Received <day> <Month>, <year>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

ARTICLE TYPE

VDES R-Mode: Vulnerability Analysis and Mitigation Concepts

Francisco Lázaro | Ronald Raulefs | Hannes Bartz | Thomas Jerkovits

Institute of Communications and Navigation, German Aerospace Center (DLR), Germany

Correspondence
Francisco Lázaro Blasco
Institute of Communications and Navigation, German Aerospace Center (DLR), Muenchnerstr. 20, 81245 Wessling, Germany. Email: Francisco.LazaroBlasco@dlr.de

Abstract

VDES R-Mode aims at providing a contingency maritime positioning and navigation system when the operation of Global Navigation Satellite Systems (GNSS) is disrupted. However, VDES R-Mode, similarly to GNSS, can itself also be subject to different types of attacks, such as jamming or spoofing. In this paper, we evaluate the vulnerabilities of VDES R-Mode, and discuss the effectiveness and cost of different types of countermeasures. The outcome of this cost-benefit analysis is a recommendation to introduce authentication for the navigation messages of R-Mode using the TESLA protocol.

KEYWORDS:
VDES; VDE; AIS; R-Mode; Authentication; TESLA;

1 | INTRODUCTION

Navigation and timing information is of key importance in maritime electronic information systems. Global Navigation Satellite Systems (GNSS) services provide the mariner with accurate navigation and timing information. However, GNSS services can be disrupted either due to a malfunction or due to intentional attacks. In 2017 the UK government analyzed the impact on society of the loss of GPS for five days[1]. The conclusion for the mariner was that there is currently no backup system available and the only option left to him is to look out of the window. In fact, this dependence of GNSS in the maritime domain had already been identified by International Association of Lighthouse Authorities (IALA)[2], which led to the development of requirements for a contingency system to GNSS. The contingency system is foreseen to mainly use terrestrial infrastructure to support the vessel in port approaches, coastal areas and, where possible, also the open sea. Currently, a system called Ranging-Mode (R-Mode) is being investigated as a contingency system to GNSS. R-Mode is planned to be deployed either as a stand-alone system in the MF band or in the VHF band, being offered as a navigation service in the VHF Data Exchange System (VDES).

VDES is currently standardized as a communication standard by the International Telecommunication Union (ITU) and is expected to be published as ITU-R M.2092-1 by 2022. The standard is built on the IALA guideline G1139[3]. VDES can provide various services including VDES R-Mode. Given the fact that VDES R-Mode is provided as a service, and that positioning is only possible when being within the so-called ranging coverage of several shore stations, R-Mode will only be available in coastal regions in which a VDES infrastructure is present. VDES R-Mode is planned to provide a lower positioning accuracy than GNSS, in particular the positioning accuracy is planned to be in the range between 10 and 100 meters, depending on the environment and distance to the base stations[2]. However, the great added value of VDES R-Mode is available in coastal areas and beyond when GNSS is not available.

Nowadays, maritime communication devices utilize the existing Automatic Identification System (AIS) by broadcasting various data content beyond the original intended positioning reports. In the near future, when VDES is deployed, vessels will be able to communicate with each other as well as with the maritime authorities using the larger capacity offered by VDES. Furthermore, maritime authorities will also be able to deploy new services, of which R-Mode is a prominent example. This will represent a major step forward in e-navigation, and will enable, among others, a safer navigation for unmanned vessels. At the

time of writing this article, VDES is planned exhibit only exiguous security features, partly motivated by the scarce bandwidth available. In particular, in VDES it is foreseen to consider Elliptic Curve Digital Signature Algorithm (ECDSA) to sign the bulletin board messages sent by shore stations. The bulletin board allocates the communication resources of the operating area of the terrestrial VDE base station. Additional security features are not planned, although security protocols can be used at higher layers.

In this paper we analyze the security of VDES R-Mode. In particular, we describe different known vulnerabilities of R-Mode and analyze which countermeasures can be used to mitigate them. Among the different countermeasures, navigation message authentication using the Timed Efficient Stream Loss-Tolerant Authentication (TESLA) protocol seems to be the most promising, due to its good trade-off between complexity and security.

The paper is structured as follows. Section 2 provides an introduction to VDES and an overview of its security capabilities. Section 3 describes VDES R-Mode, a positioning and navigation service offered through VDES. Section 4 presents a security assessment of VDES R-Mode. Section 5 describes non-cryptographic and cryptographic countermeasures that are applicable for VDES. Section 6 discusses the countermeasures and provides a future proof path for VDES R-Mode. Finally, Section 7 presents the conclusions.

2 | OVERVIEW OF VDES AND ITS SECURITY CAPABILITIES

2.1 | Overview of VDES

The VHF Data Exchange System (VDES) is composed of three subsystems and described in the IALA guideline G1139[3]: Automatic Identification System (AIS), Application Specific Messages (ASM) and VHF Data Exchange (VDE). AIS was developed in the 1990s[4]. Although several message types are supported, the most common AIS messages are reports of the current position and speed of a vessel, whose main purpose is collision avoidance. AIS is mandatory for large (class A) vessels[4], although it is also widely used voluntarily by smaller vessels (class B). AIS also supports coastal base stations operated by local shore authorities, which, among others, serve to gather navigational information. Furthermore, it is possible to receive AIS signals from satellites and airplanes[5,6], despite the fact that the system was not originally designed for this purpose. Given that AIS is overloaded in some regions, a parallel communication system, ASM, was introduced to offload AIS by carrying the so-called application-specific messages, which are not related to collision avoidance. Finally, the third component, VDE, offers a new communication channel to transmit arbitrary data, which represents a big step towards a fully digitized maritime VHF band. Besides bidirectional shore-to-ship and ship-to-ship communication, VDE foresees also communication over LEO satellites in regions which are located too far away from the shore control stations. VDES assumes a half-duplex operation for the terminals, i.e., the terminals are able to transmit and receive, but not simultaneously. As a consequence, and considering the fact that collision avoidance has priority over all other services, ASM and VDE must in general make sure that they do not interfere with AIS. In other words, VDES terminals are not allowed to transmit at time slots in which an AIS transmission is expected.

Figure 1 shows an overview picture of potential maritime entities interacting with each other by using the different VDES subsystems. The increased capacity of digital maritime systems will allow the emergence of new applications. Among others, the terrestrial VDE (VDE-TER) link will allow vessels located in coastal areas in the vicinity of shore-stations to communicate control data with the ship owner. In the open sea, or in coastal areas in which no VDE infrastructure is present, satellite VDE (VDE-SAT) will allow vessels to share crucial information about the on-board cargo with customers, or vice-versa. Furthermore, VDE-SAT could prove very useful to provide vessels with updated ice-chart maps[7].

2.2 | VDES Security Capabilities

The International Maritime Organization (IMO) has recently issued some guidelines on maritime cyber risk management[8], as well as a resolution[9] encouraging national administrations to address cyber security in their safety management systems. When dealing with cybersecurity, one usually distinguishes three different aspects: availability, confidentiality and integrity/authenticity. Availability has to do with the system staying in operation and providing the services it was designed for. Confidentiality consists of concealing information from (unauthorized) third parties, while allowing the legitimate recipient of the information to recover it. Finally, the purpose of (data) authentication is to ensure that the information was generated by a trusted source, and that it has not been modified on the way to its recipient (this latter aspect is usually known as data integrity).

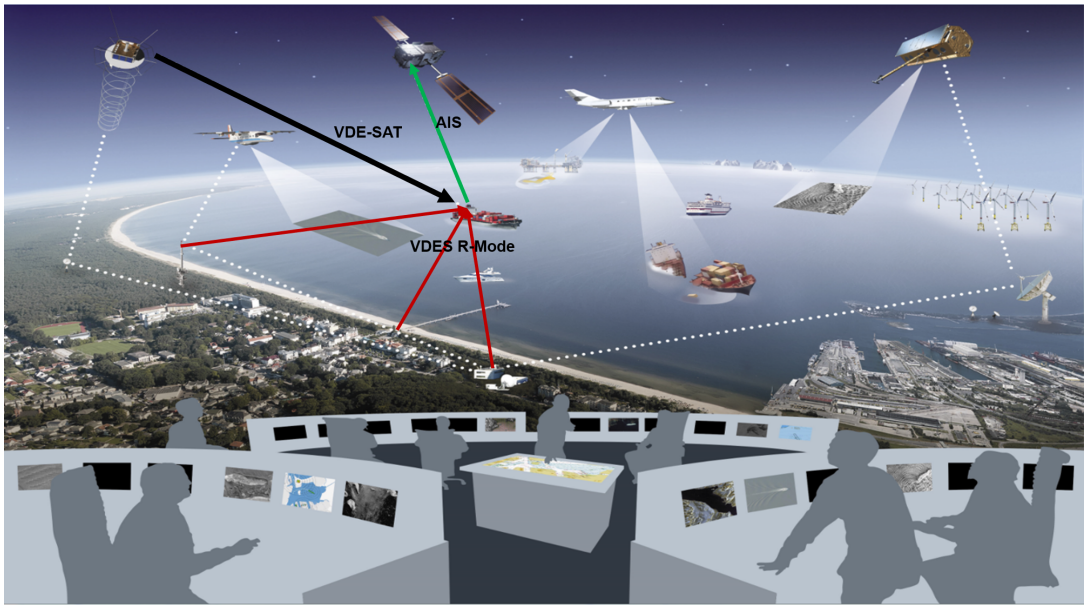

FIGURE 1 Interacting maritime entities: VDES R-Mode provided by three base stations on shore in red, and in addition in black a VDE-SAT link provides navigation data. Further, the vessel broadcasts its AIS message to an AIS satellite.

