JRC Scientific and Technical Reports

Hardening of GNSS based trackers
Final Report

An Exploratory Research Project on
options concerning more resilient position reporting devices

U. Kröner,
C. Bergonzi, J. Fortuny-Guasch, R. Giuliani,
F. Littmann, D. Shaw, D. Symeonidis

EUR 24390 EN - 2010

JRC
EUROPEAN COMMISSION

ipsc
Institute for the Protection and Security of the Citizen

The mission of the JRC-IPSC is to provide research results and to support EU policy-makers in their effort towards global security and towards protection of European citizens from accidents, deliberate attacks, fraud and illegal actions against EU policies.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Address: TP 051, Joint Research Centre, Via E. Fermi 2749, 21027 Ispra (VA), Italy
E-mail: Ulrich.Kroener@jrc.ec.europa.eu
Tel.: +39 0332 78 6719
Fax: +39 0332 78 9658

http://ipsc.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC 58733

EUR 24390 EN
ISBN 978-92-79-15878-0
ISSN 1018-5593
doi:10.2788/97633

Luxembourg: Publications Office of the European Union

Printed in Italy

Table of Contents

LIST OF FIGURES **5**

LIST OF TABLES **6**

INTRODUCTORY CHAPTER **7**

ABSTRACT 7
INTRODUCTION 7
WHY SECURITY IN THE CONTEXT OF GNSS TRACKERS IS IMPORTANT 8
The VMS regulatory environment as a special case *8*
Modern VMS devices and their abuse potential: A fisheries perspective *9*
AN ABSTRACT GNSS-BASED TRACKER 10
A FORMALISATION OF VULNERABILITIES AND DEFENCES 11
RECENT IMPROVEMENTS REGARDING VMS DEVICE SECURITY IN EU MEMBER STATES 13
PARALLELS BETWEEN THE VMS DEVICES AND THE AIS TRANSPONDERS 15
PARALLELS WITH THE EUROPEAN DIGITAL TACHOGRAPH 18

CHAPTER I: DEFENDING AGAINST FAKE GNSS SIGNALS **21**

CHAPTER INTRODUCTION 21
Scope *21*
A layman's overview of GNSS *21*
Options for attacking a GNSS receiver *24*
Options for defending a GNSS receiver *33*
With prohibitive costs, nearly all GNSS receiver defences can be overcome *68*
EXPLORE THE POSSIBILITIES OF USING GALILEO ENCRYPTED GNSS SIGNALS 71
Using the Open Service (OS) *71*
Using the Commercial Service (CS) *72*
Using the Public Regulated Service (PRS) *73*
Using other Galileo Services *74*
Conclusion on the use of Galileo encrypted signals *74*
POSSIBILITIES IN USING GPS SIGNAL SIMULATORS 74

CHAPTER II: PHYSICAL SECURITY **77**

PROTECTING SMALL VOLUMES AGAINST PHYSICAL INTRUSION 77
Why physical security in the context of GNSS trackers is important *77*
What is wrong with existing seals? *78*
Passive and active monitoring seals *78*
Defending against tracker removal *79*
The RFID bolt seal *80*
An industry standard active volume enclosing seal *82*
"Mechanical anti-evidence seals" *84*
The "time trap" seal *86*
The "tie dye" seal *88*
A challenge-response seal with some "time trap" features *89*
Active monitoring using permanent magnets and magnetometers *95*
Surrounding the tracker's electronics with potting material *95*
CONCLUSIONS ON PHYSICAL SECURITY 97

CHAPTER III: DEFENDING AGAINST SIDE CHANNEL ATTACKS **98**

EMANATIONS SECURITY 98
SIDE CHANNEL ATTACKS 99
OPERATION OUTSIDE OF TEMPERATURE OR VOLTAGE SPECIFICATIONS 100
OPTIONS CONCERNING CONDUCTED INTERFACES 101
Definition and background *101*
Wrapping the conducted data interface inside of the protected volume *102*
Securing external physical data interfaces *102*
Options regarding the power supply cable *103*
BUFFER OVERFLOW EXPLOITS 104

CHAPTER IV: SECURING THE TRACKER'S POSITION REPORTS 106

AUTHENTICATING MESSAGES TO AND FROM THE GNSS TRACKER 106
Scope *106*
Inmarsat-C service providers *107*
On the secrecy of the Inmarsat-C transmission protocol *107*
Present-day use of Inmarsat-C *108*
Proposed changes to the tracker message *112*
Migration to globally available broadband *114*
Remote firmware updates *115*

CONCLUDING CHAPTER: A PROPOSED HARDENED GNSS TRACKER 117

DESIGN, PRODUCTION, AND COSTS 117
LIFECYCLE 119
Production site *119*
Configuration *119*
Installation *120*
Operation *120*
Measures upon reaching any alarm condition *121*
Decommissioning and Refurbishment *121*
USE 121

ANNEX: DUAL ANTENNA GPS RECEIVER DEFENCE PROTOTYPE 123

ABSTRACT 123
INTRODUCTION 123
MATERIALS AND METHODS 123
Algorithm 1: Build undifferentiated carrier phase *124*
Algorithm 2: Antennae-epoch differencing *125*
Interpretation *125*
RESULTS 126
DISCUSSION 129
Hardware *129*
Software *130*
Binomial Bayesian inference of posterior probability of spoofing *130*
Algorithm 3: Binomial Bayesian inference of posterior probability of spoofing, based on slope spreads, for static scenarios (non-kinematic) *131*
Algorithm 4: Extension of algorithm 3 to a kinematic environment, by dividing observations into slots *132*
Alternate algorithm for kinematic environments *132*

GLOSSARY 135

REFERENCES 141

List of Figures

FIGURE 1: AN ABSTRACT GNSS BASED REAL-TIME TRACKER 10
FIGURE 2: SCHEMATIC LEVELS OF ATTACK: 1) INCOMING, 2) PHYSICAL, 3) INFORMATION TECHNOLOGY, 4) ELECTROMAGNETIC, AND/OR 5) OUTGOING 12
FIGURE 3: THE SATLINK ELB 2010 IR SPECIAL TYPE, COURTESY OF SATLINK. THE NOTCHES ON EACH SIDE HOUSE CONNECTION POINTS FOR SEALS 13
FIGURE 4: THE SKYWAVE DMR 800L ANTENNA (LEFT) IS PUT IN THE ANTI TAMPER ENCASING (RIGHT). THE ENCASING HAS TWO SECURITY SEAL HOLES. COURTESY OF SATLINK 14
FIGURE 5: THE "SLAP & TRACK" MT 3300 TERMINAL, COURTESY OF SKYWAVE MOBILE COMMUNICATIONS, OTTAWA, CANADA 15
FIGURE 6: "*SCHEMATA OF A TYPICAL RECORDING EQUIPMENT (EXTRACTED FROM ISO-16844-3)*" [NORDVIK07] 18
FIGURE 7: *A TYPICAL MS AND ITS INTERNAL COMPONENT* [NORDVIK07] 19
FIGURE 8: ANATOMY OF A GPS L1 C/A NAVIGATION FRAME. 22
FIGURE 9: SIGNAL MODULATION. THE NAVIGATION MESSAGE BITS (AT 50 HZ) ARE ADDED MODULO 2 (XOR) WITH THE C/A GOLD CODE (1.023 MHZ), YIELDING THE COMBINED SIGNAL (BLUE). 22
FIGURE 10: CARRIER SIGNAL ANATOMY. THE BLACK GRAPH REPRESENTS THE IN-PHASE COMPONENT OF THE GPS L1 C/A CARRIER SIGNAL (1.57542 GHZ). THE BLUE SQUARE WAVE REPRESENTS THE C/A CODE MODULO THE NAVIGATION MESSAGE (1.023 MHZ). THE CARRIER PHASE IS SHIFTED BY 180° AT THE VALUE BOUNDARY, REPRESENTING THE BINARY PHASE SHIFT KEYING (BPSK) MODULATION TECHNIQUE USED IN GPS. 23
FIGURE 11: LAYOUT OF A GPS RECEIVER-SPOOFER, COURTESY OF HUMPHREYS AND LEDVINA. 28
FIGURE 12: DUAL ANTENNA DEFENCE WITH COMMON CLOCK. THE RAW PVT OF ONE OF THE OUTPUTS IS PUT THROUGH AN EXTENDED KALMAN FILTER AND PRESENTED AS GPS POSITION TO THE USER 36
FIGURE 13: DUAL ANTENNA DEFENCE WITH SEPARATE CLOCKS AND COMMERCIAL GPS OEM RECEIVER COMPONENTS. THE RAW PVT OF ONE OF THE OUTPUTS IS PUT THROUGH AN EXTENDED KALMAN FILTER AND PRESENTED AS GPS POSITION TO THE USER 38
FIGURE 14: DUAL FREQUENCY GPS RECEIVER DEFENCE, INSPIRED BY THE CROSS-CORRELATION METHOD, LAID OUT IN REFERENCE [WOO00] 38
FIGURE 15: GPS/INS DEFENCE. BOTH THE RAW PVT AND THE RAW INS DATA ARE FILTERED BEFORE CROSS-CHECKING TAKES PLACE. HYPOTHESIS TESTING DIMINISHES FALSE ALARM RATES, AT THE COST OF INCREASING TIME TO ALARM. 40
FIGURE 16: COVERAGE OF LORAN-C CHAINS AS IN 2006. DOD GENERAL PLANNING, 6 JULY 2006, NGA REF PLANXGP, PAGE 10-1. PUBLIC DOMAIN. 42
FIGURE 17: GPS/LORAN DEFENCE. THE LORAN RANGE CALCULATION CAN BE AUGMENTED BY TECHNIQUES DESCRIBED IN REFERENCE [LO09]. THE RESULTING RANGES ARE FED INTO A CROSS-CHECKING UNIT. 43
FIGURE 18: DATA LATENCY ILLUSTRATION, AS IMPLIED BY OPTION 1: DIFFERENCE BETWEEN SPOOFED (DELAYED) DATA BIT BOUNDARIES VERSUS THE AUTHENTIC SIGNAL, COURTESY OF T. HUMPHREYS. 44
FIGURE 19: AN ILLUSTRATION OF A CORRELATOR ARRAY (X AXIS) VERSUS SIGNAL POWER (Y AXIS). COURTESY OF T. HUMPHREYS. 47
FIGURE 20: ANATOMY OF A GPS NAVIGATION FRAME. 49
FIGURE 21: CRYPTOGRAPHIC DEFENCE BASED ON ESTIMATION OF W-BITS, METHOD OF M. PSIAKI. COURTESY OF M. PSIAKI, PERSONAL COMMUNICATION. 55
FIGURE 22: A HYBRID PSIAKI-LO DEFENCE, USING SYNTHETIC P(Y) SIGNAL SAMPLES. THE METHOD COULD BE APPLIED TO THE GALILEO PRS SIGNAL *MUTATIS MUTANDIS*. 59
FIGURE 23: VMS DEVICE MOUNTED ON A RAILING (PROCESSED IMAGE). 79
FIGURE 24: VMS DEVICE MOUNTED ON A TUBE ABOVE THE SHIP'S CABIN, AFTER A PHOTO BY M. BÖRJE. 80
FIGURE 25: RFID CHIP ON TOP OF A HANDHELD DEVICE; TOP RIGHT SHOWS PARTS OF A THRANE MODEL 3026, WITH AN ALLEN STANDARD BOLT. 80
FIGURE 26: AN RFID IS INSERTED INTO A HOLE DRILLED, IN A SINGLE OPERATION, THROUGH THE BASE AND A CLOSING BOLT OF THE UNIT, CONSTITUTING AN RFID SEAL. 81
FIGURE 27: SECURE ENCAPSULATED MODULE™, A W. L. GORE SECURITY PRODUCT, COURTESY OF W. L. GORE. 82
FIGURE 28: SECURE PLUG-ON MODULE™, W. L. GORE A SECURITY PRODUCT, COURTESY OF W. L. GORE. 83
FIGURE 29: *TIME TRAP DISPLAY AFTER THE CONTAINER IS OPENED. THE TIME THAT THE CONTAINER WAS OPENED […] IS PERMANENTLY DISPLAYED, ALONG WITH THE TWO LETTER HASH ("RF" IN THIS CASE) CORRESPONDING TO THAT TIME* [JOHNSTON06-2]. COURTESY OF THE VULNERABILITY ASSESSMENT TEAM, ARGONNE, USA. 86
FIGURE 30: *TIE-DYE BOLT SEAL (RIGHT) AND ITS READER (LEFT). THE TWO HALVES OF THE BOLT SEAL SNAP TOGETHER THROUGH A HASP. THE SEAL CAN BE OPENED BY HAND WITHOUT TOOLS SIMPLY BY PULLING THE TWO HALVES APART.* [JOHNSTON06-2]. COURTESY OF THE VULNERABILITY ASSESSMENT TEAM, ARGONNE, USA. 88
FIGURE 31: GASKET SET PROVIDED BY W.L. GORE. ONE EURO COIN ADDED FOR SIZE REFERENCE. 98
FIGURE 32: FIRST PART OF POSITIONING REPORT. EACH ROW REPRESENTS ONE BYTE. 109
FIGURE 33: SECOND PART OF POSITIONING REPORT. EACH ROW REPRESENTS ONE BYTE. 110

FIGURE 34: DUAL ANTENNA SYSTEM WITH COMMON CLOCK 124
FIGURE 35: SATELLITE MOVEMENTS IMPLY CARRIER PHASE CHANGES 125
FIGURE 36: EXPERIMENT 1 ON 29/SEP/2009, ENDED 16:38 GMT, DURATION 268 SECONDS 126
FIGURE 37: EXPERIMENT 2 ON 01/OCT/2009, ENDED 13:53 GMT, DURATION 74 SECONDS 127
FIGURE 38: EXPERIMENT 3 ON 01/OCT/2009, ENDED 14:09 GMT, DURATION 301 SECONDS 127
FIGURE 39: EXPERIMENT 4 ON 02/OCT/2009, ENDED 08:33, DURATION 341 SECONDS 128
FIGURE 40: VISUALISATION OF AN INDOOR TYPE SCENARIO. THE PHASE DIFFERENCE IS CONSTANT BETWEEN ANTENNAE ("DELTA ANTENNA") FOR TWO SPACE VEHICLES WITH INDICES *I* AND *J*, AT TWO DIFFERENT TIME VALUES. THE ALGORITHM TAKES DIFFERENCES WITH RESPECT TO TIME AND SPACE VEHICLES (DELTA_EPOCH_ANTENNA_SATELLITE), WHICH WILL YIELD A NEAR-ZERO VALUE FOR AN INDOOR SCENARIO. 133
FIGURE 41: VISUALISATION OF AN OUTDOOR SCENARIO. THE PHASE DIFFERENCE IS VARIES BETWEEN ANTENNAE ("DELTA ANTENNA") FOR TWO SPACE VEHICLES WITH INDICES *I* AND *J*, AT TWO DIFFERENT TIME VALUES. THE TERM DELTA_EPOCH_ANTENNA_SATELLITE WOULD BE MUCH LARGER IN ABSOLUTE THAN IN THE INDOOR SCENARIO. 134

List of Tables

TABLE 1: SPOOFING HARDWARE OPTIONS 27
TABLE 2: MEACONING HARDWARE OPTIONS 29
TABLE 3: ADVERSARY'S SHIELDING OPTION 31
TABLE 4: ADVERSARY'S CHOICE FOR SINGLE OR MULTI-FREQUENCY OPERATION 31
TABLE 5: ADVERSARY'S CHOICE FOR SINGLE OR MULTI SIGNAL OPERATION 32
TABLE 6: SOME IDEAS THAT PROBABLY WILL NOT (DURABLY) DETER SPOOFING ATTACKS 34
TABLE 7: GNSS RECEIVER DEFENCE MEASURES 68
TABLE 8: FAO VMS POSITION REPORT, SOURCE [FAO1] 109
TABLE 9: POLL FILE SPECIFICATION. LSB AND MSB STAND FOR LEAST AND MOST SIGNIFICANT BYTE. 111

Introductory Chapter

Author: U. Kröner

Abstract

Civilian GNSS based real-time tracking systems are presently used in a number of fields, such as the fisheries Vessel Monitoring System (VMS), the maritime Automatic Identification System (AIS), and the transportation of dangerous goods. Such trackers are commonly composed of a GNSS receiver module and a communications module for transmission of positions. GNSS-based trackers are vulnerable to tampering. Modelling such trackers, one identifies multiple ways that an adversary could use to introduce false tracking information. This document summarises an array of vulnerabilities and options for hardening such trackers, such as against fake GNSS signals, physical tampering, side channel attacks, and the substitution of position reports. It presents several cost-effective tracker designs.

Introduction

Civilian GNSS based real-time tracking systems are presently used in a number of fields, such as the fisheries Vessel Monitoring System (VMS), the maritime Automatic Identification System (AIS), the transportation of dangerous goods, in Mobile Asset Management, and in Law Enforcement. Projected toll road billing systems and "pay as you go" car insurances also plan to use GNSS based trackers [Spiegel09]. These devices record and report their positions. Transmissions usually occur in an expedient manner, implying near real time. The reporting intervals are determined by the application domain. The tracker sends its position messages to a public or commercial organisation that analyses such messages. Therefore, nearly all such trackers are composed of a GNSS receiver module and a communications module for transmission of positions.

Every such GNSS-based tracker is vulnerable to tampering, and gains from tampering depend on the application domain. While the unsophisticated adversary may use a metal barrier to block radiofrequency wave reception and emission, this document is more concerned with any surreptitious introduction of false tracking information, and the mitigation of such risk. With sufficient development work on a Software Defined Radio (SDR) platform, an adversary could engineer an illegal civilian GPS spoofer that circumvents classical GPS receiver integrity checks. The intelligent abuse of SDR further constitutes a looming threat to the authenticity of the outgoing position message. Furthermore, trackers also constitute embedded computing systems that have input/output ports, and as such are vulnerable to hacking techniques such as the "buffer overflow" exploit. If trackers then use cryptography to authenticate their outgoing message, they are vulnerable to "side channel attacks". Last but not least, it is possible to physically tamper with the tracker. Physical tampering can be broadly divided into two types, namely breaching the housing of the tracker, or removing the tracker from the asset.

This document reviews various options to harden GNSS based trackers against many exploits, and is organised as follows. The rest of this introduction further motivates the necessity of hardened GNSS trackers for civilian use, then classifies possible attack vectors, and finally touches on existing initiatives concerning various trackers. The first chapter investigates GNSS spoofing and meaconing, as well as several countermeasures. The second chapter discusses options regarding physical vulnerabilities and countermeasures such as volume sealing and removal protection. The third chapter explores aspects of side channel attacks and of the buffer overflow exploit. The fourth chapter shows

how communications can be secured, and specifically addresses low bandwidth satellite communications. The conclusion presents two possible architectures for hardened GNSS trackers that could be used in some application domains.

In terms of risk management, this document points out technological vulnerabilities of GNSS trackers. The document argues that because of the emergence of new technological vulnerabilities, and because of the continued presence of economic motives to conduct attacks on some trackers, the risks may need to be re-assessed per application domain. The document gives tools and options to mitigate technological risks. A proportionate technological response is presented to secure future generations of GNSS trackers. This response is conceived in terms of software, hardware, and the costs that pertain to these. Parallels with other application domains are drawn as appropriate. The document occasionally mentions other types of risk, such as those that can be exploited by social engineering. These types of risk however are excluded from the scope.

There are several useful complements to the present research, in risk management terms alone. First, a formalised risk assessment could be conducted per application domain. Furthermore, a set of control mechanisms can be devised, and incident response plans can be drawn up.

Why security in the context of GNSS trackers is important

The VMS regulatory environment as a special case

As mentioned further above, there are different kinds of trackers (e.g. Vessel Monitoring System, AIS) used in different application areas (maritime safety, fisheries monitoring, future road tolling, future commercial applications). Below the application area of the Vessel Monitoring System (VMS) is discussed in greater detail. Further below, one finds that some of the issues surrounding the VMS are shared in other application areas, such as road safety.

The VMS uses on-board devices that transmit position reports. In the European Union, the VMS was put into Community law in Commission Regulation (EC) No 1489/97[1].

However, many older VMS boxes are easy to circumvent, their electronics boards making it obvious where the GPS receiver feeds data into the rest of the unit [Nav01]. Many “first generation” devices, such as the widespread Thrane & Thrane 3020C VMS system, were particularly open to manipulation. This led to the development of specialised tampering kits, which were for sale on the black market [Nav05 pg 16].

Subsequently, Commission Regulation (EC) No 2244/2003 repealed Commission Regulation (EC) 1489/97. The new regulation states that *Member States shall adopt the appropriate measures to ensure that the satellite-tracking devices do not permit the input or output of false positions and are not capable of being manually over-ridden*, the *master of a Community fishing vessel shall ensure that the satellite-tracking devices are fully operational at all times* and prohibits *to destroy, damage, render inoperative or otherwise interfere with the satellite tracking device*.

At time of writing, these texts were subject to reform, in particular in the framework of the Reform of the fisheries control system. The recently passed Council Regulation (EC) No 1224/2009 refers to the adoption of further detailed rules.

[1] As mandated by Council Regulation (EEC) No 2847/93.

Outside of the EU, other Regional Fisheries Management Organisations (RFMOs) have adopted regulations[2] that also state that the VMS boxes need to be tamper resistant.

Modern VMS devices and their abuse potential: A fisheries perspective

First, when mentioning "adversaries", their "attacks", "exploits" and their possible "economic gain", it is important to clarify that these terms primarily refer to prospective activities of a small minority of individuals, namely those who have the least intent of complying with the law. It is generally accepted that in many populations, a small minority of individuals will systematically attempt to circumvent the law, another small minority will systematically attempt to comply with the law, and the great majority are situated somewhere in between. It follows that if enforcement is not effective, in that offences are not sufficiently detected and/or punished, then the great majority of individuals may start considering non-compliance (especially if non-compliance has economic benefits). On the other hand, if enforcement is effective, then compliance with regulations typically increases[3].

Applied to the specific case of fisheries monitoring, as a response to both the manipulation attempts and the new regulatory requirements outlined by Commission Regulation (EC) No 2244/2003, several newer VMS equipment (like the Thrane model 3026 "mini-C" and the Satlink ELB2010IR) attempt to make wilful non-compliance more difficult. While these devices most likely will help to deter a number of people, it remains to be seen whether they will stand the test of time. In particular they could face adversaries skilled in defeating seals [Johnston06], or equipped with GPS spoofing devices [Humphreys08].

In turn, when the adversary can circumvent the VMS devices without the authorities noticing, then illegal, unreported and unregulated (IUU) fishing typically takes place. It should be noted that in fisheries, compromising a VMS device is just one of many ways to engage in such activities. However, if illegal fishing could be conducted under the cover of a near perfect alibi, provided by a functioning but misled VMS device, then that would constitute a "worst case scenario" for fisheries enforcement.

Furthermore, IUU fishing is characterised by *strong economic incentive,* and by some estimates *accounts for up to 30% of total catches in some important fisheries* [IUU-WP]. IUU fishing creates several problems. It induces fishing mortality in addition to the one induced by legal fishing. IUU fishing creates an uneven economic playing field, which in turn tempts fishermen to follow a bad example. IUU fishing completely disregards regulations, and therefore disproportionally damages fragile marine ecosystems [IUU-WP]. IUU fishing also has detrimental effects on the livelihoods of artisanal fishermen off the coast of developing nations such as Senegal [BBC05]. It should be understood that poorer developing nations usually cannot enforce restrictions on fisheries within their economic zone. In summary, the winners of IUU fishing are the dishonest fishermen engaging in it, and the losers are the ecosystem, the inhabitants of developing nations, and any other fishermen.

But what sort of sums could an individual vessel owner from the EU gain by performing illegal fishing? During previous work [kroen09], two examples of illegal fishing and black landings were found, which for the involved vessel owners respectively added corporate turnover of 150 000 GBP over a single season for multiple smaller vessels, and 3.4 million GBP over two years for a very large

[2] Such as the "Extended Commission for the Conservation of Southern Bluefin Tuna", or the ones falling under the "Alaska Region Vessel Monitoring System Program". Above regulations refer to the use of secure housings and the application of outside tamper seals.

[3] This can *inter alia* be observed in road safety, with the compliance of drivers regarding speed limits. In Belgium, compliance in Flanders is higher than in Wallonia, mostly due to the high density of cameras in the former. To increase compliance, the European Commission asks Member States to proceed with the installation of an effective net of control measures [ETSC06].

vessel (the "Altaire III", measuring 74 m). With a 29% profit margin after taxes [BBC05-2], and given that a fishing season typically lasts for a few months, it stands to reason that the illicit profit, for a very large EU vessel performing illegal fishing, could well be in the order of magnitude of several thousand Euros per day.

In order to put these monetary amounts into perspective, the average Italian automated telling machine (ATM) issued 2.5 million USD in the year 2004 [BIS04, tables 11 and 13]. This amounts to 1.95 million EUR per year at 2004 average rates, or to about 5000 EUR per day.

Concerning the VMS devices, given the stakes involved, and the emergence of new technologies able to defeat them ([Humphreys08], [Sluiman10]), security enhancements may soon turn out to be necessary for larger vessels and/or high value fisheries. In turn, any advances made in the context of VMS devices could be rolled out to other types of GNSS trackers where security is of prime importance, such as the ones used to transport dangerous or expensive goods.

Returning to examining GNSS trackers in a more general way, the next subchapter examines trackers at an abstract level.

An abstract GNSS-based tracker

This document subsequently considers a GNSS-based tracker, which is introduced as follows:

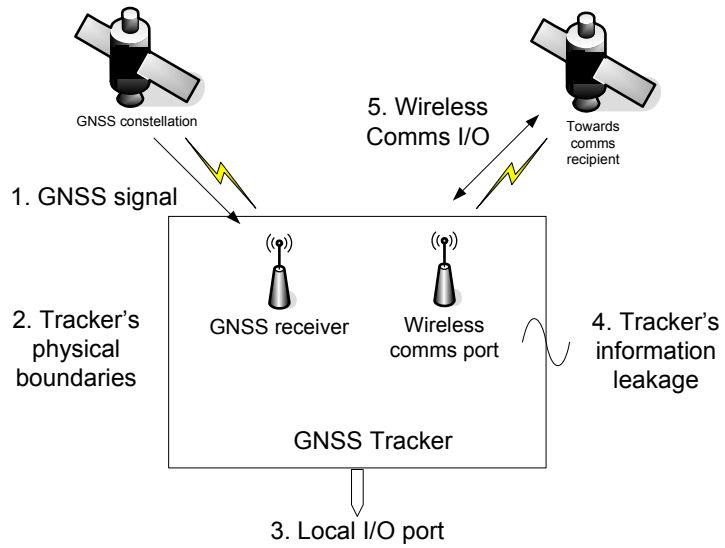


Figure 1: An abstract GNSS based real-time tracker

The GNSS based real time tracker is characterised by:

1. A GNSS constellation (GPS) emits a signal that is interpreted by the GNSS receiver inside of the tracker;
2. The tracker is installed on the asset that is to be tracked, and sealed to it. The tracker has physical boundaries, meaning that the electronics is contained in a box. (The tracker may be distributed in several boxes, however in this paper the focus shall be on trackers

that are contained in a single physical volume, as the use of multiple volumes makes physical protection unnecessarily complicated.)

3. Most trackers have an I/O port[4], such as a USB port for privileged access by authorities. Many trackers can be interacted with, by messages sent over the air[5].
4. Nearly every tracker draws power from an external source. Trackers are more or less well shielded against radiated information leakage.
5. The tracker finally transmits its information to a communication network. While land-based trackers tend to use GSM networks, maritime devices must rely on communication satellite coverage[6].

The underlying abstraction is that all trackers occasionally communicate information to a controlling entity. Also, all trackers are attached to an "asset", and have an "end user", which is a person or entity responsible for managing the asset.

Note: This document occasionally discusses VMS messages, which will refer to satellite communications to/from the authorities. This discussion is without loss of generality, since there must always be an exact mechanism for the controlling instance to recuperate the tracks of a GNSS tracker. The exact mechanism by which this occurs (real-time, delayed, using RF or conducted input/output) is less important than the general principle that any GNSS tracker eventually makes such tracks available.

A formalisation of vulnerabilities and defences

This chapter examines the possibilities that an adversary has to break such a system, and the conceptual characteristics of a system that would resist tampering to a higher degree.

Formally, in order to tamper with the tracking device, an adversary must pick at least one attack scenario:

1. Calculate and broadcast a false GNSS signal ("spoof", see glossary), or retransmit a real but delayed GNSS signal ("meacon", see glossary);
2. Physically breach the tracker, then exploit the direct access to the hardware; or remove the tracker from the asset to be tracked.
3. Gain illicit and privileged access the software on the inside of the device. Most devices need some form of computer interface, conducted or by radio-frequency. This interface can be subverted, e.g. by using the "buffer overflow exploit", to gain unauthorized privileged access [Gerg05];
4. Perform information leakage side channel attacks: read RF emissions, use timing attacks[7], or use power analysis in order to infer secrets. Operate the device outside of voltage or temperature range in order to achieve device failures that reveal secret information. Apply a strong electromagnetic field to disturb sensitive parts of the device (such as a hardware random number generator using thermal noise);

[4] This is for instance required for "Safety of Life at Sea" (SOLAS) equipment, and some VMS devices are used in such a way. Other VMS devices permit the use of their communications satellite connections to send emails.
[5] This feature is present on most VMS devices for the purpose of position polling.
[6] The AIS constitutes a special case as it creates ship-to-ship ad-hoc networks for collision avoidance.
[7] If cryptographic software is used carelessly, then "side channel attacks" [SCA-WP] can break cryptographic systems such as those based on AES, which in theory may seem impenetrable. An example, complete with source code, is provided in [Bernstein05].

5. Tamper with the satellite communications. More specifically, in order not to raise suspicion by the authorities, an adversary benefits from being able to substitute, i.e. spoof the outgoing satellite communications message. Satellite-based VMS devices offer the possibility of receiving configuration messages via communications satellites, the so-called "polling messages"[8]. The adversary could also forge these configuration messages.

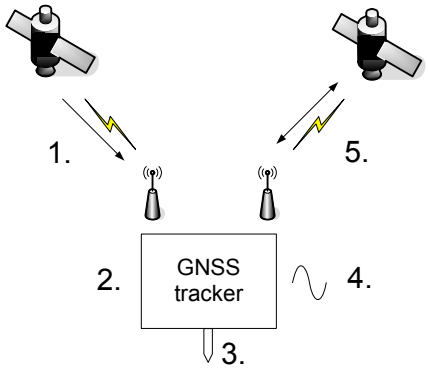


Figure 2: Schematic levels of attack: 1) incoming, 2) physical, 3) information technology, 4) electromagnetic, and/or 5) outgoing

If one counters each of the above weaknesses (1) to (5), one obtains the characteristics of a tracker that presents increased resilience.

This in turn means that one needs to implement the following[9]:

I. Fake GNSS signal mitigation: To make sure that the GNSS signal is both real and live, implement a defence against fake GNSS signals, and send both GNSS health status and a GNSS timestamp in each outgoing tracker message.

II. Physical security: to design a tamper-respondent and physical volume seal, devoid of exploitable holes in the surface. On breach an actuator must erase any secret the adversary could use to clone the tracker. Also the tracker must be properly sealed to the asset to be tracked;

III. Eliminate or secure local I/O interfaces. Security would entail

   a. At a hardware level: monitoring of bad voltage levels. Deny service and cancel any I/O at such levels.

   b. At a software level: any input must guard against buffer overrun protection, and must come from an authenticated user.

IV. Limit information leakage, meaning to reveal as little details as possible through electric, electromagnetic, or timing information. This includes masking any undesired RF emission, masking the power consumption profile, permitting only a limited number of operations per unit of time, including sensors to monitor for operation outside of voltage or temperature range, and adding a battery that provides power during the occasional power failure. Furthermore, the

[8] One of these polling messages is the message that alters the VMS device's device ID, meaning the DNID and the member number [Mrag02 appendix 6].

[9] One can consider the victim GNSS tracker to be a VMS device used for fisheries, but these attack scenarios can be transposed to other trackers as well.

outgoing message should include “health status” information regarding e.g. battery, input voltage, and temperature.

V. Wireless communications security: In parallel to (III), any satellite communications I/O needs to be carefully secured. This concerns both hardware and software. The outgoing message must contain a detailed time stamp, and would be digitally signed or encrypted. This makes sure that the message is both real and live. Should the tracker be “pollable” by satellite, as is the case for most VMS devices, then the input signals from the satellite must be secured in a similar manner. The device must then guard against buffer overruns as in (III), and reject any input for which the authenticity cannot be ascertained.

Recent improvements regarding VMS device security in EU Member States

Member States are aware that the efficacy of the VMS, including that of the VMS on-board devices are a necessary part of their own Fisheries Monitoring and Control strategy, as well as a necessary part of the EU Common Fisheries Policy.

Hence, Ireland (2008) and Cyprus (2009) have pro-actively introduced newer and more secure VMS devices. Some of the components of these newer systems are relevant to present research. Therefore these will briefly be explored below.

The Irish FMC decided in 2008 to renew their installed VMS device base. The VMS terminal combines a Satlink model ELB 2010 IR “blue box” with an external Skywave DMR800L active antenna for the Inmarsat ISatM2M service. The antenna itself is encased in an anti tamper adapter. The system is designed such that for the fisherman, the device is “fit and forget”: it is installed once and then maintenance-free.

The Satlink model ELB 2010 IR is a customised version of the ELB 2010. These differ in that the ELB 2010 IR:

- has a power input, an antenna input, but no data input,
- has an internal tamper sensor switch, which when activated will alert the authorities,
- Contains a re-chargeable battery,
- Has two connection points for seals, one on each side of the cover.

Figure 3: The Satlink ELB 2010 IR special type, courtesy of Satlink. The notches on each side house connection points for seals

The antenna is housed in an anti-tamper adapter that is sealed with two commercial security seals.

Figure 4: The Skywave DMR 800L antenna (left) is put in the anti tamper encasing (right). The encasing has two security seal holes. Courtesy of Satlink

All of the four security seal holes will be fitted with commercially available seals. While commercial seals have their weaknesses [Johnston06], such practice is clearly better than no sealing at all.

In addition to the physical security mentioned above, the VMS device of the Irish authorities has a number of reporting features that are of general interest to Fishing Monitoring and Control:

- Features that are present on many other VMS devices, and are also implemented on this device, are:
    - Geo-fencing used for port approach, precise effort zone reporting, marine conservation areas, and areas with other special reporting requirements
    - Reports on antenna blockage
- GPS data logger: Under normal mode of operation, the device records its GPS position once every 15 minutes, these position data are then retained for three months, and in this time frame the FMC can download them remotely.
- Sleep mode:
    1. If the main power is removed[10], the terminal switches to a battery-based operations mode called "sleep mode".
    2. When that happens, it immediately sends a report about this event to the FMC.
    3. In this mode, the terminal will wake up once every hour, and checks its position and velocity using its GPS receiver.
    4. If it has significantly moved from its previous position, or has significant speed, then it immediately leaves the sleep mode, sends an event report to the FMC, and will perform VMS position reporting once every hour, until power is restored or the battery runs out.
    5. If it has stayed at its last position and speed remains insignificant, the terminal will send a report once every 24 hours.
- Stillness mode:
    1. This mode is similar to "sleep mode" except that the main power is available, and that the GPS receiver is powered permanently[11].
    2. If the vessel has not moved from its previous position, and has no significant speed for a certain period of time, it enters stillness mode and sends a report about this event to the FMC.
    3. If it has stayed at its last position and speed remains insignificant, the terminal will send a report once every 24 hours.
    4. Otherwise the VMS device resumes standard reporting, and sends a matching event report to the FMC.

[10] Fishing vessels routinely use this mode when they are docked in a harbour, and switch off the mains. Very occasionally, this mode could be triggered in event of electrical issues at sea.

[11] Fishing vessels routinely use this mode when they are docked in a harbour and do not switch off the mains, or when shutting down the vessel over night at sea.

Economically speaking, the Irish authorities provide the terminal to all fishermen free of charge, and this includes a three-year maintenance contract. The Irish authorities pay for all messages that they request but are not strictly required by law (such as effort zone reporting, increased message frequency upon port approach, downloading of 15-minute GPS positions).

For the Irish vessel owners, in addition to not having to pay any one-time fees, the airtime costs are reduced by up to 50% when compared with some previously used VMS devices. In addition to economic benefits, fishing companies have access to their own VMS positions via the Internet.

Moving from Ireland all the way across the EU, Cyprus has decided to install a new terminal on all fishing vessels having a length of less than 15 metres. For that, the Cypriot authorities use a novel VMS terminal from the company Skywave Mobile Communications Inc. of Canada, the model MT 3300. This terminal fully complies with the VMS regulation, and has its origins in the road transport industry[12].

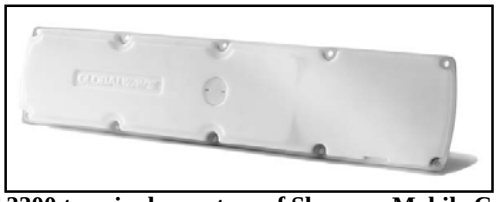


Figure 5: The "Slap & Track" MT 3300 terminal, courtesy of Skywave Mobile Communications, Ottawa, Canada

It has the following features

- Fully autonomous: Does not require a power supply. Includes a non-rechargeable battery that lasts for 2 years with 1 message per hour.
- Configurable over the air, by using Inmarsat messages.
- Can receive Inmarsat polls and can react by sending an immediate real-time GPS location
- No on/off button, however a serial port is attached that permits the unit to be reconfigured.
- Economical: Costs 700 EUR to buy, 110 EUR to renew the battery, and 40 EUR per month for messaging.

