Moreover, the quantity of vendor-specific NMEA talkers is observed. Therefore, the second approach relied on vendor-specific messages.

Table 2. Distribution of type of NMEA talkers across the detected hosts

Talkers (Description)	Host Count (%)	Talkers (Description)	Host Count (%)
GP: GPS receiver	4897 (99%)	GB: BeiDou receiver (China)	4 (0%)
Vendor-Specific	1939 (39%)	CC: Computer - Programmed Calculator	2 (0%)
GN: Combination of multiple satellite systems	1595 (32%)	II: Integrated Instrumentation	2 (0%)
GL: GLONASS receiver	1558 (32%)	DF: Direction Finder	1 (0%)
BD: BeiDou receiver (China)	115 (2%)	VW: Velocity Sensor, Speed Log, Water, Mechanical	1 (0%)
GA: Galileo receivere	74 (2%)	SD: Depth Sounder	1 (0%)
AB: Independent AIS Base Station	13 (0%)	YD: Transducer - Displacement, Angular or Linear	1 (0%)
WI: Weather Instruments	10 (0%)	PQ: Quectel Quirk	1 (0%)

Relying on several online resources, I was able to identify some device types known to emit the most common vendor-specific messages based on their TalkerID code. Additionally, I used the National Vulnerability Database (NVD) published by NIST [25] to find possible Common Vulnerabilities and Exposures (CVE) by using the identified device information. I also recorded the CVE's risk ratings that are encoded using the Common Vulnerability Scoring Scheme (CVSS). Table 3 show the identified device types, the number of hosts that emits them, and possible CVEs associated with these devices.

NMEA service In this analysis, I focused on the most observed messages, ports, and countries to stand on the status of NMEA service worldwide. This information is helpful to the cybersecurity community to manage risks related to NMEA.

Regarding message types, I have observed 4 types of AIS messages; ABVDM, AITXT, AIVDO and AIVDM, 41 NMEA messages that are specified in the NMEA-0183 standard [2], 34 messages following the standard specifications for vendor-specific messages, and 29 messages that have no specified description in the standard, however, they have a format similar to NMEA. Table 4 reflects the most observed messages, all of which are standard NMEA messages, brief description, and the number of hosts that emit them. Note that 83,25 % of the observed hosts emit at least two different NMEA messages together. Each

Table 3. The identified device types and some of their possible vulnerabilities

Messages	Description	Host Count (%)	Possible CVEs (CVSS)
Most common: PMTKAGC, PMTKGALM, PMTKGEPH ,PMTKTSX1	MediaTek MTK chipsets	1662 (97,6%)	CVE-2020-13841 (9.8) CVE-2020-13842 (7.8)
PSTT	Saab Systems position receiver	35 (0,9%)	None
PCPTI	Cradlepoint Router	28 (0,7%)	
PLEIR	LEICA GPS receiver	21 (0,5%)	
PTNL	Trimble GNSS Receiver	3 (0,1%)	CVE-2012-5053 (4.2)
PQXFI	Qualcomm chipset	1 (0,0%)	CVE-2021-1965 (9.8) CVE-2021-1955 (7.5)

message provides different valuable information for the Cyber-Physical tracking process. Among the messages in the table, GGA and RMC messages together provide the most amount of information including time, position, speed, heading, and others. Therefore, they are great candidates for fingerprinting maritime components.

Table 4. Top 10 observed NEMA messages emitted through the Internet

Message	Description	Host Count (%)
GGA	GPS Fix Data including position and time information	4815 (96%)
RMC	Recommended Minimum Navigation Information including position, time, speed, and heading.	4145 (83%)
VTG	Track made good and Ground speed	4019 (81%)
GSA	GPS Dilution of precision (DOP) and active satellites	3142 (63%)
GSV	Satellites in view	3077 (62%)
GLL	Geographic Position - Latitude/ Longitude	63 (1%)
ZDA	Time & Date	34 (1%)
GNS	Fix data	16 (0%)
DBT	Depth below transducer	14 (0%)
GST	GPS Pseudorange Noise Statistics	13 (0%)

Regarding used ports, I discovered 92 ports used for transmitting NMEA messages, the majority of which are transmitted through two TCP ports, specifically, port 7000 (50%) and port 50100 (45%). This indicates that scanning these ports alone would cover 95% of the entire NMEA presence on the Internet. Furthermore, I have observed that approximately 10.5% of hosts emitting NMEA messages over a certain port have more than one other port open ranging between 2 to 100 ports including ones used for other services such as HTTP and

Message Queuing Telemetry Transport (MQTT). These other services might as well have their own vulnerabilities. However, my analysis didn't pursue this issue any further.