Despite the existence of guidelines related to cybersecurity, VDES makes very limited use of cryptographic mechanisms for cybersecurity protection. In particular, confidentiality is so far not considered in VDES, i.e., all information transmitted is public (unless the information is protected using end-to-end encryption in a higher layer protocol, and this is in any case application specific). However, VDES does foresee to protect the authenticity of some messages. In particular, in terrestrial VDE the bulletin board[1] broadcast by base stations is foreseen to carry a digital signature[2]. In case the signature verification fails the message is flagged as not authenticated. It is important to remark that, in order to be able to verify the digital signatures, ships must have access to a Public Key Infrastructure (PKI), a concept which is explained in the next section.

It is not clear whether confidentiality will play a big role in VDES in the near future. So far the communication between the maritime authorities and the maritime traffic users is a public broadcasting service, and hence, data is meant to be public and sent in the clear (without encryption). However, having the means of establishing confidential communication channels could be of interest in many applications. For example, it could allow ships to provide confidential information such as a passenger list to a national authority while complying with the EU data protection rules.

The situation for authenticity is different, as it has been made visible by several spoofing[3] attacks in AIS [10], which can have dire consequences. A possible countermeasure to prevent such attacks is applying digital signatures to the transmitted packets, as demonstrated in the testbed of Wimpenny et al. [11]. Provided that terminals have access to a PKI, this approach is very effective. However, it results in a massive bandwidth expansion, since a single digital signature is several times larger than most of the messages exchanged in AIS. Thus, given that AIS operates already close to saturation in several regions, it is not clear whether such a solution is feasible in the near future. Bearing in mind that the security relevance of AIS is higher than that of VDE, a possible option could be using VDES to carry the authentication data of AIS messages. The first issue with such a solution is regulatory, whereas AIS is mandatory for Class A vessels, VDE is not. Thus, unless VDE is made itself mandatory for Class A vessels, authentication for AIS would be optional, and AIS would remain vulnerable to spoofing attacks. Even if this regulatory constraint is overcome by making VDE mandatory, one would still have the issue that sending the authentication for AIS over VDE would consume a large part of the bandwidth of VDE, at least in highly loaded regions. As a consequence, the already limited capacity of VDE would in practice diminish even further, making it difficult, if not impossible, to deploy further e-navigation services over VDE, and this was actually the original purpose of VDE. Another important aspect is complexity.

[1]The bulletin board defines the physical channels and MAC channels (the so called VDE slotmap) to be used in the service area of the control station which is sending the bulletin board.

[2]The signature scheme proposed is the ECDSA, with public keys of 256 bits and the message digest algorithm is SHA-256.

[3]Spoofing consists basically of injecting fake packets in the system. A more detailed treatment of spoofing attacks can be found in Section. 4

Verifying a digital signature requires considerable computing time (see Section 5.2.3 for details). Hence, having to verify the signature from all messages would require using expensive hardware, and low terminal cost is a key factor for the acceptance of VDES. In a nutshell, due to the scarce frequency resources of VDES and the high complexity of digital signature schemes, protecting authenticity by adding a digital signatures to all AIS or VDES messages does not seem a viable option. Thus, if one wishes to deploy a service with authenticity and or integrity guarantees over VDES, one can introduce *ad hoc* countermeasures to protect its authenticity.

2.3 | A Public Key Infrastructure for the Maritime Domain

The basis of any cryptosystem with support of digital signatures is a Public Key Infrastructure (PKI). However, there are multiple ways in which a PKI can be implemented and many of them are not feasible in the maritime domain (or in VDES) due to the political/regulatory constraints, the scarce bandwidth available and also due to the fact that ships sometimes are offline (with no live access to the PKI) for long periods of time. The CySiMS project[12], funded by the research council of Norway, had as objective the development of security solutions for the maritime domain, and among other aspects, outlined a PKI structure for the maritime domain. In particular, the proposed solution was employing X.509 certificates, which allow to tie a public key to its owner. This standard defines a hierarchical structure, so that trust in the system emanates from a root (or central) Certificate Authority (CA), whose role could be played by an international organization such as IMO. The root CA can issue (and revoke) certificates to national authorities, who can in turn issue certificates to other parties (ships).

In most applications of X.509, the certificates are retrieved online. However, in the maritime domain it cannot be assumed that ships always have (live) access to the PKI. For example, ships in the high seas lacking a satellite connection may have no connectivity at all with the PKI. Furthermore, even if a connection with the PKI can be established, in VDES it would not be possible/efficient to continuously request and download certificates from the PKI due to the limited bandwidth available to the system. The solution proposed in the CySiMS project is having a local cache for X.509 certificates in every ship. This local cache could be updated, for example, whenever a fast internet connection is available (maybe when the ship is docked). This cache is projected to have a size in the order of hundreds of megabytes. Furthermore, changes in the cache are expected to be slow, so that very limited bandwidth is required to keep the cache up to date.

Establishing the PKI proposed in the CySiMS project would have many advantages. First of all, it would enable VDES to use digital signatures for the bulletin board as foreseen in the standard. Second, building on the local cache of certificates ships could not only authenticate the messages they exchange, but also establish confidential (encrypted) communication sessions. This could be achieved either by directly encrypting the data with the public key of the intended recipient, i.e. using asymmetric encryption, or by running a key-exchange protocol and then making use of symmetric cryptography (e.g. AES). The latter method is more convenient when exchanging larger chunks of information. In spite of these important advantages, implementing such a PKI could prove to be challenging in the maritime domain. In particular, this solution requires every ship to have a so-called PKI unit, which stores the certificates and runs the cryptographic algorithms. This would increase the complexity, and thus the cost, of VDES terminals, and low-cost is expected to be a critical factor for the adoption of VDES. Nevertheless, in the opinion of the authors, the advantages in this case outweigh the disadvantages: establishing a PKI is the only way forward to make VDES cyber resilient and future proof.

3 | VDES R-MODE OVERVIEW

The VDES R-Mode System intends to provide a contingency Positioning, Navigation and Timing (PNT) system for maritime shipping. The operational concept foresees using VDES R-Mode when a disruption to Global Navigation Satellite System (GNSS) services on-board a ship occurs. Furthermore, when both GNSS and VDES R-Mode are available, it is possible to assess the PNT integrity by comparing the estimates obtained through both systems.

As already mentioned, VDES R-Mode is offered as a navigation service within VDES. In particular, VDES R-Mode requires sending two different messages over VDES. The first type is a so called navigation message. This message contains different navigation data, such as information about the location of the VHF antenna of the VDES shore stations to determine the distance, information about the (atomic) clocks of each shore station to track the changes in the timing between multiple transmissions of each station, and it also indicates which ranging sequence out of multiple predefined ones will be transmitted by the different stations. The second message type is a ranging sequence, i.e., a packet containing a known sequence. This sequence is used by

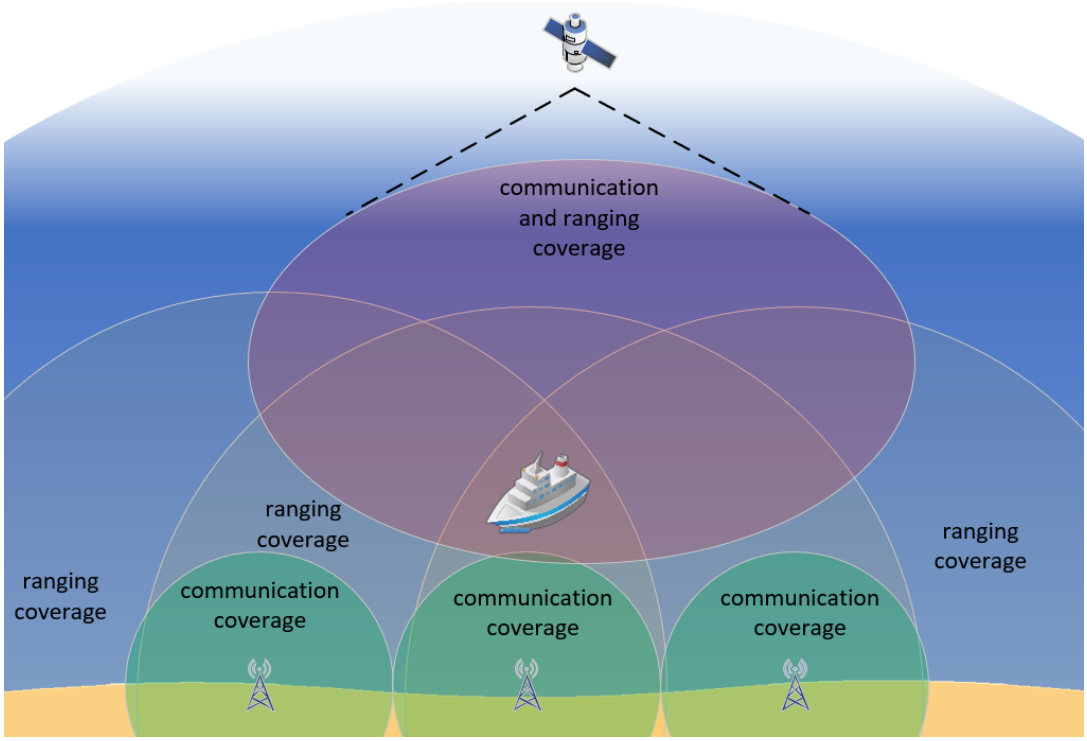

FIGURE 2 R-Mode diagram. The communication coverage of shore stations is considerably smaller than its ranging coverage. In this example the vessel is within the ranging coverage but outside the communication coverage of the shore stations. However, the vessel is within the communication and coverage radius of the VDE satellite, and this allows him to position itself.

the vessels to estimate their position. VDES R-Mode supports different ranging sequences, see IALA guideline G1158 for more details [13].