In addition Cyprus decided to seal the device to the boat, using a commercially available twist seal. While such seals are no panacea [Johnston06], a commercial seal is a clear improvement over no sealing at all.

Parallels between the VMS devices and the AIS transponders

Larger transport vessels, passenger vessels, and fixed maritime platforms have to be fitted with Automated Identification System (AIS) transponders as per IMO SOLAS Chapter V rules. A basic summary of the AIS transponders can be found at the Wikipedia [AIS-WP]:

> *AIS transponders automatically broadcast information, such as their position, speed, and navigational status, at regular intervals via a VHF transmitter built into the transponder. The information originates from the ship's navigational sensors, typically its global navigation satellite system (GNSS) receiver and gyrocompass. Other information, such as the vessel name and VHF call sign, is programmed when installing the equipment and is also transmitted regularly. The signals are received by AIS transponders fitted on other ships or on land based*

[12] This is not entirely surprising. The tracking devices used in the road transport industry have other similarities with the VMS device, see the chapter below on Parallels with the European Digital Tachograph.

systems, such as VTS systems. The received information can be displayed on a screen or chart plotter, showing the other vessels' positions in much the same manner as a radar display.

The AIS standard describes two major classes of AIS units:

- *Class A – mandated for use on SOLAS Chapter V vessels (and others in some countries).*
- *Class B – a low power, lower cost derivative for leisure and non-SOLAS markets.*

Strictly speaking, it is not entirely correct that *Automatic Identification System (AIS) is a short range coastal tracking system* [AIS-WP], since the AIS transceivers spontaneously create networks between ships, as soon as they are within VHF communications range, owing to the particularities of their communications protocols. AIS information is therefore usually exchanged whenever a transponder is within VHF communications range of another transponder.

The AIS class transponders communicate up to *26 different types of messages* [AIS-WP], and each message type is re-sent after a particular time interval. Some of these contain the exact position (WGS 84 longitude and latitude) of the vessel.

The AIS Class A and Class B transponders are different in other aspects, which are relevant in this context.

- The Class A transponders commonly obtain their position and speed data from the ship's internal systems (using the NMEA 0183/IEC 61162 protocol, physically linked to a ship-wide data communications medium, such as a coaxial cable). Each Class A transponder also contains a GNSS module (usually GPS or GLONASS), but under normal operations, this GNSS module is just used to obtain a precise time reference, needed for the VHF communications protocol.
- By contrast, many currently available commercial AIS class B transponders obtain their position and speed information exclusively using a built-in GNSS module. These same transponders commonly provide data output, but have no corresponding data input.
- The Class A transponders usually send a broad range of the 26 different AIS message types. The Class A participates in a SOTDMA network (self-organising time division multiple access) where each party knows which other party transmits information at what point in time. Class A messages can therefore occupy several slots of 256 bits each. In theory, a single class A message could be made so large as to hold authentication data (see below).
- By contrast, class B transponders send outgoing VHF messages using a protocol called CSTDMA (carrier sense time division multiple access), which implies that Class B messages tend to occupy only a single slot[13] of 256 bits [AIS-WP]. The Class B message types of 14, 18, 19, 24, and 25 [AIS-USCG] are therefore either too short for authenticated data, or are fixed in their structure, or both.

An AIS Class B transponder can therefore be considered as a GNSS tracking system with the following particularities:

1. It communicates outgoing position messages, and receives incoming position messages from peers, using VHF. It does not usually require global data connectivity; instead it transmits data in line of sight.
2. It is commonly composed of three physical housings. These are the main unit, usually housed in the cabin, an active GPS antenna, and the VHF transmission antenna. The antennae are connected to the main unit by connector cables.

It follows that for hardening purposes, authenticated position reports could be sent using other means, by using communications satellites in real time, or using e.g. UMTS for deferred transmission.

[13] The only exception to the "single slot" rule is the "Extended class B" report of two slots, which a Class B only sends if it is polled by an AIS shore base station.

An AIS Class A transponder shares the above particularities (1.) and (2.) of the Class B transponder, but there are several differences.

- It does not use the position and speed determined internally, but rather it receives its position and speed information from ship internal communications. However, for hardening purposes, the data it receives could be cross-checked against the position and speed reported by its internal GNSS module.

Therefore, an AIS Class A transponder could in theory attempt to reserve 3 slots, and then transmit an authenticated message of at most 768 bits. Of those bits, exactly 512 are needed for the digital signature using asymmetric cryptography, at a security level of 128 bits. Transmitting authenticated position reports using AIS is thus bandwidth-intensive, and may not be acceptable in regions with dense vessel traffic.

Suppose that a vessel with a Class A transponder attempts to pretend that it is located at point X, with dense traffic, while it is in fact located near Y. To transmit a fake report from Y pretending to be at X, its AIS Class A transponder would then pick an AIS transmission slot. But in the area X where it pretends to be, this transmission slot may already be in use. However, this may not be an issue for the adversary, as the satellites receive messages from a wide swath with a cross-section of about 5000 km, and so the satellites cannot reliably cross-check the individual transmission slots. (Note: if the adversary pretends that he is e.g. 5000 km from his true location, then he may be trivially detected by S-AIS.)

To mitigate the issue of spoofed AIS Class A messages, the transponders could therefore instead transmit authenticated position reports using other means of communication.

Parallels with the European digital tachograph

In the European Union, the Digital Tachograph (DT) serves to enforce legal limits on road truck driving times and speed limits, and making it difficult for transport companies to impose abusive working conditions. These restrictions improve road safety, and create a level playing field within the truck driver industry. While transport driver unions are generally favourable towards the DT, the transport companies and independent drivers sometimes attempt to manipulate the devices to their advantage.

The DT is introduced and fully defined by Regulation (EC) No 2135/98 and its Annexes. The regulatory environment was updated by Regulation (EC) 561/2006.

The DT is composed of a Motion Sensor (MS) unit, which is embedded in the gearbox, a Vehicle Unit (VU), and a passive cable linking the MS unit and VU:

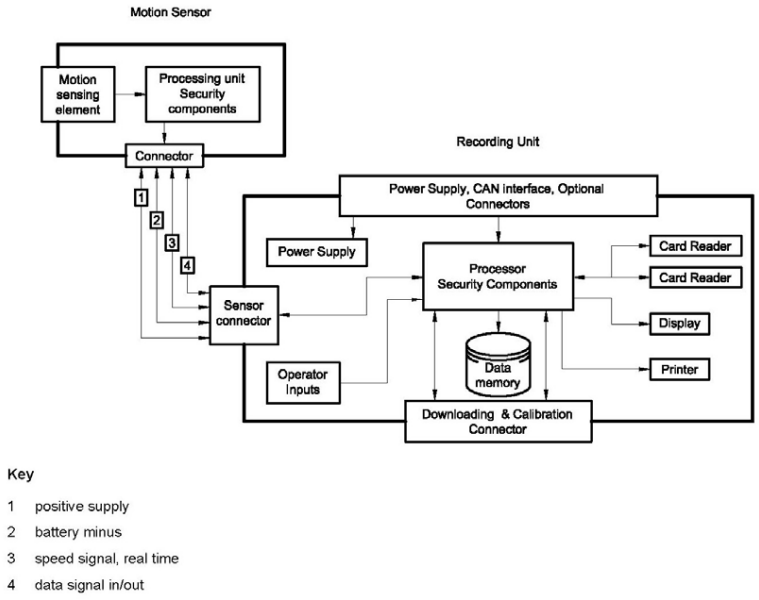


Figure 6: "*Schemata of a typical recording equipment (extracted from ISO-16844-3)*" [Nordvik07]

A few of the security features of the DT system, and ancillary infrastructure, are

- **Full use of "smart card" chips**: All parts of the DT system are secured by "smart cards", which have been pioneered in the banking and digital broadcasting world.
- **Full use of cryptography and PKI**:
    - The drivers, transport companies, authorised workshops, and inspectors all have their own chip card. Each chip card includes asymmetric cryptography, and is certified by a Public Key Infrastructure.
    - The MS unit and the VU are "paired", in that they use mutual cryptographic authentication, prior to the MS unit sending speed information to the VU unit.
- **Complete access logs**: When any of the four above parties inserts his chip card into a VU, the VU immediately records the access.
- **Tamper-evident seals**: The MS unit, the VU, and the cables are sealed using tamper-evident seals.

In summary, the DT uses a well-thought-out and modern system of security measures. However, that has not kept crafty engineers from defeating the system.

Just like in fisheries, cheating is economically advantageous and sometimes performed by unscrupulous individuals and companies. Cheating on the DT also created its own niche market, with a comparatively high degree of sophistication and creativity: In the reference [Nordvik07], JRC colleagues enumerate four different ways to cheat the DT on-board device, each one exploiting a different weakness. An exploit known to the public consists of attaching a magnet to the gearbox, which then upsets the Hall Effect sensor inside of the MS unit.

Figure 7 depicts a standard MS unit. The two metal pins on the very right connect to the Hall Effect sensor in the electronics. The left part of the electronics attaches to the cable that leads to the VU.

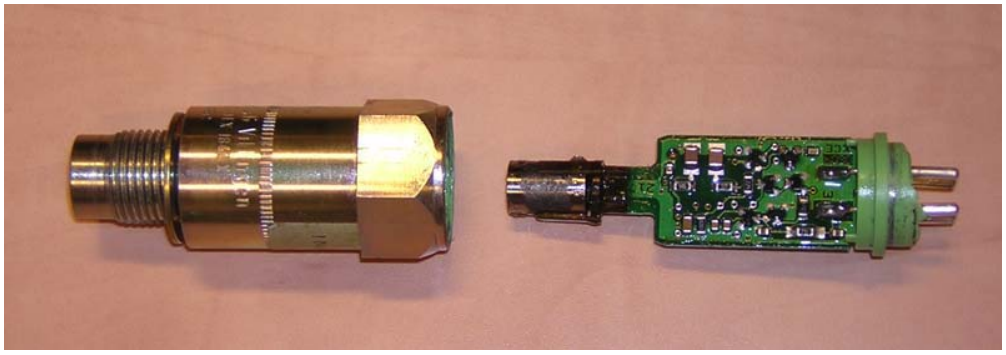


Figure 7: ***A typical MS and its internal component*** **[Nordvik07]**

JRC/IPSC colleagues from the Seals and Identification Lab (SILab) have also consulted with physical sealing specialist R. Johnston, and agree with his conclusions that physical seals represent a weakness of the system. In order to mitigate some of the risks, "anti-evidence" tamper seals [Johnston06] are considered as a research avenue.

Some of the parallels to the VMS devices are that:

- **Industry limited by EU law**: In both situations, an electronics device mandated by EU law is used to impose restrictions on motorised means of transport in a particular industry, with the intent of creating a level playing field;
- **Propensity to cheat the system**: While the great majority of stakeholders play by the rules, a small minority of dishonest operators will try to cheat the system to their economic advantage. If this is not addressed, then more stakeholders will be "stray from the path" laid out by the legislator;
- **Existence of a number of loopholes to cheat the system**: Both the VMS and the DT can be undermined by a number of tricks, which are either legal or belong to a grey area. An incomplete list for the VMS is given below:
    - Some fishing vessels from certain Flags of Convenience engage in pirate fishing. Under the current law of the seas, these skippers can refuse their vessel to be boarded, if they are found in open seas. There are private companies that cater to the needs of pirate fishermen, by providing supplies or off-loading the catches;
    - Other fishing vessels belong to "joint venture" companies from different states; commonly one is an EU member state and the other is not. These vessels then change flags in order to exploit two different quota, however the economic operator remains the same;
    - Other fishing vessels that have the VMS are known to engage in "area misreporting", where they engage in fake fishing manoeuvres, only to be able to declare their catch as coming from the area where they executed the fake manoeuvre;
    - Still other fishing vessels have unexplained power supply "interruptions" that disable the VMS device on board. While this is explicitly forbidden by EU law, some power interruptions must clearly be tolerated in order to account for real problems.

  An incomplete list for the DT is given below:

  - Trucks from some third countries outside of the DT regulation could conceivably cheat the system by pretending to transport goods to/from that third country, while actually transporting goods inside of the EEA (European Economic Area);
  - Some truck drivers fake power interruptions to the DT, in order to run various exploits. Again, some power interruptions must be tolerated in order to allow for real problems;
  - Some transport companies make use of "light vehicles", to which the current regulation R (EC) 561/2006 does not apply.
- **Existence of cheating hardware**: Devices have been found to be in use, or can be bought, that make it possible for dishonest operators to abuse the system.
- **Devices composed of two housings**: All DT devices consist of a sensor housing that must be ruggedized, a cable, and a central unit. The same is true for many VMS devices (but not for all of them).

Some of the differences between the DT and the VMS are:

- **Primary mission**: The DT's primary mission is to enforce road safety, attempting to save people's health and lives. The VMS device's primary mission is to control that the fishermen do not access marine reserves and real-time closures, attempting to save the fishing industry from overfishing its resources;
- **Acceptance**: The transport driver unions, in great majority, view the DT system favourably, while the VMS does not enjoy a similar status among the fishermen;
- **Economics of cheating**: Cheating on a single tachograph will not make you an extra six-figure euro amount per year. By contrast, owners of very large pelagic vessels could illegally reap such profits, by defeating their on-board VMS;
- **Real-time properties**: The DT logs speed by attaching to a gearbox, and the log can only be downloaded in non-real time. The VMS tracks position and velocity by using GPS, and directly transmits this information in real time;
- **Possibility to limit the device to one housing**: Recently, many VMS device designs have emerged that make use of a single housing. It may therefore be possible to require that new VMS devices be composed of a single unit. But unless the DT no longer attaches to the gearbox, the DT will remain a two-component system: The extreme environment near the gearbox prohibits the presence of slots for insertion of chip cards.

What do the lessons of the DT imply for the VMS? Clearly, if over-exploitation of fish stocks remains an issue in which the VMS plays a key part, then regulators would need to close loopholes, and hardware vendors would need to increase security on the VMS devices. Under such a scenario, from the parallels between the DT and the VMS device, one would expect the situation concerning the VMS to unfold as follows:

- If future VMS devices do not become an obstacle to occasional illegal fishing, or if various loopholes continue to be exploitable, then one would expect little activity in the area of specialised VMS cheating hardware;
- If the VMS device represents a serious obstacle to unscrupulous fishermen, then it is safe to assume that they will collude with talented and equally unscrupulous workshops, in order to defeat their VMS device;
- Should electronic cheating not be economically viable, then "social engineering" may become the method of choice for such fishermen.

In conclusion, this document claims it is reasonable to assume that a small minority of fishing vessel owners will engage in sophisticated electronic cheating, if that is a viable option to "work around" their VMS devices. This would parallel the present-day situation in the domain of road transport.

Chapter I: Defending against fake GNSS signals

Author: U. Kröner

Chapter introduction

Scope

If one wants to design a GNSS receiver that detects and/or resists attempts at manipulation, one first needs to dress an inventory of means available to an adversary.

In this chapter, following an introduction into GNSS vulnerabilities, this document will review the various possibilities that an adversary has, to spoof or meacon GNSS signals.

Once the possibilities of the potential adversary are explored, the receiver hardening measures apt at mitigating the various attacks will be considered.

Finally, this document presents a few cost-effective receiver designs, which when combined, are comparatively expensive to subvert.

A layman's overview of GNSS

Before exploring vulnerabilities, some principles of how a GNSS functions must first be clarified. The summary presented below is neither accurate nor complete, but is given with the intent of introducing some of the principles of GNSS.

With the above caveat in mind, there are several Global Navigation Satellite Systems (GNSS), such as the US' existing Global Positioning System (GPS), the EU's fledging Galileo, Russia's nearly restored GLONASS, and China's upcoming COMPASS.

Grossly and perhaps overly simplified, GPS works as follows[14]:

1. GPS uses a constellation of about 30 Medium Earth Orbit (MEO) satellites.
2. Each GPS satellite transmits a repeating navigation message containing timing, satellite specific information, and constellation specific information. The navigation message repeats every 12.5 minutes, contains 37 500 bits in total, and is composed of 25 message frames. The signal is transmitted at relatively weak power levels. To quantify, GPS signals in the L1 frequency band of 1.57542 GHz leave each satellite antenna at a power level of 500 watts, or 27 dBW.

[14] Sources: [GPSPOWER], [GPS-WP], [GNSS-WP], [GPS-SIG-WP], [Kowoma], [Borre07]

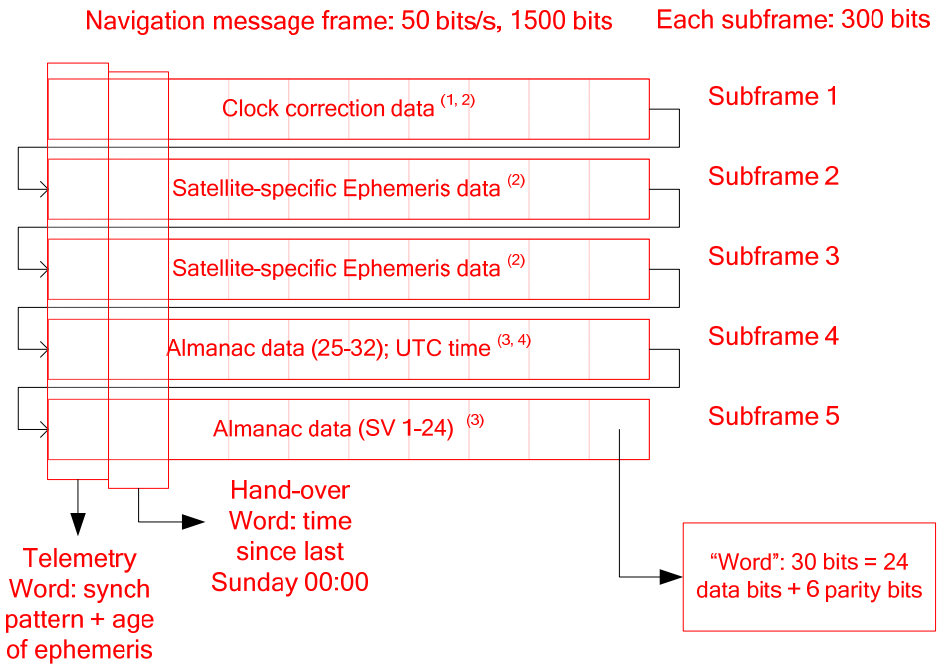


Figure 8: Anatomy of a GPS L1 C/A navigation frame.
Each frame takes 30 seconds to transmit, and is composed of 5 logically separated subframes. Each subframe is composed of 10 "words". Each word contains 24 bits of payload data, and 6 bits of parity data. The first two words in each subframe are identical.

(1) Second degree polynomial coefficients used to calculate the satellite clock offset
(2) Repeats every navigation frame (30 seconds)
(3) Repeats every 25 frames (12.5 minutes): Different frames contain different parts of a longer message
(4) Contains Almanac, ionospheric correction data, UTC (Universal Time Coordinates) time, other data.

3. The commonly used civilian "L1 C/A" signal is encoded as follows: a navigation message (50 bps or bits per second, or 20 ms per bit) is added[15] to the satellite specific C/A code[16] (1.023 Mbps, on a sequence repeating every millisecond), and the result is then modulated onto the carrier signal of the L1 frequency. The resulting civilian channel has a bandwidth of 2 MHz.

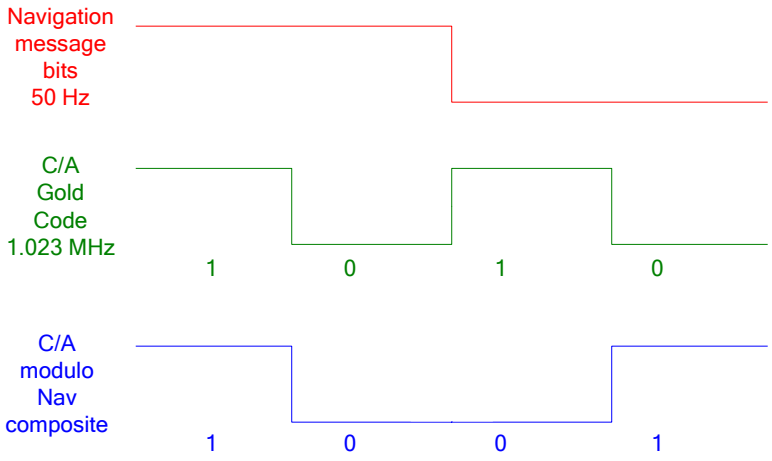


Figure 9: Signal modulation. The navigation message bits (at 50 Hz) are added modulo 2 (XOR) with the C/A Gold Code (1.023 MHz), yielding the combined signal (blue).

[15] The contents are fed into an XOR gate.
[16] The C/A codes are so-called Gold codes, after Dr R. Gold. The codes have certain mathematical properties that make them easier to line up.

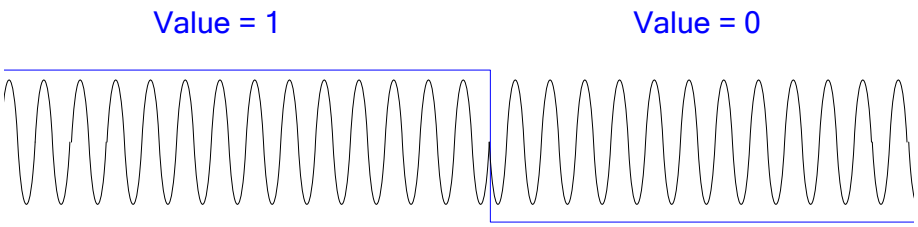


Figure 10: Carrier signal anatomy. The black graph represents the in-phase component of the GPS L1 C/A carrier signal (1.57542 GHz). The blue square wave represents the C/A code modulo the Navigation message (1.023 MHz). The carrier phase is shifted by 180° at the value boundary, representing the Binary Phase Shift Keying (BPSK) modulation technique used in GPS.

4. When the signals from each channel reach the GPS receiver after about 21 000 km of transit, they have lost about 182 dBW. So GPS signal power levels measured at the receiver (at -155 dBW[17]) are beneath the "thermal noise floor" (for the C/A signal of bandwidth 2MHz, the Johnson-Nyquist thermal noise floor is -141 dBW). Physics dictates that thermal noise ("Johnson–Nyquist noise") is generated by any type of electronics equipment. Other types of noise may also be present.
   Furthermore, because of the Doppler shift, signals can be shifted in frequency by up to +/-10 kHz. Any shift in frequency is accompanied by a proportional change in the time it takes to receive one C/A code sequence.
5. Before applying any signal processing to the GPS signal, a GPS receiver uses a radio frequency (RF) front-end to filter and amplify the received waveform. The waveform is then converted from the carrier frequency (e.g. L1) to "baseband"[18]. Then the front-end takes numerical samples.
6. After it has converted the signal to numerical samples, the GPS receiver interprets the resulting signal. Because of the GPS signal being weaker than the thermal noise floor, a typical[19] GPS receiver is unable to read out the individual C/A signals directly, so it must infer it using statistical signal processing. Therefore, before any other consideration, the receiver always accumulates a set of signal samples before attempting to read them out. Under normal conditions, one millisecond of signals is accumulated, which is precisely the time it takes for a single C/A signal to repeat itself. This accumulation delay is relevant to some topics presented further below. Because thermal noise has an approximate Gaussian distribution, the accumulated samples amplify the signal-to-noise ratio of the GPS channels, but not the thermal noise.
   In order to correlate the samples, the receiver searches in a 3-dimensional matrix consisting of satellite number, Doppler frequency offset step, and code phase step. Each one of these matrix entries is termed a "search bin". In other words, for each possible satellite and in frequency steps of 500 Hz accounting for the Doppler shift offset, and for each possible offset between the code replica and the actual code, the receiver correlates the received sample with its replicas. The receiver knows that it has acquired a satellite code when, inside one search bin, the absolute correlation power for a given satellite's C/A code reaches a pre-determined threshold
7. Once the receiver finds the correlation peak for a given signal, it is said to have acquired a satellite. It then knows that this satellite is present, and what the Doppler frequency is. It then "locks" its C/A code replicas in both the frequency and time domains (technically it parameterises its Phase and Delay Lock Loops). These lock loops are then used to continue the tracking of the satellite's signal. The receiver then recovers a first navigation message bit. Any correlation parameters are then stored, in order to simplify the read-out of further navigation message bits. If

[17] The guaranteed ("nominal") power level is between -158.5 and -160 dbW, depending on the literature one consults. The difference of about 5 dB is a tolerance margin that allow the satellites to progressively lose transmit power until their end-of-life.
[18] Baseband refers to a frequency band centred at zero Hertz.
[19] In this context, "typical" should be interpreted as those having omnidirectional antenna(e).

the satellite changes the sign of the navigation message bit, then the receiver will see this because of sharp and characteristic changes in the correlation amplitude.

8. The receiver repeats the above process for the other navigation message bits. On consumer grade commercial receivers, this is performed in parallel for many satellites, usually for up to 12. After 30 seconds the receiver should have a useable portion of the navigation message.
9. The receiver then feeds the data (consisting mostly of the navigation message containing time and approximate satellite constellation or "almanac", the correlation process results, and the status of the C/A tracking loops) into the range calculations: Given the positions of the satellites, the transmitted time, the delay between the signal's emission and reception, and physical constants such as the speed of light in vacuum, the receiver calculates its "range"[20], or distance from each satellite.
10. Once it has four or more such ranges, it proceeds to solving an equation with four unknowns[21], which will yield the receiver's latitude, longitude, height above ground, and time. A low-cost GPS receiver will give position reports that, in 95% of all cases, are accurate to within about 10-12 metres[22].

The above explores the US' GPS system. However note that other systems work according to the same principles, hence the above characterises other GNSS systems or other GPS signals *mutandis mutatis*.

As suggested above, GNSS is a complicated and wide subject matter. First, GPS was originally developed for the US Military, which uses another signal structure in addition to the L1 C/A. Then, even for a "simple" L1 C/A civilian receiver, there are many complications to this simplified description, such as relativistic distortions, multipath, atmospheric effects (ionosphere and troposphere), and errors in the satellite's positions/clocks, to name a few. Last but not least, there is the topic of GPS position enhancements for the civilian sector. In order to gain a full understanding of GNSS, the interested reader is invited to consult the GPS reference literature, such as "Global Positioning System: Theory and Applications (Volume I and II)" [GPS-REF].

Options for attacking a GNSS receiver

Outline of GNSS vulnerabilities

GNSS, such as GPS, has a number of vulnerabilities, some of which are jamming, spoofing, and meaconing. The summary on the latter, presented below, is neither accurate nor complete, but is given with the intent of introducing some of the principles surrounding vulnerabilities of GNSS.

- **GNSS jamming** causes the GNSS signal not being available in situations where it normally should be. Jamming is usually caused by an electronic device that creates signal noise, which interferes with the signal in the GPS frequency band(s). Because GPS signals are received at -155 dBW of power, very little transmission power is needed to deny GPS service at any location

[20] The ranges are often called "pseudo-ranges" in GNSS parlance, to reflect that they were obtained by the following, for each acquired satellite:

Pseudorange = travel time * light speed + constant offset

In other words, the initial pseudo-ranges are all relative to each other. These pseudo-ranges are afterwards corrected for errors, such as the imprecision in the satellite's and receiver's clock, and the effects of slower propagation in the presence of water vapour (in the troposphere). The constant offset is determined later, yielding the precise receiver time and eventually the position.

[21] Specifically, most receivers use a mathematical process known as trilateration.

[22] GPS receivers are able to line up their C/A code replicas with the actual signal to about 1% of one C/A code chip. Each C/A code chip takes 1 microsecond, implying a C/A range of 300 metres, yielding a nominal C/A signal arrival imprecision with a standard deviation of 3 metres in the receiver. This nominal accuracy is diminished by a few factors, the dominant one being the ionospheric effect, which causes another error with a standard deviation of 5 metres. Cumulatively, all factors yield a standard deviation of 6.7 metres.

within line of sight of a jamming beam. The jamming signal then blocks GPS signal acquisition, in that the GPS unit cannot correlate the GPS signals.

- The interference that causes the jamming can be unintentional, such as those created by an insufficiently shielded TV antenna [Petovello08]. If the TV antenna emits signals at 525 MHz or 788 MHz, then it may well interfere with the L1 frequency, since the latter then represents the 3rd and 2nd harmonic of the TV antenna. The power levels of such an unintentional jamming are typically orders of magnitude higher than those of the GNSS service. However, as experience has shown, the task of determining the origin of the signal may not be trivial.
- The interference can also be intentional. For instance, in a 2008 exercise, the General Lighthouse Authorities (GLHA) of the United Kingdom and Ireland performed a GPS jamming trial on the effect of GPS denial. For that they used a jamming unit provided by the UK Ministry of Defence (MoD), which at 1.5 W of radiated power was capable of jamming GPS over a 30-kilometer envelope [Grant09].
  - Another subclass of intentional jamming deserves additional attention. A low-power jammer [Ledvina10] would involve an adversary intentionally broadcasting hundreds of GNSS signals[23]. The adversary constructs the signals in such a way that the range calculations in the receiver do not yield a useable position or time. If one makes the assumption that the victim receiver cannot tell which signals are authentic, then the receiver would pick up to 12 signals from a set of signals that are mostly controlled by the adversary. This then implies that the receiver cannot calculate an authentic GNSS time and position fix. This vector of attack permits the transmission of signals at power levels beneath the thermal noise floor, which implies that the origin of the signal cannot be determined with the equipment to be used to trace the unintentional GNSS jamming (mentioned above). This vector of attack can readily be implemented using software-defined radio (SDR) techniques.

In literature, the jamming signal power is denoted as "J", whereas the thermal noise is denoted as "N", and a large J/N indicates that the ambient power reading is above normal, which (in the case of jamming at high power levels) means that an additional power source is present. To avoid saturation, many GNSS receivers compensate for stronger ambient power in the "Automated Gain Control" (AGC) stage of their RF front-end, and this has consequences. First, a larger AGC attenuation negatively impacts any other signal, including the GNSS signal. Second, receiver manufacturers could economically introduce J/N meters into their receivers, yielding a superior design with a commercial edge ([Ward07], [Langley08]). However, such features are almost never implemented.

- **GNSS meaconing** causes the receiver to track a signal that was recorded at a different location. Meaconing is the delayed re-broadcasting of recorded signals. For a GPS meaconing exercise, an adversary needs equipment that records the signals at the intended GPS band. The equipment then later re-radiates ("re-plays") the captured signals such that the GPS receiver processes them. The receiver then "naively" calculates position and time, which will then correspond to the recorded GPS signal, instead of corresponding to the live GPS signal.
- **GNSS spoofing** causes the receiver to track a signal that emanates from simulation equipment. For a spoofing exercise, an adversary needs a GNSS signal simulator, which can be a commercially available testbed for GNSS receiver validation, or could also be a custom hardware/software solution, as the one developed at Cornell University (Humphreys, Ledvina, et al) with the intent of studying vulnerabilities and developing countermeasures. The latter type of equipment, henceforth named the "GPS receiver-spoofer" is presented in reference [Humphreys08]. Conceptually it performs the following steps:

[23] For GPS, the receiver only knows 31 different C/A codes. However, the same C/A code could be transmitted multiple times with different Doppler shifts.

A. Acquires the GPS signal as laid out in "overview of GNSS" steps 5-10, thereby aligning its own C/A tracking loops with the actual GPS channels, reading out the navigation messages, and inferring any other data that a GPS receiver typically requires.
B. Feeds its own GPS receiver data, obtained in the previous step, into a GPS spoofer module, having GPS spoofing channels (one channel per spoofed satellite).
C. A spoofer control module is able to shift (delay or advance) the spoofing channels in time. Each time shift implies a range shift. Inside of the spoofer module, all of the range shifts are always calculated such that a corresponding position/time calculation is solvable, but would yield a false position or time.
D. By a process that closely mirrors the one of the actual GPS constellation (steps 3 to 4), it synthesises a replica of the L1 signal. In its 2008 set-up, the receiver-spoofer does not align this replica to the actual GNSS carrier signal that arrives at the victim receiver, leading to "phase trauma", explained further below. (Humphreys is experimenting with precise carrier phase alignment, which in turn would enable the device to radiate a suppression signal that cancels out the real GNSS signal at the victim receiver [T. Humphreys, personal communication].)
E. In a first phase, the receiver-spoofer broadcasts the actual GPS position and time, while slowly growing the false signal's power to marginally higher power levels (towards -150 dBW) than typical ambient GPS power levels. In a second phase, it introduces progressive time shifts in the spoofer channel, as mentioned in (C).
F. Let the "victim GPS receiver" be any GPS receiver that receives the GPS receiver-spoofer's signal in the first phase, and is subsequently affected by the time shifts. This receiver will then obtain the false position, as intended by the receiver-spoofer in (C).
G. In parallel with the false GPS signal, the real GPS signal may or may not be present. If it is present, then there will be two correlation peaks, but because current receivers were not built with this possibility in mind, one of the peaks is simply ignored after the receiver "locks (i.e. parameterises its Phase Lock Loop and Delay Lock Loop).

Note: As explored previously, high-power jamming will overwhelm the receivers at the RF front-end automated gain control (AGC), thereby attenuating the GNSS signals as a whole, ultimately denying GNSS service. This kind of crude jamming can readily be detected by a J/N meter, which measures RF power levels. But unless the attacker takes special precautions, low-power jamming, meaconing and spoofing could also be detected by the same J/N meter: Under normal conditions, all of the GPS signals together are stronger than the thermal noise floor, even though each individual satellite's signal is below the noise floor [Borre07, pg 65]. A fake GPS signal must be at least as strong as all of the real GPS signals combined, otherwise the receiver would probably lock onto the real signal instead. With at least twice the transmission energy being present, the J/N meter would indicate a corresponding increase of 3dB, which is a figure that is unusual under standard circumstances.

The remainder of this document will focus on meaconing and spoofing. These are of particular interest to the maritime domain, as they would enable a dishonest individual to disguise a vessel's true position. However civilian GNSS timing applications are also sensitive to GNSS spoofing, and this is an even bigger threat than maritime location spoofing. Timing applications are used to synchronise e.g. electrical power grids, GSM networks, and emergency services pagers.

Spoofing attacks

First, an adversary could make use of GPS spoofing hardware. In line with reference [Humphreys08], such hardware is classified as follows:

Spoofing hardware options	Characteristics / Costs
"GNSS Signal Generator"	- *Using a central unit and an emitter antenna to emulate a GNSS constellation, without the input of recorded GNSS data.* - *Can mislead ordinary GPS receivers.* - *No initial development effort, as equipment can be bought.* - *Hardware is available, but expensive (see subchapter "Possibilities in using GPS signal simulators").*
"GNSS Receiver Spoofer"	- *Using a GNSS receiver, a central unit and an emitter antenna to induce offsets into the signal of a real GNSS constellation. This type of device is described in reference [Humphreys08], which the authors of that report then built, but with the intent of studying countermeasures.* - *Such a device can spoof GNSS receivers that do not implement counter-measures.* - *The pre-requisite to build a GNSS receiver-spoofer is thorough knowledge of GNSS receiver technology.* - *Additional spoofer module requires at least two man-years of development effort [T. Humphreys and B. Ledvina, personal communication].* - *Deemed "most potent type of attack", where potency is the product of likelihood and damage potential [Humphreys08]* - *Present GPS receiver spoofer [Humphreys08] can align its signal with the GPS code phase; future work could enable carrier phase alignment with the possibility of cancelling the live GPS signal.* - *Present GPS receiver spoofer unit has material costs of about 1500 USD as of 2008 [Humphreys08].*
"Sophisticated GNSS Receiver Spoofer"	- *Multiple interlinked GNSS receiver spoofer nodes.* - *Targeted at multi-antenna GNSS receivers that have the same number of antennae, which employ the "angle of arrival" defence against spoofing, and where it is physically possible to feed each antenna a separate fake GNSS signal.* - *Requires significant up-front development effort, in addition to the "GNSS receiver spoofer": In order to overcome the "multi-antenna defence", the carrier phase differences between different antenna elements need to be controlled. One could achieve the above by using a common oscillator for all emitter circuits.* - *Each additional node increases the per-unit costs of this type of equipment, by roughly the amount of a single GNSS receiver-spoofer.*

Table 1: Spoofing hardware options

Both the "GNSS receiver-spoofer" and each "node" of a sophisticated GNSS spoofer follow the layout given below.

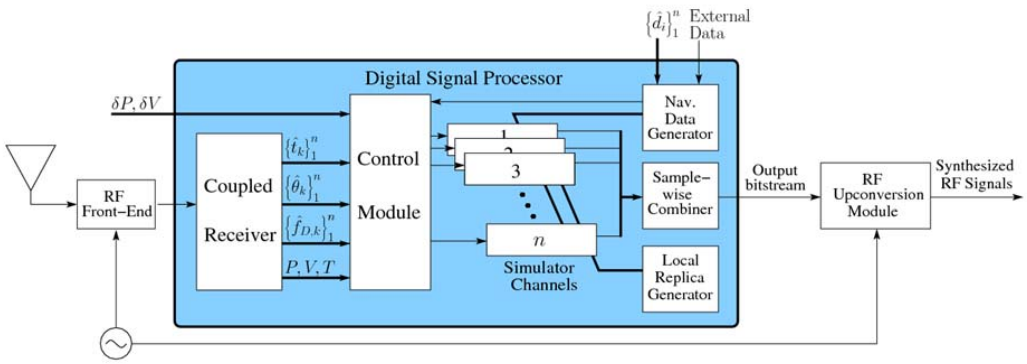


Figure 11: Layout of a GPS receiver-spoofer, courtesy of Humphreys and Ledvina.