Regarding countries, in total 66 countries have hosts emitting NMEA messages. The majority of hosts have IP addresses registered in Brazil (78%), Argentina (3,3%), Spain (3,2%), Japan (2,9%), Morocco (2%), and United States (1,7%). Although the high number of NMEA messages in Brazil is unexpected, I have not investigated the reason behind any further in this paper.

Maritime components identification As mentioned in section 1, one of the objectives of this paper is to investigate the ability of attackers to fingerprint maritime components and identify their cyber-related information as well as physical information during the reconnaissance stage to aid during further stages of cyber attacks. In this section, I present my method and results for the identification of maritime components with observed presence on the internet, identify their cyber-related information (IP, ports, data, and vulnerabilities), and track their physical location to obtain combined cyber-physical records of the components. My approach relies on the following assumption, a host is considered a maritime component under two conditions, its communicated coordinates are located at a sea area or it is emitting an NMEA message with a talker that is an AIS base station.

Based on that, I developed an algorithm that will parse the NMEA banner data for each host, and obtain valid coordinates information (latitude and longitude) from either RMC or GGA messages. Then I utilized an algorithm provided by Karin [12] to check if these coordinates belong to a sea or land area. Moreover, if an AIS base station Talker ID is observed, a component is labeled as a possible maritime component. The results of the algorithm are depicted in Table 5. I have detected 331 possible maritime components, obtained their cyber as well as physical information. To evaluate my algorithm. I have manually and randomly verified some of the obtained results. I have randomly chosen 10 detected land positions, 10 sea positions and checked; using Google Maps, if they are accurately labeled. The results suggests that my algorithm returns valid results. Further development and evaluation are expected for future work by utilizing vessel tracking services to correlate the cyber and physical properties.

Table 5. Results of the maritime component fingerprinting algorithm

Maritime Component?	Rational	Count
No	No Evidence	159
No	Land position	4502
Yes	Sea position	325
Yes	AIS Base Station	6

Comparison between scanners In this section, a comparison is presented for the two most referenced scanners, namely, Shodan and Censys concerning the scanning process in this paper. Table 6 depicts a summary of this comparison. Shodan provided the best possible results for analysis due to a sufficient subscription plan. Therefore, this comparison doesn't reflect the actual utility of each scanner. However, it justifies the focus of the analysis on the records collected from Shodan.

Table 6. Comparison between NMEA queries between Shodan and Censys

	Scanner	**Initiated Queries**	**Messages Detected**	**Collected Records**
	Censys	182	52	9582
	Shodan	12206	53	22726
	Shodan more	**Censys more**	**Same results**	**Both 0**
# of Messages	34	18	5	99

4 Discussion and Limitations

Discussing the risks of the Cyber-Physical tracking process can be conducted by considering the risks of the associated *ATT&CK* techniques. *ATT&CK* [29] is a common knowledge repository for observed cyber adversarial behaviors. The presented Cyber-Physical tracking process in this paper emulates an adversarial behavior that includes several techniques indicated in the *ATT&CK* framework. The relevant techniques to this paper are i) Gather Victim Host Information (T1592) [20], ii) Search Open Technical Databases (T1596) [22], and iii) Search Open Websites/Domains (T1593) [23]. In my approach, I have fine-tuned the scanning process by searching open websites and domains as resources for identifying information such as vendor-related information. Also, I have searched technical databases such as Shodan, Censys, and NVD. Moreover, I have gathered the victim host information such as IP address, ports, possible device type as well as possible vulnerabilities. The IP addresses and ports can later be utilized for subsequent adversarial techniques to gain initial access to the victims' networks or impact the operations of the emitting devices. Initial access might later be achieved through External Remote Services (T1133) [18]. Considering that the NMEA messages were detected from the internet, this indicates that each emitting device has at least one external-facing remote service that is remotely accessible. Moreover, as mentioned in Section 3.6, 10.5% of the detected NMEA-emitting hosts have between 2 to 100 remote services open. If any of these services has a vulnerability that can be remotely exploited, it may lead to enabling the attacker to gain an initial foothold to the connected network. Additionally, attackers may attempt to inflect impact through remote Network

Denial of Service (T1498) [19] in the case that the NMEA talker is susceptible to such vulnerability.