Each VDE-TER shore station allocates the data resources by time-division multiplexing access (TDMA) using a frame lenth of 1 s. Each shore station transmits navigation messages with a periodicity of 60 s and ranging sequences with a periodicity of 1 s. In order to provide a good positioning accuracy, VDES R-Mode requires neighboring shore stations to coordinate so that a shore station is silent in slots in which neighboring shore stations are transmitting their ranging sequence, see Fig 3 .

Shore stations are planned primarily taking into account their communication coverage in order to provide a continuous coverage in waters close to the shore, while preventing too much overlap between neighboring shore stations. However, the ranging coverage of a shore station is considerably larger than its communication coverage, since ranging boils doing to detecting the presence of a *known* sequence in the received signal, which is an easier task than recovering unknown data from the received signal (i.e, it can be achieved at a lower signal to noise ratio). In order to be able to access the VDES R-Mode service, a vessel needs up-to-date navigation data, and must thus receive the navigation messages error free with high probability. Thus, the vessel must be within the communication coverage of at least one VDE-TER shore station or alternatively within the communication coverage of a VDE-SAT satellite. Furthermore, the vessel must be within the ranging-coverage of at least 3 VDE-TER shore stations or VDE-SAT satellites, since it needs to compute its distance to at least 3 transmitters in order to be able to estimate its position. Fig. 2 depicts a scenario of one vessel that is receiving navigation data by a VDE-SAT link and is in range of multiple base stations to estimate its position. As shown in the figure, the satellite coverage area is significantly larger than that of the shore stations. Hence, VDE-SAT can be used to increase the service area of VDES R-Mode.

From an operational point of view, VDES R-Mode intends to operate as a contingency system for GNSS and has therefore, weaker performance requirements than GNSS. This includes the update rate of the position estimate as well as the accuracy requirements. The accuracy requirements range from 10 m for harbor approaches to 100 m on the open sea. Another difference to GNSS is that there is no continuous transmissions from each base station. This requires to relax the update rate of the position estimation. Finally, VDE-TER is a communication system and the remaining time slots unused by other base stations are foreseen for data communications.

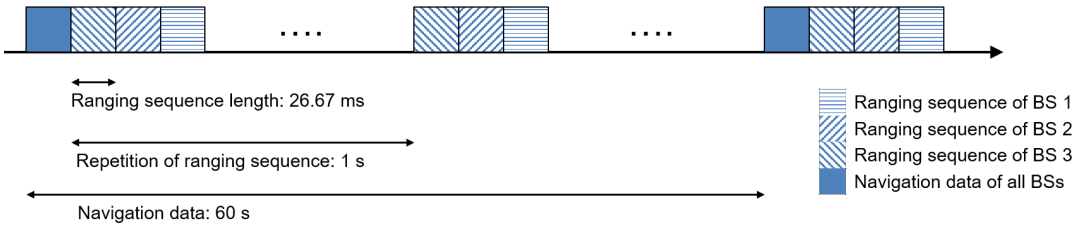

FIGURE 3 Time slot allocation of the ranging sequences of three different base stations and the navigation data from eithr one base station or a VDE satellite.

4 | A SECURITY ASSESSMENT OF R-MODE

VDES R-Mode, like any other navigation system, is vulnerable to different intentional and unintentional attacks on the radio frequency signals used to transmit the navigation information (navigation messages and ranging sequences). First of all, given the fact that R-Mode will not be a standalone system, but rather relies on VDES, it is worth remarking that R-Mode can also be attacked indirectly by attacking VDES. For example, a successful attack on VDES which disrupts completely the service will also make R-Mode unavailable. In the following we analyze different types of attacks on VDES R-Mode.

4.1 | Jamming

Jamming attacks, broadly speaking, consist of the transmission of undesired signals (interference) in a given part of the radio frequency spectrum. Thus, jamming results in a reduction of the signal to noise and interference ratio (SNIR) and thereby in a reduction in the service quality. In the case of R-Mode, the SNIR decrease results into an increment of the positioning error. Note that jamming can be both intentional and unintentional. One speaks of intentional jamming when an attacker is deliberately generating the interference in order to disrupt a given service that makes use of the RF spectrum. However, jamming can also be, and frequently is, unintentional. In this case, a third party, that is external to the system, unintentionally generates interference, be it through malfunction of its communication systems or through negligence. In general, jamming attacks are relatively easy to carry out. They can also be mitigated, at least in part, relying on non-cryptographic countermeasures (see Section 5).

When dealing with unintentional interference (jamming), one commonly distinguishes between in-band and out-of-band interference. In-band interference originates from systems which operate in the same frequency band. In some parts of the world, the VDE frequncies are currently used for (analog) voice communication. At a future point in time, these frequencies will be assigned to VDE. After that, it could well happen that some vessels accidentally enable the voice channels creating in-band interference to VDE. Out-of band interference originates from systems operating at frequencies outside the VDE communication bandwidth, whose transmission power spills into the VDE communication bandwidth, commonly due to harmonics or intermodulation products. In the case of VDE, the neighbor frequencies bands are used for AIS, ASM and (terrestrial) mobile communication systems. Additionally, the IMO (International Maritime Organization) NCSR subcommittee warned that electromagnetic interference (EMI) effects of light-emitting diode (LED) lighting systems that are co-located on board maritime vessels caused failure of AIS reception on vessels [14]. VDE, in contrast to AIS, uses forward error correction at the physical layer. Thanks to this, VDES R-Mode has some resilience to jamming/interference.

4.2 | Spoofing

The second class of attacks, which are generally known as spoofing attacks, deals with all attacks in which an attacker carefully generates bogus signals which are interpreted by the receiver as valid, but which contain fake information which causes the system to malfunction. For example, in the case of R-Mode, a spoofing attack would consist of an attacker generating fake navigation messages and/or ranging sequences. Receivers process these signals as if they were legitimate, obtaining timing and positioning information which can greatly differ from the real ones, and being unaware of it. This can have worse consequences than a jamming attack where the service is rendered unavailable.

In GNSS systems each transmitter transmits continuously a stream of data containing navigation messages and the ranging sequences. Hence, an attacker usually aims at spoofing the whole data stream from a GNSS satellite. In contrast, VDES R-Mode

relies on the TDMA access scheme of VDE which dedicates some time slots to navigation messages and ranging sequences. As a consequence, navigation messages and ranging sequences can be spoofed separately in VDES R-Mode.

Let us focus first on spoofing attacks on the navigation message. The navigation message in VDES R-Mode contains several fields of information about the transmitter locations, the clock drift of the transmitters, the ranging sequence sent by each transmitter and the slot used by each transmitter. A spoofing attack on the clock drift field results in an increased positioning estimation error. A spoofing attack on the ranging sequence field or the field, which indicates which slots are used by the VDES R-Mode transmitters, results, in general, in an interruption of the PNT service, i.e., the receiver will most likely not be able to compute a valid position. The spoofing of the transmitter locations deserves a more detailed analysis. An unsophisticated spoofer may simply inject random data in the navigation message. This will result in vessels obtaining a random position estimation. Such an attack might be easily identified by the logic of the VDES R-Mode receiver. However, a more sophisticated attacker may carefully choose how to spoof the data field containing the transmitter positions. For example, it might change the transmitter locations in a way so that receivers in a particular area end up with a position estimation shifted 300 m to the south, making vessels strand in a sandbank. In other words, by carefully choosing the transmitter locations, a spoofer could induce vessels to believe that they are at a different position, and this could have dire consequences for the safety of the maritime traffic. Typically, spoofing attacks on the navigation messages can be hindered relying on cryptographic countermeasures (see Section 5).