Naïve spoofing attacks

In order to delimit the realm of possibilities of the sophisticated adversary, it is also beneficial to consider some poorly constructed attacks, which could be detected with a relatively low level of sophistication.

- **Large transmission area**: First, any adversary would do well to target his attack such that the signal is not visible outside the intended target area, using e.g. a directive antenna. Any signal that radiates far beyond the intended area will tip off other users that their GPS units are not working. This is because the radiated signal carries one, and only one, Position- Velocity-Time (PVT) solution, which will be picked up by any affected receiver. Thus, if an imprudent adversary calculates a PVT and then radiates it as a 1 watt signal from higher ground, then this signal would take control of the tracking loops of virtually all GPS receivers within line of sight, which will then report exactly this PVT to the user. Even casual users will notice this as a gross mistake in their GNSS equipment.
- **Use of pseudolites**: Second, for GNSS receiver testing purposes, one often installs GNSS pseudolites around a larger area of e.g. tens of kilometres. But it is impractical for an adversary to follow such a pattern, in a hope to upset GNSS receivers in above area. This is because signal strength always decays with the square of the distance between sender and receiver. However, the decay curve for pseudolites is much steeper than for real satellite signals, as the distance is that much closer. This is known as the pseudolite "near-far problem" [Wang02]. If a GNSS receiver monitors the relative signal strength of the various satellites, then an adversary can use a "pseudolite" configuration only for a limited target area, and many GNSS receivers outside that area would be impaired, e.g. when the signal power will trigger attenuation in the Automated Gain Control of the RF front-end.

This concludes the section about spoofing attacks.

Meaconing attacks

If instead of spoofing, the adversary is using meaconing, he needs to capture a real signal to begin with. He then has two possibilities for broadcasting the signal: Either he uses the signal immediately, or he first records, then replays the signal later. These two possibilities differ by the time period between capture and use.

Meaconing hardware options	Characteristics
"Replay Meaconing"	- *Recording a GNSS signal to a mass storage device, then re-broadcasting it.* - *Time between recording and re-broadcasting is subject to the physical transport of the mass-storage device.*
"Live Meaconing"	- *The theoretical recording a GNSS signal from another location, and transfer of signal by telecommunication, then near instantaneous re-broadcasting of signal.* - *Time between recording and re-broadcasting would be dominated by the data transfer delay. This delay could be less than a millisecond, when using point-to-point broadband communications.*

Table 2: Meaconing hardware options

Briefly, the theoretical "live meaconing" attack would be conducted as follows: At position A, the GNSS waveform would be downconverted and sampled. Any GNSS receiver capturing that waveform will believe it is located at position A. The samples would then be transmitted by wide-band radio link (such as the GigaBeam Gi-Linx point-to-point models, or a model from the Gi-Flex series [GIGABEAM]) to the accomplice performing illicit activities at position B. In theory, the transmission bandwidth is sufficient, since each single GPS L1 C/A signal transmits at 1 Mbps, maximum 12 signals are present, and the transmission bandwidth of above devices are rated at 54 Mbps or above. Once received at position B, the samples would be upconverted and the resulting waveform would be radiated, essentially reconstructing the GNSS signal of position A. The signal delay would be the sum of a fixed processing delay and a variable propagation delay. While the fixed processing delay requires further investigation to be determined, a 30 km distance always implies a propagation delay of 100 microseconds.

This concludes the section about meaconing attacks.

Other characteristics of a GNSS attack

Independently of whether the adversary uses spoofing or meaconing, he must broadcast the fake navigation signal either with or without the presence of the actual live GNSS signal.

Shielding option	Characteristics
"Open air broadcast" with "code phase alignment"	- *The real GNSS signal is not blocked. The fake GNSS signal co-exists with the real GNSS signal. Both signals have power levels within 3 dB of each other.* - *There is a "vestigial GNSS signal", which corresponds to the real GNSS signal.* - *This vector of attack is explored, for the purpose of research, by the GPS receiver-spoofer described in reference [Humphreys08]: The device aligns the GPS code signal with the GPS constellation, resulting in the victim receiver initially calculating the correct time and position. The GPS receiver-spoofer his however unable to do carrier phase alignment.* - *If the spoofer does not synchronise the carrier phase, then when the spoofing attack starts and stops, the GPS receiver experiences "phase trauma" effects, akin to ionospheric scintillations[24] (see further below for a brief introduction to "phase trauma"). Such scintillations are characterised by sudden variations in carrier phase and carrier amplitude. This can cause the receiver to lose carrier lock [Humphreys09].* - *This option would not incur additional per-unit costs beyond the GPS receiver-spoofer.*
"Shielded broadcast"	- *The real GNSS signal is blocked, and the fake GNSS signal is radiated in its place, at the correct power level, so there is no "vestigial GNSS signal".* - *This option requires the purchase of some radar absorbent material (additional per-unit costs see below), and some upfront engineering, but both are cheap when compared to the per-unit cost of a GNSS receiver spoofer.* - *This option is more feasible if the operator of the vehicle is an accomplice in the GNSS signal attack. This could well be the case for fishing and container vessels.* - *As the spoofing attack starts and stops, the victim receiver will experience some "phase trauma" (see further below for a brief introduction to "phase trauma"). This trauma would either resemble an ionospheric "deep fade" (if the spoofer is switched on, just after having blocked the real signal), or a sudden variation in carrier phase (if the victim receiver, for a brief time, sees both the GNSS constellation and the spoofed signal).*

[24] These two RF phenomena are so similar that they imply shared hardware for their respective study: the Cornell "GRID receiver", which led to the GPS receiver-spoofer, was originally developed to investigate ionospheric scintillation phenomena. The GRID receiver is still used for such studies.

"Carrier phase alignment" and "Suppressed vestigial signal"	- *The real GNSS signal is suppressed by the GNSS receiver-spoofer, so there is no "vestigial GNSS signal"* - *This requires that the GNSS receiver-spoofer must emit a "cancellation signal", which destructively interferes with the real GNSS signal.* - *This requires carrier phase tracking and precise positioning vectors down to "1/8 of a cycle" [Humphreys08], which for the L1 frequency translates into knowing the vectors between all three antennae with a precision of +/- 2 cm.* - *Because the carrier phase of the spoofer is aligned with the carrier phase of the real signal, there is no "phase trauma" phenomenon that a victim receiver could detect (see further below for a brief introduction to "phase trauma").* - *This option also requires significant up-front development costs, estimated at one man-year [T. Humphreys and B. Ledvina, personal communication].* - *This option would not incur additional per-unit material costs.*

Table 3: Adversary's shielding option

As for "shielded broadcast" material costs, 250-300 EUR should be sufficient to cover about 10 m^2 of surface with radar absorbent tiles, so as to absorb radiation at 1-1.5 GHz with sufficient attenuation [J. Fortuny-Guasch, personal communication].

This concludes the section about shielding options.

Furthermore, independently of any of the above, the adversary has the choice between single-frequency and multi-frequency operations.

Number of frequencies	Characteristics
"Single frequency"	- *The adversary emits one or more fake GNSS signals, but all in a single frequency band.* - *This vector of attack is explored, for the purpose of research, by the GPS receiver-spoofer described in reference [Humphreys08]: The device operates a single frequency, L1, in which it broadcasts only the C/A signal.*
"Multi frequency"	- *The adversary captures and emits fake GNSS signals in several frequency bands.* - *Concerning the cost of spoofing, for each additional frequency, the adversary could use a separate GNSS receiver-spoofer, implying that spoofing costs increase with the number of frequencies. Alternatively, the adversary could turn a single frequency platform into a multi-frequency receiver-spoofer, with additional development cost for a platform hosting several receiver/emitter boards and related real-time requirements. The required processing power could come from a GPGPU. Alternatively the adversary needs to wait several years for hardware processing power to increase to the point of being able to handle the processing load.* - *Concerning the cost of meaconing, multi-frequency devices are more expensive than their single-frequency counterparts, and this is quantified in the chapter "Possibilities in using GPS signal simulators" below.*

Table 4: Adversary's choice for single or multi-frequency operation

An adversary can also choose to create not just one fake signal, but signals for several satellite constellations:

Number of signals	Characteristics
"Single signal"	- *The adversary emits a fake GNSS signal from a single satellite navigation system. Other navigation system signals would be unaffected.* - *This always implies that the adversary chooses the "single frequency" attack mode as well.* - *This type of attack can be explored for research purposes by the GPS receiver-spoofer described in reference [Humphreys08], which is able to spoof the GPS L1 C/A signal only:*
"Multi signal"	- *The adversary would emit ore than one GNSS signal. This may imply the use of "multiple frequencies" (e.g. all of the Galileo Open Service on bands L1, E5a, and E5b) or only of "single frequency" (e.g. both of GPS L1 C/A and Galileo Open Service on L1).* - *It can be asserted that the spoofer's computing hardware requirements are a linear function of the GNSS code chip rate [T. Humphreys and B. Ledvina, personal communication]. As a corollary, if a signal (GPS L2C or Galileo on E5b) has a code chip of 10230 Mcps as opposed to the 1023 Mcps of the GPS L1 signal, then that in turn requires a platform with roughly 10 times the processing power. At those levels, an adversary needs different hardware such as a FPGA, GPGPU or PC platform, and/or more efficient software, and/or waiting for hardware to become cheaper*[25]. - *If the adversary decides to use one spoofing device per frequency band, then the cost of spoofing also increases with the number of GNSS systems emulated.* - *Alternatively, if the adversary turns a single signal platform into a multi-signal receiver-spoofer, then that would again require additional processing power and development costs for demanding real-time requirements.* - *Concerning the cost of meaconing, multi-signal devices are more expensive than their single-signal counterparts, and this is quantified in the chapter "Possibilities in using GPS signal simulators" below.*

Table 5: Adversary's choice for single or multi signal operation[26]

Note: The "Multi-Signal" and "multi-frequency" categories are not independent from each other. For any attack, the number of frequencies is less or equal than the number of signals. It may however also be useful to differentiate between "multi-frequency" and "multi-signal", because not all "multi-signal" attacks are created equal: A spoofer that emulates the GPS L1 and L2C civilian signals (multi-frequency) does not have the same requirements as a spoofer that emulates GPS and Galileo, but only on L1 (single-frequency).

If an adversary chooses to emulate signals for the GPS L1 C/A and the GPS L2C signal, he would need two GNSS receiver-spoofer devices.

- The first device would be the GPS receiver-spoofer "as-is" in order to emulate the GPS L1 C/A; the device would cost 1500 USD in materials.
- The second model would target the GPS L2C signal, and could be hosted e.g. on a "desktop replacement" laptop running a real-time Linux Operating System [Ledvina, personal communication]. Costs should be in the vicinity of 1000 USD for the laptop, and 1000 USD for a USRP device fitted with RX and TX daughterboards [EttusUSRP1], or an equivalent device from another vendor.

[25] Moore's Law stipulates i.a. that the computing performance, at the same cost, doubles every 2 years. So a 10-fold increase in processing power is roughly 6 ½ years away.

[26] In the table, the GNSS systems are lumped together with their respective Space Based Augmentation Systems. For instance, GPS is augmented by WAAS and EGNOS.

The total material unitary cost of spoofing both the GPS L1 C/A and the GPS L2C signal can therefore tentatively be situated at 3500 USD.

This concludes the section on single signal versus multiple signal operation.

Any attack vector must at one point choose how to operate. The options that the adversary has at his disposal is therefore a four-dimensional array given by

```
adversary options := (spoofing XOR meaconing)
                        AND shielding option
                        AND number of frequencies
                        AND number of signals
```

Generally, if none of these options is economically sensible, then one would expect that the adversaries resort to other means, such as social engineering (e.g. bribes), or modifying the trackers at the manufacturer plant, or challenging any evidence of an attack in court. However, a defender is still encouraged to setting up technological and logistical means, in order to make an attack on the GNSS tracker itself uneconomical.

With the above in mind, counter-measures will be discussed, which can be implemented on a GNSS receiver module to detect spoofing.

Options for defending a GNSS receiver

There are several possibilities for GNSS receiver spoof detection. These have been raised as early as 2001 by the Volpe institute [Volpe01]:

- Some measures, applicable to standard GPS receivers without (much) additional hardware, such as measuring the ambient power level (J/N) and the carrier to noise density ratio (C/N_0).
- Some other “classical” measures would mean larger investments and hardware changes in GNSS receivers, such as adding a second antenna, comparing with inertial sensors, or reading out LORAN signals.
- Various new techniques, based on cryptography, could be rolled out in order to secure GPS and for Galileo.

In addition, during their JRC visit in December 2009, Humphreys and Ledvina [personal communication] pointed out several defences (“data latency bit”, “vestigial signal”) that are related to issues that any GPS receiver-spoofer must overcome in its design.

The rest of this chapter examines each of these defences in turn, and how they impair various attacks.

First, as an introduction, and in order to hint at the complexity of the subject, two ideas are examined below that are somewhat flawed.

Naïve design	Idea *Flaw*
Requiring that VMS boxes use multiple GNSS systems (GPS and Galileo), but all centred at the same frequency (L1)	- The adversary would need to spoof multiple GNSS systems, all at once, but in the same frequency range. - The adversary would need to emulate different encoding techniques, such as BPSK used in GPS, and BOC(2,2) used in Galileo. - Single frequency multi-signal GNSS receivers (for e.g. GPS, EGNOS and Galileo) will soon enter the mass market, and will be a commodity item. - *If a DSP is used, then by Moore's Law, in 2-3 years, signal processors will become powerful enough to run a multi-signals single frequency GNSS receiver-spoofer on a single hardware platform. At this point, a successful attack only requires software modifications to account for e.g. Galileo Open Service on L1.* - *Alternatively an adversary could implement a GNSS spoofer on a General Purpose Graphical Processing Unit (GPGPU).* - *A skilled adversary may not encounter a significantly higher workload when emulating more than one modulation technique, such as BOC and BPSK. The adversary would focus his development on synthesising a believable composite signal, in essence building on existing synthesising algorithms, such as the one presented in reference [Humphreys08].*
Requiring that VMS boxes be mounted on long poles so as to be difficult to access	- In order to block or interfere with the antenna, the adversary would need to install a device on top of the antenna, in what is fundamentally a dynamic environment. - *Long poles require tethering cables for resisting engine vibrations and storms on high seas.* - *For safety reasons, any part of a boat needs to be accessible at all times.*

Table 6: Some ideas that probably will not (durably) deter spoofing attacks.

At present other ideas will be explored.

Classical options for defending a GNSS receiver

A single-band single antenna GNSS receiver can be enhanced by a set of checks on the GNSS signal. These are summarised in reference [Scott03]:

- "Use J/N meter to check for above normal energy levels" and "Monitor C/No meter for consistency / unexpected C/No given J/N".
    - A J/N meter measures RF power levels, at the automated gain control stage. It is relatively easy to include such a meter in the design of a GNSS receiver ([Langley08], [Ward07]).
    - Alternatively to implementing a J/N meter, one could monitor the energy readings obtained via the RF front-end's Automatic Gain Control (AGC) [D. Akos, personal communication].
    - A C/No meter measures the carrier to noise density, after the automated gain control stage, in this case of GNSS satellite signals. Many commercial GNSS receivers output C/No figures.

  - The relatively simple checks above effectively preclude an adversary from using an "open air broadcast", since the presence of two signals with equal energy effectively increases the RF power by 3 dB, which can readily be detected.
  - An adversary could defeat such measures by a "shielded broadcast" or by "vestigial signal suppression".
- "Deep acquisition to look for weak, real signals".
  - Some form of filtering is used to excise the dominant GNSS signal from the RF samples. The filtered samples are then fed into a separate set of GNSS correlators. "Deep acquisition" can be implemented by the "vestigial signal defence", which is explored further below.
  - An adversary could defeat such deep acquisition by a "shielded broadcast" or by "vestigial signal suppression", in case of which no significant signal vestiges would be present.
- "Compare 'Watch Time' with 'Signals Time' (Most signal generators can't synchronize with GPS time)", and "Continuity checks in time and position":
  - This measure prevents an attack which implies a "gross" jump in the navigation solution's time, where "gross" means that it can be detected by being unrealistic with respect to a good crystal oscillator. TCXOs can be as precise as +/- 1 ppm under real-life conditions involving accelerations, vibration, and temperature differences [Maxim08]. Aging can be compensated by "pulling".
  - This measure will prevent the adversary from using "meaconing" and "GNSS signal generators", in their currently available form, as most of them only foresee "canned scenarios" [Scott03].
  - Unfortunately, this measure does not detect the "GNSS receiver-spoofer" or the "Sophisticated GNSS receiver-spoofer", both described in reference [Humphreys08].
  - However, when this measure is combined with signal authentication, then spoofing becomes quite challenging. This combination will be explored when discussing Navigation Message Authentication (NMA).
- "Large residuals, particularly in differential correction channels", and "Receiver Autonomous Integrity Monitoring (RAIM), and Fault Detection & Exclusion (FDE) type functions".
  - Differential GNSS (such as DGPS) is supposed to augment the accuracy of a GNSS constellation. For instance, Space Based Augmentation Signals (SBAS), such as WAAS and EGNOS, fall into this category.
  - One has to consider that spoofing a GNSS constellation is more difficult than spoofing the corresponding SBAS signal. One also has to consider that, if either the GNSS or the SBAS are spoofed, the receiver will notice that the SBAS range has a large residual. But without aiding, the receiver cannot know which system (SBAS or GNSS) is the source of the problem. For instance, at the edge of the SBAS coverage, the SBAS ranging code may naturally produce large residuals. Therefore the receiver would do best to rely on RAIM and FDE to exclude the range with the largest error.
  - The correct use of RAIM and FDE in a receiver would imply, for the adversary, that he needs to simulate the GNSS rather than the SBAS signal. This is detrimental to the attacker, since it is more difficult to simulate a GNSS rather than an SBAS signal.
  - RAIM and FDE type measures do not detect the "GNSS receiver spoofer" or "Sophisticated GNSS receiver spoofer" [Humphreys08], neither do they detect a "meaconing" type attack.

Beyond the measures that can be used for single-antenna single-frequency GNSS receivers, there are other measures discussed in reference [Scott03], which require additional hardware:

- "Consistency with other navigational sensors".
  - The use and implications of Inertial Navigation Sensors are described further below.

- "Monitor phase difference between multiple antenna elements (All signals shouldn't come from the same direction)".
    - A separate paper describes replicating earlier results in reference [Montgomery09], but using open source software and commodity hardware. This paper is in the annexes of this document.

Beyond the measures cited above, it is also possible for the GNSS provider to add cryptographic protections. These somewhat more durable measures are explored further below.

Multi-antenna defence

In order to check that all signals are not coming from the same direction, one option is to "monitor phase difference between multiple antenna elements".

It is possible to set up a dual-antenna receiver as shown in reference [Montgomery09]. For studying spoof defences, the design of ref. [Montgomery09] was replicated, by fitting a single "Universal Software Radio Peripheral" [EttusUSRP1] with two DBSRX daughterboards [EttusUSRP2]. Two commodity antennae complete the receiver hardware. Both antennae sample the L1 frequency, and are driven by a common clock.

Since this GPS receiver uses software defined radio techniques, GnuRadio [GR1] and GPS-SDR [Heckler-GPSSDR] were installed on a Linux PC platform. GPS-SDR was modified to cause it to send certain observables to a separate programme. That separate programme, labelled "gps-cpm" for Carrier Phase Measurement, performs epoch-antenna differencing[27], as explained in reference [Montgomery09].

The resulting multi-antenna defence system can be summarised as per the drawing below:

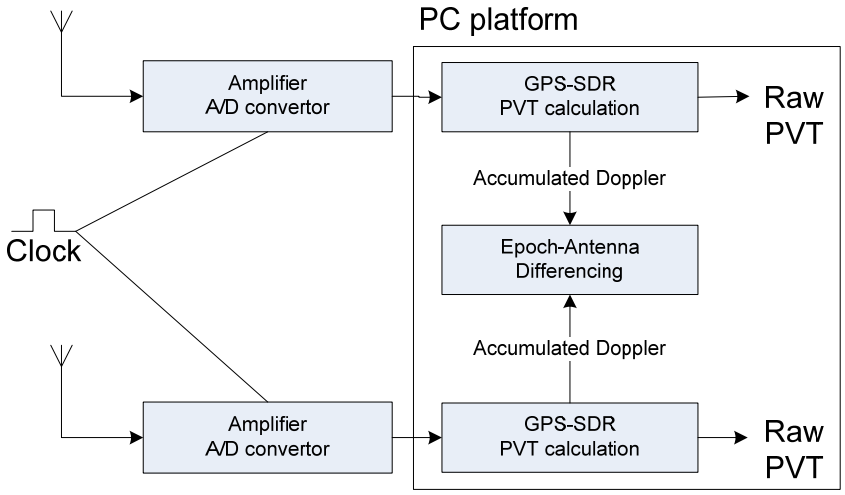


Figure 12: Dual antenna defence with common clock. The raw PVT of one of the outputs is put through an Extended Kalman Filter and presented as GPS position to the user

For each antenna, baseband samples feed into a GPS process, and each process forwards the accumulated Doppler drift. The Doppler drift is then used to determine the carrier phase, or its change with time, which is different for each satellite. This is a process not unlike stereophonic hearing.

[27] which is a form of single differencing

For the exact method, further algorithmic details, and results of the calculations, please consult the appendix.

Due to the L1 wavelength of 19 cm, the antenna elements of the defence need to be spaced apart by at least 10 centimetres. Longer distances improve time to alarm, but make for a larger ("bulkier") volume to protect. Note that any distance under 50 cm usually implies "mutual coupling" between the two antenna elements, however this effect can be accounted for, and hence be mitigated, by calibration [Backen07 chapter 5].

This defence benefits from additional physical shielding. In reference [Humphreys08], the authors state that an Angle of Arrival type defence can be thwarted by an attack of an equivalent number of GNSS receiver-spoofers, which need to be phase-locked. Each receiver-spoofer module would compute and send the signal that each defending antenna would expect to perceive, in order to create an illusion of carrier phase change. But that sort of attack is possible only if the adversary can send individual signals to the defending antenna elements. If one constructs a physical volume seal around a volume of a cube of about 30x30x30 cm, and places the two antenna elements well within that volume, then the adversary cannot send separate signals to each antenna element. The adversary is then forced to deploy a GNSS receiver-spoofer where each antenna element sends an individual GNSS satellite signal.

In other words, the adversary is faced with the challenge of building a "Sophisticated GNSS receiver spoofer" in order to overcome this defence, as any single-source GNSS signal simulator or meaconing device will be detected. Overcoming a dual antenna receiver defence with physical shielding is quite difficult, to the point that the literature occasionally calls it "unassailable" ([Scott03]). This may well imply that the attack becomes un-economical to the adversary to build such a platform, at least for most applications.

The same defence can also be put together using two separate GPS processes, with separate clocks. Such a design may be less expensive to implement, if one can obtain two commercial GPS OEM receiver boards that include the accumulated Doppler in their outputs. Humphreys and Ledvina [personal communication] mentioned that, when one uses the accumulated Doppler outputs, and in order to cancel out any errors due to differences between the two clocks, one must replace the "epoch-antenna differencing" by an "epoch-antenna-clock" differencing. Such algorithms are e.g. laid out in reference [Rizos]. It helps to observe that the naturally occurring quartz clock drift is disciplined by the GPS time in any GPS receiver.

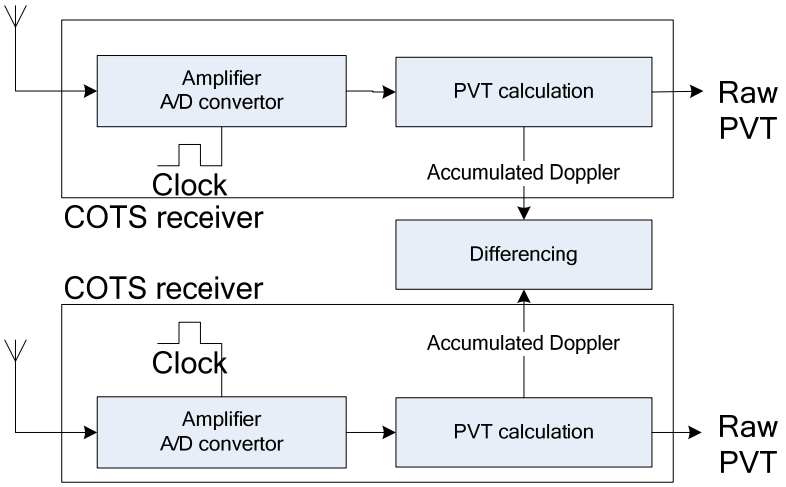


Figure 13: Dual antenna defence with separate clocks and commercial GPS OEM receiver components. The raw PVT of one of the outputs is put through an Extended Kalman Filter and presented as GPS position to the user

The defence can also be used in moving environments. For a full discussion on how to use the defence on a moving platform, refer to the paper in the annex.

In terms of economics, OEM kits from UBlox (LEA-5T) or NovATel (OEMV family) are available for about 200 EUR each, of which one would need two. The accumulated Doppler output would need to be processed on a main board, which would perform the carrier phase differencing operations and would also send an alarm in case of detection.

In summary, a dual-antenna defence would be available today and would provide adequate world-wide protection against GPS receiver-spoofers and meaconing devices. It may cost about 500 EUR per unit.

Multi-frequency defence

Multi-frequency GNSS receivers are available since the advent of dual frequency GPS L1/L2 receivers, which are commonly used in geodesy and ionospheric studies. One could use a dual-frequency receiver, or two single-frequency receivers that can each operate on one of the frequencies, and cross-check the navigation solutions from both signals. Ideally one of the signals has a chip rate much higher than 1023 Mcps[28]. Commercial dual-frequency receivers are available, but mainly because of their lower production volume, cost several hundred Euros per piece as OEM products.

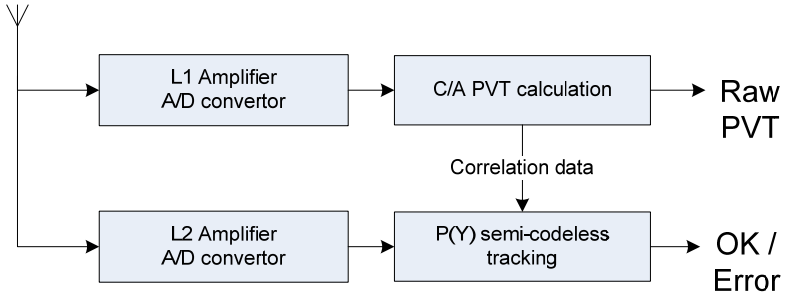


Figure 14: Dual Frequency GPS receiver defence, inspired by the cross-correlation method, laid out in reference [Woo00]

The GNSS signals that operate at 10230 Mcps already provides for some defence, because the adversary is forced to use a more powerful platform to synthesise the signal. Dual-frequency receivers routinely cross-check the GPS L1 signal against the GPS L2C signal, and infer ionospheric properties in the process. Aberrant values or fluctuations in the ionospheric terms could hint at the presence of a spoofer that emits GPS L1 and L2C signals that are not synchronised. If an adversary seeks to overcome this defence, then he needs to build a "multi-frequency" and "multi-signal" spoofer that

[28] While civilian GNSS signals in the L1 band currently have chip rates of 1023 Mcps, other civilian signals have higher chip rates. The GPS L2C signal on L2 at 1227.60 MHz and the Galileo Open Service at E5a (1176.45 MHz) and E5b (1207.14 MHz) all have a 10230 Mcps chip rate [GPS-SIG-WP] [Gal02].

emulates the properties of the ionosphere in a believable fashion. In turn that raises the bar for the adversary.

Therefore, a multi-frequency multi-GNSS receiver defence could be implemented today, and would work world-wide. It would provide "some" protection against an adversary, where "some" is a function of the adversary's sophistication. It would cost a few hundred EUR per unit.

Using INS sensors to defend a GNSS receiver

At present, the addition of an INS sensor to a single-frequency, single-antenna GNSS receiver is being considered. Consistency checks of GNSS versus with Inertial Navigation Sensors (INS[29]) are indeed an option. Such verification techniques can use both accelerations as well as calculated positions, and can rely on INS hardware proposed by commercial companies. Integrated designs also exist in industry ([Xsens], [Septentrio]). Software that compares both GNSS and INS measurements however must allow for certain error tolerance margins, both concerning position and accelerations. (Concerning the former, such systems suffer from "integration drift" when they calculate positions). A spoofer would need to be aware of those tolerance margins in order not to raise suspicion. Therefore a GNSS Receiver Autonomous Integrity Monitoring system, using accelerometers, gyroscopes, and magnetic sensors in a coordinated fashion could be considered a possible additional option for a GNSS receiver.

Applied to the VMS, a simple measure would compare the "yaw" (rotation rate along the Z axis) obtained from the INS, with the heading changes calculated from the GNSS. Another simple check is to compare the absolute heading obtained by the magnetic sensor (which is a digital compass) with the heading calculated via GNSS. (Arguably, for higher latitudes, one would need to compensate for the angle between the magnetic pole and the geographical pole. Also, the earth's magnetic pole changes unpredictably, requiring system updates.) The spoofer could still use magnetic coils in order to mislead the INS' magnetic sensor, but this requires mounting and appropriate software.

An unscrupulous fisherman could still attain forbidden areas, despite INS cross-checks, at least in theory. Near the beginning of a trip, when he is beyond coastal line of sight and cruising to the fishing grounds, the GNSS receiver-spoofer manages to introduce gradual changes between the real position and the spoofed position. These imply accelerations, which must be beneath the detection threshold of the INS. The undetected accelerations are intended to cause cumulative effects later in the journey. The GNSS-INS cross-checks will not identify such a measure.

But the above fraud scenario requires careful planning and execution, as the adversary cannot disable the GNSS receiver spoofer during the illegal trip, and because the distortions can only be introduced at a limited rate. In summary, the argument is made that such cross-checks strongly diminish the temptation of opportunism: a skipper cannot decide to "make a quick extra buck" at an arbitrary point in his trip.

The INS defence has an added benefit: during a GNSS outage, a GNSS/INS combined receiver can perform "dead reckoning", which enables it to keep tracking the location. During a GNSS outage, position accuracy would then be a function of both the INS equipment's quality and the time since the last GNSS lock. On fishing vessels, GNSS outages are more frequent than one may think. One factor that causes them is that a ship has various radio devices that can interfere with GNSS. This then leads to aberrant VMS position reports, which are unintentional, and are frequently observed by FMCs for fishing vessels. Cutting down the number of such aberrant position reports is a challenge [personal

[29] Source: Wikipedia http://en.wikipedia.org/wiki/Inertial_navigation_system

communication with Scottish and Dutch FMC members]. Once that is achieved, it would become much easier to spot malicious interference.

When implementing a GNSS/INS defence, in terms of economics, there is a trade-off between inertial sensor accuracy and price, which impacts detection rates. In 2009, a higher quality GPS-INS anti-spoof defence comes at considerable unit costs. However high quality "solid state inertial guidance systems are about 5 years away" [R. Johnston, personal communication].

To quantify present prices, in May 2009, an XSens INS solution, as an OEM evaluation kit, costs 2100 EUR per piece, including the embedded software that compares input from various sensors. The DMU02 from Silicon Sensing is available for about 450 USD per unit. Judging from the price of various consumer electronics products, lower quality devices based vibrating gyros would cost perhaps 20 EUR per unit.

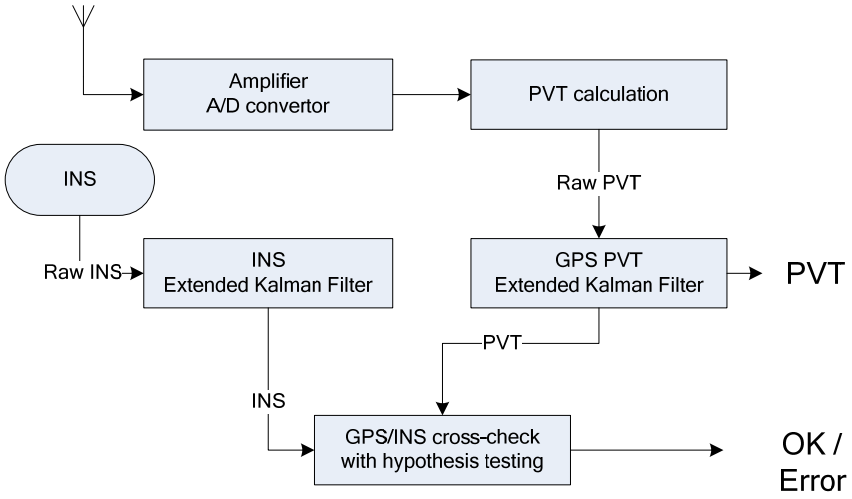


Figure 15: GPS/INS defence. Both the raw PVT and the raw INS data are filtered before cross-checking takes place. Hypothesis testing diminishes false alarm rates, at the cost of increasing time to alarm.

In summary, a GPS/INS solution is an adequate and world-wide defence against GNSS location spoofing. Additional costs depend on the trade-off between price and quality of the INS equipment. Looking forward, the price/performance ratio for quality INS equipment is projected to improve drastically over the next five years [Johnston, personal communication].

One also observes that an INS device can improve the performance of a GNSS tracker on maritime vessels, where GNSS outages can be frequent, which some FMCs [personal communication] believe to be a leading cause of aberrant position reports. VMS trackers with an integrated INS would therefore have a double benefit for FMC operators: they reduce the possibility of spoofing, and mitigate the effect of GNSS outages on VMS position reporting.

This defence should be considered if use of inertial navigation sensors is warranted by the application domain. That would apply for fisheries monitoring and the tracking of dangerous goods.

LORAN

Next, consider the LOng Range Aid to Navigation (LORAN) in its European version, EuroFix. The UK General Lighthouse Authorities (GLHA) are implementing eLORAN. By contrast, the US

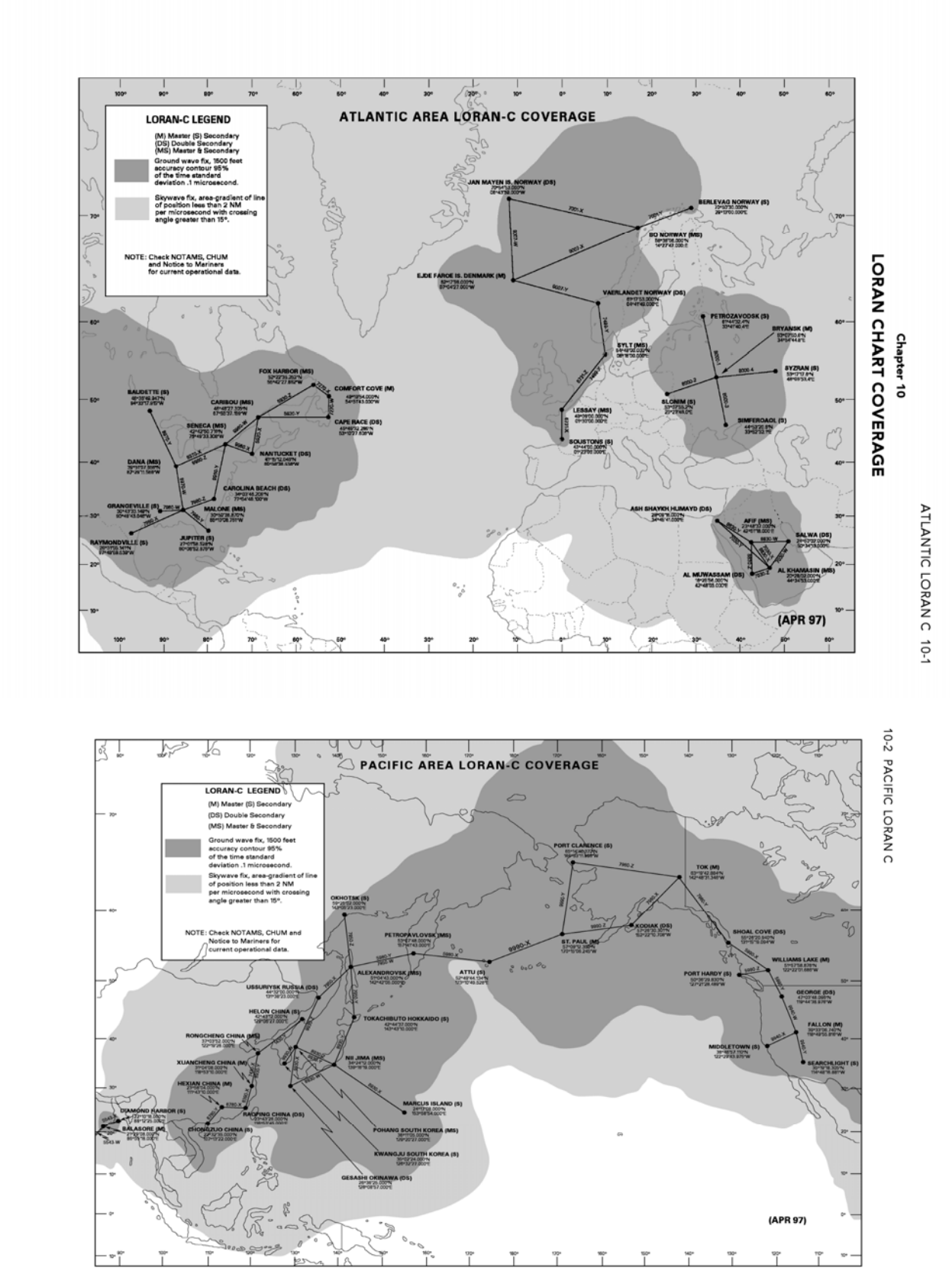
administration has decided to switch off LORAN in 2010 [Loran10] [USCG10], so statements made below apply to areas outside of the US and Canada, as made evident by the coverage chart below.