The *ATT&CK* framework refers to the difficulty of mitigating the techniques T1592, T1596, and T1593 as they are performed outside the scope of the defensive capabilities of organizations. However, efforts to minimize the availability and sensitivity of data to external parties are suggested. The *ATT&CK* framework mentions the very high occurrence and associated false positive rate of such activities. This is reflected in my work as the ability to scan is always possible even without a proper subscription. Additionally, a certain false positive rate is expected due to the passive scanning approach. The obtained results might not reflect the actual status of hosts. However, the obtained NMEA messages from the banner data include several fields containing the Coordinated Universal Time (UTC) the information was captured which can reflect the freshness of the scanning result. Additionally, it has been reported that both Shodan and Censys reflect updates within 24 hours [3].

Similar adversarial techniques have been observed by the cybercriminal group "Sandworm Team" [21] during the development of the NotPetya attack. The Sandworm team searches open websites and databases for information to craft credible spearfishing emails [4]. This incident highlights the utility of such available resources to attackers and the necessity to investigate such threats in different domains including maritime.

The maritime sector is witnessing a digital transformation era leading to expected drastic changes in technology, business models and operations [10]. A new operational model for future maritime components has been communicated by members of the classification society in maritime, specifically, the Norwegian organization DNV. The operational mode is called auto-remote; autonomous as possible and remotely controlled when needed [7]. Tam and Jones [30] have discussed the unique cyber-physical opportunities in specific geological locations when considering futuristic unmanned ships. The authors indicated the utility of such opportunities to pirates adopting cyber attack techniques. Therefore, the associated risks of the demonstrated approach in this paper are increased when considering the auto-remote operational mode.

The field of IoT vulnerability scanning is recent and growing [1]. Also, the field of Cyber-physical tracking is scarce as very limited artworks have discussed it. However, I argue that immediate actions are needed for demonstrating the feasibility and possible impacts of such activities. This is important to support the ongoing efforts for improving cybersecurity in maritime. Therefore, I acknowledge the following limitations in the proposed approach and discuss the rationals for dealing with them:

- The choice for utilizing Shodan is only to present a proof of concept for the proposed approach. Other scanner tools might provide different results. For instance, it has been communicated by Li et al [13] that ZoomEye scans over 1.2 billion devices compared to only 0.4 by Shodan. Therefore, evaluating the proposed approach using ZoomEye and other scanner tools is considered for future work.

– The scope of the utilized NMEA messages in the fingerprinting process is limited by the discussed messages in the NMEA-0183 standard [2] as well as the comprehensive documentation of NMEA protocol by Raymond [26]. Other messages that are not documented or that have not appeared in the search might exist but are still undetected. However, the results suggest that only a few of the NMEA messages types constitute the majority of the detected messages which might render the impact of any missing messages insignificant.

5 Conclusion

Cyber security in the maritime domain is a growing area of interest due to the undergoing digital transformation. The maritime infrastructure is consuming additional digital components including IoT and Industrial IoT [15, 5]. The future of the maritime domain includes new modes of operation (e.g. auto-remote) that require increased connectivity and reduces human presence around maritime components. Risk management activities in maritime have been proposed by the International Maritime Organization (IMO) and documented in Resolution MSC. 428(98) [6]. Such activities include analyzing the threat landscape and continuous improvement of defenses.

This paper supports the efforts in this direction as it demonstrates an offensive capability that can be conducted by attackers to gain tactical advantages by combining cyber and physical information regarding maritime components and utilize them during the development of more directed cyber-physical attacks. I have presented a new approach for scanning and identifying cyber-related information for NMEA emitting devices. The scanning process yielded in identifying 4942 hosts emitting NMEA messages the majority of which using ports 7000 and 50100. I have also identified several device types and expected vulnerabilities. Such information aims to capture the status of NMEA service worldwide to attract attention toward improved cybersecurity.

Additionally, I have proposed a new approach for detecting maritime components that are connected to the internet. The algorithm utilized information collected from IoT scanners of hosts emitting NMEA messages some of which include position information. The algorithm detected 331 maritime components that are connected to the internet. It identifies their location, speed, and other physical properties in addition to their IP addresses, ports, and other cyber properties. Such components could be susceptible to cyber-physical attacks. In summary, I argue that the Cyber-Physical tracking process constitutes a threat against the detected maritime components and I urge the maritime community to consider the outcome of this work.