An attacker may also spoof the ranging sequences. Note that the ranging sequences carry no data, i.e, they are deterministic. Thus, in a spoofing attack on the ranging sequences, the spoofer transmits the (known) ranging sequences carefully selecting the transmission time. This way the spoofer can induce vessels to believe that they are at a different position. Depending on the level of sophistication, different spoofing attacks on the ranging sequences can be distinguished. An unsophisticated spoofer may simply transmit the ranging sequence within the slot in which it is expected by the receiver, without tightly controlling the transmission time, which would require a local clock with a high stability. This will in general result in a high position estimation error at the receiver, and could be detected by the logic of the receiver. For example, if the timing is not carefully controlled, the position estimation might well result in an altitude well above the sea level, which can be easily detected. A more sophisticated spoofer may have access to a highly stable local clock, and could carefully tune the transmission time in order to induce a position error of his choosing. In order to overcome the logic of the receiver, the attacker will also need to carefully select its transmit power in a realistic way, i.e., so that closer transmitters are receiver with a higher power than those being farther away. Such an attack could easily induce vessels to believe that they are a different location. However, if the attacker uses a single transmitter, multi antenna techniques could be used to detect the attack (See Section 5). Finally, a very sophisticated spoofer having access to a network of multiple transmitters could use a different transmitter to transmit each ranging sequence, deceiving even a receiver with multiple antennas. Such an attack would require a large investment in infrastructure from the attacker, but it would be easier to carry out on VDES R-Mode than in GNSS, since in a GNSS system the attacker would need to transmit its signals using drones and mimicking the geometry of the GNSS system. Compared to navigation message spoofing, ranging sequence spoofing is in general more difficult to carry out, especially if one wants to carefully induce the receiver to believe that he is in a different position. The additional difficulty comes from the fact that the spoofer needs to have access to a good local clock. Deceiving a receiver with multiple antennas implies additional difficulty for the attacker. Finally, it is important to stress that standard cryptographic techniques cannot be used to protect the ranging sequences, since they do not carry any data (see Section 5).

4.3 | Meaconing

In the literature one sometimes finds a third class of attacks, meaconing, which lies halfway between jamming and spoofing[15]. Meaconing attacks involve intercepting the transmission of the navigation signals, storing them, and replaying them at a later point in time. Meaconing lies halfway between jamming and spoofing, because it does not imply sending random-like interference, as it is the case for jamming, but it also does not allow the attacker to transmit arbitrary information, since the attacker is limited to reproducing signals which have been transmitted at a previous point in time by a legitimate transmitter.

In the simplest example of a meaconing attack on VDES R-Mode, the attacker must first store the VDE signal containing VDES R-Mode transmissions (navigation messages and ranging sequences). At a later point in time, the attacker will replay the stored signal. A terminal which receives this signal and uses it to compute its signal relying on pseudo-ranging will be induced to believe that he is at the position in which the signal was stored by the attacker. Note that in the previously described attack the VDE signal stored will contain ranging sequences from multiple VDES R-Mode transmitters. If the attacker uses a single

transmitter to replay the signal, a receiver equipped with multiple antennas can easily detect the attack. However, a receiver equipped with a single antenna may easily be fooled by the attack.

More elaborate meaconing attacks are possible. For example, the attacker might store the VDE signal, and later only transmit some of the TDMA slots in the transmission, or combine multiple recorded signals. For example, the attacker may only decide to transmit TDMA slots containing navigation messages, and not those containing ranging sequences.

5 | POSSIBLE COUNTERMEASURES

In general, the countermeasures to protect a navigation system such as R-Mode can be split into two different classes: cryptographic and non-cryptographic countermeasures [16]. Cryptographic countermeasures are based on the use of cryptographic protocols and can be used to protect the confidentiality and authenticity of the navigation signals transmitted by the shore stations / satellites. Non-cryptographic countermeasures encompass signal processing (e.g, zero forcing techniques against jammers using multiple antennas or using SNIR measurements to estimate the positioning accuracy) as well as relying on a local clock and/or an inertial system to authenticate the navigation signals by checking its plausibility.

In the following sections we provide a description of the different types of countermeasures.

5.1 | Non-Cryptographic Countermeasures

Different non-cryptographic countermeasures can be used by a receiver to protect itself against different attacks. In the following we will describe the most important classes of countermeasures.

5.1.1 | Signal- and State-Analysis Techniques

A fist class of non-cryptographic countermeasures are the so-called signal- and state-analysis techniques, which were described in detail by Günther [17]. Generally speaking, these techniques consist of creating a model about how a set of parameters changes with time. These parameters are then monitored in the received signal and compared with the predictions of the model. If the prediction matches the model, the navigation signals are considered to be authentic, otherwise, if the measured parameters substantially differ from the ones predicted by the model, the receiver assumes it is the victim of an attack. Parameters that can be tracked are the positions of the receiver and the transmitters as well as their clock offsets and the delay, frequency and phase of the received signals. The simplest example of a signal- and state analysis technique would be a receiver that simply tracks his own position. Under the assumption of a given maximum vessel speed, one can derive a model about how fast the receiver position can change with time. Thus, if at a given point in time the estimated position jumps to a distant point, the receiver can raise a flag indicating that something is not right (it is possibly suffering a spoofing attack, for example).

In a GNSS context one usually distinguishes two types of receivers, sanp-shot receiver and tracking receivers [18]. A tracking receiver is locked to the GNSS signal, tracking continuously the estimated frequency, delay and phase of the signal. In contrast, a snap shot receiver does not continuously monitor the received signal, instead it intermittently takes a *snap-shot* (a continuous window of samples) and processes it to estimate its position. Provided that the receiver initially locks to the authentic signal, tracking receivers provide a higher protection than snap-shot receivers against spoofing and meaconing attacks.

In VDES R-Mode receivers are forced to operate in a snap-shot fashion because the navigation signals are not a continuous stream (as in GNSS) but intermittent transmissions in a TDMA access scheme. In particular, every transmitter transmits its ranging sequence once every second fitting its transmission within one VDE slot. Thus, in VDES R-Mode it is not possible to continuously monitor the frequency and phase and delay of the signal. However, the receiver may still monitor its own position, the transmitters' positions, the received power from each transmitter, and possibly other parameters.

The protection offered by signal- and state-analysis techniques varies depending on the moment in which the spoofing/meaconing attack takes place. If the attack starts at a point in time in which the receiver already has an estimate of its model parameters (receiver and transmitters estimated positions, etc) the receiver is likely to detect the attack since once the attack starts the parameters of the model will start deviating from their expected values. However, in case the attack starts when the receiver is performing a cold start and has no information about its parameters, a receiver relying solely on signal- and state-analysis techniques will be unable to detect the attack.

5.1.2 | Antenna-Array Techniques

A second class of non-cryptographic countermeasures are antenna-array techniques, also known as multi-antenna techniques. As their name indicates, these techniques require vessels to be equipped with multiple antennas. In general these techniques rely on signal processing techniques that allow to exploit the spatial dimension of the received signals in order to detect or counteract different kinds of attacks.

The simplest of these techniques is relying on direction of arrival (DOA) measurements to estimate from which direction every VDES R-Mode signal is received. The receiver can then carry out a plausibility check based on the DOA measurements to determine whether a signal is authentic or not. For example, such a technique is highly effective against spoofing and meaconing attacks in which a single transmitter tries to spoof signals from multiple transmitters. In this case all VDES R-Mode signals come from the same direction, which is easily detected. In general, the DOA measurements can be used in conjunction with the previously described signal-and state-analysis techniques, so that the receiver keeps track of the direction from which each signal is received.

A higher level of protection can be achieved relying on smart-antenna techniques. In this case the receiver does not only measure the direction of arrival of the signal, instead it relies on adaptive beamforming, steering the antenna in a desired direction. Many different such techniques exist. Antenna-array techniques are very powerful to combat jamming attacks. In most cases attackers use a single or few transmitters to *jam* the system. In this case, the receiver can try to suppress the jamming signal by modifying its radiation diagram so that it exhibits zero gain in the direction in which the jamming signal(s) are received (zero-forcing).

5.1.3 | Inertial Systems

An inertial system is a navigation system which uses accelerometers (motion sensors) and gyroscopes to measure the orientation and velocity and provide an estimation regarding how the position changes with time. Inertial systems usually employ three gyroscopes and accelerometers to measure angular velocity and linear acceleration in orthogonal directions. Inertial systems, when provided with an initial position obtained externally, can be used to obtain an (absolute) position estimate, i.e., they can be used as a navigation system. However, when used standalone, inertial navigation systems suffer from integration drift. This basically means that the small errors in the measurement of the acceleration and angular velocity add up with time, so that the position estimate degrades with time.

When combined with another PNT system such as R-Mode, inertial system can serve two purposes. First of all, the estimates provided by the two systems, R-Mode and the inertial system, can be combined together in order to provide an improved position estimation. This is sometimes known as sensor fusion, and it has multiple advantages, the main one possible being that it allows to overcome (short) outages of the navigation system. For example, if the VDES R-Mode signal is blocked by an obstacle (e.g., an island), the inertial system allows to update the position even during the outage of VDES R-Mode. Additionally, the inertial system can serve as a trustworthy external reference to detect anomalies in the position estimation through VDES R-Mode, and hence to detect spoofing/meaconing attacks. For example, if the course and speed obtained through VDES R-Mode consistently differs from the estimation of the inertial system during several seconds (several consecutive measurements), an alarm could be raised to indicate that possibly a spoofing attack is taking place.

The cost of inertial systems varies depending on the precision. Obviously, the more precise (the smaller the drift) the inertial system is, the better protection it will provide against attacks on VDES R-Mode. Unsophisticated spoofing/meaconing attacks can be easily detected using consumer grade inertial measurement units with cost in the order of 10 $. However, in a much more sophisticated attack, a spoofer, knowing the position of a vessel, might generate a spoofed signal so that the estimated position slowly drifts away from the real position. Detecting these type of sophisticated attacks may require much more accurate (and expensive) inertial measurement units. But one can argue that such attacks are unlikely.