LORAN covers many of the vulnerabilities of GPS [Jacobson07], and is considered hard to jam or spoof [Lo09]. In order to detect LORAN spoofing, there are counter-measures that one can implement in a LORAN receiver [Lo09].

Improvements in eLORAN mean that within the area of three LORAN stations, position can be determined with an accuracy that rivals unaided civilian GPS [Lo07]. In addition, eLORAN may include an authentication message that further increases the difficulty faced by a spoofer [Lo09].

If (e)LORAN receivers would be built with above upgrades in mind and could be used in maritime, then an adversary would need to start a resource-intensive research project to overcome such an impediment. This would be unattractive, meaning that the adversary would instead pursue other avenues. Implementing a LORAN receiver does not require a lot of hardware and is technically feasible using software defined radio (SDR), as evident by recent discussion on GNURadio [GR1] bulletin boards. A GNSS-eLORAN combined receiver would therefore be a cost-effective hardening measure, if it were not for the fact that the coverage of LORAN is not global.

However the major issue with LORAN is that the coverage never was global, even before the American and Canadian decisions to switch off their stations. LORAN never covered the southern hemisphere. Therefore, for a global maritime application, a LORAN-based defence may require complementary measures.



Chapter 10
LORAN CHART COVERAGE
ATLANTIC LORAN C 10-1

10-2 PACIFIC LORAN C

Figure 16: Coverage of LORAN-C chains as in 2006. DoD General Planning, 6 July 2006, NGA ref PLANXGP, page 10-1. Public domain.

The above images depict availability of a LORAN groundwave fixe (95% accuracy 500 metres) in the darker tones, and availability of a less accurate LORAN skywave fix in the lighter tones.

Furthermore, if at least 3 LORAN stations are available to the receiver, then it is possible to cross-check a LORAN-determined position against a GNSS-determined position. If two LORAN stations are available, then the two signals from the stations intersect in two points, one of which must correspond to the GNSS position. If only one LORAN signal is available, then the GNSS position should be situated on a circular band around the LORAN station. If these checks do not compute, then the receiver knows beyond reasonable doubt that GNSS spoofing occurred, and can take appropriate measures.

Within the European Union, Council Decision 92/143/EEC mandates the implementation of a LORAN system. However, in 2000 "Loran-C/Eurofix policy clarifications at European level" were deemed missing by one member state [Jorgensen00]. This then led to subsequent implementation delays.

A defence based on an integrated LORAN receiver can be schematised as follows:

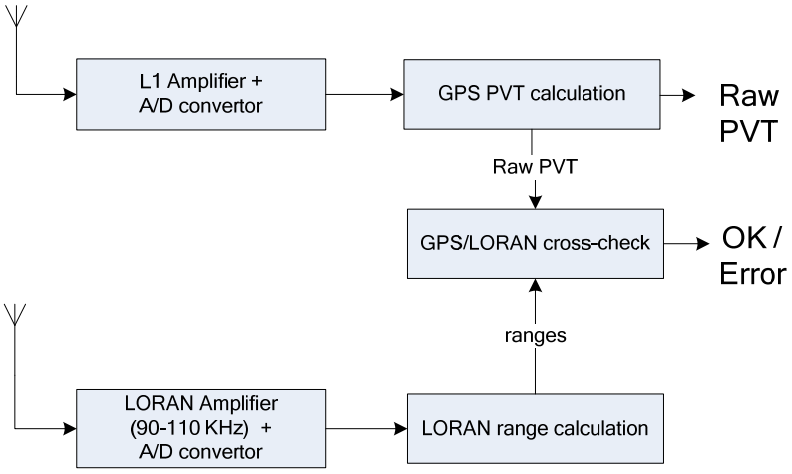


Figure 17: GPS/LORAN defence. The LORAN range calculation can be augmented by techniques described in reference [Lo09]. The resulting ranges are fed into a cross-checking unit.

In conclusion, a GNSS-LORAN receiver would be a sensible hardening measure, if the application area is limited to LORAN coverage, e.g. in European maritime and land-based applications. Unitary costs for a LORAN receiver front-end are low. LORAN is well explored, and detection rates should be easy to quantify following the work in references [Lo07] and [Lo09]. For maritime monitoring, further coverage of the middle and southern Atlantic, and the Indian Ocean may be of interest. Further LORAN coverage and eLORAN modernisation would require policy making at an international level, such as convincing the US and Canadian governments to reverse their earlier decisions to dismantle their stations. It would also require EU policy decisions and implementation by Member States.

Data bit latency defence

A GNSS receiver-spoofer would first read out the data messages of individual GPS satellites. For that it needs to accumulate at least 1 ms of signal samples (or a multiple thereof), in order to raise the GNSS signal above the thermal noise floor. Once it is confident of the data message bit received, it then replays the same data messages on its spoofer channels, altering the transmission times as desired.

Each GPS satellite transmits a navigation message that repeats itself every 12.5 minutes. As stated e.g. in reference [Humphreys08], for a GPS receiver-spoofer, it is difficult to retransmit the correct navigation message in exact real time. First, the ephemeris component is subject to change every 2 hours, which affects an adversary who desires to spoof a GPS receiver for a longer time period. Second, satellites that are rising over the horizon also pose a challenge to a GPS receiver-spoofer: the spoofer has to know, in real time, what information the satellite is transmitting.

The navigation message's data bits change according to several patterns:

1. The counters and clocks (age of ephemeris, time since last Sunday, UTC time) all change in successive messages, but in a predictable fashion.
2. *The ephemeris is updated every 2 hours* [GPS-WP].
3. *12.5 minutes are required to receive the entire almanac from a single satellite* [GPS-WP] [Weston99].

A GNSS receiver-spoofer essentially has the following options for dealing with the unpredictable parts of the data message:

1. **Delay**, meaning to "data bit latency": *The simplest approach is to pass data bits to the spoofer channels as soon as they can be reliably read on the incoming GPS signals. Naturally, this approach results in a delay in the arrival time of the spoofing data bit as compared to that of the true data bit at the target receiver's antenna. The delay is most conveniently made an integer ("n") number of 1-ms C/A code intervals. Clearly, such a delay is undesirable in a spoofer because a target receiver could be designed to watch for such a delay and thereby detect a spoofing attack.* [Humphreys08]
2. **Improvise**, meaning to perform "data bit prediction" and/or "data bit fabrication": Nearly all data bits can be predicted. For those which cannot be predicted (and the spoofer knows which), impose a partially fake data message: *fill the unpredictable data bit segments with arbitrary data bits and adapt the* [rest of the message including the checksums] *accordingly.* [Humphreys08]
3. **Determine**
    a. Use external aiding to obtain the values bits, perhaps by fetching the data from a network of accomplices, or downloading it from the NASA's web site, but the latter imposes latencies for the actual signals;
    b. Read out the message directly, using directional antennae or beamforming techniques, but this implies further engineering and the use of comparatively expensive hardware;
    c. More speculatively, obtain an algorithm that predicts the bits with high accuracy, but such algorithms are currently unknown.

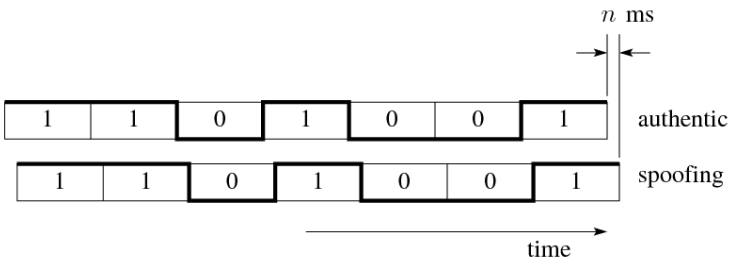


Figure 18: Data latency illustration, as implied by option 1: Difference between spoofed (delayed) data bit boundaries versus the authentic signal, courtesy of T. Humphreys.

If option 1 or 2 is used, then this has implications for the GPS receiver-spoofer design. On its spoofing channels, the GPS receiver-spoofer would send delayed (option 1) or partially fabricated (option 2) data for each spoofed GPS satellite, including the ephemeris. But when the adversary is spoofing a position on open waters, then the spoofing attack would need to cease at a given point in time,

otherwise inspectors could detect the setup. At that time, in order to avoid the "vestigial signal defence" (explored further below), the position of the vehicle must be consistent with the actual range of the GPS satellite's signal. This requires that the spoofed and real position need to "*rendez-vous*" on a given curve at a given time during an illegal trip.

Once the adversary switches off the spoofer equipment, two conditions can arise:

- If the GNSS receiver-spoofer previously sent a data bit with a delay (i.e. he chose option 1), then the "*n* ms" delay will be cancelled, meaning that the transmitted data bit will seem to be shortened by the same *n* ms. To defend, the GPS receiver needs to *continuously monitor bit lock* [Humphreys08].
- If the spoofer previously fabricated one or more data bits (i.e. he chose option 2), then to defend, the GPS receiver could notice that the ephemeris changed, with a frequency different from 2 hours. It is therefore *vulnerable to detection at the 2-hour ephemeris update boundaries* [Humphreys08].

Possible implementations of the "data bit latency" defence could:

1. Check that the navigation message bits do only change at the 20 millisecond boundary ("monitor the bitlock")
2. Check the navigation message checksums for inconsistencies.
3. Determine the 12.5 minute and 2 hour boundaries at which each individual satellite can change its navigation messages.
4. Check that the ephemeris is updated only at 2 hour intervals.
5. Inside every 12.5 minute window, check that the changes in the navigation message, as transmitted every 30 seconds, are limited to satellite clock and almanac changes. Check that the satellite clock drift is realistic. Check that the almanac is reported in 25 different parts, as consistent with GPS normal operations, and that the changes in the almanac only occur in 12.5 minute intervals.

Note: It is possible that this method does not ensure 100% detection. Indeed, there may be time windows and matching coordinates, where a spoofer can be switched off, but without provoking any of the two conditions above. Additionally, a "shielded broadcast" or "carrier phase aligned" GPS receiver-spoofer could work around this defence by introducing noise into GNSS channels that mimics atmospheric disturbances.

The defence can be made more potent by coupling it with a high-precision, battery-powered TXCO clock signal. This makes it possible for the receiver to determine the 2 hour and 12.5 minute GPS navigation message windows well in advance.

This defence requires further research to quantify the false positive and misdetection rates. To be effective for spoof detection, this method, and other single-frequency single-antenna tests, must quantifiably distinguishing between an attack by a GPS receiver-spoofer, and naturally occurring ionospheric scintillations (see "phase trauma" below). It is not entirely clear how single-frequency single-antenna tests can by themselves distinguish between these two effects. However, there are two key differences between scintillations and the effects of a GPS receiver-spoofer:

1. Scintillations would affect neighbouring GNSS users as well. So if multiple users were to report on their respective GNSS signal qualities, a control centre may be able to tell the difference between an occurrence at a single site (could be a spoofer) or many occurrences at neighbouring sites (more likely a scintillation).
2. Spoofing generally affects GNSS channels from all satellites, and attempts to generate a different navigation solution with a controlled ranging error. Generally, scintillation does not impact all channels simultaneously, and causes uncontrolled ranging errors [Humphreys, personal communication].

In summary, once the detection rates have been successfully quantified, the Data Bit Latency defence could be a sensible hardening measure for any GNSS application world-wide that involves location or timing services. This is because the adversary would need to implement a solution to work around this defence, and these workarounds are currently considered to be a challenge [Humphreys08]. This heightens the bar for the adversary. The defender's costs are negligible, since the method can be implemented using receiver firmware changes.

Phase trauma

In presence of a real GNSS signal, the effects of atmospheric scintillation and of switching on a GNSS simulator are quite similar. The effects of scintillation on a GNSS receiver can be simulated by *forcing phase and amplitude variations in the output of a GPS signal simulator* [Humphreys09].

The term "phase trauma" is occasionally used to refer to the way in which GNSS signal interference upsets a receiver's tracking loop. This is a subject of research in relationship with the way in which the Ionosphere[30] disturbs navigation signals[31].

> *Radio wave scintillation, the temporal fluctuation in phase and intensity caused by electron density irregularities along the propagation path, stresses a GPS receiver's carrier tracking loop, and, as severity increases, can lead to navigation bit errors, cycle slipping, and complete loss of carrier lock* [Humphreys09]

Ionospheric scintillations are characterised by stronger effects near the equator, and mainly in those regions, *deep power fades (> 15 dB) accompanied by abrupt, approximately half-cycle phase transitions*. In some receivers, the same ionospheric scintillations can also provoke errors in reading out the data bit [Humphreys09].

What does this have to do with GNSS spoofing or meaconing? If a GNSS simulator's carrier signal is not aligned with the GNSS constellation's signal, then the two different signals must physically interfere with each other. Depending on how the phase between the two signals, the resulting composite signal can have different (stronger or weaker) amplitude and/or a different phase, than either two of the original signals.

Therefore, a GNSS spoofing or meaconing device, when it is not "carrier phase aligned" with the constellation's signal, will generate interference that looks much like scintillation. The interference will provoke phase and/or amplitude changes, and thereby will stress the victim receiver's tracking loops[32] and/or may cause (temporary) loss of lock for some satellites.

In conclusion, scintillation and signal simulation are intimately related.

Vestigial signal defence

If the adversary uses an "open air broadcast" scenario, then the receiver can potentially detect the presence of two GPS signals, corresponding to the fake and the real signal. While it may be difficult

[30] Ionosphere disturbances vary with the solar cycle. When solar activity is maximal, one expects increased disruptions to GNSS.
[31] Ionospheric scintillation can be loosely compared to temperature-induced optical scintillation, so-called mirages.
[32] Since commercial receivers usually first convert the signal from the carrier frequency to baseband, they perceive "carrier signal" anomalies in the form of artefacts in the accumulated Doppler, tracked in Delay and phase lock loops (DLL/PLL).

without external aids to tell which signal is the authentic one, the very fact that two consistent GPS signals are present is indicative that spoofing (or meaconing) is ongoing.

The figure below visualises a GPS receiver's auto-correlation function during a spoofing attack. It depicts the correlators used to track a single GPS channel. The X axis symbolises *correlator taps at 81 different 0.2-chip offsets* [Humphreys08], while the Y axis illustrates the amplitude of the auto-correlation.

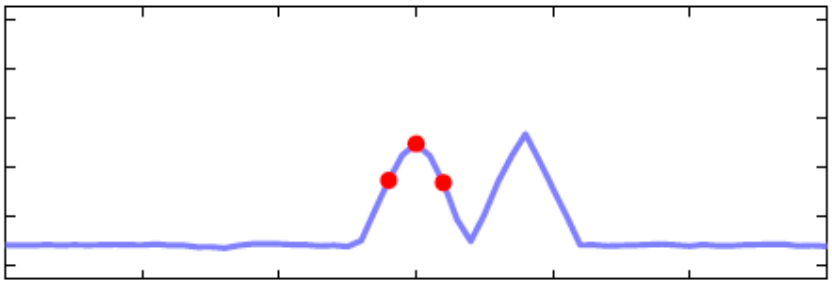


Figure 19: An illustration of a correlator array (x axis) versus signal power (y axis). Courtesy of T. Humphreys.

The three red dots represent the [code frequency's] *delay-lock loop's tracking points, which continuously attempt to align themselves so that the center point is maximized and the flanking points are equalized* [Humphreys08]. GPS receivers routinely use this procedure, the output of which feeds into the data message read-out and pseudo-range calculations.

In the figure, the centre curve is the original signal, and the curve to the right the spoofed signal. In the figure, there is a clear way to distinguish the two, but this is an artefact of the simulation exercise performed by Humphreys and Ledvina. In a real life scenario, no such difference would be noticeable, and the two peaks would both have smoothened peaks.

However, if two such peaks are present, then an attack must be ongoing: The adjacent "twin peaks" in the illustration translates into 300 metres of distance, effectively excluding multipath as a natural cause of this sort of artefact[33].

Interestingly for the adversary, Humphreys and Ledvina also found that if two GPS signals are present, then the tracked signal does not need to have the higher correlation power level.

If a receiver implements the vestigial signal defence, then it is helpful if also attempts to detect unrealistic discontinuities in position and/or time. Indeed, if the tracked position were allowed to jump by several kilometres, then the vestigial signal may no longer be within the correlator taps, and instead be located "off to the left" in the above Figure 19.

This matter is pursued further in reference [Humphreys08]:

> *The vestigial signal defense is premised on the difficulty of suppressing the authentic signal after successful lift-off of the delay-lock loop tracking points. […] Construction of an effective suppressor signal requires knowledge to within roughly 1/8 of a cycle of each authentic signal's carrier phase at the phase center of the target antenna. Such precise knowledge of carrier phase implies cm-level knowledge of the 3-dimensional vector between the target*

[33] If multipath were the cause, then this would imply that the real signal is reflected off a near perfect mirror that is located more than 300 metres away, for each satellite. The reflection would need to be persistent in time, implying a large flat mirroring surface as a natural cause. Also, each affected tracking channel would notice these two peaks, with a distinct space between them. One possible cause for this phenomenon is scintillation, but usually, not all signals scintillate such that they yield a consistent PVT solution.

antenna and the transmitter phase centers. This would be challenging except in circumstances where the receiver-spoofer could be placed in the immediate proximity of the target antenna phase center. Absent an effective suppressor signal, a vestige of the authentic GPS signal will remain in the input to the target receiver. Soon after lift-off of the delay-lock look tracking points, the vestige may be well disguised as multipath, but its persistence and distance from the spoofed correlator peak will eventually distinguish the two effects.

To detect the vestigial authentic signal, the target receiver employs the following software-defined technique. First, the receiver copies the incoming digitized front-end data into a buffer used only for vestigial detection. Next, the receiver selects one of the GPS signals being tracked and removes ["wipes off"] *this signal from the data in the buffer. This is the same technique used to remove strong signals in combating the near/far problem in spread spectrum multiple access systems, including GPS. Once the tracked signal has been removed, the receiver performs acquisition for the same signal (same PRN identifier) on the buffered data.*

The adversary could defeat this measure by a "shielded broadcast" or by a "carrier phase aligned" GPS receiver-spoofer.

Last but not least, this method also requires further research to quantify the false positive rate, especially in relationship to ionospheric scintillation. See the discussion on this topic in the section on "phase trauma".

In summary, once the detection rates have been successfully quantified, the Vestigial Signal defence could be a sensible hardening measure for any GNSS application world-wide that involves location or timing services. This is because the adversary would need to suppress the original GNSS signal, which is possible only with a "shielded broadcast" or "carrier phase aligned" spoofer [Humphreys08]. This heightens the bar for the adversary. The defender's costs are negligible, since the method can be implemented using receiver firmware changes.

Considerations for kinematic environments

The sea is a kinematic environment, both because of the presence of waves, and because the vessel changes directions. Applied to maritime GPS location spoofing, if the vestigial signal is to be cancelled in order to perform illicit trips ("carrier phase alignment" and "suppressed vestigial signal"), then the adversary must deal with this kinematic environment, which frequently upsets the 3D vectors between the antennae, and hence requires constant monitoring to maintain the effectiveness of the GPS suppression signal. It is possible to overcome this difficulty by coupling the GPS receiver-spoofer to an inertial sensor, and this would require either one of the following:

- Purchase of a real-time 3D kinematic tracking system (for example those in references [Xsens] or [Septentrio], but others also exist), OR
- Development of equivalent technology, implying engineering research and development effort.

In both cases, the adversary would then need to integrate the real-time 3D kinematic tracking into a GPS receiver-spoofer, with the latter able of providing carrier phase alignment. In literature, due to its applications in avionics, GPS/INS integration is a well-covered subject, so quite a few GNSS engineers have such knowledge. Also none of the above equipment is considered "sensitive technology".

Navigation Message Authentication (NMA)

In reference [Scott03], L. Scott defines two GNSS signal authentication enhancements that could be used in the civilian sector:

1. "Data Message Authentication", also known as "Navigation Message Authentication" (NMA) in reference [Hein07]
2. "Public Spreading Code Authentication" (PubSCA)

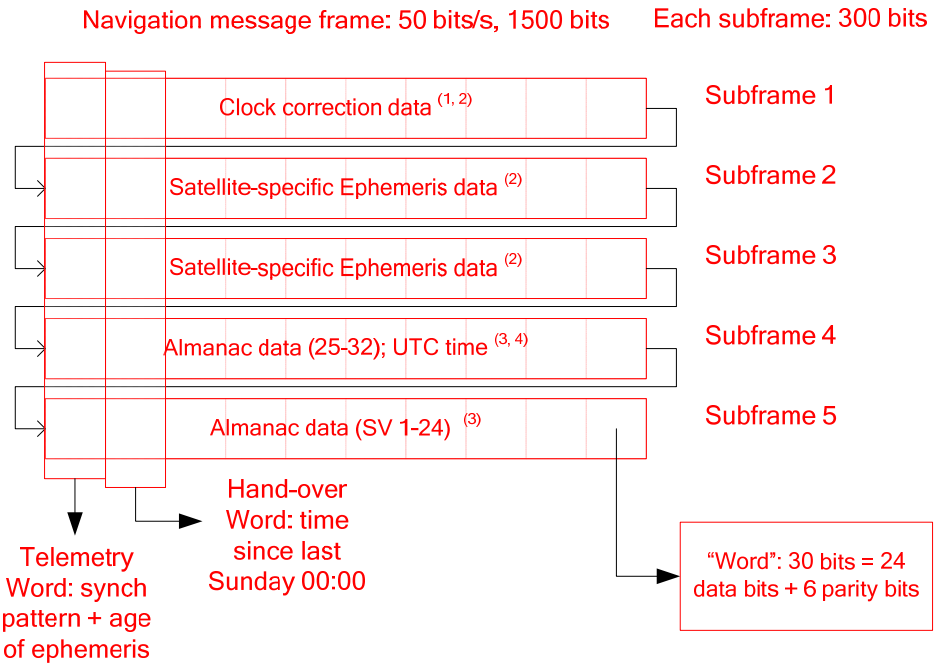


Figure 20: Anatomy of a GPS navigation frame.

NMA works by adding a public signature, created by a corresponding private key, at the end of every navigation message. G. Hein writes [Hein07]:

> *NMA denotes the authentication of satellite signals by means of digitally signing the modulated navigation data. […] The navigation message consists of one or several data blocks D_N, containing the orbit and clock parameters, and of one or several data blocks containing the digitally signature D_S. The recipient receives the data blocks D_N and the signature blocks D_S via the modulated data bits on the ranging signal. After receiving a complete message, the recipient can authenticate it by means of the […] publicly known validation key.*

If one compares these requirements with Figure 20, one notes that additional authentication data must be broadcast per satellite. Hence NMA implies a change to the GNSS infrastructure.

In addition, NMA suggests the use of a public key infrastructure (PKI), which in turn suggests signed software updates to update the public keys and other parts of the receiver. L. Scott writes [Scott03]:

> *For added security, the distributed public key should also be signed by a Certificate Authority (CA) in order to validate the public key to the user community. This keeps a spoofer from publishing a public key of his own. A CA should sing any downloadable software updates as well.*

A carelessly implemented NMA defence can be overcome "by comparatively simple means" [Hein07]:

> *The spoofer receives the authentic navigation signal of the satellite and bitwise reads out the navigation message. The bit stream of the valid navigation message is transferred to the signal*

> *generator, which modulates the cryptographically correct data on the forged navigation signal. […]As the spoofer merely interchanges the signal transmission time and sends an identical copy of the navigation message, the receiver is not able to detect the forgery using cryptographic methods.*

In summary, an attack on NMA achieves its goal (of spoofing the victim receiver) simply by replaying the navigational messages with an associated delay, misleading the victim's ranging calculations, and hence leading the victim receiver to calculate a PVT solution that is controlled by the adversary. In a receiver-spoofer, each spoofer channel would need to introduce a variable delay when transmitting the messages.

But the receiver-spoofer is not just subject to variable delays in the individual channels, but also to an additional constant delay for all channels. This is required because the receiver-spoofer must for each channel ([Hein07]):

> *Process the satellite signal and read out a single bit of the navigation message. This time delay is roughly approximated by the reciprocal transmission rate, thus, on the order of 10 milliseconds.*

The delay may shrink by an order of magnitude if the adversary can read out the signal directly, by use of a directive antenna array, or a multichannel beamformer. But without such equipment, a GNSS retransmission time delay would be associated with a time reference jump of at least 10 milliseconds.

Thus, in the event that an adversary succeeds in spoofing NMA, the victim receiver would, at one point, be tricked into accepting both a constant delay (corresponding to the time required to read out the navigation message) and a set of variable delays per satellite (which permits the spoofer to indicate a false position to the receiver). In the case of a cryptographically authenticated navigation message, this implies some sort of jump in the GNSS time reference.

Therefore, the above "jumps" are in principle detectable, but in practice this depends on how accurate the receiver's clock is, and how long ago the receiver was last tracking an accurate GNSS timing signal. Together these two factors determine a time drift "window of acceptance" [Humphreys08]. But how large is this window? Industry sells temperature-controlled OEM quartz clocks (TCXO) for as little as 40 EUR, with maximal clock drift of 1 part per million (ppm) in either time direction, under realistic conditions of vibration, temperature differences, and accelerations. (Vendors compete on improving key performance metrics, therefore it is reasonable to assume that future models will have superior characteristics.) With a 1 ppm drift rate, for a 10 ms drift to seem just about plausible to a receiver "time jump" detection algorithm, the adversary must first stop the GNSS receiver from tracking any GNSS timing signal for 2.7 hours, before imposing the fake GNSS signal.

Though the above loss of GNSS coverage may be unusual, it is no *prima facie* evidence of an ongoing attack. It would therefore be useful if trackers were to report loss of GNSS coverage. Fortunately, some existing GNSS tracking devices, such as the VMS, include provisions for warning the control centre of exactly such events.

Because of this vulnerability of NMA, G. Hein argues for complementary anti-spoofing measures, such as in the case of a *"participant of a GNSS-based tolling system"*. In essence, this means that the Galileo receivers for such purposes should *"consider a combination with one or more non-cryptographic anti-spoofing methods"*

1. monitoring the receiver for sudden clock biases, *"of the reciprocal transmission rate, thus, on the order of 10 milliseconds"* [Hein07, pg 73][34],
2. *"monitor the absolute power of received signals as well as their signal-to-noise ratio against unexpected behaviour"* [Hein07, pg 76],

[34] This delay could be as low as one millisecond.

3. *"detect spoofed signals by monitoring the direction of signal origin using phased array antennas/ receivers"*, which is explored e.g. in reference [Montgomery09] and in the annex to this document on an open source dual antenna defence,
4. *"Sensors such as inertial measurement units […] or compasses".*

The options 3 and 4 imply additional hardware, and are defences in their own right. Concerning option 2, Humphreys and Ledvina have already shown [Humphreys08] that it can be overcome by the adversary. The remaining option 1 leads to an effective defence when coupled with NMA.

An extension to option 1 is proposed, which may prove effective in several application areas, notably in maritime: If by law, the GNSS receiver must have view of the sky during normal operations, then one can prolong the 2.7 hour time period cited above, by about two orders of magnitude. One would require the use of a rechargeable battery. When main power is cut, then at intervals determined by an Exponential distribution[35] with $1/\lambda$ chosen at e.g. 60 minutes, the GNSS receiver circuit starts up, using only the battery. It would then fully acquire the GNSS navigation signals[36], download fresh ephemeris data from satellites in view, adjust its clock, and shut down immediately. This would repeat until the battery is exhausted. Depending on the relationship between λ, battery capacity, and GNSS module power drain the above process can go on for several days or weeks. This defence would be guaranteed to work as long as the GNSS module has view of the sky during each cycle. If as in road transport, the view of the sky is not guaranteed during each cycle, then one can compensate by adapting the value of λ.

In summary, a receiver that uses Navigation Message Authentication (NMA), when coupled with an accurate clock and backup battery, constitutes a strong and world-wide defence against GNSS spoofing. Apart from these relatively simple additional hardware requirements, a NMA type defence can be implemented with firmware changes in the GNSS receiver. However, NMA itself requires infrastructure changes at the GNSS provider. However, while NMA will not be put into GPS any time soon [Logan07], the Galileo system is considering NMA on its Open Service, but only for a later phase after its initial go-live [ESA09]. It is therefore unlikely that the NMA defence would be available in the near future.

If this defence was available, then it would be an optimal trade-off in terms of cost versus protection (except if the application domain warrants the use of Inertial Navigation Sensors).

Public Spreading Code Authentication (PubSCA)

G. Hein writes [Hein07]: *Another approach to preventing spoofs employs the proposed public spreading code authentication (PubSCA) described in the paper by L. Scott* [Scott03]*. This method expands navigation message authentication by adding another security feature. Besides the digital signature of the navigation data, additional* [Spread Spectrum Security Codes (SSSC)] *codes are inserted into the ranging code in fixed time windows.*

In fact, the SSSC are based on the concept of "delayed authentication with symmetric keys", which is conceptually related to the "Timed Efficient Stream Loss-Tolerant Authentication" (TESLA) algorithm.

The TESLA paper [Perrig02] succinctly explains the core concepts:

> *Broadcast authentication requires a source of asymmetry, such that the receivers can only verify the authentication information, but not generate valid authentication information.*

[35] The Exponential distribution has the "memorylessness" property, which means that the adversary always faces the same probability that the device will start up.
[36] The factor λ is chosen such that it is unlikely that the receiver's window of acceptance widens beyond 10 ms.

TESLA uses time for asymmetry. We assume that receivers are all loosely time synchronized with the sender — up to some time synchronization error delta [which is zero for GNSS], *all parties agree on the current time. Here is a sketch of the basic approach:*

- *The sender splits up the time into time intervals of uniform duration. Next, the sender forms a one-way chain of self-authenticating values, and assigns the values sequentially to the time intervals (one key per time interval). The one-way chain is used in the reverse order of generation, so any value of a time interval can be used to derive values of previous time intervals. The sender defines a disclosure time for one-way chain values, usually on the order of a few time intervals. The sender publishes the value after the disclosure time.*
- *The sender attaches a MAC to each packet. The MAC is computed over the contents of the packet. For each packet, the sender determines the time interval and uses the corresponding value from the one-way chain as a cryptographic key to compute the MAC. Along with the packet, the sender also sends the most recent one-way chain value that it can disclose.*
- *Each receiver that receives the packet performs the following operation. It knows the schedule for disclosing keys and, since the clocks are loosely synchronized, can check that the key used to compute the MAC is still secret by determining that the sender could not have yet reached the time interval for disclosing it. If the MAC key is still secret, then the receiver buffers the packet.*
- *Each receiver also checks that the disclosed key is correct (using self-authentication and previously released keys) and then checks the correctness of the MAC of buffered packets that were sent in the time interval of the disclosed key. If the MAC is correct, the receiver accepts the packet.*

One-way chains have the property that if intermediate values of the one-way chain are lost, they can be recomputed using later values. So, even if some disclosed keys are lost, a receiver can recover the key chain and check the correctness of packets.

The original TESLA algorithm applies to packets sent over an IP connection. In the GNSS context, the packets are sent using the RF spectrum, at a very low power level that is not directly readable. Furthermore, each SSSC block (at the end of the 1 millisecond GNSS message packet) is calculated separately.

How is this relevant? This feature further impedes a GNSS spoofer because he cannot just read the SSSC block by ordinary “correlation”. By design, if one simply correlates the SSSC blocks, without first feeding the SSSC RF samples through a de-spreader that uses the one-way-chain keys, one ends up with white noise. In turn, the one-way chain keys cannot be calculated without having received the disclosed key, which the sender will do after a disclosure delay. L. Scott [Scott03] proposes a single authentication message every 5 minutes.

An anti-spoof GNSS receiver would therefore use SSSC by

1. Storing SSSC signal samples that cannot yet be interpreted.
2. Reading the NMA digital signature block that the GNSS infrastructure sends at the end of the navigation message.
3. Verifying the authenticity of the digital signature as in NMA.
4. Using the NMA signature bits to generate a local copy of the SSSC bits. The NMA bits are fed into a PRNG as a seeding value, and the PRNG yields the SSSC bits.
5. Cross-correlating the stored signal samples with the local copy of the SSSC bits.

An adversary is therefore faced with the following problem [Scott03]:

Without access to the digital signature, the spoofer must read the SSSC directly from the transmitted signal, which [is] buried below thermal noise. […] The spoofer needs a high gain

> *antenna to successfully read the SSSC chips directly. Because such a high gain antenna is also very directive, multiple high gain antennas or a multichannel digital beamformer is needed to successfully receiver SSSC chips directly from all satellites in view. […] Such an antenna* [dish] *would be about 30'' (75 cm) in diameter and have a 20dB bandwidth.*

In any kinematic environment, movements (changes of heading, jerks and vibrations) further complicate the picture for directive antennae. With high directivity, even a small offset in a pointing angle of the above antennae means that the GNSS signal cannot be read. Some method must be used to nullify the influence of motion on the signal read-out. For parabolic antennae, this could be a gimbal or a swivelling robotic Steward platform, but both require further stabilisation measures that operate in real time. An alternative to the use of directive antennae is the use of a beamforming array, coupled with an INS in order to adapt the beamforming parameters, also in real-time. With beamforming arrays, samples can be buffered, and this would help the beamforming process. But any buffering would contribute to the processing delay, which the victim receiver could measure. In conclusion, while the above are possible in theory, high-gain beamforming systems in dynamic environments are very challenging to implement in practice, and costs would be commensurate with that challenge [Fortuny, personal communication]. So it is still true that s*uch a spoofer architecture is highly impractical.* [Scott03].

If spoofing is no longer viable, an adversary could theoretically use "live meaconing": The adversary would sample a GNSS signal at point A, transmit the signal samples to point B using wireless broadband point-to-point connectivity [GIGABEAM], and reconstitute a signal at point B. It is for that reason that a receiver protected by PubSCA also benefits from an accurate clock, as explored above in the chapter on the NMA defence.

In summary, a receiver that uses Public Spreading Code Authentication (PubSCA), when coupled with an accurate clock and backup battery, constitutes a strong and world-wide deterrent against GNSS spoofing. Apart from these relatively simple additional hardware requirements, a PubSCA type defence can be implemented with firmware changes in the GNSS receiver. However, PubSCA requires major infrastructure changes at the GNSS provider [Scott07], therefore the defence would not be available in the near term.

If this defence were available, then it would be an optimal trade-off in terms of cost versus protection (except if the application domain warrants the use of Inertial Navigation Sensors).

Psiaki method: "Cryptographic defence based on estimation of W-bits"

In reference [Psiaki09], Psiaki proposes a semi-codeless method that could be used for GPS spoof defence. The method is subject to a patent application. It relies on algorithms to estimate the W bits, which are the secret bits that the US Military uses to encrypt the P code to yield the P(Y) code. Each satellite transmits its own P(Y) signal, and uses its own "W bit" sequence.

The method requires hardware that is able to perform semi-codeless P(Y) processing. Even though it is possible to implement the method using single-frequency L1 receivers, Psiaki mentions that

> *[…] The calculation […] is already implemented in dual frequency civilian GPS receivers that use semi-codeless techniques and the process that is called "soft-Decision Z-Tracking".* [Psiaki09]

As presented in the above reference, the method foresaw a wide-band receiver. However, recent advances imply that the method could also use less expensive narrow-band receivers [Psiaki, personal communication].

There are algorithms that can be used on a receiver with a wire antenna to estimate the W bits with a probability strictly greater than 50%. The same algorithms can be used on high-gain antennae, where they can attain accuracies of asymptotically approaching 100%. Accuracy is a function of the antenna gain and estimation method used [Woo00]. With an antenna gain of *33 dB for typical received P(Y) power levels and typical noise power spectral densities […] the probability of* [any given W bit being in error] *is less than* $1.3*10^{-12}$ [Psiaki09].

Psiaki observes that the above can be exploited for a semi-codeless cryptographic spoof defence.

- The equipment includes a module that can estimate the W bits, and accumulates these as a series of W_{est}. The estimate must be better than pure chance.
- A separate ground segment is created, consisting of several ground stations which have an array of beamforming or steerable antennae. Three to four stations would achieve worldwide coverage. The W bits would be estimated to near 100% using post-processing. Below these estimates are therefore labelled as "W_{true}".
- Alternatively, if the US authorities could provide *short segments of true W bits after the fact, perhaps with a 0.5 second delay* [Psiaki09], then the above ground stations would not be required.
- Periodically, at the ground segment, a sample of W_{true} are accumulated, cryptographically signed and sent using e.g. SBAS communications satellites (WAAS, EGNOS, others), to its user base. The proposed message consists of 480 W bits.
- When the user equipment receives a message containing the signed W_{true} bits, it correlates its W_{est} with the W_{true}. This will yield typical correlation strengths. Checks are performed per P(Y) channel.
- Because the W_{true} are authenticated, no adversary can fabricate the W bits, effectively excluding a man-in-the-middle attack.
- The method, as published, has a false positive rate of $4*10^{-5}$ and a missed detection probability of $3*10^{-5}$. With such detection statistics[37], no further hypothesis testing is needed; instead any positive result can be sent directly to the control centre.

The method, as described for the GPS P(Y) signal, can be transposed to the Galileo encrypted signals, or to the GPS M-Code signal. If the firmware of the tracker can be updated to accommodate algorithms for the M-Code and/or the Galileo encrypted signals, then the defence would be future-proof.

The US authorities have to allow use of "W" bits only if one wants to derive them from the US system. If one is willing to implement one's own high-gain antenna ground infrastructure, then no such permission is needed, but the system would face the increased costs mentioned above. In any case, the ground station would provide the "W bits" in delayed mode, and not all "W bits" for all satellites are needed simultaneously. According to Prof. Paul Kintner, the US Authorities would not have a problem with the delayed release of the "W bits" [personal communication].