References

1. Amro, A.: Iot vulnerability scanning: A state of the art. Computer Security pp. 84–99 (2020)

2. Association, N.M.E., et al.: Nmea0183 standard. =https://www.nmea.org/content/STANDARDS/NMEA$_0$183$_S$*tandard*(2002)
3. Bennett, C., Abdou, A., van Oorschot, P.C.: Empirical scanning analysis of censys and shodan
4. Brady., S.W.: United States vs. Yuriy Sergeyevich Andrienko et al (2020, October 15), https://www.justice.gov/opa/press-release/file/1328521/download
5. Chubb, N.: Maritime Applications for IoT ((accessed September 8, 2021)), https://thetius.com/maritime-applications-for-iot/
6. Committee, T.M.S.: International maritime organization (imo) (2017) guidelines on maritime cyber risk management. http://bit.ly/MSC428-98
7. DNV GL: Dnvgl-cg-0264: Autonomous and remotely operated ships (2018)
8. Durumeric, Z., Adrian, D., Mirian, A., Bailey, M., Halderman, J.A.: A search engine backed by internet-wide scanning. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. pp. 542–553 (2015)
9. Fiorini, M.: Maritime awareness through data sharing in vts systems. In: 2012 12th International Conference on ITS Telecommunications. pp. 402–407. IEEE (2012)
10. Fruth, M., Teuteberg, F.: Digitization in maritime logistics—what is there and what is missing? Cogent Business & Management **4**(1), 1411066 (2017)
11. Haynes, D.: Iran's secret cyber files on how cargo ships and petrol stations could be attacked (Jul 2021), https://news.sky.com/story/irans-secret-cyber-files-on-how-cargo-ships-and-petrol-stations-could-be-attacked-12364871
12. Karin, T.: Global land mask. https://github.com/toddkarin/global-land-mask (October 2020)
13. Li, R., Shen, M., Yu, H., Li, C., Duan, P., Zhu, L.: A survey on cyberspace search engines. In: China Cyber Security Annual Conference. pp. 206–214. Springer, Singapore (2020)
14. 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. pp. 796–807 (2002)
15. Maritime, T.: The Internet of Things Makes Waves on a Global Maritime Network ((accessed September 8, 2021)), https://telenormaritime.com/digital-shipping/internet-of-things-iot/
16. Matherly, J.: Complete guide to shodan. Shodan, LLC (2016-02-25) **1** (2015)
17. Miessler, D., Smith, C.: Owasp internet of things project. OWASP Internet of Things Project-OWASP (2018)
18. MITRE: External Remote Services (T1133) (2021 (accessed November 2, 2021)), https://attack.mitre.org/techniques/T1133
19. MITRE: Network Denial of Service (T1498) (2021 (accessed November 2, 2021)), https://attack.mitre.org/techniques/T1498/
20. MITRE: Gather Victim Host Information (T1592) (2021 (accessed September 8, 2021)), https://attack.mitre.org/techniques/T1592
21. MITRE: Sandworm Team (2021 (accessed September 8, 2021)), https://attack.mitre.org/groups/G0034/
22. MITRE: Search Open Technical Databases (T1596) (2021 (accessed September 8, 2021)), https://attack.mitre.org/techniques/T1596
23. MITRE: Search Open Websites/Domains (T1593) (2021 (accessed September 8, 2021)), https://attack.mitre.org/techniques/T1593
24. Munro, K.: Tracking hacking ships with shodan ais (Jan 2018), https://www.pentestpartners.com/security-blog/tracking-hacking-ships-with-shodan-ais/

25. NVD, N.: National vulnerability database (2011)
26. Raymond, E.S.: https://gpsd.gitlab.io/gpsd/NMEA.html
27. Sivkov, Y.: Transformation of nmea ship network from sensor-based to information-based model. In: 2018 20th International Symposium on Electrical Apparatus and Technologies (SIELA). pp. 1–4. IEEE (2018)
28. statista.com: Internet of things (iot) connected devices installed base worldwide from 2015 to 2025. https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
29. 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)
30. Tam, K., Jones, K.: Cyber-risk assessment for autonomous ships. In: 2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security). pp. 1–8. IEEE (2018)
31. Tran, K., Keene, S., Fretheim, E., Tsikerdekis, M.: Marine network protocols and security risks. Journal of Cybersecurity and Privacy **1**(2), 239–251 (2021)
32. Tundis, A., Mazurczyk, W., Mühlhäuser, M.: A review of network vulnerabilities scanning tools: types, capabilities and functioning. In: Proceedings of the 13th International Conference on Availability, Reliability and Security. pp. 1–10 (2018)