5.2 | Cryptographic Countermeasures

As a first broad classification of the possible countermeasures one can divide them into encryption and authentication based countermeasures [19]. An encryption countermeasure aims at protecting the confidentiality of the information sent, whereas authentication countermeasures rely on cryptographic algorithms to ensure that the information originates from a trusted source (and that it has not been tampered on the way to the receiver). In the following we provide a more detailed description of different cryptographic countermeasures.

5.2.1 | Symmetric Encryption

Symmetric encryption relies on a shared secret between transmitter and receiver (a so called secret key) in order to protect the confidentiality of the transmitted information. Symmetric encryption is used to protect all military GNSS signals as well as the Galileo public regulated service (PRS), which is intended only for governmental authorised users. Usually, GNSSs rely on Direct-Sequence Spread-Spectrum (DSSS) techniques and the role of symmetric encryption is generating an encrypted spreading code, which is then used to spread the navigation signal. This spreading code is only known to those in possession of the secret key. In the context of VDES R-Mode, symmetric encryption could be applied by encrypting the navigation messages directly (e.g., using AES) as well as the ranging sequences.

Among all cryptographic countermeasures, symmetric encryption is, in general, the cryptographic technique which can provide the best protection[15]. However is also comes at a high cost. When symmetric encryption is used, anyone in possession of the secret key can impersonate the transmitter. Thus, in order to protect the system, it is necessary to equip the receivers with tamper-resistant modules. Furthermore, it is necessary to implement procedures for the secure distribution of the secret keys, e.g. providing keys to new users or exchanging the encryption keys after a given period of time (rekeying). Such procedures are used in some civilian applications, such as pay-TV systems, but they imply a considerable increase of the terminal cost as well as the system operation cost. It is for this reason that symmetric encryption techniques are not appropriate for R-Mode, since the increase in the cost would be unacceptable for the users.

5.2.2 | Watermarking Authentication

In general, watermarking techniques consist of hiding some information (a so called watermark) in a carrier signal. In our case the carrier signal is the R-Mode navigation signal, and the watermark contains some authentication related information which allows a receiver to determine whether the signal comes from a trusted source or not. In the context of GNSS, watermarking techniques have been proposed to protect the navigation signals[20]. Systems like Galileo and GPS employ (direct-sequence) spread-spectrum techniques with large spreading factors, so that the navigation signal at the receiver lies around 20 dB below the thermal noise level. Watermarking works on the assumption that a potential attacker experiences a very low SNR.

In the following we present the working principle of a watermarking system as proposed by Kuhn[20]. Let us consider first the transmitter side with multiple transmitters. The i-th transmitter selects its pseudo random watermark s_i independently from other transmitters. All watermarks have the same time duration δ. At time t_0 all the transmitters transmit their watermarks. Hence, seen from a receiver the watermarks from all transmitters overlap at least partially. At time $t_0 + \varrho$ each of the transmitters sends a digitally signed message M_i revealing his watermarking sequence s_i. The receiver operates as follows: The receiver experiences a *low* SNR so that it is not possible to reliably decode/estimate the watermarks s_i. Instead, the receiver digitizes (samples and stores) the window of samples in which the watermarks are received. After ϱ time units, the receiver obtains the digitally signed messages M_i which contain the watermarking sequences s_i. The receiver first verifies the digital signature. If the digital signature is correct the receiver then tries to detect the watermarking sequences in the recorded window of samples by means of a correlation. If the sequence s_i is found exactly ϱ time units before the message M_i was received, the receiver can safely assume that the signal transmitter i can be trusted. On the contrary, if either the digital signature does not match or the delay deviates from ϱ, the receiver cannot trust the signal from the i-th transmitter. Fig.4 shows a graphic representation of this watermarking technique for a navigation system.

From this description, it is easy to infer that an authentication delay of ϱ time units is incurred when using watermarking. Note however that the receiver may process the ranging sequences immediately to obtain its position, and verify the authenticity later.

Watermarking techniques for VDES R-Mode present several disadvantages. Firstly, watermarking requires to store the digitized (sampled) received signal. This can require a substantial amount of memory. Second, authentication is achieved after a given delay (when the watermarking sequence is revealed). Third, watermarking relies heavily on the assumption that the watermarks cannot be reliably decoded/estimated due to the low receive SNR. For example, in GNSS one could recover the watermarking sequence (at least in theory) by employing highly directive antennas to improve its link budget. After obtaining the watermarking sequence, it is possible to compromise the authenticity of the navigation system. Thus, the underlying security assumption is that it is difficult/expensive for an attacker to use very large antennas (note that the transmitters can use different techniques to lower the SNR of the watermarking sequences[20] to increase the antenna size necessary to carry out the attack). In the case of VDES R-Mode, it is in principle possible for an attacker to get physically close to a shore-station in order to receive the ranging sequence with a very high SNR (well above the noise level), obtaining thus the watermarking sequence. Such an attack would not hold for a VDE satellite transmitting R-Mode signals. However, it would be highly preferable to adopt the same

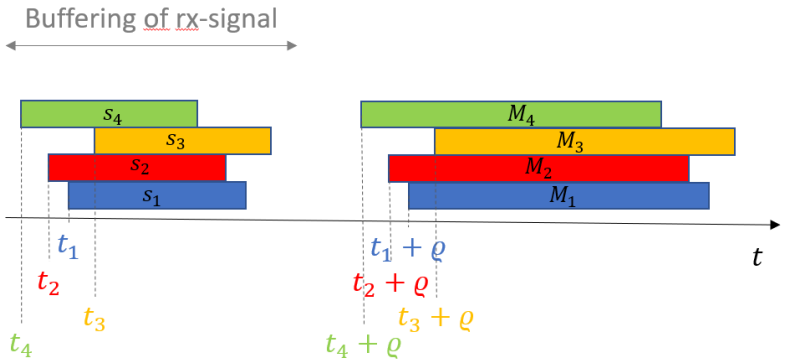

FIGURE 4 Watermarking for a navigation signal as proposed by Kuhn[20]. The transmitters first transmit their watermarks s_i which arrive at time t_i at the receiver, resulting in overlapping watermarks. At time $t_i + \varrho$ the digitally signed messages M_i arrive at the receiver. The receiver must first buffer the window of samples in which he expects to receive the watermarks. Next he verifies that the digital signatures are valid and that the signed message M_i arrived ϱ time units after the watermark s_i.

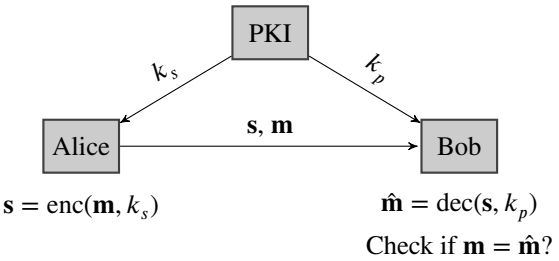

FIGURE 5 Illustration of a digital signature scheme using a PKI.

countermeasures for all communication links (shore stations and satellites). Finally, watermarking requires the receivers to be able to decode the navigation message of a transmitter in order to protect its ranging sequence. In VDES R-Mode, since the communication coverage of different transmitters barely overlaps, this would imply that in most cases only one of the transmitters can be authenticated, and not the remaining 2 or more. Thus the level of protection provided by watermarking would not be high in VDES R-Mode. The last two aspects make the use of watermarking sequences in R-Mode questionable.

5.2.3 | Digital Signatures

The best known cryptographic technique to protect the authenticity of data is the use of digital signatures which are based on asymmetric cryptography. Digital signatures are public key cryptosystems used to verify the authenticity of digital messages or documents. Assume Alice wants to send a message $\mathbf{m}$ to Bob and Bob wants to be sure that the message in fact is from Alice. The way digital signatures work is as follows. Alice creates a *digital signature* $\mathbf{s}$ by using an encryption function $\mathbf{s} = \text{enc}(\mathbf{m}, k_s)$ which is a function of the message $\mathbf{m}$ and a *secret key* k_s that is only known by herself. It is assumed that computing $\mathbf{s}$ without knowing the secret key k_s is computationally hard. Bob and everybody else have access to the message $\mathbf{m}$ and a *public key* k_p. There must exist a decryption function $\text{dec}(\cdot, \cdot)$ with which it is possible to compute $\mathbf{m}$ from $\mathbf{s}$ and k_p. That means Bob can verify that the message is from Alice by checking if $\mathbf{m} = \text{dec}(\mathbf{s}, k_p)$. Since this way everybody can distribute public keys, a trustful third party is required, the so-called PKI that distributes the public keys (see Section 2.3). The process of this scheme is depicted in Fig.5 . Different cryptographic primitives can be used to implement digital signatures, which results in different sizes for the digital signature. A good candidate is the Elliptic Curve Digital Signature Algorithm (ECDSA), which for a security level of 128 bits[4] leads to signatures of size 512 bits.

[4]This basically means that an attacker needs to run at least 2^{128} operations in order to break a signature.

TABLE 1 Comparison of the public key and signature size for the NIST Cat. I post-quantum parameters (128 Bit classical security level).