However, should a degradation of the W bits be required for national security reasons, then any degradation in the accuracy of the "W bits" would lead to changes in detection statistics (at the same message length of 480 bits), or it would lead to increases in message length to achieve similar detection rates.

[37] If the user equipment receives a P(Y) signal at C/N_0 of 45 dB-Hz, then each check requires about 480 bits per channel to achieve the above alarm rates.

Figure 21 summarises and illustrates the architecture:

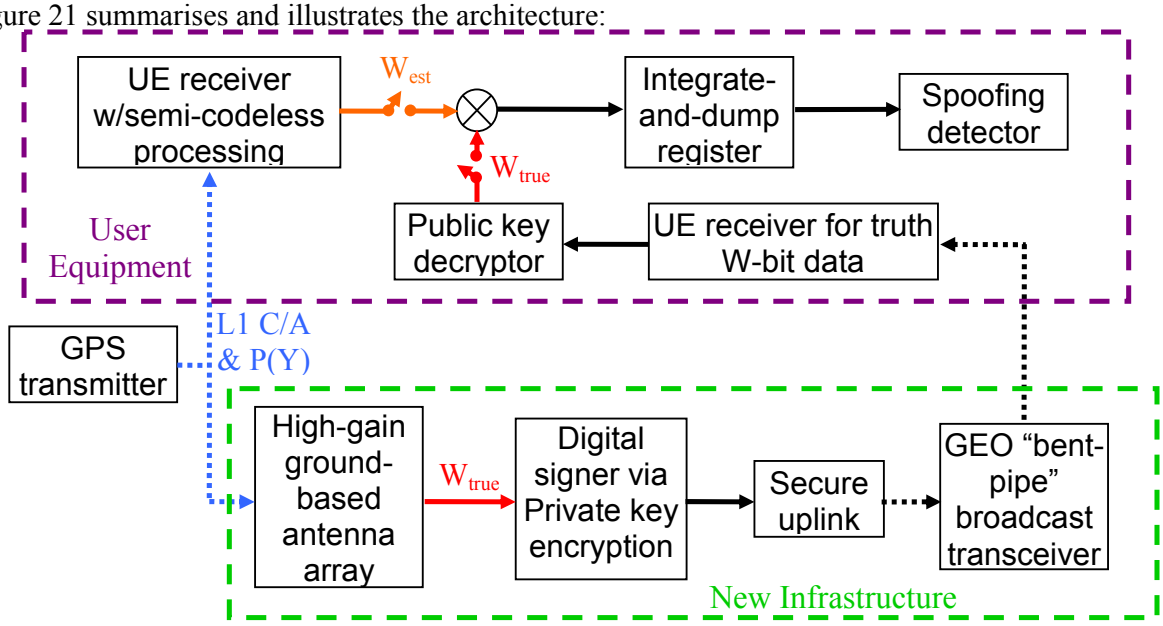


Figure 21: Cryptographic defence based on estimation of W-bits, method of M. Psiaki. Courtesy of M. Psiaki, personal communication.

How about the vulnerabilities of this method? As in the case of "Navigation Message Authentication", the "Cryptographic defence based on estimation of W-bits" technique is potentially vulnerable to GNSS meaconing: An adversary can record the full signal, as well as the W bits messages, and replay both later without the real signals being present.

It also is vulnerable to spoofing, to a limited extent. An adversary could use a technique similar to the one described in the section on "spoofing Navigation Message Authentication" in reference [Hein07]. This would require

- The use of a steerable antenna array to raise the P(Y) signals above the noise floor, otherwise the P(Y) will not be available without noticeable delay for his purposes;
- The use of digital radio frequency memory techniques [Hein07] to introduce arbitrary delays in the message sent by each satellite, in order to obtain some freedom with respect to the position to spoof;
- The retransmission of the obtained and purposefully delayed P(Y) code sequences.

The resulting processing delay is composed of an overall and a per-satellite part. The overall delay decreases with the antenna gain, while the per-satellite delay is a function of the distance between the real and the spoofed position.

Because it is in principle vulnerable to the above attacks, the method "Cryptographic defence based on estimation of W-bits" could optionally be augmented by any additional measure that also hardens the "Navigation Message Authentication". Most notably, it could be augmented with a high-accuracy clock and a backup battery.

Concerning global availability, Europe's SBAS service (EGNOS) could be used to send the W bits to the user equipment. The availability of each SBAS service determines the geographical coverage range of the "W bit" defence[38]. Regional coverage can be extended later using other SBAS services, communication satellites, or publication on the Internet.

[38] For example, the EGNOS service exclusively covers western and central Europe.

The costs are dominated by the following requirements:

- The installation of several ground stations with high-gain (steerable or beamforming) antennae, which may cost up to 30 million USD, OR
- Alternatively, one would need to convince the GNSS authorities to release some of their secret encryption bits after a given delay,
- A "bent pipe" service must be established that transmits the W bits to the user base.

According to the US authorities, the P(Y) signal will be switched off in 2021, but since the GPS M-Code and Galileo encrypted signals broadly operate according to similar principles, the defence remains conceptually valid. If in addition, the firmware of the device can be updated to accommodate algorithms for the M-Code and/or the PRS signals, then the defence would be future-proof.

In summary, a receiver that uses the "W bit defence", when coupled with an accurate clock and backup battery, constitutes a strong defence against spoofing and meaconing, which could be implemented soon. Because the method can be implemented using narrow band receivers, user equipment costs are low. The defence requires a broadcasting service, which in turn may limit the service to a regional availability. The defence may require several ground stations, each hosting a grid of directive antennae, but only if the GNSS provider does not supply the encryption bits with high accuracy. The GNSS provider could oppose concerns over the delayed publication of the encryption bits, which may make it easier for cryptanalysts to break the secret key(s) used to generate the encryption bits.

Lo method: Signal authentication using L1 samples

Sherman Lo et al propose a signal authentication mechanism in reference [Lo09-2]. It is subject to a patent application. Briefly, the concept presented in InsideGNSS [Lo09-2] would authenticate the position of the receiver, in real time, as follows:

- The user equipment samples the L1 P(Y) signal using a 15-20 MHz wideband antenna. The user equipment occasionally sends about 4800 bits (600 bytes) of samples to a reference station.
- In order to exclude a man-in-the-middle attack, the equipment must sign the sample data with a cryptographic signature in his message.
- The reference station can then use the samples and a high-gain antenna, in order to infer an offset of the user with respect to its own position. For that it needs the P(Y) signal, which it could read out with high gain antennae, or which it could calculate, if the GNSS authority gives access to the P(Y) code or the W bits on a secured and delayed basis.

In reference [Lo09-2], the L1 and P(Y) signals are just for illustration purposes. The method would just as well work with the Galileo encrypted signals or the planned GPS M-Code. The patent 20090195354 specifies both directions as possibilities for the transmission of signal samples. As with the "Psiaki method", if the firmware of the tracker can be updated to accommodate algorithms for the M-Code and/or the PRS signals, then the defence would be future-proof.

The method as presented in reference [Lo09-2] has several two key features:

1. The verification is performed not at the GNSS receiver, but in a reference station, simplifying the design of the user equipment, and
2. the reference station would only assume that a P(Y) signal exists
3. with respect to the Psiaki method, the *signal samples both reduce the communication bandwidth and increase the detection power from a given length of data* [Psiaki, personal communication]. It also yields a range, of the distance between both the user equipment and the reference station.

As originally presented, the method is unlikely to be "sensitive" to the US military, as long as the high-fidelity P(Y) signal samples are guaranteed not to leave the reference stations.

However, if signal samples leave the reference stations, then this method is just as "sensitive" as the "Psiaki method" described further above: Any broadcasting of high-fidelity P(Y) signal samples can be used to recover the W bits. As the signal to noise ratio of the samples approaches infinity, the bit error rate for estimating the W bits tends to zero. To quantify, from the diagrams in reference [Woo00], once the L2 P(Y) C/N_0 reaches 60 dB, the "soft z-tracking" and MAP methods predict the W bits with near 100% success rate.

What about the vulnerabilities of this method?

- Meaconing: A precise time stamp on the L1 data would make it possible to exclude conventional meaconing attempts, and one is left with the speculative case of "live meaconing": Any data packet transmitted from the use equipment to a reference station will inevitably have a transmission time, and this time must have a tolerance margin. The adversary could then impose a meaconed signal, if it is captured and replayed within this tolerance margin.
- Spoofing: Any spoofer would read out the L1 signal against replicas of both the civil C/A and the military P(Y) codes. For reading out the P(Y), this requires the use of directive antennae or of a beamforming array. The spoofer will need to introduce a global and a set of per-satellite delays into the spoofed signal, because of its own correlation process, as described in reference [Hein07].

Because of the vulnerabilities described above, this method becomes even more potent with the addition of a battery-powered high-accuracy clock, and algorithms for detecting jumps in the GNSS time signal.

The original paper of the "Lo method" does not specify the false alarm and non-detection rates, yet these are crucial information. This is however the subject of current research by Psiaki and others [Psiaki, personal communication].

For the purposes of near real-time GNSS spoof protection, as per the reference [Lo09-2], the costs of this architecture are dominated by the following requirements:

- User equipment capable of sampling a wider band (15-20 MHz) of the L1 signal, required in order to sample the P(Y) part of the L1 signal, which would cost an estimated 10 EUR apiece in larger volumes [Fortuny, personal communication]
- The installation of at least three ground stations with high-gain (steerable or beamforming) antennae, in order to cover a continent such as Europe [Lo09-2], where the cost rises with the signal-to-noise ratio to be achieved, and could reach several million EUR, OR
- Have the GNSS authorities disclose occasional signal samples or W bits with a certain delay. (The signal samples could optionally incorporate artificial noise to prevent any adversary from recuperating 100% correct W bits.)
- The obligation for the user equipment to secure and communicate L1 data sample sets, of about 600 bytes each, in near real time. In the maritime domain, this implies the use of global broadband services[39], such that communication costs remain subdued.
- The user equipment must safeguard and digitally sign the L1 data samples, requiring that each user device
    - host a secure controller, like an Infineon SLE88, NXP SmartMX P5S chip series, or the Dallas Semiconductor DS5000 series
    - use tamper resistant hardware to store L1 samples.

Note that for hardened GNSS trackers, tamper resistant sealing is required anyhow. Also, any tracker would do well to sign its position observations.

Unfortunately, the method is difficult to scale for real-time use as a road tolling system. In order to be an effective spoofing deterrent, one would expect that each registered vehicle would emit 4 L1 samples per day on average, spaced apart by an exponential distribution. With e.g. 33 million vehicles

[39] With 6.5 USD/MB for Inmarsat FleetBroadband, perhaps four L1 sample sets distributed randomly per day, and about 650 bytes of data for sending each sample set, effective spoof protection would come at about 50 US cents per month.

registered in the UK in 2006 [DFT06], that would amount to 132 million messages per day. Each message would contain at least 600 bytes. If sent by mobile telephony SMS messages, then such L1 sample messages would represent about twice the volume of present text messaging: A SMS message has a payload of 140 bytes, so one would need 5 SMS messages for sending a single L1 sample, amounting to 660 million SMS messages per day for such samples. By comparison, in Britain in December 2009, 9.6 billion SMS messages were sent [MDA10], or about 310 million per day. Therefore, if a road tolling system is safeguarded by sending L1 samples by mobile telephony messaging, then the quantity of messages would become a challenge to mobile phone operators. In turn, this will make such a system more expensive, as operators must subsequently recoup the costs for upgrading their networks to handle a 200% increase of SMS traffic.

In summary, a receiver that uses the "Lo method", when coupled with an accurate clock and backup battery, constitutes a strong defence against spoofing and meaconing, which could be implemented soon. As for user equipment costs, an adequate wide-band analogue-digital converter circuit would cost around 5-10 EUR apiece in high volumes [Fortuny, personal communication], and some additional processing capacity may be needed in the receiver.

The principal weakness of the method lies in the fact that it is difficult to scale to an e.g. road tolling system. The method would require some more work to determine the false alarm and misdetection rates. The method further may require several ground stations, each hosting a grid of directive antennae, but only if the GNSS provider does not supply the information in delayed mode.

Hybrid Psiaki-Lo method

As explored above, the "Psiaki method", as presented in the original paper, relies on the GNSS provider publishing the "W bits", or on the use of expensive directional antennae, neither of which is free of issues.

By contrast, the "Lo method" as presented in the original paper, presents an alternative to the "Psiaki method" but has issues with scalability.

One can create several "hybrid Psiaki-Lo methods", depending on the direction in which data are sent (from the user equipment to a reference station or vice versa), and on the nature of the data (samples versus W bits). If the method involves W bits, it is covered by a patent of Psiaki, if it involves signal samples, it is covered by patent 20090195354 of Lo et al [Psiaki, personal communication]. One of these hybrid methods has the strengths of both methods, and apparently none of the above weaknesses.

1. With respect to the "Lo method" as presented in reference [Lo09-2] one inverts the path of the L1 samples: Instead of sending them from the user equipment to the reference station, the samples are sent from the reference station to the user equipment. It has to be noted that the above patent application of Lo et al explicitly mentions the possibility of inverting the path of the samples, thereby explicitly covering the present "hybrid method". An additional EGNOS message type could be used to encapsulate such a message, making for a regional defence.
2. With respect to the "Psiaki method", one replaces the W bits in the message by actual L1 signal samples.

Psiaki plans to implement this hybrid method using budget-friendly commodity narrow-band receivers [Psiaki, personal communication].

The above approach leads to a method that has many characteristics of the Lo method. In particular, for the US military, it may be as sensitive as the original Psiaki method, since high-fidelity signal samples reveal the W bits. The Galileo encrypted signals should be considered as an alternative.

If the GPS authorities agree to transmit their samples, then the reference stations can generate "synthetic samples" for distribution. The design can be visualised as follows:

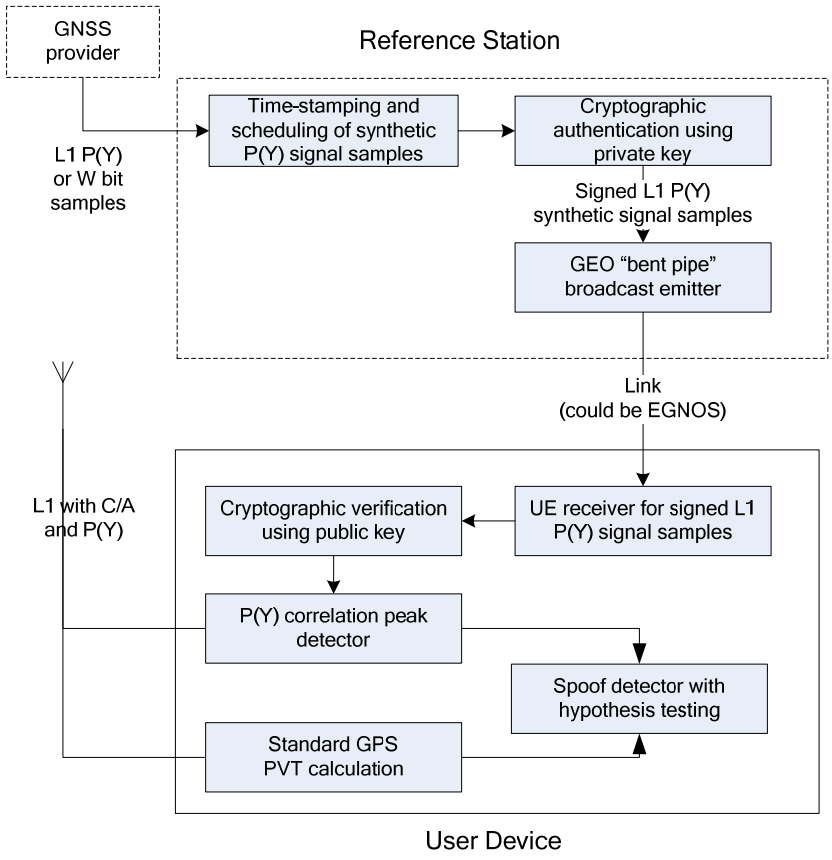


Figure 22: A hybrid Psiaki-Lo defence, using synthetic P(Y) signal samples. The method could be applied to the Galileo PRS signal *mutatis mutandis*.

Economically speaking, the method's costs are mostly a function of the reference stations.

- The costs in the user equipment are subdued, since according to research by Psiaki, no wideband receiver is necessary [Psiaki, personal communication]. If the correlation algorithm consumes too much processing power to be implemented in the receiver, then it can be hosted on a separate board, with marginally increased costs. This option may require that the tracker is connected to a power supply or larger battery, but economically this would not be an issue for tracking dangerous goods or monitoring larger maritime vessels.
- The method requires that the user equipment be able to receive signal samples from a reference station. In case EGNOS is used, then that is a feature that commodity SBAS-augmented GNSS receivers will be able to implement with software changes.
- The method requires the set-up of reference stations, which may or may not require expensive steerable antennae.

The method is just as vulnerable to meaconing and spoofing as the original "Lo method", with which it shares the detection algorithm. Therefore, the method becomes more potent with the addition of a battery-powered high-accuracy clock, and algorithms for detecting jumps in the GNSS time signal.

The detection statistics of the "hybrid Psiaki-Lo method" also should be identical to those of the original "Lo method". Psiaki performed initial research on this method, and anticipates *an ability to accurately detect spoofing in less than 1.5 seconds* [Psiaki, personal communication]. Note that with

such low times to alarm, one could afford enlarge the sampling interval by an additional order of magnitude, if that further improves the type I and type II error rates.

Areas for further research would be:
- If and how the method impacts the security of the GNSS provider's encryption keys.
- The frequency of transmission of the signal samples, and how the user equipment "knows" which signal samples to store, so as to limit the size of the extra signal buffer.
- The design of the messages that transport the signal samples, and how transmission loss affects the user equipment's detection statistics.

In summary, a receiver that uses a prospective "hybrid Psiaki-Lo method", when coupled with an accurate clock and backup battery, would constitute a strong defence against spoofing and meaconing, which could be implemented soon. As for user equipment costs, some additional processing capacity may be needed in the receiver. Just like the original "Psiaki method", this method would be easy to scale for e.g. a road tolling system. If the reference stations would receive their data directly from the GNSS provider, they would not require directive antennae.

This defence is a strong contender if neither "Navigation Message Authentication" nor "Public Spreading Code Authentication" methods are available, and if the use of inertial navigation sensors is not warranted or not deemed sufficiently strong for the application domain. For instance, this method can be used for GNSS time-keeping applications. Due to its high degree of protection, it can also be considered in the maritime domain and for the transport of dangerous goods. In that case, it can be combined with an inexpensive INS, to achieve dead reckoning ability, if desirable for the application domain.

Relationships between defensive measures and possible attacks

The above options are presently summarised, and compared against the possibilities available to the adversary.

Measure	**Characteristics** ***Trade-off***
Jamming-to-Noise (J/N) meter, cross-check unexpected J/N against C/No	1. Can detect unusual GNSS signal strengths, therefore detecting open air fake GNSS signal broadcasting. - *Low receiver unit cost: requires minor changes in the receiver's configuration.* - *No infrastructure cost.* - *Requires method to diminish the false alarm rate in case of unintentional radio-frequency interference. Otherwise, this measure may result in too many false positives to be useful.* - *Detects spoofing or meaconing that is using "open air broadcast", and any other situation where the adversary uses a high J/N or high C/N_0 level.* - *Cannot detect "shielded broadcast" or "carrier phase aligned" GNSS spoofing, or other situations with realistic J/N and C/N_0 level.*
Multi-Antenna Single Frequency Receiver	1. Several antennae tracking the same GNSS frequency can detect if GNSS signal comes from a single direction. 2. To be more effective, all antenna elements must be shielded inside of a single volume seal. Otherwise it can be defeated by a setup described in reference [Humphreys08]. - *Medium unit cost: Requires additional hardware and software, but architectures and algorithms are known.* - *No infrastructure cost.* - *Available today.* - *Can detect any a "GNSS signal generator", "GNSS receiver spoofer", and any form of "meaconing", as long as the adversary uses a single point source.* - *Can be defeated by a "GNSS receiver spoofer" that employs multiple sending antennae, which are slowly moved independently of each other.*

Measure	**Characteristics** ***Trade-off***
Single-Antenna Multi Frequency Receiver	1. A single antenna with at least two different RF down-converters, each tuned to a different frequency. 2. In its simplest effective configuration, a dual-frequency receiver would track the L1 and the L2C signal. - *Medium unit cost: Requires commercial off-the-shelf dual frequency hardware and software.* - *No infrastructure cost.* - *Available today.* - *Can detect any spoofing or meaconing using only one "single frequency".* - *Cannot detect any attack simultaneously conducted on all frequencies that it checks.* - *The L2C code signal has 10 times the frequency of the L1 code signal [GPS-SIG-WP]. With the capacity of current digital signal processors (DSP), Humphrey's and Ledvina's GPS receiver-spoofer, as of 2008, was not able to generate a spoofed version of the L2C code signal in real time.* - *However an adversary may well choose to implement a GNSS signal generator on a more powerful platform, in particular on a General Purpose Graphics Processing Unit (GPGPU).*
INS: Inertial navigation System (gyro-scopes, accelero-metres, optionally magneto-metres)	1. Requires INS hardware 2. Requires deep integration between GNSS and INS measurements - *Low receiver unit cost for a basic INS. Unit cost depends on precision of the INS. A basic INS would cost about 20 EUR per device. High-quality INS equipment costs orders of magnitude more.* - *However, INS systems' price/performance ratio is likely to improve significantly in a few years [R. Johnston, personal communication].* - *No infrastructure cost.* - *Available today.* - *On its own, can detect any spoofing or meaconing, if the relationship between the fake and real position implies measurable and implausible accelerations. This makes it much more difficult and impractical to mount an exploit.* - *Inertial measurements imply that any adversary must carefully manage the rate of change at which the fake signal's position moves away from the real signal's position. However any INS has a tolerance margin for accelerations, resulting in "integration drift", which an adversary could exploit.* - *GNSS/INS has a dual benefit: the INS' "dead reckoning" abilities make it possible to cope with GNSS outages, which in turn will reduce the number of unintentional and aberrant position reports.* - *Hardware and algorithms commercially available.* - *Vulnerable to an adversary who can program accelerations that are beneath the INS detection threshold.*

Measure	**Characteristics** ***Trade-off***
LORAN with or without modernisation (eLORAN)	1. Requires additional antenna and signal processing 2. Requires software integration between GNSS and LORAN measurements (not very difficult) - *Low receiver unit cost: Requires additional hardware and software, but architectures and algorithms are known.* - *Available, however it will not be usable in North America starting 2010.* - *No additional infrastructure cost, if eLoran would be maintained anyway.* - *On its own, can detect any spoofing or meaconing, as long as the LORAN signal can be used to determine position.* - *A poorly implemented LORAN receiver can be spoofed. Mitigation techniques can be developed, with or without the proposed LORAN message authentication. At present state of knowledge, a receiver implementing these mitigation techniques cannot be spoofed [Lo09].* - *A problem is the availability of LORAN signal, which does not cover the southern Hemisphere. Neither will it cover the western Hemisphere from 2010 onwards.* - *If using eLoran, if operating within reach of eLoran, and if properly implemented, has no known vulnerabilities. [Lo09]*
Data bit latency defence	1. Check that the navigation message bits do only change every 20 milliseconds. 2. Check the navigation message checksums for inconsistencies. 3. Determine the 12.5 minute and 2 hour boundaries at which each individual satellite can change its navigation messages. 4. Check that the ephemeris is updated only at 2 hour intervals. 5. Inside every 12.5 minute window, check that the changes in the navigation message, as transmitted every 30 seconds, are limited to satellite clock and almanac changes. Check that the satellite clock drift is realistic. Check that the almanac is reported in 25 different parts, as consistent with GPS normal operations, and that the changes in the almanac only occur at 12.5 minute boundaries. - *Low receiver unit cost: requires minor changes in the receiver's configuration.* - *No infrastructure cost.* - *Could be made available soon. Requires further research with respect to type I (false negative) and type II (false positive) errors.* - *May not render a spoofing attack impossible, but requires further theoretical and practical research by the adversary.* - *Becomes more potent when the GNSS receiver hosts a battery-powered high-precision clock, as the per-satellite 12.5 minute and 2 hour windows can be determined in advance.* - *If one adds a battery and checks for clock jumps, then that excludes meaconing type attacks, except for the speculative "live meaconing".* - *Vulnerable to an adversary in possession of high-gain antennae.*

Measure	**Characteristics** ***Trade-off***
Vestigial signal defence	1. Uses two buffers, one for primary, another one for vestigial signal detection. 2. In the vestigial buffer, the first detected signal is "wiped off", and GPS signal detection is repeated. 3. Amplitude and width of the vestigial signal are estimated. If such a signal is found, then hypothesis testing is performed with subsequent signal samples. 4. If the "vestigial signal" hypothesis is confirmed, then an alarm is raised, and the pseudo-range distance with respect to the primary signal may be calculated. - *Low receiver unit cost: requires minor changes in the receiver's configuration.* - *No infrastructure cost.* - *Could be made available soon. Requires further research with respect to type I (false negative) and type II (false positive) errors.* - *Becomes more potent when the GNSS receiver also checks for unrealistic position or time jumps.* - *Requires that the adversary either uses a "shielded broadcast" or implement "Carrier phase alignment".* - *For the adversary, "Carrier phase alignment" is a challenge in kinematic environments: the adversary must recalculate the emitter-to-receiver geometry in real time, by using gyroscopes and accelerometers.* - *Vulnerable to an adversary that can suppress the original signal.*
Navigation Message Authentication (NMA)	1. Requires "authentication of satellite signals by means of digitally signing the modulated navigation data" [Hein07]. 2. For civilian use, the signature uses asymmetric cryptography. - *Low receiver unit cost: requires minor changes in the receiver's configuration.* - *High infrastructure costs: Requires changes to the GNSS infrastructure, which itself are expensive.* - *Not available as of 2009.* - *On its own, can detect a "GNSS Signal Generator", but cannot detect "meaconing" (as it replays a real signal) or a "GPS receiver-spoofer" (as the latter simply delays interchanges the authenticated navigation messages).* - *Becomes more potent when the GNSS receiver checks for jumps in the GNSS time. This in turn requires that the GNSS receiver is powering its digital clock, and has been tracking a real GNSS signal recently (up to 3 hours ago).* - *When combined with above checks for clock jumps, can detect meaconing and spoofing, unless high-gain antennae are used to read out the satellite's signals directly.* - *Vulnerable to an adversary in possession of high-gain antennae.*

Measure	**Characteristics** ***Trade-off***
Public Spreading Code Authentication (PubSCA)	1. Requires Navigation Message Authentication (NMA) 2. Requires implementing a TESLA-like algorithm adapted for GNSS - *Low receiver unit cost: requires minor changes in the receiver's configuration.* - *High infrastructure cost: Requires changes to the GNSS infrastructure. Could be implemented at the same time as NMA.* - *Not available as of 2009.* - *On its own, can detect a "GNSS Signal Generator", but cannot detect "meaconing" or "spoofing" beyond the above.* - *Becomes more potent when the GNSS receiver checks for jumps in the GNSS time. This in turn requires that the GNSS receiver is powering its digital clock, and has been tracking a real GNSS signal (up to several years in the past).* - *When combined with above checks for clock jumps, can detect meaconing and spoofing, unless high-gain antennae are used to read out the satellite's signals directly.* - *Difference with the effectiveness of NMA is that PubSCA places a lesser constraint on the detection of clock jumps.* - *Vulnerable to some extent to an adversary in possession of high-gain antennae.*

Measure	**Characteristics** ***Trade-off***
"Psiaki method": Cryptographic defence based on estimation of W bits	1. Requires ground stations to read out the encrypted navigation message (in GPS, the P(Y), or a Galileo encrypted signal). The ground station read-out must be performed with a high gain. The same ground stations then infer the encryption bits (in GPS, the W bits). – OR – 2. Alternatively requires convincing the GNSS authorities to release the encryption bit sequence with an acceptable delay and acceptable precision. 3. Requires transmitting the W bits, in a digitally signed message, to a user community 4. The end-user receiver must be able to sample the civilian L1 and some of the military P(Y) signal. Both are sent using the L1 centre frequency, but the P(Y) is sent on a wider band. 5. The receiver performs its own W bit estimation (accuracy > 50%) and stores the result. 6. The receiver obtains the higher accuracy W bits through external aiding. 7. The receiver correlates its estimates with the higher accuracy received W bits, and constructs a pseudo-range. 8. The receiver checks the pseudo-range against the PVT solution from the non-authenticated signal. - *Low receiver unit cost: requires some changes in the receiver's configuration, as well as the reception of aiding information.* - *Medium infrastructure cost: Requires ground stations and transmission of delayed encryption bits by communication satellites, -OR-* - *Low infrastructure cost, provided that the GNSS provider can be convinced to disclose the encryption bits with some delay.* - *Not available as of 2009, but could be implemented rapidly.* - *The GNSS provider could be concerned that publication of encryption bits weakens cryptography.* - *By itself, the method is vulnerable to GNSS meaconing and GNSS receiver-spoofers.* - *Becomes more potent when the GNSS receiver checks for jumps in the GNSS time. This in turn requires that the GNSS receiver is powering its digital clock, and has been tracking a real GNSS signal recently (up to 3 hours ago).* - *When combined with above checks for clock jumps, can detect meaconing and spoofing, unless high-gain antennae are used to read out the satellite's signals directly.*

Measure	**Characteristics** ***Trade-off***
"Lo method": Signal authentication using L1 samples	1. Requires ground stations to read out the encrypted navigation message (in GPS, the P(Y), or a Galileo encrypted signal). The ground station read-out must be performed with a high gain.– OR – 2. Alternatively requires convincing the GNSS authorities to release the encrypted navigation messages with an acceptable delay and acceptable precision. 3. The end-user receiver must be able to sample the civilian L1 and the military P(Y) signal, both are sent using the L1 centre frequency, but the P(Y) is sent on a wider band. 4. The receiver sends baseband samples of the L1 signal to a reference station. These samples should be cryptographically authenticated by the receiver. 5. The reference station correlates the receivers' signal with its own estimates of the P(Y), yielding a pseudo-range. 6. The receiver also periodically sends a position and velocity message. 7. The base station checks the receiver's claimed position and velocity against the pseudo-range, as given by the receiver's own baseband samples. - *Low receiver unit cost: requires some changes in the receiver's configuration, as well as the transmission of aiding information to a ground station.* - *Medium infrastructure cost: Requires ground stations and transmission of delayed encrypted navigation message by communication satellites, -OR-* - *Low infrastructure cost, provided that the GNSS provider can be convinced to disclose the encrypted navigation message with some delay.* - *Method as put forward in [Lo09-2] has certain scalability issues.* - *Not available as of 2009, but could be implemented rapidly.* - *By itself, the method is vulnerable to GNSS meaconing and GNSS receiver-spoofers. Becomes more potent when the GNSS receiver checks for jumps in the GNSS time. This in turn requires that the GNSS receiver is powering its digital clock, and has been tracking a real GNSS signal recently (up to 3 hours ago).* - *When combined with above checks for clock jumps, can detect meaconing and spoofing, unless high-gain antennae are used to read out the satellite's signals directly.*

Measure	Characteristics *Trade-off*
"Hybrid Psiaki-Lo method": L1 signal samples are sent to user equipment, which performs P(Y) correlation	1. Requires ground stations to read out the encrypted navigation message (in GPS, the P(Y), or a Galileo encrypted signal). The ground station read-out must be performed with a high gain. – OR – 2. Alternatively requires convincing the GNSS authorities to release the encrypted navigation messages with an acceptable delay and acceptable precision. 3. Requires transmitting the encrypted navigation bits, in a digitally signed message, to a user community. 4. The end-user receiver must be able to sample the civilian L1 and some of the military P(Y) signal. Both are sent using the L1 centre frequency, but the P(Y) is sent on a wider band. 5. The receiver also obtains baseband samples of the L1 signal from a reference station. These samples should be cryptographically authenticated by the reference station. 6. The receiver correlates the reference station's samples with its own estimates of the P(Y), yielding a pseudo-range. 7. The receiver checks its claimed position and velocity against the above pseudo-range, as given by the reference station's samples. 8. The receiver periodically sends a PVT message. It includes information on mismatches between its PVT and the range calculations. - *Low receiver unit cost: requires some changes in the receiver's configuration, as well as the reception of aiding information to a ground station. Requires some additional processing power for additional correlation process.* - *Medium infrastructure cost: Requires ground stations and transmission of delayed encrypted navigation message by communication satellites, -OR-* - *Low infrastructure cost, provided that the GNSS provider can be convinced to disclose the encrypted navigation message with some delay.* - *No scalability issues.* - *Not available as of 2009, but could be implemented rapidly.* - *By itself, the method is vulnerable to GNSS meaconing and GNSS receiver-spoofers. Becomes more potent when the GNSS receiver checks for jumps in the GNSS time. This in turn requires that the GNSS receiver is powering its digital clock, and has been tracking a real GNSS signal recently (up to 3 hours ago).* - *When combined with above checks for clock jumps, can detect meaconing and spoofing, unless high-gain antennae are used to read out the satellite's signals directly.*

Table 7: GNSS receiver defence measures

With prohibitive costs, nearly all GNSS receiver defences can be overcome

If the adversary has very large resources, he can in theory build hardware that misleads nearly all civilian spoofing or meaconing defences. In other words, economics plays an important role: A successful spoof defence is achieved if the total cost to the adversary of overwhelming the defence is larger than his budget, and/or if the unitary cost for overcoming a single spoof resistant module is higher than the prospective benefit.

Expanding the budget of the adversary in an arbitrary fashion, one can arrive at a theoretical scenario in which the adversary disposes of:

- A "Sophisticated GNSS Receiver Spoofer" with stepping motors that mimic actual satellite movements.
- A Multi-frequency approach: Coverage of all available GNSS frequencies
- The direct read-out of GNSS satellites, through directive antennae or beamforming antenna array.
- A spoofer array is aligned in carrier phase with real GNSS signal. Furthermore, it permanently calculates and sends a signal that is the exact opposite to the real GNSS signal, so that the real GNSS signal is effectively cancelled ("suppressed").
- A spoofer array uses gyroscopes and accelerometers to continuously control the vectors between its receiving and emitting antennae, so that the above suppression signal works even if there is vibration and/or arbitrary rolling motion.

This scenario will be dubbed the "GNSS planetarium spoofing". Arguably, it is not very realistic to expect that a criminal enterprise build such a device for use in a kinematic environment, mainly because social engineering represents more economical way to circumvent control measures. Technological impediments to such a "planetarium spoofing" platform will be explored further below.

How does a victim GNSS receiver fare against a "GNSS planetarium spoofer"? Running through the available defences that were discussed above, here is how nearly every measure can be defeated by such an adversary:

- Checking power levels will not help, since the power level is correct.
- Verifying that the carrier phase difference matches the orbits of the satellite will not help, because it is defeated by the setup of the "planetarium".
- GNSS Public Spreading Code Authentication (PubSCA) together with an internal clock on the victim receiver may not help: When the receiver is first put into the "planetarium", the associated time jump may be less than the detection threshold. Navigation Message Authentication (NMA) is set to fail for the same reason.
- Cross-checks against inertial sensors may work, if the adversary attempts sudden movements. However, with an adversary skilled in implementing a "GNSS planetarium spoofer", one would expect movements within the tolerance range of the INS devices.
- Any cross-checks that the receiver makes against a magnetometer (digital compasses) can, at this level of sophistication, be overcome with a magnetic field generator.
- The "data bit latency defence" is useful only if the spoofer cannot read out the GNSS signals directly. If the spoofer can directly read out the GNSS signals, then there is no latency.
- The "vestigial signal defence" relies on the presence of a detectable real GNSS signal, which in the present case the spoofer cancelled by emitting the suppression signal.
- A GNSS receiver could check against LORAN signals. While Loran it is hard to spoof, it is technically feasible [Lo09]. In order to impede Loran spoofing, there are counter-measures that one can implement in a Loran receiver [Lo09]. However, the concern with LORAN is its demise in the USA [Loran10], and its limited availability, rather than its susceptibility to spoofing. Indeed, from 2011 onwards, and as a GNSS spoof defence, Loran will only be available in Europe.
- The "Navigation Message Authentication", "Spreading Code Authentication", "W-bits estimation" and other[40] cryptographic techniques could potentially stop a spoofer. But for that the receiver needs to combine this with a technique that detects sudden receiver clock jumps, as described in reference [Hein07]. Without the clock jump verification technique in the victim receiver, and with direct read-out of the satellite signals, the adversary can get past cryptographic techniques: First he introduces a clock jump into the past, corresponding to the maximum distance between real and calculated position. Then he delays or advances the GNSS signals as required, to cause the erroneous pseudo-range calculations in the victim receiver.

[40] This includes the techniques proposed by S. Lo and M. Psiaki.

A military GNSS receiver uses a cryptographic technique that relies on the calculation of the W bits. But since the military is unable to provide such devices to the civilian sector without compromising its own secrets, military receivers are outside of scope.

Anyone attempting to build a GNSS planetarium spoofer would face prohibitive development and unit costs. These costs would be dominated by several major items:

- For each planetarium spoofer unit, an orb-shaped skeleton needs to be designed built separately, containing several arcs to simulate GNSS space vehicle orbital planes, matching stepper motors, all of which is mounted on one platform that can swivel as required by the direction of the vehicle.
- On the arcs, the adversary must mount a number of "GNSS receiver-spoofer" pseudolites. If N is the number of signal wave fronts from a satellite, and M the number of satellites to emulate, then NxM pseudolites are required (unless additional research permits the use of a single emulator for N wave fronts). These emulators need to be synchronised with each other.
- The orb-shaped skeleton (with the pseudolites) must be mounted inside of a chamber. The hull of the chamber would provide protection against humidity/salinity found on e.g. the high seas.
- In case the "planetarium" uses a directive antenna array, the receiving antennae would need to be mounted on a separate robotic Steward platform that uses gyroscopic stabilisers.
- In case the "planetarium" uses a multichannel digital beamforming antenna array, the platform and the stabiliser may not be necessary. Physically speaking, in order to read out N satellites directly, one needs N+1 antenna elements, which need to be spaced apart by at least half a wavelength, which is about 10 cm for L1 and 13 cm for L2. Concerning signal processing, the adversary would need to understand and apply beamforming techniques such as MUSIC or ESPRIT in a real-time kinematic environment, implying use of gyroscope data to compensate for movement. This will result in a real-time signal processing computer. While such a device may be feasible, great care must be taken to design the system such that it does not introduce a measurable time delay in the GNSS signals. Such delays could in turn be measured by a victim receiver. To minimise delay, one needs to maximise processing power, use a real-time operating system, and perform custom development.
- As Humphreys and Ledvina plan to demonstrate [personal communication], it is possible to design a spoofer that performs carrier phase alignment with a victim antenna, in a static environment. In a dynamic environment that includes vibration, roll and pitch movements, the vector between the spoofer receiving antenna, its emitting antenna, and the victim receiver antenna must constantly be computed in real time. Otherwise the spoofer cannot suppress the original GNSS signal. This requires tight integration between each pseudolite and an INS system. While GPS/INS integration is a well-explored subject, for instance in reference [GPS-REF], such integration will add to the overall bill of the "planetarium spoofer".
- Last but not least, all inputs and outputs of the system need to be coordinated by a central unit. This unit would synchronise the signals from the pseudolites and any stepping motors of the "planetarium", and provide a user interface.

From the above, it is clear that the device by its sheer dimensions would be hard to transport and needs special mounting while in use.

While the above may constitute an interesting story line for the film industry, it is certainly unattractive in the context of illicit activities. Assuming that a criminal organisation would act in an economical and rational fashion, it would prefer other avenues[41].

In conclusion, given plenty of resources, an adversary could theoretically overcome nearly all anti-spoofing measures. Those measures that cannot be overcome may be unavailable to the civilian sector (military GPS receiver), or may rely on means that are not globally available (LoRAN). Therefore, the

[41] However, adversaries with large financial resources and large prospective gains can afford quite considerable investments, which would make excellent story lines for the film industry. South American drug cartels repeatedly prove this point by their technical aptitude in the development of submarines.

challenge is not so much to design a perfect spoof defence. Rather the challenge is identify relatively simple defences that require relatively large effort to be overcome, where the cost to the adversary is measured both "per resistant receiver" (unitary cost) as well as for the adversary's activities as a whole (total cost).

Explore the possibilities of using Galileo encrypted GNSS signals

With the NMA and PubSCA authentication architectures explored above, this chapter looks at planned Galileo activities. Galileo holds the promise of addressing GNSS spoofing, by using authenticated signals. The Galileo brochure available at ESA states that "certified Galileo services will allow authorities to confirm that fishing vessels operate only in designated areas."

However, such services are yet to be implemented, which would happen in a second phase, after initial go-live in 2013 [ESA09].

In the following subchapters, various Galileo services will be explored, in light of their potential concerning Galileo signal spoof prevention.

Using the Open Service (OS)

Authentication based on security features implemented in the Open Service was and is the subject of various debates. In 2007, G. Hein [Hein07, pg 77] stated that the Open Service civilian Galileo signal would not contain any authentication or encryption.

There have been more recent discussions regarding signal authentication on the Open Service, which have led to the "Galileo Advanced Concepts" (GAC, a FP6 project) introducing Contract Change Notice 4 (CCN4). The GAC-CCN4 work finished in September 2009. The GAC-CCN4 explored an authenticated Open Service for road toll charges, and assessed the potential of both "navigation message authentication" (NMA) and "spreading code authentication" (SCA). While NMA can be added with little changes, the introduction of SCA would require updates to the Galileo satellites themselves [ESA09]. Consensus is therefore building in the direction that such a feature will not be made available in the current development of Galileo, but could be made available in the "next generation of Galileo" [ESA09].

It has to be noted that the Galileo Supervising Authority (GSA) is also aware that an authenticated Open Service would have various uses beyond road tolling, such as [Warfield08]:

- *Multimodal transportation: authenticating the handover of goods and/or containers*
- *Authenticated timestamp: authenticate timestamp based on Galileo timing signal, for use in ecommerce, electronic transactions*
- *Monitoring resting times for truck drivers*
- *Person tracking: to keep away stalkers with a restraining order, tracking offenders on parole* and, last but not least
- *Maritime and fisheries: to prove where you were and that you did not fish in someone else's territorial waters*

The Galileo Open Service uses both the L1 frequency of 1.575 MHz and additional frequencies, which are at 1.176 (a.k.a. "E5a") and 1.207 GHz (a.k.a. "E5b") [Gal02]. The Open Service features both data and pilot channels, which are transmitted at all of the above frequencies with a 90° phase difference. Moreover the signals transmitted on E5a and E5b have a 10.230 Mcps chip rate, the same rate as the modernised GPS L2C signal. GNSS receivers could be required use and decode all of these data and

pilot channels as well, which is an implementation of the "multi-frequency defence. This would increase the difficulty of spoofing the GNSS tracker: A spoofer would need to use one transmitter board per spoofed frequency band. Alternatively, the adversary could use several spoofing devices, which would be driven by one single oscillator signal, and hence would operate in lockstep. However, the GNSS tracker would also become more expensive: multi-frequency GPS receivers cost much more than their single-frequency counterparts, because these are typically used in niche products. As an indication, dual frequency GPS boards usually start at 1000 EUR apiece.

In summary, once the Galileo Open Service reaches Full Operational Capability, it can be used to render GNSS receivers more resistant against fake GNSS signals. This is because the Open Service broadcasts additional GNSS signals (on E5a and E5b), which some receivers would then interpret. The use of such triple-frequency receivers (L1, E5a and E5b) would then force the adversary to simulate those signals as well. Once the Open Service integrates an authenticated signal structure, it may be regarded as a method of choice to defend against fake GNSS signals.

Using the Commercial Service (CS)

For authentication purposes, it appears that the commercial service would be using signal encryption[42] [ESA-GAL]. Each Commercial Service provider would use a separate key, and there would be many such providers. The keys would be updated over a secured channel. The Commercial Service was not fully specified at time of writing.

Mechanisms as the one outlined above can be used within a trusted user community, which has the financial means to contribute to the Galileo undertaking. Some potential users of the Commercial Service are geodesy institutes, electricity distribution networks, and time stamping for financial services.

Broadly, there would be two ways to use an encrypted signal in order to harden a GNSS tracker. Either one is a full member of a Commercial Service user community (such as in the Galileo CS, a geodesy institute), or one infers some information from an opaque signal without having the keys (such as at present in GPS, the "code-less" "z-tracking" enabled GPS receivers).

Using the Commercial Service through paid membership

The GNSS applications considered in this document are, by their very design, exposed to hackers. This implies that special considerations need to be addressed when implementing a hardened GNSS tracker as a full user of the Commercial Service.

If a GNSS tracker would be based on full use of the Commercial Service, then a "GNSS service provider" would use some symmetric key(s) to generate an encrypted GNSS navigation signal, to be used by the trackers. Then the resulting system has a flaw described as "break once, break everywhere" (BOBE). Without going into the details on how such flaws can be exploited, it is important to state that such systems require additional security[43]. If such a system existed and was fully broken by a successful adversary, then it cannot be fixed by rolling out new keys, since the same successful adversary could simply read them out. All affected GNSS trackers would then need to be replaced.

[42] If the Commercial Service were also to send an authenticated signal, then that would indicate a departure from present-day policy.

[43] If that is not possible, proper protection may be expensive. In the Pay TV industry, there is a niche market of security vendors, whose reason of existence is to keep up with the hacker community. Content provides are in turn forced to buy the services of these vendors, or to provide content for free.

The Galileo Commercial Service uses an additional frequency at 1.278 GHz [Gal02]. GNSS tracker receivers could be required use and decode these frequencies as well. This could lead to the development of a multi-frequency receiver that also tracks the Galileo Commercial Signal. Because of its BOBE vulnerability, such a receiver would be only marginally more secure than a multi-frequency Open Signal receiver. A multi-frequency Commercial Service receiver would likely be more expensive than its multi-frequency Open Service equivalent, because the former would be build for a smaller niche market than the latter.

Using the Commercial Service by code-less tracking

Consider the possibility to use a receiver that listens to the Galileo Commercial Service, but does not have the keys required to decode the actual signals. Then one can regard the entire encrypted CS signal as being "opaque", just like the military GPS signal, and apply the "Hybrid Psiaki-Lo method" to obtain authenticated range information. In turn, that implies that occasionally, the receiver obtains additional signal samples from a known reference station, through a satellite downlink.

Note that if the "Hybrid Psiaki-Lo method" would be authorised by the GPS or Galileo authorities, on top of their respective military or PRS signals, then that would obviate the entire concept of using encrypted CS signal mentioned above. What is written below therefore refers to a scenario in which the GPS or Galileo authorities have issues with the use of the military or PRS signals.

GNSS receivers could be strengthened by using the "Hybrid Psiaki-Lo method" on top of the Commercial Service's encrypted signal. The adversary would then be forced to obtain a very large percentage of the CS keys, in order to be able to forge a believable encrypted CS signal, which would then be used to spoof such a receiver. However, the encrypted Galileo CS signal is supposed to generate revenue for Galileo. The adversary would therefore become a paying user in an official registry, or be forced to gather the available keys illegally.

It serves to observe that in the best of scenarios, this leads to a solution similar to the "Hybrid Psiaki-Lo method", but without any sensitivity issues. On the downside, there is the risk that the Galileo Commercial Service providers would be forced into an "arms race" against adversaries illegally attempting to use the Commercial Service, but that risk was already inherent in the Commercial Service.

In summary, once the Galileo Commercial Service reaches Full Operational Capability, it can be used in a specific scenario to render GNSS receivers more resistant against GNSS spoofing. This scenario arises if both the GPS and Galileo authorities forbid the "Hybrid Psiaki-Lo method" on top of their respective military and PRS signals. Then the Commercial Service's encrypted signal could become a basic ingredient in a hardened GNSS tracker.

Using the Public Regulated Service (PRS)

The Galileo equivalent of the Military GPS, the "Public Regulated Service", according to reference [Pacific07], will have "*encrypted ranging codes*". However, the applications are designed to counteract "*several low-power jammers, or a single high-power jammer placed in a strategic location [or] a hostile use of open GNSS services*". Sectors are limited to "*defence, law enforcement, internal security, customs, emergency services, critical transport, transport of hazardous goods, energy, telecom, and strategic economic/commercial activities"*. It is therefore sensible not to deploy such Galileo PRS receivers in applications where collaboration is not ensured, such as in road tachographs and maritime fisheries enforcement.

The Galileo PRS uses an additional frequency at 1.278 GHz, and in addition uses "side lobes" of the GPS L1 band, located near 1.563 and 1.591 GHz [Gal02]. GNSS tracker receivers could be required use and decode these frequencies as well. But that would not help beyond the techniques described for the Open Service. This is because without the PRS keys, the receiver cannot interpret the contents of the signal at 1.278 GHz and at the lobes at 1.563 and 1.591 GHz.

However, an interesting application of the Galileo PRS is its inclusion in a cryptographic defence in the civilian sector, as explored further above in the "Psiaki method", "Lo method", and "Hybrid Psiaki-Lo method" subchapters.

Using other Galileo Services

There are two other Galileo Services, the Safety of Life (SoL) and the Search and Rescue (SAR) services. Concerning navigation, they do not add any additional precision.

On these other Galileo services, the ESA web site states [ESA-GAL]:

- *The Safety-of-Life Service (SoL) improves the open service performance through the provision of timely warnings to the user when it fails to meet certain margins of accuracy (integrity). […] The Safety-of-Life Service will be certified and its performances will be obtained by using certified dual frequency receivers. Under such conditions, the future Galileo Operating Company will guarantee SoL.*
- *Galileo will introduce new SAR functions such as the return link (from the SAR operator to the distress beacon), thereby facilitating the rescue operations and helping to reduce the rate of false alerts. The service is being defined in cooperation with COSPAS-SARSAT, and its characteristics and operations are regulated under the auspices of IMO and ICAO.*

The Galileo Safety of Life Service uses an additional frequency at 1.207 GHz [Gal02], which is also used by the Open Service.

The use of the SAR uplink for enforcement purposes raises ethical concerns, since any additional overhead placed on the SAR uplink may adversely impact humanitarian operations.

Conclusion on the use of Galileo encrypted signals

Concluding this chapter, it can be stated that Galileo with Navigation Message Authentication on the Open Service would be a method of choice for a hardened GNSS receiver. If that is not possible, then the "Psiaki method", "Lo method", or "hybrid Psiaki-Lo method" could be applied to the Galileo PRS signal. If in turn, the Galileo authorities forbid this sort of use of the PRS, then the encrypted signal of the Galileo Commercial Service offers an alternative.

Possibilities in using GPS signal simulators

Note: This chapter may read like a "shopping list" for an adversary. However, the aim of this chapter is to ascertain that the options of the hacker are limited, and if not, to propose additional measures.

What would an adversary be interested in? Knowing that

- The GPS receivers on board of VMS boxes are only using the GPS L1 signal,

- The GPS receivers require signals ("channels") from at least 4 satellites to calculate a "position, velocity, and time" solution,
- The adversary must use a solution that permits custom scenarios, and would much prefer a GUI to plot a fake course, cutting down on preparation and planning.

Therefore an adversary would aim to replicate multiple channels in the L1 band. He would then attempt to acquire a relatively cheap but suitable unit that satisfies the above.

A number of manufacturers were asked to submit pricing information for such products. Some of them replied with specifications and prices of GPS test beds that could be used.

Device Manufacturer	Cost	Features
Multichannel GPS/WAAS Signal Simulator (GPSS-CA-002-04) Center for remote sensing, www.cfrsi.com	55 000 USD	Simulates signals for up to 12 GPS satellites. Able to generate civilian navigation signals at L1, L2C, and WAAS. User can specify a scenario for the receiver's PVT solution, and can plot trajectories with a graphical user interface. No Galileo, but a Galileo product is available separately.
STR 4500 APM Instruments, www.apminstruments.it	45 000 EUR + software	Simulates signals for up to 12 GPS satellites. Signals in the L1 band only. If the user buys the "Simplex 45" software, he can specify scenarios for the receiver's PVT solution; for that he needs a string of NMEA 183 messages. No Galileo; a Galileo enabled product is available at about twice the price.
NS 600 Multi-Channel GPS Simulation OLinkStar, www.olinkstar.com	18 750 USD	Can generate composite GPS L1 C/A code with 12 channels Simulates any movement of GPS receiver at configurable location and time User cannot without further assistance specify a scenario for the receiver's PVT solution. However the device offers a graphical user interface. No Galileo available.
NI GPS Simulator National Instruments, www.ni.com	34 464 EUR	Simulates signals for up to 12 GPS satellites. Signals in the L1 band only. No Galileo.
GPS Spoofer Cornell University	**NOT FOR SALE, DIS-SEMINATION FORBIDDEN**	Simulates and spoofs GPS signals as needed. Signals in the L1 band only. No Galileo yet.

Some manufacturers propose more expensive devices with additional features. These were deliberately left out from the above table, because they would have added no further information.

Note: Other manufacturers presented simpler simulators, which are able to simulate only one satellite. The cheapest such unit was the “Gnss Signal Simulator Model 49003” from Chroma ATE, www.chroma.com.tw, at 7000 USD, but each such unit can only be used to simulate a single satellite. If a constellation of 4 of those devices would be used to spoof a GPS receiver, then that would still require further software development. Such development would aim to transform a simulated position into a concerted set of navigation messages coming from the constellation.

Chapter II: Physical Security

Authors: U. Kröner, F. Littmann, C. Bergonzi

Protecting small volumes against physical intrusion

Why physical security in the context of GNSS trackers is important

Concerning the special case of fisheries enforcement, in the introduction, the regulatory environment and the economics of illicit fishing activities were discussed. Even though some effort has already been made in the direction of building more secure VMS devices, the conclusion holds that the security of VMS devices may yet need to be strengthened.

As an example, the case of the Thrane 3026 "mini-C" warrants closer examination. With this widely used model, Thrane introduced the following design:

- A single unit which houses all parts,
- The circuit boards between the GPS and the main processing unit are tightly integrated,
- Communication between the main processing board and the Inmarsat terminal is encrypted,
- The main board is shielded against RF analysis.

However, the mini-C can still be opened, analysed, altered, and re-assembled. It will in principle continue to work under such circumstances. In addition, the model 3026 can receive configuration messages, both via the local interface and via RF signals, which can alter or disable some of its functionality [TT1] [TT2]. So although the model 3026 is more secure than the older model 3022, it is only tamper resistant to some degree, and has no tamper-evident seal.

As seen in the introduction, there are other developments that also validate the thesis that VMS device security is important: Ireland and Cyprus are renewing their installed base of VMS devices, and in each case, physical sealing with tamper-evident seals is being performed. The VMS device to be rolled out in Ireland has additional security features. The VMS device introduced in Cyprus does not require a power supply, which increases reliability.

Generally, with regards to attacking an electronic device, any successful physical breach fundamentally compromises its security: It should be clear that once an adversary can perform reverse engineering, the security of the device is fundamentally broken. For instance, reverse engineering enables the production of circuits that impair proper functioning of electronic devices, and the manufacture of unlicensed/unapproved duplicates, so-called "clones" or "ghost devices".

In order to mitigate reverse engineering risks, one could identify and secure a sufficient set of individual components of the tracker. This leads to the concepts behind the "Trusted Platform Module": *a computer chip (microcontroller) that can securely store artefacts used to authenticate the platform* [TPM]. If one would secure individual components, then one might as well take advantage of "Trusted Computing", as defined by industry.

A different approach is to secure the tracker to be protected using a volume enclosing seal. This is the approach that was chosen in previous work [Kroen09]. This path may lead to more cost-effective solutions, if indeed it is less expensive to secure the overall tracker than every single component.

In addition to sealing a volume, one needs to seal the tracker to the asset. Otherwise, it may be too easy to remove the tracker, and to put it temporarily at some other premise.

What is wrong with existing seals?

Currently used seals, even high-technology versions, are relatively easy to defeat. Physical security expert R. G. Johnston states that [Johnston06]:

> *We intensively studied 244 different seal designs […] half are used for what can be considered critical, high-security applications […] In our tests, the average attack time on each seal was 1.4 minutes, with a median value of only 43 seconds.*
> *[…]*
> *To make better seals, one has to understand why existing seals are so easy to defeat. Their Achilles heel appears to be not so much detecting unauthorized access, as securely storing the alarm condition, or the fact that trespassing has been detected, until such time as the seal can be inspected. With current seals (even electronic ones), it is simply too easy for an adversary to hide or erase the alarm condition, or to replace the seal with a fresh counterfeit that shows no evidence of tampering.*
>
> *One way to deal with this weakness is to invert the problem: At the start, when we first install a seal, we store information in or on it that unauthorized access has not yet been detected. We call this information the anti-evidence. These devices can be mechanical, but more often are electronic. Once the seal detects trespassing, it instantly erases its anti-evidence. At inspection time, the absence of the anti-evidence indicates tampering has occurred.*

Passive and active monitoring seals

For the purpose of this document, a "passive seal" is defined as one that needs to be checked by an inspector. In other words, the GNSS tracker does not actively report on the status of the seal.

The reference [Kroen09] also mentions "active monitoring", which was defined as: *An active system implies an extension of functionality so that attempts at tampering would be detected remotely and reported to authorities in a timely manner.*

Active monitoring could be applied for securing a tracker's housing (volume enclosing seals). It could also be applied to ensure that the tracker stays mounted in place (mounting seals).

Examining various volume enclosing seals, there is one "proven design" that would make it possible for the tracker to notice tampering and to react with an automated "tamper response". A "proven design" would be defined as one that is used in e.g. banking and defence domains.

What are the requirements for banking and defence applications? If one desires to devise a physical volume enclosing seal that is useable in the banking or defence sector, one needs to be approved by the following organisations at the following levels:

- US "National Institute of Standards and Technology" (NIST) "Federal Information Processing Standard" (FIPS) 140-2 on "Security Requirements for Cryptography Modules", level 3 or 4. [FIPS140-2]
- US National Security Agency (NSA) type 1.
- UK "Government Communications Headquarters" (GCHQ) "Communications-Electronics Security Group" (CESG), "Baseline, Enhanced, and High Grade Cryptographic Products"
- German "Zentraler Kreditausschuss" (ZKA)

The above requirements constitute a (highly) sufficient level of certification. At first glance they would look "over-engineered" for tracker applications. However, what matters is to explore the economics of fitting the GNSS trackers with seals certified at that level, versus the profitability of certain types of exploits performed by the adversary.

Below several seal designs are presented. Some are purely passive. Others can be viewed as either passive or active. Integrated into a passive monitoring design, they require periodic inspections. As components of an active monitoring system, they require counterparts in the GNSS tracker that read out values from the seal, and transmit the status to a control centre.

Defending against tracker removal

GNSS based trackers are usually fitted onto the asset to be protected using simple bolts. This opens up a possible exploit, the removal and intermediary storage of the tracker at another location. For instance, an adversary could remove the tracker from the asset, place it elsewhere (together with a battery pack, if needed), perform illicit activities, and return later to pick up the tracker. This "removal vulnerability" requires further consideration; otherwise it weakens overall security.

Taking the domain of fisheries monitoring as an example, the VMS devices tend to be mounted on fishing vessels in two distinct ways, illustrated by the pictures below:

Figure 23: VMS device mounted on a railing (processed image).

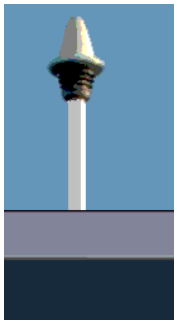


Figure 24: VMS device mounted on a tube above the ship's cabin, after a photo by M. Börje.

Some of the solutions proposed below can be used to defend against the adversary attempting to remove a tracker.

The RFID bolt seal

During previous work, colleagues at of the Seal and Identification Lab pioneered the use of an Allen standard bolt, fitted with a glass capsule containing a Radio-Frequency IDentification (RFID) chip.

As one can see in Figure 25, such RFID devices have an appropriate form factor.

Figure 25: RFID chip on top of a handheld device; top right shows parts of a Thrane model 3026, with an Allen standard bolt.

In order to retro-fit the RFID chip into above VMS device, the installer would remove one of the screws, *re-drill its hole, and then to tap a new thread adequate for an M6 (6mm) bolt [..] Once this is accomplished, a hole 2.2mm in diameter is drilled through both the PVC base and the M6 bolt in a single operation. The RFID is slipped into the hole so that it is lodged within the hole crossing the length of the bolt. Because the RFID is longer than the bolt is wide, it overflows into the mass of the base of the unit. It is then held permanently in place by an injection of epoxy that plugs the newly drilled, perpendicular hole.* [Kroen09].

Once installed, the RFID chip will be located inside of the VMS device, in a manner depicted in Figure 26.

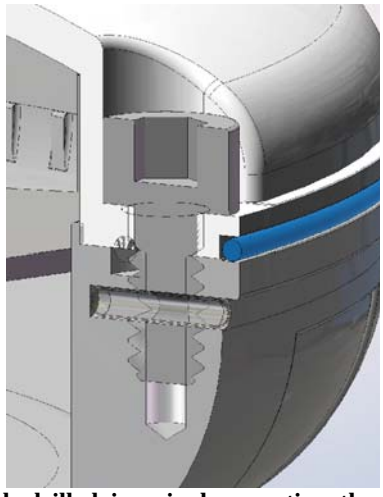


Figure 26: An RFID is inserted into a hole drilled, in a single operation, through the base and a closing bolt of the unit, constituting an RFID seal.

The thus obtained "RFID seal" is a physical seal, which uses RFID to communicate its status. Communication takes place in the presence of an RFID reader. In the above set-up, the RFID chip would be powered by the reader[44].

This kind of approach has the following advantages:

- It solves both the problem of sealing the tracker's volume, and of sealing the tracker to the asset.
- It can be retro-fitted to many trackers that have already been deployed.
- It costs a few Euros per tracker.

The approach has a number of weaknesses, some of which are:

- It is passive and requires regular inspections with specialised RFID reader equipment.
- RFID *security is a very active research field for a few years, with more than 200 scientific papers published since 2002* [RFID-WP]. One of the areas of research is "tag cloning", where the adversary reads out the chip's unique identifier, and then loads the ID into a programmable tag [RFID08]. This can be mitigated e.g. by using a "challenge-response" mechanism, implying the use of cryptographically-enabled RFID chips. However, c*ryptographically-enabled tags typically have dramatically higher cost and power requirements than simpler equivalents, and as a result, deployment of these tags is much more limited* [RFID-WP].
- Hence the use of an "RFID bolt seal" would benefit from the introduction of "challenge-response" mechanisms with cryptographically-enabled RFID chips. These also need to be defended against side-channel attacks. Otherwise the adversary can abuse programmable RFID tags to create a RFID clone.
- Given the seal operates under the presence of vibration, care must be taken in how it is installed, otherwise the RFID capsule may break under normal operating conditions.
- Inspectors could pretend that they have inspected the seal, instead of actually inspecting it. A mitigation technique is to use cryptographically enabled RFIDs that make use of a challenge-response protocol.
- Finally, if many users would agree to sabotage their tracker's seal from time to time, they would create a high number of tampering alarms, which in turn would psychologically de-sensitize the authorities to tampering.

[44] In RFID terminology, when a battery-less RFID tag is powered by an RFID reader, this is called a "passive RFID tag". This "passive RFID tag" is not the same as a "passive seal". The latter, in the context of this document, refers to a seal that must be inspected manually. An "active RFID tag" can be part of a "passive seal": the RFID tag has a battery attached to it, yet the seal itself requires manual inspection. Conversely a "passive RFID tag" can be part of an "active seal": The RFID tag is located near an RFID reader, and that reader then forwards its result to e.g. a satellite uplink.

In summary, as a measure for tracker volume sealing and removal protection, an RFID seal remains interesting as long as the RFID chip includes a cryptographic key and a challenge-response mechanism.

Passive and active implementations: By integrating an RFID reader into the GNSS tracker, one obtains an active monitoring seal.

Protection against opening and removal: This method protects both against opening of the tracker and the removal of the tracker from the target asset.

Comment: RFID chip should include a cryptographic key and a challenge-response mechanism.

An industry standard active volume enclosing seal

The (to the authors' knowledge) sole OEM vendor of volume enclosing seals for banking and defence applications seems to be the Scotland branch of the company W. L. Gore. The company sells the "GORE™ Anti-Tamper Physical Security for Electronic Hardware" products [Gore09]. The Secure Encapsulated Module is used e.g. in IBM's cryptographic coprocessors such as IBM 4758 model 002. Some of its products can be used for physical shielding that fits present purposes. The products are each composed of a sensor film, linked to a voltage difference detection circuit. The circuit is wrapped inside of the sensor film. Two similar products, the "Secure Encapsulated Module" and the "Secure Plug-on Module", are useable for present purposes of volume sealing. Both use the same type of material with different levels of security.

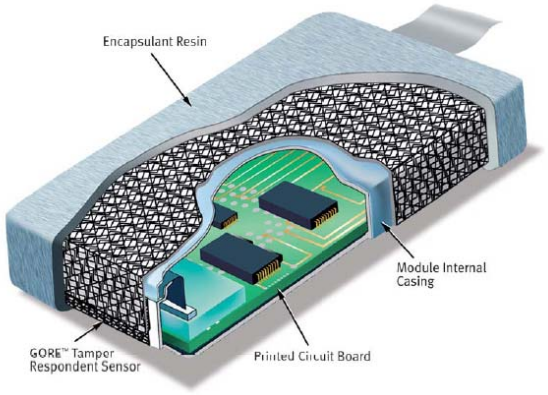


Figure 27: Secure Encapsulated Module™, a W. L. Gore security product, Courtesy of W. L. Gore.

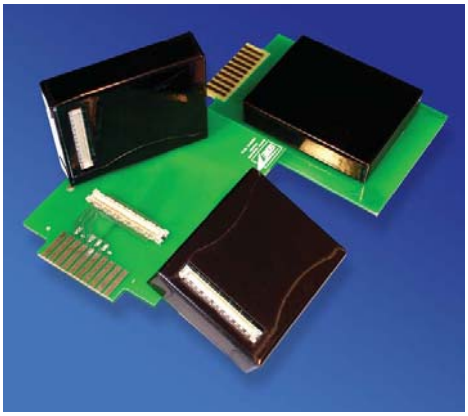


Figure 28: Secure Plug-on Module™, W. L. Gore a security product, Courtesy of W. L. Gore.

Both products make use of *a matrix of conductive ink tracks to shield encryption codes and other sensitive data. If the enclosure is tampered with, critical data is automatically "zeroed out," rendering the data unusable. Systems are available with multi-level protection against physical intrusion such as puncture, chemical attacks, and laser penetration.* [Gore09].

The Secure Encapsulated Module can be encased in a special polyurethane-based elastomeric resin. This encasing adds a further hurdle to the adversary, as it is mechanically difficult to remove the resin without accidentally breaking the underlying volume enclosing sensor.

Concerning the price tag, a secure encapsulated module test bed with 10 sensors, module casing, and polyurethane resin, is available for about 1100 GBP. The sensors used in the proposed test bed are somewhat smaller than the ones that would be needed for wrapping a dual-antenna GPS receiver. However the volume discounts for larger quantities would partially compensate for the more expensive larger sensors. One can therefore estimate the additional manufacturing cost of the volume enclosing seal to be 150-200 GBP per tracker. Depending on the tracker's application domain, the costs for such a seal could be justified.

This approach has the following advantages:

- The security level is that of high-security devices encountered in the banking or defence sectors. This known design facilitates the approval for various security levels (e.g. FIPS 140-2 level 3 and above, and other standards).
- It is based on "anti-evidence" [Johnston06], as keys are erased on tampering.
- Active monitoring immediately alerts the authorities.
- If one adds a backup battery, the resulting tracker is difficult to power down.
- If one adds a temperature sensor, the tracker can also be defended against a nitrogen bath.
- Because of active monitoring, because the sensor is placed inside of a box that is not usually opened, and because it is a proven concept used by the defence sector, the users cannot just collude in order to sabotage their trackers' seals, as with the other options. If they would proceed in that way, all of them could be sanctioned.

The approach has two matters that need to be addressed:

- It does not by itself solve the problem on how to seal the tracker to the asset. However, the sensor can become part of a bolt-type seal, designed in such a way that the opening of the bolt would cause disruption of the ink layers, enabling detection by the voltage comparator circuit.
- The seal is designed not to be opened during inspections. Hence end-of-life inspections would be a useful complement to the present hardening method.

Active implementation: Once a breach is detected, the tracker would report the breach.

Protection against opening and removal: This method primarily addresses volume protection, but can be integrated into a bolt-type seal. The latter can become part of a removal defence.

"Mechanical anti-evidence seals"

"Mechanical anti-evidence seals", is a term used at the Vulnerability Assessment Team of Argonne, IL. These seals mechanical devices, which have features that can be used as "memory" and are difficult for the adversary to re-construct.

The following sealing concept relies on a mechanical anti-evidence sealing device [R. Johnston, personal communication], which is verified by comparing photographs taken at mounting and at inspection time.

The process can be summarised as follows:

1. The tracker is secured to the vessel using conventional bolts. Furthermore,

   The bolt heads, which should be in the field of view of the camera, can be covered in clear epoxy with sparkling particles to detect if they have been removed and then reapplied. A thin […] plastic sheet could protect the [epoxy] from damage, UV aging, [and salt].

   The plastic sheet would be replaced on inspection. Alternatively an UV resistant rubber sleeve can be deployed over the coated bolt heads. (The bolt heads would be laid bare before taking the pictures as described below.)
2. Select an arbitrary point on the vessel for taking a photograph of the tracker. Determine a 3D vector and zoom angle for shooting the picture. (This could be achieved by soldering a standard camera thread mount onto a point on the asset, and engraving a line that designates camera orientation.)
3. Take one photograph after initial set-up. Store picture in a safe location at the control centre.
4. Take one photograph at each inspection. One may want to ask the inspector to take a newspaper and put it next to the tracker, in order to ascertain the date of the picture, so as to be certain that the picture was actually taken ("anti-gundecking" feature).
5. At the control centre, perform before-and-after "blink comparison" as described by R. Johnston.

The "blink comparator" relies on the ability of the brain to detect motion. Two images containing the same items, supposedly found at the same location, are shown rapidly in succession. Any item that has moved between the two images will be immediately detected as "motion"[45].

This method is both inexpensive and provides tamper resistance:

- As the sparkly flakes provide a unique and random pattern, breaking the epoxy, removing the tracker, re-installing the tracker, and re-applying the epoxy will yield a different pattern, hence this will be easy to detect.
- An adversary could remove the tracker together with parts of the asset's hull, perform illicit activities, and finally reassemble the hull. However this would require a soldering job with millimetre precision in position, as well as high angular precision. This soldering job may need

[45] This "blink comparison" technique was pioneered by astronomers when looking for planets, leading to the discovery of Pluto in 1930.

to be performed while the asset is still in motion. This is the case for most maritime application domains.

Another possible work-around for technologically skilled adversaries is to compose a fake image, using image processing tools such as “Photo Shop”. However, the inspector would be taking the photographs, hence this would require a colluding inspector. This in turn is a “social engineering” vulnerability, something that is outside of the scope of this document.

The aforementioned defence is not limited to passive removal detection, but can also be used as a passive volume protection seal. For that, one would apply the epoxy as a plug around a sealing hole.

Depending on the geometry of the GNSS tracker, this defence can be used for retro-fitting purposes.

Passive implementation: This method is devised as a passive monitoring seal.

Protection against opening and removal: This method protects both against opening of the tracker and the removal of the tracker from the target asset.

The “time trap” seal

Another sealing mechanism was recommended by expert R. G. Johnston for the present study [R. G. Johnston, personal communication]. A “time trap” device [Johnston06] would be placed inside the GNSS tracker.

The “time trap” device combines a microprocessor, a battery, a numeric key, a clock, a visible LCD screen, and a sensor used to detect intrusion. The device is first set up by loading it with a key, and the LCD screen is off. It is then armed and placed in the volume it protects, before closing the volume.

> *If the seal determines that the container has been opened (it doesn’t care whether by good guys or bad guys), it turns on its liquid crystal display. The screen then shows the time that entry took place and a […] code generated using the key, [the time], and a […] hash function. [..] The seal’s microchip erases [any] key in microseconds when the seal first detects intrusion.* [Johnston06]

Information leakage is always a concern, since it can be used by the adversary to build a cloned device. Hence the “time trap” minimises information leakage. Before the physical opening of the volume, the “time trap” will not release any information: The screen is switched off, and is inside of the volume. Once the volume is opened, it will display just one particular piece of information.

Figure 29: ***Time Trap display after the container is opened. The time that the container was opened […] is permanently displayed, along with the two letter hash (“RF” in this case) corresponding to that time*** **[Johnston06-2]. Courtesy of the Vulnerability Assessment Team, Argonne, USA.**

The “time trap” could be integrated into a GNSS tracker as follows. At inspection time, the inspector opens the GNSS tracker (which when using this design, should permit relatively easy opening, using e.g. a mechanically secured clasp or buckle), and reads out the time and hash code on the LCD of the “time trap”. The inspector then rearms the trap by pushing a button and noting the key, which is a value encoded on two bytes.

A basic “time trap” also *monitors the battery voltage and seal temperature to detect attacks on the seal* [Johnston06-2], for detecting extreme cold or heat attacks. Extreme cold both disables a seal, and maximises the “data remanence” phenomenon.

At additional cost, the Time Trap […] can monitor […] additional sensors simultaneously. When multiple sensors are used, they are polled in a random […] order so that an adversary cannot predict when a given sensor will be read by the seal. Sensors that can be used with the Time Trap include commercial, solid-state Hall Effect magnetic sensors, color sensors (as with the Tie-Dye Seal) [see

below], *tilt detectors and accelerometers (for stationary assets), ultrasonic motion detectors, and passive infrared detectors.* [Johnston06]

Volumetric alarms such as those used in the "time trap" require at least one sensor. This could be a light sensor, or could consist of two motion detectors. In the latter case two identical detectors are placed on the inside of each part of the cover. An additional circuit then takes the difference between the outputs of each sensor, and feeds the result into the microprocessor that reacts to tampering.

Some trackers operate without an external power supply, yet some detectors also use energy when in use. Under such a scenario, a trade-off is required between higher device autonomy and device security. The trade-off is that more frequent sensor checks make the device more secure, but diminish the battery life expectancy, implying more frequent inspections and replacements of the seal. The proper trade-off depends on the adversary's expected profile. For defending against a hobbyist adversary, sensing can occur each second. But organised crime and state sponsored adversaries could very well have access to equipment in the 100 000 EUR range, able to perform attacks that take on the order of a millisecond[46]. If that is the case, one could adapt the sensing frequency towards the same order of magnitude. One could also encase the electronics of the device in solid epoxy resin potting material.

In addition, if the tracker could meet a skilled adversary, one needs to deal with the issue of data remanence. The tracker should then make sure that it can wipe a sufficiently large portion of the key, even in the absence of the principal power source, by using energy from a smaller battery or from a capacitor. Wiping would include specific algorithms to deal with "data remanence", not unlike those implemented in e.g. the UNIX `shred` command.