Authentication Scheme	Public Key Size (byte)	Signature Size (byte)
ECDSA	32	64
CRYSTALS-Dilithium[23]	896	1387
FALCON[24]	896	666
Rainbow[25]	58800	66

Digital signatures per-se do not offer protection against replay attacks. This means that an attacker can record a message with a valid signature at time t and replay it at time $t + \Delta$, and the replayed message passes the signature check (it is considered authentic). The standard procedure to protect a digitally signed message from a replay attack is adding a timestamp which allows receivers to detect if a message has been recorded and replayed. Digital signatures together with time-stamping can be used to safeguard the authenticity of navigation messages, however they cannot be used to protect the ranging sequences, since these carry no data (they are deterministic sequences).

A straightforward application of digital signatures is signing every single navigation message. This implies that the signature must be appended to each navigation message which results in a considerable communication overhead (512 bits per signed message assuming ECDSA signatures). Moreover, the verification of ECDSA signatures is computationally complex. For example, state-of-the-art software implementations of ECDSA running on a modern ARM processor need between 50 to 200 ms of processing time in order to verify a signature.

A possible alternative is not signing every message, but signing together n messages, which reduces the overhead by a factor n. However, this technique has two disadvantages. The first one is an increase in the latency, since the receiver must wait until all n messages are received before it can verify the signature. The second is an increase in the authentication error rate, since the signature can only be verified if all n navigation messages are received error free. Assume now that each message is received erroneously independently from other messages but with identical probability of error p_e. In this case, signing n messages together leads to an authentication error rate

$$p_a = 1 - (1 - p_e)^n \approx np_e$$

where the approximation is tight when the product np_e is small ($\lesssim 10^{-3}$).

Quantum-Resistant Digital Signatures

The recent developments in quantum computing constitute a serious threat for today's asymmetric encryption and authentication schemes like RSA and ECDSA. Once large-scale quantum computers can be built, most of the current asymmetric cryptosystems can be broken in polynomial time by Shor's factorization algorithm. In order to protect the long term security of the data that is transmitted and stored today, new quantum-resistant cryptosystems need to be developed. Currently, there is an ongoing standardization for quantum-resistant encryption and authentication schemes at the National Institute of Standards and Technology (NIST)[21]. The third-round digital signature finalists[22] consist of the two lattice-based digital signature schemes CRYSTALS-DILITHIUM[23] and FALCON[24] as well as the multivariate digital signature scheme Rainbow[25]. These schemes are currently the most promising general-purpose algorithms for public-key digital signature schemes[22]. A comparison between the key and signature sizes of the third-round finalist and ECDSA for a classical security level of 128 Bit (NIST Cat. I) is given in Table 1. As it can be observed from the table, post quantum digital signatures lead to larger public key and signature sizes as compared to ECDSA.

5.2.4 | TESLA Broadcast Authentication Protocol

The Timed Efficient Stream Loss-Tolerant Authentication (TESLA) protocol is an efficient, low-computing broadcast authentication protocol that extends to many receivers and tolerates the loss of packets[26]. Since insertion of malicious packets in many broadcast networks can pose a risk, the receivers want to ensure that their received packets originate from a trusted source. When dealing with point-to-point communication, the authenticity (and integrity) of a message is usually protected using a so-called Message Authentication Code (MAC), which is a (hard to forge) tag that gets appended to the exchanged messages. MACs are generated using a secret key which is shared by the two communication endpoints. Employing MACs in broadcast communications does not provide secure source authentication, since it requires that all the receivers have access to the secret key, and having access to the secret key makes it possible to generate valid MACs. Thus, anyone in possession of the secret key can *forge*

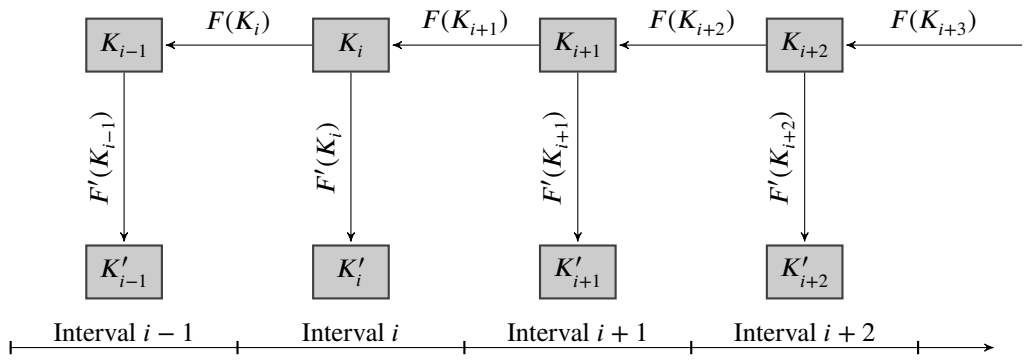

FIGURE 6 Illustration of the key chain used in the TESLA protocol.

the MAC of a message. Obviously, applying a digital signature to each data packet provides secure broadcast authentication, but it comes with a high overhead, both in terms of bandwidth and signing and verification time. The core idea of TESLA is that the sender attaches a MAC to each packet, which is computed using a (secret) key known only to itself. The receivers in turn buffer the received messages, without being able to authenticate them. Shortly afterwards, the sender discloses the (secret) key and the recipients are able to authenticate the packet. This way, a single MAC appended to every message suffices to provide broadcast authentication, resulting in lower overhead and less computation time when compared to the use of digital signatures. One requirement of the TESLA protocol is a loose time synchronization of the receiver with the senders clock. In particular, the receiver needs to know an upper bound on the local time of the sender. Furthermore, TESLA requires receivers to buffer a number of packets, which results in an authentication delay, which is commonly in the order of one round trip delay between the sender and receiver.

The authentication in the TESLA protocol is based on a one-way key chain. A one-way key chain is build by using a one-way function $F(\cdot)$ such that a key K_i is computed from another key K_{i+1} as $K_i = F(K_{i+1})$. Each key K_i is associated with a time interval of uniform duration, indexed by i. Furthermore, using another one-way function $F'(\cdot)$ a key K'_i is derived from K_i as $K'_i = F'(K_i)$. If a message M_j is sent during the $i-th$ time frame, the sender uses the key K'_i to compute a MAC, $\text{MAC}(K'_i, M_j)$ of the packet P_j. Additionally the packet P_j contains the key $K_{i-d}$from d time intervals in the past. The parameter d is called the *packet disclosure delay* and is specified beforehand along with the duration of the time intervals and the maximum length of the one-way key chain. The packet is then given as

$$P_j = \{M_j || \text{MAC}(K'_i, M_j) || K_{i-d}\} \tag{1}$$

TESLA works as follows: When a receiver receives a packet P_j the key K_{i-d} can be used to determine i. Assuming a loosely synchronized clock with the sender, the receiver can check a latest possible time interval l that the sender could currently be in. If $l < i + d$, then the packet is considered to be safe. The receiver has to reject all packets that are not safe, i.e., those that contain keys that already have been disclosed and are publicly known. Packets for which $l < i + d$ contain keys that, at this moment, can only be known by the sender. When receiving a packet P_j sent in the time interval i, the receiver cannot yet verify the authenticity and needs to add the collected information to a buffer until it learns the key K_i which is after d time intervals.

The use of TESLA to authenticate navigation messages in GNSS was first proposed by Wullems et al.[27]. Recently, TESLA has been proposed to protect the authenticity of the Galileo navigation messages by Fernández Hernández et al.[28]. In Galileo, each transmitter (Galileo satellite) transmits navigation messages containing navigation information about itself but not to any of the other satellites. In the proposal by Fernández Hernández et al.[28], all the Galileo satellites build a single TESLA chain allowing navigation messages from different satellites to cross-authenticate each other. The main benefit of this approach is a reduction of the authentication error rate (or a decrease in the time between authentications). Consider for simplicity the case in which the receiver sees 4 satellites. If each transmitter uses its own independent TESLA chain, the receiver must receive the 4 MACs and the 4 corresponding keys error free in order to successfully authenticate the service. However, when a single chain is used, it is enough to receive the 4 MACs and one key free of errors. This feature is particularly useful when the 4 satellites are not received with the same link quality, as it could be the case in an urban environment where 1 satellite is seen with high elevation, whereas the remaining 3 satellites are received with a somewhat lower elevation (and thus with a higher bit error rate). The disadvantage of this scheme is that the TESLA key-chain becomes longer by a factor S, which corresponds to the (maximum) number of senders. This implies also an increase in complexity.

Using a single TESLA chain is a good solution for GNSS. However, this approach does not seem a workable alternative for R-Mode. First of all, the communication coverage of the multiple transmitters barely overlaps. If transmitters only send its own navigation information in the navigation message, most receivers would only receive the navigation message from one transmitter and would not be able to compute its position. Thus, transmitters must provide navigation data relative to other transmitters in the neighborhood. In practice this eliminates the aforementioned advantage in terms of authentication error rate (or time between authentications). Second, while the maximum number of transmitters in Galileo is at most 40, the number of transmitters in R-Mode will be in the order of the hundreds or even thousands, making a single chain prohibitively complex. Hence, for R-Mode a better solution seems to let every transmitter provide navigation data relative to all its neighboring transmitters and have each transmitter build its own TESLA chain. If the VDE satellite is to act as a R-Mode transmitter, it will need to select the list of neighboring shore-stations dynamically depending on the area that the satellite is illuminating at that time.

Recommended TESLA Parameters

Let us now try to narrow down the application of TESLA to R-Mode and propose tentative values for its parameters. We will work on the assumption that each transmitter builds its own TESLA chain. We start by analyzing the key-size. A key-length of 128 bits is recommended to guarantee security at least over the next 10 years[29], and is inline with what has been proposed for Galileo[28]. For most applications the MAC length ranges from 80 to 160 bits. For a MAC size of b bits, and assuming that an attacker tries to guess randomly the MAC, the probability that a random MAC is correct is 2^{-b}. For example, a 10 bits MAC would be guessed correctly why a probability $1/1024$. Furthermore, one needs to take into account, that the consequences of a single right guess would not be dramatic. In order to substantially compromise R-Mode, an attacker would need to guess correctly several consecutive MACs[28], which becomes extremely unlikely. Hence, using a MAC length of 80 bits or larger seems unnecessary for R-Mode. A MAC length ranging between 15 and 20 bits would already provide a good level of security.

Additionally, it is necessary to sign the first key of a TESLA chain using a digital signature. We will assume that ECDSA with a signature size of 512 bits is used. Note that a vessel entering the communication coverage of a shore-station would not be able to authenticate any navigation message until it receives a valid signature. For this reason a digital signature would need to be transmitted periodically. Here a similar approach as that proposed for Galileo[28] could be used. In particular, each navigation message could transport a small fragment of the digital signature. For example, if each navigation message carries a 16 bit fragment of the digital signature, after receiving 32 consecutive fragments (error free) a receiver would be able to verify the signature. Obviously, this fragmentation of the signature results in an increase of the authentication error rate. For example, a receiver which has lost the n-th fragment would need to wait until this exact fragment is retransmitted again and received error free. A straightforward countermeasure to decrease the authentication error rate due to fragmentation would be employing erasure coding. This would of course increase the overhead, but even a slight increase of overhead (using a high-rate code) could yield a considerable performance improvement.

Let us now analyze the complexity of TESLA. In a straightforward implementation of TESLA for R-Mode, the authentication delay would correspond to the time between two consecutive navigation messages, which is planned to be 60 s. Regarding complexity, two different scenarios must be considered. In the first one a ship enters the coverage area of a *new* shore-station / satellite. If the vessel receives the n-th key K_n, it must apply n times the one way function $F()$ as well as verify the digital signature. The computation of the one way function is fast. In particular if we assume that SHA-256 is used as one-way function $F()$, its execution on a modern ARM platform takes about 25 μs[30]. If a new TESLA chain is initiated every month, and assuming an inter-message time of 60 s, in the worst case (last message of the month), one needs to compute $F()$ 43,200 times, which takes 1.08 s. In addition one needs to verify the signature, which for ECDSA takes around 200 ms. This yields a total computing time of 1.28 s. In the second scenario, the ship has already verified the chain in the past. Hence, it does not have to verify the signature. Instead it needs to compute $F()$ a number of times n, where $n-1$ corresponds to the number of navigation messages it has missed. For example, if we make the (pessimistic) assumption of the vessel being in the edge of the communication coverage, so that only 1 out of 10 navigation messages are received error free, this yields a computing time of 250 μs. For a ship that has not missed any packets, we have $n = 1$ and the computing time is 25 μs. Hence, except for the initial chain verification, TESLA is several orders of magnitude less complex than ECDSA.

6 | DISCUSSION

Let us start by analyzing the protection provided by the different countermeasures against different types of attacks, which is summarized Table 2:

- Non-cryptographic Countermeasures. In General, they are the only countermeasures that can provide some protection against jamming attacks. And they provide a medium level of protection against spoofing and meaconing.
  - Signal- and state-analysis techniques. The provide a low protection against jamming, since they may only detect its presence. They provide a medium level of protection against spoofing and meaconing attacks, since these techniques can detect simple attacks but they may fail to detect sophisticated attacks.
  - Multi-antenna techniques. The provide the best protection against jamming, since they can be used to spatially filter out the jamming signal. The provide a medium protection against spoofing and meaconing. Simple attacks with a single transmitter may easily be detected relying on multi-antenna techniques, but more sophisticated attacks with multiple transmitters could be difficult to detect.
  - Inertial System. They provide a low protection against jamming since they can only protect against short outages caused by jamming. Inertial systems provide a medium level of protection against spoofing and meaconing since they can detect simple attacks. However, more sophisticated attacks might not be detected by an inertial system (specially if it is a low cost system exhibiting a large drift).
- Cryptographic Countermeasures. They are powerless against jamming, since jamming attacks the system before any cryptographic measures can be applied. However, they are very effective against spoofing attacks. Their effectiveness against meaconing attacks depends on the specific type of cryptographic countermeasure. In general the protection against spoofing is limited due to the fact that in VDES R-Mode the navigation signals are not continuously transmitted, and hence the navigation message and the ranging sequences can be attacked separately.
  - Symmetric Encryption. It provides the best protection against spoofing and meaconing attacks.
  - Watermarking authentication. It provides a high protection against spoofing (both the navigation messages and the ranging sequences are protected) and low protection against meaconing.
  - Navigation message authentication. It provides medium protection against spoofing (the navigation messages are protected, but no the ranging sequences). The protection against meaconing is low.

Table 3 provides a summary of the cost associated to the different countermeasures. Non-cryptographic countermeasures affect only the terminal cost, since they do not require any infrastructure at system level. However, cryptographic countermeasures do require some system infrastructure to rely on. Among the non-cryptographic countermeasures, signal- and state-analysis techniques have the lowest terminal cost, since they do not require additional hardware. Inertial systems have a medium terminal cost, since they require additional hardware, an inertial measurement unit, although this hardware can have low cost. Finally, multi-antenna techniques have the highest terminal cost since they require installing multiple antennas in the ship. Among the cryptographic countermeasures, authentication implies in general a lower cost than symmetric encryption, since a PKI is somewhat less expensive than a complete secure key distribution system. As for the terminal cost, authenticating only the navigation message has the lowest cost, followed by watermarking. Symmetric encryption results in expensive terminals, whereas for non-cryptographic countermeasures, the cost depends on the exact measures which are implemented.

Keeping in mind that a low cost is a key factor for the acceptance of R-Mode, the best way forward seems introducing mandatory navigation message authentication, and possibly some optional non-cryptographic countermeasures. The main advantage of non-cryptographic techniques is that they can be applied at the receiver, and can be thus optionally applied by some receivers. A low cost receiver could apply some simple signal- and state-analysis techniques, a medium cost one could additionally include a low-cost inertial system, and a high end receiver could additionally apply multi-antenna techniques.

Given the fact that VDES R-Mode is conceived as a contingency system for GNSS, it is logical to compare the vulnerabilities and countermeasure for both systems. Following our suggestion for VDES R-Mode, we compare GNSS and VDES R-Mode assuming that navigation message authentication is used in conjunction with signal- and state analysis techniques at the receiver. A first difference between GNSS and VDES R-Mode is the geometry, i.e., in GNSS we have transmitters on board of satellites whereas VDES R-Mode relies mainly on shore stations complemented by few satellites. Hence, for an attacker it is more difficult to produce signals that mimic the geometry (angle of arrival) in a GNSS system, since this requires probably relying on transmitters on fast flying drones. A second and more important difference has to do with the channel access. In GNSS systems we have transmitters sending their navigation signals in a continuous transmission. Thus, one can combine the periodic transmission of authenticated navigation messages and signal- and state-analysis techniques with a tracking receiver to protect the whole

TABLE 2 Protection provided by the different countermeasures to the different kind of attacks.

Attack	Countermeasure					
	Non-Cryptographic			Cryptographic		
	S&SA Analysis	Multi-Antenna	Inertial Sys.	Symm. Encrytion	Watermarking	Nav. Mess. Auth.
Jamming	low	high	low	none	none	none
Spoofing	medium	medium	medium	high	high	medium
Meaconing	medium	medium	medium	medium/high	low	low

TABLE 3 Cost of the different countermeasures

Cost	Countermeasure					
	Non-Cryptographic			Cryptographic		
	S&SA Analysis	Multi-Antenna	Inertial Sys.	Symm. Encrytion	Watermarking	Nav. Mess. Auth.
Terminal	low	high	medium	high	medium	low
System	none	none	none	high	low	low

transmission. At a high level of abstraction, we can say that such a receiver marks as authentic whatever belongs to the same continuous transmission which conveyed an authentic navigation message. However, this is not possible for VDES R-Mode. Since the navigation signals are transmitted using the TDMA access scheme of VDE, we have an intermittent transmission. Thus, a snap-shot receiver must be used, which renders signal-and state-analysis techniques less powerful.