In his original design, Johnston proceeds as follows:

> *Once the seal detects that the container has been opened (by either the good guys or the bad guys), it immediately erases both the secret key (in a few µsecs) used by the hash algorithm and parts of the hash algorithm itself (in a few msecs).* [Johnston06-2]

Conceptually, the "time trap" design has the following advantages:

- The adversary cannot break, and subsequently forge a seal ("anti-evidence", [Johnston06])
- The inspector must inspect the "time trap", meaning that he cannot just pretend that he did inspect the seal ("anti-gundecking", [Johnston06]).
- Anyone can open the tracker, but the inspector will know if he was the first to open it, and this evidence has forensic qualities.
- The hardware costs a few Euros.

The approach must however carefully evaluate the following:

- The design does not by itself solve the problem of how to seal the GNSS device to the asset. However, the "time trap" concept can be re-used in settings that permit the use of bolts.
- The issue of "data remanence" may need to be solved even against a skilled adversary. Johnston uses byte-sized keys and degenerate hash algorithms [Johnston07], both of which can be wiped quickly. Small keys and degenerate hashes put the onus on the choice of algorithm, which may have far-reaching implications in terms of cryptanalysis.
- Finally, if many users would agree to sabotage their GNSS trackers from time to time, they would create a high number of tampering alarms, which could psychologically de-sensitize the authorities to tampering.

[46] R. Johnston et al write [Johnston07, footnote 24]: *Attacks on electronic circuits of less than a millisecond are quite possible based on making foreign connections to the electronic circuitry, disconnecting the power supply, or even by shooting the equivalent of a bullet through the correct location to break an electronic connection or a security bit.*

Protection against opening and removal: As initially conceived, the time trap seal protects only against opening of the tracker. However, if properly miniaturised, the design can be put into a bolt.

Passive and active implementations: As initially conceived, the time trap seal is a passive device requiring manual inspection. However, by wiring the time trap circuit into the GNSS tracker, one obtains an active monitoring seal. In case of a bolt design, the tracker would need to perform active monitoring through Near Field Communications.

Comment: The time trap seal requires a separate sensor circuit, such as a motion sensor, light detector, or colour detector, which detects that a breach occurred.

The "tie dye" seal

The "tie dye" seal [Johnston06-2], could also be used to seal the GNSS tracker to the asset. Moreover, it can also guard against opening of the tracker. It requires an additional reading device to check the status of the seal. In its original form, it is presented as a bolt-type seal:

> *Recently, small, inexpensive solid-state color sensors with remarkable color resolution have become commercially available. These perform precise color measurements that were previously available only with expensive colorimeters or spectrophotometers. For example, the TAOS TCS230 color sensor outputs RGB color values.*
>
> *The color sensor is placed inside the hollow body of the seal and rigidly mounted. A white LED is used to provide illumination inside the seal. This does not need to run continuously, but can instead be turned on a random, unpredictable times so that a color spectrum can be measured intermittently (thus extending battery life).*
>
> *The inside of the seal is painted with a complex varying color pattern, not unlike the "tie-dye" T-shirts popular in the 1960's. Because this interior color pattern is so complex, it is difficult for an adversary to counterfeit it in order to try to defeat the seal.*
>
> *At startup time, the reader chooses the […] anti-evidence, then communicates them to the seal. For inspecting the seal, the reader is again plugged into the seal.*

Figure 30: ***Tie-Dye Bolt Seal (right) and its reader (left). The two halves of the bolt seal snap together through a hasp. The seal can be opened by hand without tools simply by pulling the two halves apart.*** **[Johnston06-2]. Courtesy of the Vulnerability Assessment Team, Argonne, USA.**

Conceptually, the TAOS TCS230 colour sensor (and its recent in-pin replacements, the TCS3200 and TCS3210) is an active light detector. Unlike conventional light detectors, it does not detect the absence or presence of light, but emits its own light at various frequencies, then checks the properties (spectral intensity) of the reflected light. Unlike classical light sensors, such colour sensors cannot be defeated by opening the seal in complete darkness.

If the colour sensor detects an opening at any given time, the "tie dye" seal erases a secret numerical key inside of its electronics. As with other anti-evidence seals, the deletion of the key indicates that unauthorised access took place. As the secret key is very short, its erasure can proceed in a few microseconds.

The above section on the "time trap" discusses the frequency at which the sensor operates, versus battery life expectancy, inspection requirements, and expected adversary profile. Everything that was written in that context also applies to the "tie dye" seal.

A "tie dye seal" type approach has the following advantages:
- The adversary cannot break, and subsequently forge a seal ("anti-evidence", [Johnston06])
- The inspector must inspect the seal, meaning that he cannot just pretend that he did inspect the seal ("anti-gundecking", [Johnston06]).
- The hardware costs a few Euros.
- The design solves the problem of how to seal the tracker to the asset, and of how to secure the interior of the tracker.

The approach must however carefully evaluate the following:
- The issue of "data remanence", which is difficult to solve against a well-funded adversary (e.g. organised crime). For details see the discussion concerning the "time trap".
- Users could agree to sabotage the tracker by exceeding operating temperatures, causing extreme vibration (hitting the seal hard to trigger the alarm), or to use electromagnetic pulses to create a high power level inside of the tracker. This would create a high number of tampering alarms, which could psychologically de-sensitize the authorities to tampering.

Protection against opening and removal: As initially conceived, the tie dye seal protects only against opening of the tracker. However, if properly miniaturised, the design can be put into a bolt.

Passive and active implementations: As initially conceived, the tie dye seal is a passive device requiring manual inspection. However, by wiring the tie dye circuit into the tracker, one obtains an active monitoring seal. In case of a bolt design, the tracker would need to perform active monitoring through Near Field Communications.

Comment: The time trap seal requires a separate actor circuit, which acts upon breach, such as the time trap.

If used in a bolt, this seal is "passive" in the sense that the tracker itself does not report tampering. If used inside of the tracker, it can be integrated into an active sealing concept.

A challenge-response seal with some "time trap" features

Further above, this document explored the "time trap" and "tie dye" seals. Johnston already stated [Johnston06] that the sensor used in the "tie dye" seal can be used together with the principle of the

"time trap" seal. A seal that combines both features can be incorporated into a single design. Below, a design is presented that is useable both for volume protection and sealing bolts.

Briefly, the present design would use the "tie dye" sensor and inner coating as breach detector, would be useable in a marine environment because it would not have any opening or metallic connector, would send messages via satellite or Near Field Communication (NFC), and would use applied cryptography to authenticate its communications.

In summary, the modified time trap would be constructed by using the following components:

- A secure controller, which is capable of
    - Public key cryptography. Ideally it should support Elliptic Curve Cryptography, because the private keys are smaller than those used in RSA, and hence can be wiped faster
    - Operating a clock
    - Near Field Communications, if the device is used in a bolt.
- The time trap algorithm.
- A light sensor. The light sensor can be
    - a standard light sensor,
    - an active colour sensor proposed in the "Tie Dye seal", with the inside of the device coated in a random pattern.
- A temperature sensor.
- A battery, which is either rechargeable or primary, depending on other design considerations. If used in an autonomous bolt design, one should foresee a non-rechargeable battery of adequate capacity, which is safe to use, has long shelf life, and can operate in a large temperature range. Lithium Poly Carbon Monofluoride or Lithium Thionyl Chloride batteries[47] may be good candidates.
- If used to protect a GNSS tracker, the GNSS tracker should incorporate a communications terminal (satellite or GSM).
- If Near Field Communications are used, and the seal receives energy from a reader, a voltage sensor would monitor the energy received.
- Optionally, a second smaller battery or capacitor, which is not located near the main battery, and is connected to the secure controller independently of the main battery. If needed, the secure controller draws on this secondary power source for wiping its private key.
- Optionally, an external power connection, the voltage of which would be monitored.
- Optionally, epoxy potting material encases the device's electronics.

The sealed device acts as follows:

- At the manufacturer plant, the secure controller generates a public/private key pair. The control centre obtains the public key.
- Under normal operating conditions, the secure controller sends messages that are signed by its private key. The messages contain a time stamp, a cyclical counter, battery status, temperature status, volume seal status, and optionally a challenge sent by the control agency.
- To stop the adversary from collecting a large amount of signatures, the device always delivers the message with a delay of e.g. 1 second. During this delay, the device pretends that it is "busy" and will not accept additional requests.
- The temperature sensor alerts the device to extreme temperatures or extreme temperature changes.
- The light sensor monitors the volume of the device.
- Upon encountering abnormal conditions, the device memorises a signed alarm message, and erases its private key. It then keeps sending the "stale" signed alarm message.

[47] Lithium Thionyl Chloride and Lithium Poly Carbon Monofluoride batteries have a very long shelf life and high energy density. They are hard to deplete, even at temperatures that usually self-discharge other Lithium batteries.

- Upon inspection, the inspector verifies the physical integrity of the device, and queries the device to send a time-stamped and signed message. If the time stamp is not the current time, if the signature is missing or does not compute, or if the device does not react, then the device is disassembled and the failure reason would form the basis of an investigation.
- Upon end of life conditions, such as routine battery depletion, an inspector disconnects the device and installs a fresh one.
- Finally the used device is sent back to the manufacturer for refurbishment.

Below the changes with respect to the original design are motivated and the design is further detailed. First, the commonalities between such a design and the original "time trap" and "tie dye" seals are:

1. The use of a battery to power the electronics of the seal. The battery power level is checked. If the battery's voltage drops beneath a given voltage, indicating extreme cold or battery depletion linked to age, then the recommendation would be that the seal wipe its key.
2. The use of a temperature sensor, as mentioned in the original article on the "time trap", in order to mitigate extreme temperature attacks. Upon encountering a typical cold attack scenario[48], the seal wipes its key.
3. The "time trap" element of the seal includes a real-time clock that keeps the UTC time.
4. The detector could be an active light detector, such as the one used in the "tie dye" seal, as well as the idea to coat the interior of the seal using a unique and random colour pattern.
5. In order not to draw too much battery, an active detector would only operate intermittently. The detector's operation intervals would be determined "at random", using a pseudo random number generator (PRNG)[49]. To deter the adversary, there should always be a real possibility that the next operation occurs within one millisecond. The battery is to be chosen accordingly.
6. The use of the "time trap" concept, where the device hosts an internal key that is used to compute a hash function, using the current time as input to the hash function. Also, the internal key is wiped on breach.
7. The seal communicates authenticated data via digital communications.

This design is then different from the design presented in reference [Johnston06] in a number of points.

1. Replace the secret key and secret hash algorithm by a private key and a public hash algorithm.
2. If used as an autonomous bolt, replace the time trap's LCD screen by NFC radio frequency communications between the seal and the reader apparatus. Alternatively, if used as a volume protection seal for the tracker itself, the seal affects a single bit in the outgoing position message, the so-called "tamper bit". This change goes beyond a mere technicality, as it changes the conceptual design of the "time trap".
    - The original "time trap" is placed into the volume to be protected, and will only display valuable information once, when the device detects that the volume has been opened. If the "time trap" were to permanently display the time and the hash value, then it would release so much information that cryptanalysis could be successful: Cryptanalysis would be feasible under such a setting: The hour and minute represent about 10.5 bits of entropy, the key is coded on two bytes, and the device uses one of a set of ("degenerate") hash algorithms. In order to prevent cryptanalysis, the original "time trap" design reveals as little information as possible.
    - By contrast, the present design can allow the broadcasting of information for a long time before its integrity is threatened by cryptanalysis: The length of the message is at least 72 bits, and the device signs the message using an ECDSA algorithm and

[48] Sub-zero temperatures combined with a steep negative gradient, e.g. of more than 1 degree per second.
[49] To make this PRNG unpredictable, it should be implemented using a Secure PRNG algorithm, feeding off a private key that is different from the private key used for the hash algorithm, as per commonly understood best practices.

Elliptic Curve private key. If necessary, cryptanalysis can be made more difficult by adding further random bits in each message.

3. If used as an autonomous bolt, the design does not require that the inspector open the seal. Instead, the design communicates one sort of message if the seal is intact (the "OK message"), and another sort of message if the seal is broken (the "KO message"). (Alternatively, if used as a volume protection seal, the tamper bit takes over the role of the OK and KO messages, as the modified position message of explored further below would communicate time and signature anyway.)
4. If used as an autonomous bolt, an inspector verifies integrity using a separate reader apparatus. In that case, the seal is "passive" in the sense that it needs to be inspected.
   In order to use this design for "active monitoring" purposes, one would integrate the reader apparatus into the tracker itself.

Below is a discussion on the reasons leading to these changes.

1. First, the original "time trap" design requires an inspector to generate and write down a "secret key", which he must then communicate to his inspectorate, and must not share with other parties. The key is a 2-byte value, and it is then fed to an equally secret hash function that computes the hash value seen on the display in reference [Johnston06]. Given the high amounts of value that the present seal must protect, and the fact that an inspector could deny having passed the secret to the end user, such circumstances make for unnecessary temptation.
   Therefore the original "time trap" design proposed by Johnston would be modified, by introducing asymmetric cryptography, by using a publicly known hash function, and by implementing the above on a secure controller, such as the e.g. Infineon SLE88 series, NXP SmartMX P5S series, or the Dallas Semiconductor DS5000 series[50]. This would base the design on authentication mechanisms common to secure controllers: Anyone can verify that the hash function is correctly computed, but nothing except the tracker loaded with the private key can compute the hash function. In addition, the "time trap" would successfully have wiped the private key, meaning that all knowledge used to compute the hash is lost. Elliptic Curve Cryptography (ECC)[51] offers shorter key lengths than the more classical RSA Cryptography. A secure controller that supports ECC at the hardware level would be required[52]. Shorter key lengths (usually) translate into quicker key wiping times[53].
2. Second, and only for use as an autonomous bolt, the present design would make use of radio frequency to communicate with a seal reader apparatus (the latter can be integrated into the GNSS tracker or not). In the original design, the reader and the seal communicate by using a 3.5mm phono plug. But at sea, any physical connector port is vulnerable to salty water and harsh environment. Wireless communications would therefore be preferred. Johnston rightly argues that wireless communications can use a lot of battery power. This could be mitigated by using "Near Field Communications" (NFC), which draws much less power than for instance BlueTooth™.
   If this design is incorporated into a bolt, then in order to spare the bolt's battery, communications could be entirely powered by the reader electronics (so-called "passive NFC"). Such design is common with e.g. secure passports, where the "Single Wire Protocol" is used, with the reader powering the radio frequency communications terminal linked to the secure controller, also known as "passive mode"[54]. However, the "passive mode" opens up

[50] A secure controller is needed anyway because the GNSS tracker must communicate authenticated position messages to the control centre.
[51] ECC is hardware-implemented on the Infineon SLE88 series, making it available as a "drop-in" replacement for RSA. Given the competitive market in chip cards, one expects that other vendors also offer similar features.
[52] ECC is supported by some but not all secure controllers, for instance the Dallas Semiconductor models may not support ECC.
[53] This would however need to be experimentally verified.
[54] For instance, the Infineon SLE88 series features a version that implements the "Single Wire Protocol", and since such controllers are used in e.g. secure passports, the "passive mode" should be available.

another avenue of attack, which is related to the reader apparatus emitting too much or too little power. This risk must be duly addressed by examining the specifications of the secure controller, and implementing a voltage sensor circuit if needed. An adversary can also try to guess the secret key by running a long series of read-outs of the device. In order to mitigate such attempts, the secure controller would maintain a "message counter", which is incremented on every message sent[55], and may also artificially slow down the device's response time (such that it is not possible to obtain more than 1 message per second).

3. Third, and only if used as an autonomous bolt, the apparatus would not need to be opened in order to be examined. For that to work, one first defines two types of messages:
    - The "OK" message, which is calculated upon the reader requesting a read-out from the bolt, and which is an up-to-date version of the message. In particular, the clock represents current UTC time, and none of the terminal states has been reached.
    - The "KO" message, which has been computed during an alarm, prior to wiping the private key. After wiping the key, the message is stored permanently in the chip's EEPROM. The exact problem can readily be identified, because the UTC time no longer represents current time, and one of the terminal states has been reached.

    In both cases, a digital signature field authenticates the message's contents.

The device's message contents are:

- The time, computed in a 32-bit field as defined in the UNIX time command.
- The counter that informs the inspector on how many read-outs were already performed, encoded on 24 or 32 bit.
- A challenge sent by the reader apparatus, which is a number on e.g. 16 bit.
- A pseudo-random bit sequence of length N, which is added for the sole purpose of increasing difficulty of cryptanalysis. Keeping with cryptographic best practices, the sequence should be generated using a separately chosen private key and a cryptographically secure pseudo-random number generator (CSPRNG).
- The battery status, encoded e.g. on 2 bits (high, medium, low, battery was disconnected). If the battery status was "disconnected", then this value represents a terminal state, and should no longer be altered.
- The temperature status, encoded e.g. on 2 bits (too cold, cool, warm, too hot). If the temperature status was "too cold" or "too hot" at one point, then optionally, this value could represent a terminal state, and would then no longer be altered.
- The light detector sensor status, encoded on 1 bit (OK, breached). If the detector status was "breached", then this value represents a terminal state, and should no longer be altered.
- The signature field over the above fields, computed using e.g. ECDSA (Elliptic Curve Digital Signature Algorithm).

Seals that implement this concept have a unitary hardware cost in the order of about 25 EUR. Secure chip card controllers cost a few Euros apiece at high volume[56]. With the batteries, the advanced light sensor, and an emergency capacitor included, but development effort counted separately, several thousands of these seals should not exceed 25 EUR apiece.

This modified seal design would have the following benefits:

- Can be used as a volume protection seal. In the latter case, it can be integrated with the GNSS tracker's electronics.

[55] This message counter should take 24 or 32 bits. With 24 bits, the adversary is forced to make 1.6 million attacks before the counter overflows and restarts at 0, while with 32 bits, 4 billion such attempts are needed. An inferior design would implement the message counter on only 16 bits. Such a choice would mean that after 65536 attempts, the counter cycles back to its original value. Yet an adversary can accomplish this number of attempts in a single day.

[56] For instance, the SLE88CFX4000p in DSO-20 packaging would indicatively cost 5.55 EUR apiece at a quantity of at least 2500 pieces [personal communication from an authorised Infineon vendor]

- Can be used in a bolt type seal, where it would use wireless communications.
- Combines “time trap” and “tie dye” seal features.
- Is very hard to clone, since a private key is hosted on a secure controller, which is part of the seal.
- In a bolt type seal, the “OK” and “KO” messages from the seal have forensic qualities.
- In a bolt type seal, prevents the adversary from repeatedly read out the seal in order to guess the private key. By each read-out, the “OK” message counter is increased. Very high message counter values will then alert the inspector that the adversary tried the above process.
- The design mitigates the “data remanence” issue, owing to the shorter key lengths of Elliptic Curve Cryptography.
- When used for volume protection, the tracker can immediately alert authorities when the alarm condition is reached.

The approach must however carefully evaluate the following:

- The design is relatively complex. It calls for thoughtful implementation, a high degree of inspector training, and requires specialised reader equipment.
- Should the “bolt” version of this seal use NFC, then that requires that any “NFC reader” be located in close proximity (about 10 centimetres). If the tracker monitors its own bolt (active implementation), then installing the tracker directly on a steel rail (as in Figure 23) may yield best results. Otherwise, if the tracker is installed on a vessel on a tube above the cabin (as depicted in Figure 24) the lower part of the tube cannot be monitored, since it will be beyond the reach of NFC communications. For tube-mounted designs, one should therefore consider BlueTooth as an alternative to NFC, if active monitoring is desired.
- Should the “bolt” version of this seal be used in an active monitoring scenario, then the “NFC reader” inside of the tracker needs to know the public key of the smartcard in the bolt, so that the “reader” can verify the authenticity of the “OK” and “KO” messages. This implies that the “reader” must be pre-loaded with the public key of the smartcard. But the “bolt” may need to be replaced if its battery wears out or if it is damaged. In turn, that requires that the “reader” apparatus can receive updated public keys for the bolt, transmitted in a secure mode. This suggests an architecture similar to the one discussed in “Remote firmware updates” (page 115). It also suggests some form of Public Key Infrastructure, perhaps similar to what is being done in the area of the European Digital Tachograph (briefly presented on pp 18).
- As with the “tie dye seal”, the “bolt” version of this seal design could become the target of colluding users attempting to sabotage it.

Protection against opening and removal: This design can be used in the GNSS tracker housing, or can become part of a bolt.

Passive and active implementations: When guarding against opening, the seal is wired into the GNSS tracker, and influences one bit in the position message, therefore this is an active monitoring seal. In case of a bolt design, for active monitoring, the GNSS tracker would need to interact with the bolt through Near Field Communications. Otherwise the bolt can be inspected periodically in a “passive monitoring” system.

Comment: This design combines “anti-evidence” with known secure controllers and known cryptographic mechanisms.

Active monitoring using permanent magnets and magnetometers

The next sealing concept, provided by R. Johnston, relies on the use of a magnetometer. The apparatus can be summarised as follows:

1. Install a magnetometer inside of the GNSS tracker.
2. *Glue a strong magnet to the ship frame near the [GNSS tracker], and one or two to the [GNSS tracker]. Removal of the cone would be detected by a change in magnetic field in real-time, as would separating the whole [tracker] unit from the [asset].*

While surveying a magnetic field in real-time provides a formidable defence, care needs to be taken to exclude "false positives". In maritime applications, one often encounters moving metallic objects such as steel cables. On vessels, it is common to use welding equipment for legitimate repairs, and metal arc welding creates strong magnetic fields. Many vessels still use compasses as a backup for navigation. In addition, if one installs the apparatus beneath a rotating radar beam, such as used in ships' radars, any metallic parts of the beam will be subject to electromagnetic induction, generating electricity that potentially could adversely affect the radar's functioning. Hence this technique has issues in the maritime domain, rendering its adoption therein rather difficult.

This concept can also defend the volume of the tracker. To achieve this, one places two magnetometers inside of the volume, close together, on each part of the housing. Then one calibrates the alarm on the difference of the signals from both magnetometers. The housing would be designed such that it can only be opened by twisting the halves in opposite directions.

This defence conflicts with the use of some INS units, namely those which have magnetometers to determine the direction of the North Pole. Yet such INS units can be part of a defence against fake GNSS signals.

Even though it may have its liabilities, this method affords some of the most effective active removal protection to date. It would therefore be a method of choice.

Protection against opening and removal: This design is targeted at preventing the removal of a GNSS tracker from the asset to be tracked.

Active implementations: The purpose of this design is to alert authorities in real time of any attempts to remove the GNSS tracker from the underlying asset.

Comment: Not well suited for the maritime domain. Cannot be combined with INS devices that are augmented by magnetometers.

Surrounding the tracker's electronics with potting material

The FIPS 140-2 [FIPS140-2] standard would consider a GNSS tracker to be a "multiple chip standalone cryptographic module". In order to reach level 3 of that standard, one of the options is to pot it *with a hard potting material (e.g., a hard epoxy material) that is opaque within the visible spectrum.*

There is an alternative in the FIPS 140-2 standard for level 3, namely *the cryptographic module shall be contained within a strong enclosure such that attempts at removal or penetration of the enclosure will have a high probability of causing serious damage to the module (i.e., the module will not function).* One supposes that the above type of damage refers to physical damage, not key zeroisation, which is a more specific vocabulary that the FIPS 140-2 standard addresses in other sections.

The use of such potting material brings several advantages:

- It provides a cost-effective redundant protection barrier,
- Some adversaries employ small projectile weapons in order to separate the detection circuit from its power supply [R. Johnston, personal communication]. This design makes such an exploit difficult, since it slows down the projectile and also because the shock waves from the projectile risk damaging the rest of the module.
- It delays any attempts at physical tampering, because the adversary needs to dispose of a portion of the potting material,
- In disposing of the potting material, there is a significant risk of the device being damaged beyond repair,
- It insulates the device's core against sudden temperature fluctuations, and therefore renders a "cold attack" more difficult for the adversary,
- The potting can be sprayed with a random colour pattern, for combination with the above "tie dye" seal.

The design has the following liabilities

- Units cannot be physically refurbished by the manufacturer, in the sense that it is impractical to remove the potting material in order to change a single component that has been "potted".
- Any battery is either potted with the rest of the electronics, making it unreachable, or is not protected by the potting material. To mitigate, one can include a smaller backup battery or a capacitor in the epoxy resin, and a larger battery outside of the resin.
- Any light sensor guarding the cover needs to be mounted outside of the epoxy resin potting.
- The epoxy resin potting is either permeable to the RF spectrum at 1.5 GHz, in case of which one can mount the required antennae inside of the resin, or the antenna(e) must be installed outside of the potting, with connection wires reaching the inside of the resin.

Note that for obtaining FIPS 140-2 level 4, it is furthermore required to add a volume enclosing seal that monitors the cover of the GNSS tracker. This volume enclosing seal can be any of the seals previously presented. The "multiple chip standalone cryptographic modules" have different requirements from the "Multiple-Chip Embedded Cryptographic Modules", in which a "tamper detection envelope" is required. The latter strongly suggests the use of a product such as the sensor developed by W. L. Gore, see above.

Protection against opening: An epoxy resin potting would be used in the GNSS tracker housing.

Passive implementation: This design guards the electronics against further analysis, if the hull of the device has been penetrated. Ridding the device of the potting material is a time consuming and risky process.

Comment: This design would be used to comply with FIPS 140-2 level 3 and above. It has some other desirable consequences.

Conclusions on physical security

Several techniques to provide for physical security were reviewed above.

- The passive RFID Bolt seal, which provides good security in case the RFID itself uses cryptography
- The time trap and tie dye seals, which can be considered complementary components in an active volume protection seal
- The W.L. Gore active volume enclosing seal, which is expensive but provides adequate security
- Passive monitoring using a low-cost mechanical anti-evidence seal
- Active monitoring using permanent magnets and magnetometers, providing excellent protection against removal at an affordable price. However the monitoring of magnetic fields may present issues, especially in the maritime domain.

The above are only examples of possible sealing mechanisms. Creative and/or simple concepts are likely to continue emerging.

For retro-fitting existing vessels, any passive monitoring using mechanical anti-evidence seals represent good potential, with low cost and comparatively good security. This concept would seal the tracker to the asset, but also safeguard it against physical opening. An inspector periodically checks the seal; for maritime applications that could take place once per year.

For new GNSS tracking applications, one could consider “active monitoring” seals, where the tracker communicates its status to the inspectorate. If compliance with FIPS 140-2 level 3 or above is desired, then an epoxy resin potting would be required. Such seals may then need less frequent physical inspection at the same level of security. However physical inspections are still necessary, otherwise inspectors will only have the illusion of control.

Chapter III: Defending against Side Channel Attacks

Author: U. Kröner

Emanations Security

Electromagnetic leakage is commonly mitigated through a series of measures.

There are two classes of electromagnetic emissions from an electronic device. These emissions are either from conducted signals, where they are sometimes considered as electromagnetic interference (EMI), and emissions of radiated signals that would constitute radio-frequency interference (RFI).

Both EMI and RFI have known mitigation measures. Mitigation is required by law, and international standards for both types of measures are given by the "International Special Committee for Radio Interference" (CISPR). A whole niche industry exists that specialises in the manufacturing of such mitigation measures. Common techniques include EMI enclosures, coated interiors and gaskets.

The picture below represents a set of gaskets:

Figure 31: Gasket set provided by W.L. Gore. One Euro coin added for size reference.

Beyond the realm of CISPR, one finds the broad field of Emission Security or Emanations Security (EMSEC), sometimes referred to as TEMPEST. *Compromising emanations are defined as unintentional intelligence-bearing signals which, if intercepted and analyzed, may disclose the information transmitted, received, handled, or otherwise processed by any information-processing equipment* [TEMPEST-WP].

Known EMSEC standards are [TEMPEST-WP]:

- NATO SDIP-27 Level A – C: These standards vary with respect to the distance between the device under test and the equipment that measures the emissions. For GNSS trackers, it must be noticed that the adversary would have immediate access to the vicinity of the device, implying that NATO SDIP-27 Level A would be the standard to implement. Level A is also the strictest level available. However, quite a set of consumer goods are "Level B certified".
- NATO SDIP-29 governs the installation requirements for aforementioned NATO SDIP-27.
- AMSG 799B defines NATO zoning procedures.

All of the above documents are classified.

Arguably one of the difficult issues to solve is the mitigation of RF emissions for the tracker as a whole. Consider a hardened GNSS tracker composed of a secure controller, an embedded computing platform, a GPS receiver, an INS, a communications satellite transceiver, a volume seal sensor, a battery, and several other components. The wires between these components could reveal secret information about the system that an adversary could exploit, so this calls for radiofrequency shielding. However, at least one wire of the tracker must be able to receive and emit radiofrequency waves in the 1.5 GHz spectrum, otherwise GPS and satellite communications are not available. Reconciling these conflicting requirements can be a challenge, especially if the tracker would need to comply with standards such as NATO SDIP-27 Level A. Level B may be more attainable, since some general purpose computers are certified for use in the equivalent BSI[57] zone 1, where BSI *Zone 1 [..] correspond[s] to Equipment TEMPEST Zone 1 […] of SDIP 29 (former AMSG-799)* [BSI10]. Some comfort may be derived from the notion that if the tracker as a whole complies with Level B, and if it employs a secure controller that employs additional shielding, then it becomes difficult to eavesdrop on the secure controller without opening the device. This difficulty may well imply that an adversary would need to use expensive equipment in order to perform eavesdropping. If that is the case, then a "Level B" certification may offer an optimal compromise between economic costs and level of deterrence.

Side channel attacks

Closely related to the issue of EMSEC, one finds the issue of side channel attacks. The "side channel" refers to the information leakage of a cryptosystem that stems from its implementation, as opposed to from its mathematics.

There are several main forms of side channel attacks, listed at reference [SCA-WP], all have the aim of inferring the private or secret key(s) used:

- **Timing attacks**: monitor the time that cryptographic calculations take. Timing attacks can be potent when run on the same processor that also runs the underlying crypto graphical calculations. For an example of an attack on AES with source code, consult reference [Bernstein05],
- **Power Analysis**: monitor power absorbed during cryptographic calculations. Simple power analysis monitors electrical activity over time, while differential power analysis uses statistics,
- **TEMPEST** attacks, see above,
- **Acoustic** cryptanalysis, which is not applicable in the present cases,
- **Differential fault analysis**, attempting to introduce errors into computations by measures such as high temperature, radiation, and/or magnetic fields. Some of these measures are outlined further below in the discussion on operation "outside of voltage" and "outside of temperature". Note that GNSS trackers benefit from special environmental protection features, as required for FIPS 140-2 level 4 [FIPS140-2] certified cryptographic modules.

In relation to the GNSS tracker presently discussed, the above would suggest the following measures:

- **Absence of exposed conducted interfaces**: Conducted input/output interfaces represent a comparatively easy way to compromise such a device, by using EMI scans, by operating the interface outside of its specifications, or by trying to bypass the embedded software with a "buffer overflow" attack. Also FIPS 140-2 level 3 and above requires that access to a maintenance interface implies wiping of any private or secret key in the device [FIPS140-2]. It is therefore preferable to have future hardened GNSS trackers without exposed conducted interfaces. These and other measures are further detailed below in the chapter "Options concerning physical interfaces". Such a measure would then also restrict any EMI scan. The only wires out of a device may then be the ones connecting it to the power supply.

[57] BSI: Bundesamt für Sicherheit in der Informationstechnik

- **Use of verifiable features implemented by secure controllers**:
    - When a device only presents physical wires for the power supply, it is worth noting that many secure controllers have "dual rail logic", which mitigates the risk of the adversary using power or timing analysis. For instance, chips from the Infineon SLE88 series, NXP SmartMX P5S series, or Dallas Semiconductor DS5000 series present such features. These can be used, as long as one continues to investigate and discover the potential vulnerabilities.
    - The secure controllers also mitigate the problem of concealing the device's identity. Manufacturers of secure controllers have made considerable effort to secure their chips against an adversary extracting any private keys. This opened many application areas to secure controllers, such as credit cards[58], digital TV smartcards, and secured passports. In addition, the designs proposed further enclose the secure controller inside of the hardened GNSS tracker, by means of volume protection. Therefore, while it may theoretically be possible for the adversary to obtain a private key stored in the hardened GNSS tracker, it is certainly not the easiest way to render the tracker useless. For instance, removing the tracker may be easier.
    - Secure controllers commonly are wrapped in coatings that mitigate RF emissions[59].

The JRC has a unit that performs TEMPEST type tests. The relevant expertise represents a large body of knowledge that goes beyond the scope of the present document.

With the issue of electromagnetic leakage outlined, attention will be shifted to the issue of volume sealing.

Operation outside of temperature or voltage specifications

Concerning temperature, many secure controllers contain an integrated temperature sensor. The primary concern is that the adversary uses extreme cold to disable the volume protection of the GNSS tracker, then performs a physical attack.

If FIPS 140-2 level 4 certification is required, then environmental failure protection (EFP) measures must be present.

One option is that the secure controller could be programmed to check for these thresholds.

1. An absolute lowest temperature threshold, for instance -45 °C, beneath which most electronics components commonly stop working.
2. The gradient of temperature changes with respect to time, combined with absolute temperature. If an object is thrown into liquid nitrogen, one would expect temperature descending by several degrees per second, at temperatures below zero. Such quick descents cannot occur on the high seas, and presume an enveloping source of extreme cold.
   Note that an adversary could elect to cool the GNSS at a slower rate than the above gradient. However, upon reaching the above minimal operating temperature, the detection logic would take appropriate action.
3. An absolute highest temperature threshold, for instance +100 °C, above which most electronics components commonly stop working.

[58] As a cautionary note, in 2010 possible attacks against the EMV credit card payment systems were blatantly exposed. This problem was known for a long time; however Industry had not dealt with it. Details are available e.g. at http://www.schneier.com/blog/archives/2010/02/man-in-the-midd_1.html

[59] For example, the SLE88 series include an "active shield", covering the requirement FPT_PHP.3 "Resistance to physical attack". It is unlikely that the competition lags far behind.

This higher threshold is designed to prevent computing faults that could be exploited using differential fault analysis, a form of side channel attack.

If the inside of the device is potted in epoxy resin, the temperature gradient cannot be arbitrarily steep, which means the tracker has more time to erase its secret key.

But what sort of action should be taken? Further below, a secured protocol for the exchange of GNSS position messages between a tracker and a control centre. This secure protocol uses both asymmetric (long term) and symmetric (short term) keys. Upon encountering an alarm condition, the short term key should be wiped immediately. But what about the asymmetric keys, which supposedly are protected by the secure controller? While it is very expensive to obtain the private key from the secure controller, this document argues that the adversary could "re-use" the secure controller in a pirate device, effectively creating a "clone" of the original tracker, which the adversary then controls. Therefore it is advisable to erase any secret or private key upon meeting an alarm condition.

Another remark can be made concerning the above temperature thresholds. Some application areas may require operating in areas of extreme cold. This naturally occurring cold may then be a problem, in that it would cause false positives with the above temperature detector. For such application areas, one could include a heating resistor in the device, which will operate if an external power supply is present, and if the temperature descends below a certain threshold. However, if an epoxy resin potting is also used, then its temperature expansion characteristics must be checked against the heating capacity of the resistor, so that excessive physical stress is avoided.

Concerning voltage, it is the position of this publication that it is preferable not to include digital input/output wire lines with the GNSS tracker. If an external power supply is required, the supplied voltage should be monitored. If the voltage drops below a threshold, a circuit can block the power supply and switch to backup power. Such an "out of voltage" protection circuit hence benefits from the presence of a backup battery.

Indeed, other parts of a hardened GNSS tracker also require such a battery:

- Any light, colour, and temperature detectors,
- The clock of any "time trap" seal,
- The GNSS module can be hardened against spoofing and meaconing attacks, for which one needs a battery-powered precision TCXO clock.

Therefore, the addition of a suitable rechargeable or primary backup battery would be strongly recommended in any hardened GNSS tracker.

Options concerning conducted interfaces

Definition and background

Let "conducted interfaces", be any wires that the GNSS tracker exposes. The power cables constitute a special conducted interface, therefore the "conducted data interface" and the power cables will be considered separately.

Many GNSS trackers have conducted data interfaces. These can commonly be queried for the current GPS position, or permit the user to use the tracker's communication link as a gateway. In maritime applications, if the asset value is low, then it is frequent that the skipper uses the tracker's communication link in order to e.g. send emails or perform phone calls. This has economical reasons: If both the tracker and the user would have separate communication links, then that commonly would incur an additional subscription fee.

Wrapping the conducted data interface inside of the protected volume

In order to create a hardened GNSS tracker, it makes sense to suppress any exposed conducted data interface. The reasons are:

- Given economical considerations, a hardened GNSS tracker makes sense mostly for high asset values[60]. The higher the asset value, then less the cost of communications matters[61]. Hence one can suppress the conducted data interface in such cases.
- By definition, the lower the asset value, the less security is a concern. Yet if the asset value is low, then the additional communication costs have more weight with respect to the overall finances. A rational adversary considering illegal activities will consider the initial investment into illegal devices versus the possible profit, and the possible risk of getting caught.
- In maritime applications, the less wealthy skippers may still need to rely on a satellite terminal provided by e.g. a GNSS tracker. If the skipper is not wealthy, then the asset value tends to be low. Hence it makes sense to leave the conducted data interface intact in such cases.

In summary, a hardened GNSS tracker only makes sense for high-value assets, and if the asset value is high, then there is little need for exposed conducted data interfaces.

Taking away the GNSS tracker's conducted data interface begs the question on how to commission the tracker. For instance, any Inmarsat-C system needs to be loaded with a DNID and a member number. Without a conducted data interface, this sort of information needs to be loaded at the manufacturer site, or programmed remotely. Indeed, for many Inmarsat-C products, there is a special configuration message ("polling service") for loading the tracker with a DNID and member number.