As we have seen, different techniques exists to authenticate navigation messages. Namely, the application of a digital signature to each navigation message, the application of a digital signature to n messages together, and the use of TESLA. Table 4 compares these 3 options. In terms of overhead the best solution is TESLA, with an overhead in the range of 150 to 180 bits per message. We can see how signing every navigation message yields an overhead of 512 bits per message, whereas signing 1 out of n messages reduces this number to $512/n$. In terms of computation time, the most complex solution is signing every message, which requires 200 ms of computing time per message, signing only 1 out of n messages improves this figure by a factor n, leading to a $200/n$ ms. However, the best solution in terms of complexity is TESLA, with a computing time 3 orders of magnitude below, namely 250 μs (here made the pessimistic assumption that only 1 out of 10 navigation messages are correctly received). However, in a cold start, when a vessel enters a new coverage area it has never seen before, TESLA requires (in the worst case) 1.28 s of computing time. For the signature schemes the computing time is the same as for the verification of a navigation message from a known shore-station. In any case, one should remark that this step is carried out seldom. Regarding the (worst case) time to authenticate, and assuming no message losses, a digital signature is the best scheme with a time equivalent to 1 Navigation Message Inter-arrival Time (NMIT) ,i.e. the time between two consecutive navigation messages, which is foreseen to be 60 s. Signing n messages together requires to wait n NMITs. TESLA in turn shows a good performance with 2 NMITs (here we assumed the packet disclosure delay equals the NMIT). Finally, regarding robustness to (packet) losses, digital signatures and TESLA are both very resilient, whereas signing together n messages has a low robustness to losses. In view of this comparison, the most promising scheme is TESLA, due to its low complexity and resilience to losses. Its only major drawback is the fact that, in the worst case, one needs to wait for 2 NMITs in order to authenticate the service. However, we believe this is reasonable in a navigation environment.

In a nutshell, after our assessment of the security of R-Mode, we believe the best option to secure it is to use TESLA to authenticate the navigation messages. One should remark that this requires the implementation of a PKI. Given the particularities of the maritime domain (intermittent connectivity and low bandwidth), a good way forward seems implementing the proposal from the CySiMS project described in Section 2.3. Without a PKI it is impossible to build a secure future-proof system. Finally, another point to keep in mind regarding future-proofness is the fact that the currently used signature schemes are vulnerable to attacks using quantum computers and the Shor's algorithm. For this reason, it is of tantamount importance to consider a PKI that supports a future upgrade to so-called post-quantum (or quantum resilient) signature algorithms, and eventually also to increase the key and MAC sizes for TESLA.

TABLE 4 Comparison of the direct application of signatures and TESLA for navigation message authentication.

	dig. signatures	1-out-of-n dig. sign.	TESLA
overhead per nav. message [bits]	512	$512/n$	150-180
computation time per authenticated nav. message	200 ms	$200/n$ ms	250 μs
computation time for cold start	200 ms	$200/n$ ms	1.28 s
Time to authenticate	1 NMIT	n NMIT	2 NMIT
Robustness to losses	high	low	high

7 | CONCLUSIONS

In this paper, we have studied the vulnerabilities of R-Mode, the contingency maritime positioning and navigation system that is currently undergoing standardization through the International Association of Lighthouse Authorities (IALA). After analyzing the different types of attacks that the system can be subject to, we have analyze the effectiveness of different cryptographic and non-cryptographic countermeasures. Keeping in mind the fact that low terminal and system operation cost are of key importance in order to foster the acceptance of the system, a promising solution seems authenticating the navigation messages of VDES R-Mode using the TESLA protocol. Additionally, some receivers might optionally adopt some non-cryptographic countermeasures to increase the level of protection.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

BIBLIOGRAPHY

References

1. Sadlier G, Flytkjær R, Sabri F, Herr D. The economic impact on the UK of a disruption to GNSS. tech. rep., Innovate UK; 2017.
2. IALA . R-129 GNSS VULNERABILITY AND MITIGATION MEASURES. tech. rep., IALA; 2012.
3. IALA . G1139 – THE TECHNICAL SPECIFICATION OF VDES. tech. rep., IALA; 2019.
4. Recommendation ITU-R M.1371-0 . Technical characteristics for a universal shipborne automatic identification system using time division multiple access in the VHF maritime mobile band. *ITU Tehcnical Report* 1998.
5. Clazzer F, Lázaro F, Plass S. Enhanced AIS receiver design for satellite reception. *CEAS Space Journal* 2016; 8(4): 257–268.
6. Plass S, Poehlmann R, Hermenier R, Dammann A. Global maritime surveillance by airliner-based AIS detection: preliminary analysis. *The Journal of Navigation* 2015; 68(6): 1195.
7. Safar J, Shaw G, Grant A, et al. GNSS Augmentation using the VHF Data Exchange System (VDES). In: ION. ; 2018.
8. IMO MSC-FAL1/Circ.3 . Guidelines on Maritime Cyber Risk Management. *IMO* 2017.
9. IMO MSC 98/23/Add.1 . Resolution MSC.428(98): Maritime Cyber Risk Management in Safety Management Systems. *IMO* 2017.

10. Balduzzi M, Pasta A, Wilhoit K. A security evaluation of AIS automated identification system. In: ACM. ; 2014: 436–445.

11. Wimpenny G, Safar J, Grant A, Bransby M, Ward N. Public Key Authentication for AIS and the VHF Data Exchange System (VDES). In: ION. ; 2018: 1841–1851.

12. Frøystad C, Bernsmed K, Meland PH, Rødseth OJ, Nesheim DA. Using digital signatures in the maritime domain. *CySiMS Project Technical Report D2.2* 2017.

13. IALA . Guideline 1158: VDES R-Mode. tech. rep., IALA; 2020.

14. SUB-COMMITTEE ON NAVIGATION C, SEARCH , RESCUE . REPORT TO THE MARITIME SAFETY COMMITTEE NCSR, 7/23. tech. rep., IMO; 2020.

15. Ioannides RT, Pany T, Gibbons G. Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques. *Proceedings of the IEEE* 2016; 104(6): 1174–1194.

16. Wesson K, Rothlisberger M, Humphreys T. Practical cryptographic civil GPS signal authentication. *NAVIGATION: Journal of the Institute of Navigation* 2012; 59(3): 177–193.

17. Günther C. A Survey of Spoofing and Counter-Measures. *Navigation, Journal of the Institute of Navigation* 2014; 61(3): 159–177.

18. Enge PK. The global positioning system: Signals, measurements, and performance. *Int. J. of Wireless Inf. Networks* 1994; 1(2): 83–105.

19. Margaria D, Motella B, Anghileri M, Floch JJ, Fernandez-Hernandez I, Paonni M. Signal structure-based authentication for civil GNSSs: Recent solutions and perspectives. *IEEE Signal Processing Magazine* 2017; 34(5): 27–37.

20. Kuhn MG. An Asymmetric Security Mechanism for Navigation Signals. In: Springer. ; 2004: 239–252.

21. Post-Quantum Cryptography. https://csrc.nist.gov/projects/post-quantum-cryptography; .

22. Alagic G, Alperin-Sheriff J, Apon D, et al. Status report on the second round of the nist post-quantum cryptography standardization process. *US Department of Commerce, National Institute of Standards and Technology* 2020.

23. Ducas L, Kiltz E, Lepoint T, et al. CRYSTALS–Dilithium: Algorithm Specification and Supporting Documentation. Round-2 submission to the NIST PQC project. 2019.

24. Fouque PA, Hoffstein J, Kirchner P, et al. Falcon: Fast-Fourier lattice-based compact signatures over NTRU. *Submission to the NIST's post-quantum cryptography standardization process* 2018.

25. PQCRainbow. https://www.pqcrainbow.org/; .

26. Perrig A, Canetti R, Tygar J, Song D. The TESLA Broadcast Authentication Protocol. *RSA CryptoBytes* Summer/Fall 2002; 5: 2-13.

27. Wullems C, Pozzobon O, Kubik K. Signal Authentication and Integrity Schemes for Next Generation Global Navigation Satellite Systems. *European Navigation Conference, (ENC-GNSS)* 2005: 1–10.

28. Fernández-Hernández I, Rijmen V, Seco-Granados G, Simon J, Rodríguez I, Calle JD. A Navigation Message Authentication Proposal for the Galileo Open Service. *Navigation, Journal of the Institute of Navigation* 2016; 63(1): 85–102.

29. ECRYPT-CSA . Algorithms, Key Size and Protocols Report (2018). *H2020-ICT-2014 – Project 645421* 2018.

30. Bond B, Hawblitzel C, Kapritsos M, et al. Vale: Verifying high-performance cryptographic assembly code. In: USENIX. ; 2017: 917–934.

31. Lázaro F, Raulefs R, Wang W, Clazzer F, Plass S. VHF Data Exchange System (VDES): an enabling technology for maritime communications. *CEAS space Journal* 2019; 11(1): 55–63.

32. Recommendation ITU-R M.1371-2 . Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile band. *ITU Tehcnical Report* 2007.

33. IMO SN.1/Circ.289 . Guidance on the Use of AIS Application-Specific Messages. *IMO* 2010.

34. Koch P, Gewies S. Worldwide Availability of Maritime Medium-Frequency Radio Infrastructure for R-Mode-Supported Navigation. *Journal of Marine Science and Engineering* 2020; 8(3): 209.

35. Šafář J, Grant A, Williams P, Ward N. Performance Bounds for VDES R-mode. *Journal of Navigation* 2020; 73(1): 92–114. doi: 10.1017/S0373463319000559