Apart from commissioning, the maintenance and refurbishment of the GNSS tracker needs to be considered as well. If the conducted data interface is inside the protected volume, then a tracker can still take advantage of having this interface for refurbishment at the manufacturer's site. This enables the manufacturer to re-use expensive electronics from a repairable or accidentally opened device. In this context, it makes sense as laid out in FIPS 140-2 level 3, namely to let the tracker erase its private key when the interface is powered up. This can then be combined with housing the cabled interface inside of the protected volume.

In summary, if a hardened GNSS tracker is devised, a conducted data interface is best placed inside of the device, where it is protected by a volume enclosing seal.

The next section examines what can be done in case the conducted data interface must remain accessible during normal operations.

Securing external physical data interfaces

If the conducted data interface must be kept accessible at all times, then it must be secured against out-of-voltage operations, and against buffer overflow type attacks.

One can argue that if a conducted data interface is kept, it is probably required in order to maintain the device. This concept is however incompatible with FIPS 140-2 level 3 requirements.

[60] In fisheries enforcement, a secured VMS device makes sense mostly for larger fishing vessels aiming for valuable target species. Also, a larger vessel will exert higher fishing pressure, which means it ought to be subject to more control.

[61] In fisheries, because large vessels fishing for high-value species are usually quite profitable, they can afford other communication means on board with relative ease, such as Inmarsat BGAN ("FleetBroadband"). Therefore they usually do not rely on their VMS devices to communicate with the shore.

- The FIPS 140-2 standard requires, from level 3 onwards, that any physical maintenance interface to a "module" use identity-based authentication. Commonly this is achieved by the maintenance staff presenting a digital certificate that is trusted by a Certificate Authority that is also trusted by the device.
- The same standard also requires that upon any access to the physical maintenance interface of the "module", all plaintext secret and private keys shall be zeroized.
- By "module" one should consider the tracker as a whole, including any components on top of a secure controller containing the keys and performing any authentication. If one does not, then it is possible for an adversary to leave the secure controller intact, subvert the components "around" the secure controller, and then use the secure controller's signature functionality for his own purposes.

Given that the FIPS 140-2 standard was carefully engineered, the above would imply that it is highly inadvisable to expose conducted data interfaces for a hardened GNSS tracker, since such a tracker should adhere to the concepts behind the FIPS 140-2 level 3 specifications (in other words, be engineered "in the spirit of" FIPS 140-2 level 3).

Note that previously it was established that the tracker reports its position using wireless communications, and that the key used to sign its position message must be discarded if a breach is suspected.

Finally, note that at least one data interface (albeit not necessarily a physically cabled interface) must exist in any case, or the tracker will be unable to report. Securing any wireless data interface is the subject of chapter IV.

Options regarding the power supply cable

A recurring concern in some application areas is that some tracker users disable the GNSS tracker when it seems convenient to them[62]. This then causes the control centre not to receive any data.

In fisheries enforcement, the non-receipt of position data should be followed up using the available community legislation, as Commission Regulation (EC) 2244/2003 specifies that

- *"the master of a Community fishing vessel shall ensure that […] the power supply of the satellite tracking devices is not interrupted in any way."* (Article 6)
- *"When the FMC of a flag Member State has not received data transmissions […] it shall notify the master […]. If, in respect of a particular vessel, this situation occurs more than three times within a period of one year, the flag Member State shall ensure that the satellite tracking device of the vessel in question is checked. The Member State concerned shall investigate the matter in order to establish whether the equipment has been tampered with."* (Article 12)

Therefore, skippers who routine cut the power supply to the VMS device will attract unwanted attention, and this is part of Community law.

However, it is notoriously difficult to prove that any malfunctioning was due to other circumstances than technical failures. If this happens on a routine basis, one can expect that the control centre staff become progressively de-sensitized to such issues.

Therefore the following options are proposed:

- The use of a rechargeable battery, which can withstand power cuts for at least several hours, or
- The use of a non-rechargeable battery, perhaps based on Lithium Manganese Dioxide, which could function for up to two years.

[62] In fisheries enforcement, it is not uncommon that a VMS device's power supply fails while the vessel is located close to a marine protected area.

If such a battery is included into a GNSS tracker, it would require the transmission of a battery charge level in the tracker's message, such that the authorities are alerted before the device is about to stop functioning.

As mentioned in the introduction, in the context of fisheries enforcement, Cyprus announced in 2009 that it is installing VMS devices that do not require a power supply. Instead, this type of VMS device uses a long-term, non-rechargeable battery. This VMS device is non-configurable, and transmits VMS messages every hour, on a permanent basis, until the battery is exhausted. At this point, the VMS device is sent back to a service centre for refurbishment. The approach taken by Cyprus shows that the option of using a VMS device without a power supply is realistic.

If the device is to be operated in extreme cold, a heating resistor may be needed to avoid operation outside of temperature range. This requires an internal heating resistor, and in turn a larger battery. The tracker's position message would then also provide additional information to the control centre about the use of the heating resistor. When picking a battery size, one needs to consider that in most application areas, users will flee from extreme cold (less than -40 °C air temperature)[63]. Hence everyone wins if the tracker reports extreme cold, since it enhances both the user's safety and the tracker's security.

Buffer overflow exploits

As mentioned in the introduction, buffer overflow exploits are comparatively easy to implement [Gerg05]. As with other security, the adversary must only find one particular exploit to subvert the device, forcing the defender to consider all possible entry points into the system.

In the document "A Comparison of Buffer Overflow Prevention Implementations and Weaknesses" [Silberman04], the authors compare many different commercial products that protect against buffer overflow exploits. The document concludes that a combination of kernel hardening measures and compiler protections afforded the best protection.

Embedded operating systems used in e.g. secure controller chips are often Common Criteria "Evaluation Assurance" certified at Level 5 or more (EAL5 to EAL7). Such an OS represents a strong defence, even though it is not formally proven Operating Systems will ever be perfect.

The Linux kernel represents an alternative to closed EAL-certified systems. It is freely available, and is among the best researched Operating System kernels. Some variants are EAL4 certified usually relying on the National Security Agency's SELinux. Kernel and compiler hardening measures are available for free (i.e. on the "OpenWall" or "grsecurity" web sites), and embedded Linux systems can be found in several mainstream consumer markets (such as mobile phones, PDAs, internet routers, satellite TV set-top boxes). It is therefore reasonable to assume that a hardened embedded Linux kernel, together with compiler defences, is a good choice for hardening a GNSS tracker against buffer overflow exploits.

But are OS and/or compiler hardening measures sufficient on their own? Security literature repeatedly points out that this would be an inferior approach. Instead, individual protection measures must become part of a cultural shift in a security minded organisation[64]. For an account of measures to

[63] When such cold weather actually occurs, the vessel risks being trapped in sea ice, therefore the skipper would head to port immediately.

[64] This remark is also valid in other contexts. For instance, this document also advocates that a complete risk assessment be performed for GNSS tracking applications, and that GNSS tracker hardening measures on their own are beneficial but not sufficient.

implement, one can e.g. consult the reference [Taylor05]. Without these measures, the organisation is likely to fare poorly against hackers. For a brief account of hacker culture, consult the reference [Bratus07].

Even if a security aware organisation uses best available technology and practices to protect GNSS trackers, it can be that someone finds exploits that can only be fixed by the deployment of updates. These would preferably use “firmware over the air” (FOTA) technologies to minimise logistic overhead. For GNSS trackers, such updates imply using encrypted communications. For extra security, one could consider encrypting the update more than once, e.g. by the manufacturer and by a control centre.

The possibility of updating the device incurs at least two additional risks

- The risk that the adversary abuses the update service in order to alter the firmware at will. This can be mitigated by encrypting the software updates.
- The risk that the software update, due to whatever reason, renders the device unusable. In that case the device must be refurbished by the manufacturer.

However, the above risks are in principle preferable to the other risk, namely of having a device that cannot be updated, coupled with the danger that exploits will be found and made public.

Concerning costs, embedded Linux OEM systems are available today for less than 50 EUR each, as evident by various commercial Internet access ADSL routers that use Linux.

Chapter IV: Securing the tracker's position reports

Author: U. Kröner

Authenticating messages to and from the GNSS tracker

Scope

The scope of this chapter is to explore how low-bandwidth messages could be authenticated, and if any special considerations are warranted for migration to higher bandwidths. As an example for a low-bandwidth scenario, and without loss of generality, this chapter examines the existing Inmarsat-C protocol position messages, in such a way as to harden it against spoofing and delayed retransmission. In this chapter, spoofing shall refer to the imitation of an Inmarsat-C position message by a device other than the GNSS tracker. It is possible that the adversary performs an "identity theft", meaning he steals the cryptographic secrets of the GNSS tracker, but for this chapter, that would be out of scope.

The Inmarsat-C protocol is particularly relevant in the area of maritime position reporting, since it dominates both the Vessel Monitoring System (VMS) and the Long Range Identification and Tracking (LRIT) markets. Furthermore, if the Automated Identification and Tracking (AIS) is to be secured, one option would be to fit the AIS terminals with a satellite communications transceiver, and have them send authenticated position messages, in parallel to the existing VHF communications. Inmarsat services are also used for land-based applications. For instance the "Slap & Track" MT 3300 terminal (Skywave Mobile Communications) uses the IsatM2M service.

The Inmarsat-C position reporting uses two types of messages, which are (as documented in reference [TT1])
- The periodic[65] position report, sent by the GNSS tracker.
- The "poll request", sent by the control centre

Note: The interested reader will have noted that there are other types of Inmarsat-C reports. These reports would need to be changed in a similar manner to the "position report". The affected Inmarsat-C reports are:
- Time Of Position Report
- Power Up Report
- Power Down Report
- Antenna Disconnection Report
- Antenna Blockage Report
- Enter Sleep Mode Report
- In Sleep Mode Report
- Leave Sleep Mode Report
- Fix Time Begin Report
- Reserved
- Enter Reduced Transmission Mode Report
- In Reduced Transmission Mode Report
- Leave Reduced Transmission Mode Report

[65] Whether this period is once ever six, two or single hour impacts not the protocol but only the economics of data transmission. Economical aspects are explored further below.

There also is an acknowledgement to the "poll request", the "poll acknowledgement response". This message type is not frequently used and can be replaced by the position reporting message.

Inmarsat-C service providers

In theory, any service and protocol that is globally reachable can be used. These are explored in the annex B of document [Kroen09], and are

- **Euteltracs** – the European manifestation of the U.S. Qualcomm Omnitracs system, based upon geostationary satellites, which commands the largest part of the land transport market.
- **Globalstar** – System based upon low-earth orbiting satellites that covers most of the world's land mass and extends into coastal waters and beyond.
- **Iridium** – based upon low-earth orbiting satellites with global coverage
- **Inmarsat Fleet 33** – maritime voice and data system based upon geostationary constellation
- **Inmarsat-C** – maritime data system based upon geostationary satellites that enjoys the most substantial portion of the VMS market and provides a choice of two services: an unformatted message or a pre-formatted, maximum 32-byte data report. This latter service is the commonly used in VMS.
- **Thuraya** – voice and data system, based upon geostationary satellite and covering Europe, northern Africa and the Middle East.
- A more recent addition is **Inmarsat BGAN** (Broadband Global Area Network). Based on three geostationary satellites, it achieved global coverage as of February 2009 (except for the east coast of Greenland). Data throughput can reach 492 kbps.

In fisheries enforcement, Inmarsat-C has a dominant market position. Many VMS terminals, such as those sold by the market leader Thrane & Thrane, use Inmarsat-C, or to be more precise, its "Position Reporting" feature in Inmarsat parlance [inmC].

How would the installed base of Inmarsat-C devices evolve? Because the installed base is large, it has a certain inertia. One therefore expects that the present-day preference for Inmarsat-C will continue in the near future, but that in a decade or two, fishermen will have universal access to globally available broadband. A discussion on migration to broadband is included near the end of this chapter.

On the secrecy of the Inmarsat-C transmission protocol

The Inmarsat-C transmission protocol is based on a "Slotted ALOHA", the exact implementation of which is kept secret. The above notwithstanding, the companies that are interfacing with Inmarsat terminals are aware of many of the specifics of the underlying protocol.

On "security by secrecy", Bruce Schneier, book author and involved in the creation of many cryptographic algorithms, writes [schneier04]:

> *The argument that secrecy is good for security is naive, and always worth rebutting. Secrecy is beneficial to security only in limited circumstances, and certainly not with respect to vulnerability or reliability information. Secrets are fragile; once they're lost, they're lost forever. Security that relies on secrecy is also fragile; once secrecy is lost there's no way to recover security. Trying to base security on secrecy is simply bad design.*

Amateur radio hackers are indeed able to decode (read) Inmarsat-C messages [Sluiman10]. Once the actual encoding mechanism becomes public knowledge, the fabrication (writing) of counterfeit

Inmarsat-C messages is facilitated. In turn, if such messages can be fabricated by the adversary, then any message transmitted via Inmarsat-C is fundamentally unreliable.

Present-day use of Inmarsat-C

Inmarsat-C has two modes of operation. The Inmarsat-C **data report** is made up of 15 byte packets, up to a maximum of three. There are, then, 45 total bytes, but 32 user-available bytes when headers and checksums are accounted for. An Inmarsat-C **message** is made up of 256 bit packets, up to a theoretical total of 32 k bytes [R. Gallagher, personal communication].

Many vendors implement position reporting using the Inmarsat-C **data report**. The Inmarsat-C **messages** are interesting when discussing the proposed changes that secure the Inmarsat-C data reports. This document first examines several Inmarsat-C data reports, with the aim of making them safe against spoofing and retransmission.

Each Inmarsat-C data reports is split into one or more data packets. Each data packet contains 15 bytes. The data reports are split into packets according to the following rules:
- The first packet contains 5 bytes of header data, 8 bytes of the data report, and 2 bytes of checksum.
- Any subsequent packet contains a 1-byte header, 12 bytes of the data report, and 2 bytes of checksum.
- In a position report, there is exactly one such subsequent packet, for a total of two packets.

Below the anatomy of some Inmarsat-C data reports is explored in more detail, with respect to its contents and its entropy. The discussion on entropy is useful in the context of digital signatures.

Inmarsat-C “position report”

This is the first type of Inmarsat-C reports presently examined. It is commonly transmitted once every six hours for LRIT, and once per hour for VMS. It is also sent in response to a “polling request”. The report identifies the vessel, the position, the speed, and the course.

For fisheries monitoring, the document [fao1] recommends the following VMS position report with course and speed:

Recommended optimized VMS position report

Field	Expression	Data
Hemisphere	North or South	1 bit
Degrees of latitude	whole number 0 to 90	7 bits
Minutes	whole number 0 to 60	6 bits
Fraction of minute	multiples of 0.04	5 bits
Hemisphere	East or West	1 bit
Degrees of longitude	whole number 0 to 180	8 bits

Minutes	whole number 0 to 60	6 bits
Fraction of minute	multiples of 0.04	5 bits
Speed	number with resolution of 0.2 knots	8 bits
Course	whole number 0 to 360	9 bits

Table 8: FAO VMS position report, source [fao1]

As mentioned previously, Thrane implemented Inmarsat-C reporting using two packets of 15 bytes each. This is illustrated in the relevant documentation [TT1]:

2.1.2 First part of Positioning Report

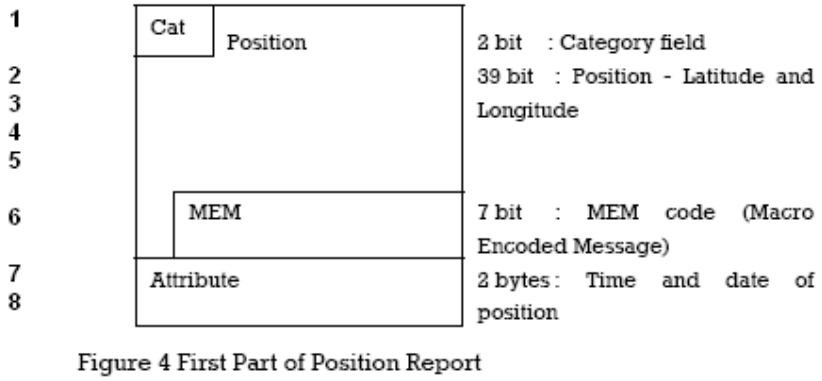


Figure 4 First Part of Position Report

Figure 32: First part of positioning report. Each row represents one byte.

The first Inmarsat-C data packet contains
- a 5 byte header,
- a 2 bit category field, neglected for the purpose of this document,
- the position, encoded as recommended by FAO,
- a MEM code, which is constant for position reports,
- a time attribute, encoded on 2 bytes
    - 1 bit for the month (0: this month; 1: next month)
    - 5 bits for day of month
    - 5 bits for hour of day
    - 5 bits for the minute, where the minute is given in units of 2 minutes
- a 2 byte checksum.

This first packet contains position entropy (39 bits[66]) and date/time entropy (16 bits) for a total of 55 bits.

[66] One can argue that the position entropy is not 39 bits, because not all positions are equally likely, since the position is a function of the fishing grounds. This argument is evaluated further below.

2.1.3 Second Part of Positioning Report

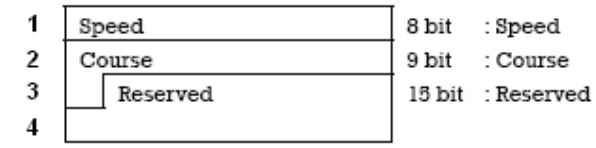


Figure 5 Second Part of Position Report

Figure 33: Second part of positioning report. Each row represents one byte.

The second data packet contains
- a 1 byte header,
- the speed on 8 bits in multiples of 0.2 knots,
- the course on 9 bits, effectively ranging from 0 to 359 degrees,
- some reserved information, the contents of which are unused and set to zero,
- a 2 byte checksum.

This second packet contains speed and course entropy for a total of 17 bits.

From the above, one concludes that
- For the second packet, out of 15 bytes, 7 bytes are used, including a 15 bit reserve field that is unused in practice.
- That leaves 7 bytes + 15 bit = 79 bits for digital signatures, without any impact on the economics of the position report transmission.
- Should one not use the 15 bit reserved area, then one still has 64 bits available for digital signatures, without changing the underlying economics.

Therefore, authenticating the position report messages should in principle be possible.

Inmarsat-C “position report polling request”

At infrequent intervals, for verification purposes, a control centre may request a position from a particular tracker, by sending a “position report polling request” Inmarsat-C data report. The tracker replies by sending the “position report” described above.

Since a control centre will often polling requests in batches, many “position report polling requests” are commonly lumped into a “poll file”. In the document [TT2], one finds the following specification for a poll file:

Byte	Information	Remark
0	Header Length	Header size including this byte
1	LES ID	Compressed Inmarsat format: Two bits for Ocean and 6 bits for ID, e.g. 5FH for LES ID 131
2	LSB of DNID	
3	MSB of DNID	
4	Member Number	
5	Sub Address	
6	Command	Command parameter from poll packet. This includes the Acknowledgement bit (80h) if this was included. […]
7	Sequence Number	A LES poll packet identifier. This value is returned in an acknowledgement.

Table 9: Poll file specification. LSB and MSB stand for Least and Most Significant Byte.

Given the above, it is likely that the "position report polling request" occupies 8 bytes.

Note: In order to cause the tracker to send an immediate "position report", as documented in references [TT1] and [TT2], the control centre needs to send a "position report polling request" with

- Field 5 (sub-address) = of 1
- Field 6 (Command) = 2 (for Pre-assigned data reporting) or = 6 (for Programmed Unreserved reporting)

As mentioned previously, reports are sent in data packets of 15 bytes each, with the first packet having a payload of at most 8 bytes. Therefore, the "position report polling request" is likely to contain no unused bytes. Therefore, if "position report polling requests" are to be secured, then each request is likely to be encoded on two data packets instead of one. This may double costs to the FMC for position polling, but such polling represents a negligible fraction of the total costs of operating an FMC.

There are other "polling request" messages besides the "position report polling request". These tend to configure the device's settings. For instance they can

- Change the DNID of the VMS device
- Change ocean regions for reporting
- Change the reporting interval
- Disable the "blocked GPS antenna detection" feature
- Disable the retransmission of failed reports

Any of the above represent sensitive "over the air parameter administration" commands, which the authorities would not like to see replicated by an adversary. For instance, the tracker usually sends special reports if the GPS antenna was temporarily blocked. This is to deter the adversary from putting a metal bucket on top of the tracker.

The structure and contents of these sensitive polling request messages are nearly identical to the "position report polling request". All of the "polling requests" can be secured using the same mechanism, explored further below.

Dynamics of report transmission

This subchapter discusses the process concerning routing and transmission of messages.

Transmission protocol

In practice, the trackers using Inmarsat-C transmit position reports using a connection-less mechanism. For each position report, a fixed number of maximal attempts are made to send it. The satellite receives the report, and acknowledges it. If the tracker does not receive this transmission-level acknowledgement, the report is re-sent. (Note: The exact number of attempts is largely a secondary consideration, because no matter the number of maximal attempts, the possibility of failure remains.)

Message routing

When the tracker sends a report to the MCS centre, it is first intercepted by a communications satellite. The satellite then forwards the report to another intermediary, the Land Earth Station (LES). Every few minutes, the LES bundles all reports according to their DNID. Each DNID is mapped to exactly one MCS centre[67]. The MCS centre receives the report, and identifies the boat, as a single DNID-Member combination is mapped to one (and only one) boat.

Therefore, the sender is identified by the DNID and the "member" fields, and the recipient is identified by the DNID, for the purposes of transmitting the "position report".

Proposed changes to the tracker message

Changes to the "position report"

The following method is proposed for sending authenticated "position reports".

1. Let each tracker have an asymmetric encryption key pair of sufficient strength not to be broken before 15 years, taking a 5 year shelf life and 10 year use into account. According to common recommendations[68] this would imply a RSA key size of 2048 or 3072 bits, or an Elliptic Curve key size of 224 bits. These values need to be adapted as years pass, because of Moore's law. Let the control centre know the public key of each tracker.
2. Let the control centre have an asymmetric encryption key pair of sufficient strength. Let the trackers know the relevant public key of the control centre.
3. The tracker and the control centre agree on a shared secret key, using key encapsulation or a key agreement protocol. In practice the parties could use PSEC-KEM [PSEC-KEM], protocols inspired by Menezes-Qu-Vanstone (such as the Fully Hashed MQV [FHMQV] or MQVKC [Popescu04]), by securing the key in a message sent using the Integrated Encryption Scheme (IES), or using MTI/A0 [Popescu04]. This shared key is subsequently used to sign N messages, after which the old key is zeroized and a new key is generated.
   If FIPS 140-2 levels 3 or 4 are a requirement, then any "Unified Model C(x, 2)" key agreement scheme can be envisaged. A C(x, 2) scheme ensures that *only [the sender] and the other intended party can compute the shared secret*. The FIPS 140-2 approved unified models are specified in NIST SP 800-56A [NIST07]. The variants C(0, 2) or C(1, 2) only use a single request-response cycle ("single pass"), which is helpful in low-bandwidth situations.
4. Tentatively fixing N at 168, one secret key exchange would cover about one week of position reporting and other traffic at 1 message per hour.
5. Concerning the costs of this key exchange, it should be under 1 EUR. An Inmarsat-C **message** costs about 0.60 EUR for 128 bytes [Kroen09], which would be sufficient for the tracker to send

[67] This does not mean that a MCS is limited to 256 boats, as many DNIDs can be mapped to a single MCS centre.
[68] http://www.keylength.com/en/ offers a gateway to Lenstra/Verheul, NIST, ECRYPT, and other recommendations.

an ECIES message containing a shared secret key of 16 bytes. The control centre replies with a shorter authenticated acknowledgement message, indicating good reception of the new shared secret key. A 32 byte acknowledgement message costs 0.15 EUR [Kroen09]. Other protocols based on Elliptic Curves would have message sizes of the same order of magnitude.[69]

6. Subsequently, the tracker signs each position report using a MAC such as HMAC-SHA-224. The signature is generated by including at least the 3 bytes of the DNID and member number, the 7 bytes payload of packet 1, and the first 17 bits of speed/course in packet 2.
7. In addition, the design requires additional bits in the reserved area. The "spoof bit" would change if the receiver believes that GPS spoofing is ongoing. The "broken bit" would change if the physical seal is broken. The "cold bit" indicates if the device detected unusually low temperatures or unusually fast temperature declines, as would occur with liquid nitrogen baths. The "hot bit" accomplishes the same for fast temperature increases or excessive heat. Because the tracker's design would include an additional battery, two more "battery bits" would indicate if the device is powered by the power cable (11), or is using the battery. In the latter case, the tracker checks the battery's status, and transmits 10, 01, or 00. The message authentication code would authenticate these additional bits as well. The payload of packet 2 of the tracker will therefore grow from 17 to 23 bits of plaintext.
8. Feeding the above data and the secret key into a HMAC-SHA-224 algorithm, each message will be mapped to a corresponding hash of 224 bits (i.e. the hash is larger than the message). The hash is then truncated such that it fits into the rest of packet 2, meaning it would have a length of at most 74 bits.

When using the existing use of Inmarsat-C, the above implies that sending the MAC incurs no additional communication costs.

The recipient of the message has the shared secret key, and

1. Verifies the MAC against the message, and
2. Discards any message that has unrealistic date and time values, where "unrealistic" is defined by the communication protocol latency.

In essence, the above relies on a hybrid cryptographic system with a shared secret key for message authentication, where the shared key is protected by asymmetric key pairs. The above mechanism considerably shortens the message signatures with respect to Digital Signature Algorithms (DSA): The Elliptic Curve DSA, one of the best-known DSAs that creates shorter signatures, requires at least 40 bytes of signature data, which if used would adversely impact the economics of report transmission.

The adversary's impediment is that HMAC with an adequate cryptographic hash function (such as SHA-224) is known to depend only on the size of the secret key, not on the size or entropy of the message.

Changes to the "polling request" report types

Do the "polling requests" need authentication? This depends largely on the type of the polling request.

1. The "position report poll request" is answered by the tracker, causing it to send a "position report". If an adversary transmits fake "position report poll requests", he could use the resulting messages to gather data for a brute-force attack. This assumes that the adversary prevents the control centre from receiving the position reports.
2. Other types of poll requests are in fact remote configuration messages. These are sensitive by their very nature. An adversary could reconfigure a tracker over the air, undoing the changes at another

[69] Even if the key agreement message sizes were larger, that would not be a significant problem. Consider for a moment that a key agreement protocol were needed that would inflate message sizes by a factor of five. Under that scenario an additional cost of 5 EUR per exchange per tracker would be generated. If the validity of the shared keys were then extended to one month, the costs would still be acceptable.

point in time, and then blame the tracker's supposed "odd behaviour" on a technical failure of the tracker. The main problem for enforcement is that it cannot prove the contrary.

The above points are both cause for concern. Therefore, the following method is proposed for sending authenticated "poll request" reports, which closely mirrors the proposal for position reports:

1. Assume that the tracker and the control centre have passed points 1-4 of the steps set forth in the "changes to the position report". This implies that the tracker and the control centre share a secret key.
2. Let the message *m* be the command and approximate UTC time information.
3. Calculate a MAC using the same procedure as put forth in "changes to the position report".
4. Assuming that the additional space available in the Inmarsat-C data packet is insufficient, reserve an additional data packet. Transmit UTC time information and MAC in the second data packet.

The recipient checks both the MAC and the UTC time.

The above provides sufficient security, because the impediment is the same as for position reports.

Public Key Infrastructure

The changes characterised above require the roll-out of an adapted public key infrastructure (PKI) between the trackers and the control centres. Trust flows from a Certificate Authority, which signs any public keys issued from control centres and trackers. Both parties have to periodically verify the validity of the certificates, using e.g. the Online Certificate Status Protocol (OCSP).

Migration to globally available broadband

In this document, the term "globally available broadband" loosely refers to a class of services that offer bi-directional connectivity that is at least comparable to ISDN in speed. Furthermore it should cover all or nearly all of the areas required for the messaging required by GNSS trackers.

In land based applications, wireless broadband is a reality with the emergence of GPRS and 3G services at subdued communication costs[70].

In maritime application domains, it is likely that such a migration will eventually happen, because of the desire of seafarers to remain in contact with their families, friends and associates. This is compounded with economic drivers such as the amortisation of, and competition between, satellite communications networks. As of June 2009, a brief survey of pricing indicates that airtime rates for Inmarsat FleetBroadband cost about 6.50 USD / MB.

When such broadband services replace existing maritime satellite services, the tracker data reports could be migrated to TCP/IP-based connections, which would use Transport Layer Security (TLS) with mutual authentication (i.e. both client and server side certificates) and a trusted path for certificates.

The trackers need not have their own satellite communications terminal; instead they can use a vessel-wide connectivity hub and router to decrease the tracker's running costs. Indeed emerging Inmarsat FleetBroadband routers incorporate standard 802.11 b/g/n technology for wireless connectivity. In terms of security, the price for depending on a vessel-wide router is that it introduces an additional point of failure.

[70] In 2010, the author used land-based 3G services for a one-time fee in the order of 15 EUR, and for monthly fees in the order of 20 EUR, for up to 300 hours of connectivity per month.

What are the economics of encrypted and authenticated tracker messages, when using globally available broadband? The absolute marginal cost is quite low:

- The message will be at least 19 bytes long. Adding some overhead, one may assume that 32 bytes are fed into a 128-bit block cipher encryption such as AES.
- Let each message have a full SHA-224 signature, which adds 28 bytes to the message length.
- Because the authentication code roughly doubles the length of the VMS message, the cost for a authenticated message would be roughly double that of the initial message.
- A single authenticated message would have a length of about 60 bytes. At 6.50 USD per MB, each message would cost about 0.03 US cent. Put otherwise, one US cent affords 27 authenticated messages at 60 bytes each.
- When establishing a shared secret using TLS, the client and the server exchange their certificate chains. These certificate chains create most of the traffic between both parties. A digital certificate has a size of about 1 KB. If both client and server are certified by a chain of trust of three levels, then certificate exchanges and other overhead may add up to about 8 KB of traffic. At 6.50 USD per MB, a full key establishment handshake would cost about 5 US cents.
- The shared secret key can then remain valid for a week. When used on the internet, TLS foresees timeouts of less than one hour. For low-bandwidth communications these timeouts can be made longer, such as to decrease the overhead of subsequent key renegotiation.

TLS incorporates many different key exchange protocols, notably some that are also mandatory by FIPS 140-2 levels 3 or 4.

TLS offers a key re-negotiation feature where parties that previously shared a secret key can generate a new secret key with reduced overhead. This key re-negotiation feature is vulnerable to a “man in the middle” attack, but the exploit assumes that

- the adversary is in possession of a client certificate trusted by the server , and
- during initial negotiation the server does not immediately require the certificate of the client

With mutual authentication and a trusted path for certificates, the possibilities of the adversary are limited. If still deemed necessary, the TLS re-negotiation feature can be disabled in the GNSS tracker and at the control centre.

Remote firmware updates

The argument can be made that the firmware of the GNSS tracker should be updateable, in order to mitigate the risk of buffer overflow exploits. Also, should the tracker depend on any cryptographic certificates, these may need to be replaced. If the manufacturer does not implement “firmware over the air” (FOTA) or “over the air parameter administration” (OTAPA) type functions, then maintenance staff needs to physically reach the tracker. Previous experience in fisheries pilot projects (SHEEL, CEDER) showed that it is not always easy to make firm appointments between an end user and the qualified service staff, and that these appointments are often missed.

In order to mitigate such difficulties, one could explore the possibilities offered by FOTA and OTAPA, which uses encryption. Public key authentication and/or encryption would ensure that the risk from tampering with firmware updates is minimal. It would also be preferable to have the manufacturer contact the end user using other means of communication, in order to ensure that in case of issues with the update, the end user can quickly obtain a replacement tracker. For extra security, one could consider encrypting the firmware update twice, once by the manufacturer, and once by a control agency.

Considering the economics, clearly a firmware update cannot be transmitted by e.g. Inmarsat-C or other similarly priced services. A Linux-based firmware update would have a size of about 4 MB, and would therefore be too expensive to be sent by Inmarsat-C at 10 USD/KB.

Therefore, if the GNSS tracker must use low-bandwidth communications satellite to report on position, it would need to receive firmware upgrades by other means, such as by a mobile telephony data service (UMTS), once it is within GSM network range. If instead the GNSS tracker uses FleetBroadband, then communication costs, at about 6.50 USD/MB, imply that an additional UMTS data module unnecessary.

Concluding Chapter: A proposed hardened GNSS tracker

Author: U. Kröner

Design, production, and costs

Below, two different possible designs are considered. For use in the near future, a "high-security" variant would be used on sensitive targets, while a "low cost" variant would protect lesser assets.

- **GNSS receiver, low-cost variant**: The design would require a low cost GPS receiver module, typically available for about 20 EUR apiece, augmented by an energy meter (J/N or AGC readings), receiver clock drift checks, and an IMU. This low cost variant could be used in the transportation of dangerous goods than the ones cited below, and initially in maritime application areas. It can be rolled out without further delay.
- **GNSS receiver, high security variant**: This variant adds signal authentication measures on top of the low cost variant. In order to implement the "hybrid Psiaki-Lo method", the design would require a custom GPS receiver that is able to sample parts of the L1 P(Y) signal. The method requires the deployment of a ground station infrastructure and the transmission of messages containing the signal samples. The transmission medium could be SBAS, BGAN, Iridium, or UMTS. The "high security" variant could be phased in progressively as a replacement for the "low cost" variant, to protect transportation of highly dangerous goods such as UNECE Division 1.1 (mass explosion), Division 2.3 (toxic gases), and Class 7 (radioactive material).
- **Use of an appropriate IMU**: In order to defend against GNSS meaconing and spoofing, and in order to achieve "dead reckoning" abilities, the GNSS receiver of a hardened tracker would be augmented by an Inertial Navigation Unit. While a MEMS based IMU may cost as much as 1000 EUR (appropriate for highly dangerous goods), lower quality devices based vibrating gyros would cost perhaps 20 EUR per unit (appropriate for e.g. fisheries monitoring and road toll systems). The price of MEMS INS units is projected to fall drastically over the next 5 years, so trackers built by e.g. 2020 can take advantage of increased sensitivity.
    - As a cost-effective GNSS receiver hardening measure, the low-cost design would use an IMU based on vibrating gyroscopes in the near future.
    - The data from the IMU can supplement the GPS data, in case the GPS readings are temporarily unavailable. Unavailability can occur spuriously in the maritime domain due to interference of marine electronics devices. It could also be the result of malicious GPS jamming.
    - Consider the case when the GPS receiver is tracking and indicates movement that, after filtering out false positives using hypothesis testing, is inconsistent with the IMU. Then the tracker logs the event, logs IMU and GPS position data in parallel until the IMU position error margin covers the GPS position, and sends a message containing a "GPS spoof bit" to the control centre[71]. The control centre then has the opportunity to follow up with an inspection and/or phone call.
- **Physical volume protection**: An active monitoring system could be implemented as follows:
    - The electronics of the tracker would be "potted" with epoxy resin in order to achieve FIPS 140-2 levels 3 or 4, especially for the "high security" variant. The cost of the potting material would be no more than 1 EUR.
    - The tracker would use a secure controller. This secure controller establishes two asymmetric key pairs, one for communications (the communications keys), and one for securing its log file, which also contains any alarm conditions (the log file keys). Elliptic Curve Cryptography keys are preferred because their smaller size allows for

[71] In this event, the tracker does not wipe any private keys.

quicker key destruction in case of an alarm. The public keys are signed by a Certification Authority. The secure controller would cost about 6 EUR apiece.

- The tracker would use two batteries: A smaller CMOS battery is potted near the centre of the tracker, and a larger rechargeable battery that can be removed and exchanged, is attached outside of the resin potting but inside of the protected volume. The chemistry of the larger battery would be based on NiMH, which supports temperatures of up to -40 °C. If the outside battery reaches a critical status, then an alarm condition is raised. The batteries would cost no more than 20 EUR.
- The tracker has two connecting wires to the outside, which charge the larger battery at +5V DC. If the voltage drops, a circuit switches to the rechargeable battery. This circuit would cost perhaps 3 EUR.
- A colour detection integrated circuit, such as presented in the "tie dye" seal, would detect opening by the adversary. The colour sensor may cost up to 5 EUR.
- A temperature sensor (on the secure controller) would check for abnormal heat or cold; both absolute temperatures and temperature changes would be monitored. An alarm is raised as appropriate.
- A USB port would reside inside of the protected volume, and allows full control over the tracker for the manufacturer. When the USB port is connected, an alarm is raised.

- **Removal protection**: An active removal protection can be implemented using e.g. active transponder bolt (hereafter labelled "bolt"), which in case of tampering sends a distress message to the tracker. The tracker periodically communicates with the bolt(s) to ensure that each one is still present. The authorities can send a message to the tracker, asking it to communicate with its bolt(s). Each bolt contains a secure controller with NFC or BlueTooth™ capabilities, a battery, a light or colour detector, and a temperature sensor. It would cost an estimated 25 EUR apiece.
  Note that the above is just one example for removal protection, and indeed other possibilities are sought that represent alternatives in various application domains. Also, active monitoring does not obviate the need for physical inspections.
- **Buffer Overflow protection**:
  - The tracker has an embedded Linux EAL4 compliant operating system that is hardened against buffer overflow exploits. In order to guard against yet undiscovered exploits, the firmware can be updated remotely using FOTA. Any firmware updates are encrypted and signed. Updates also contain changes to any certificate authorities, and any signed certificates of the tracker. A Linux-enabled OEM board may cost up to 50 EUR apiece.
- **Outgoing message protection**:
  - In case of **maritime applications**,
    - The tracker uses a global broadband transponder, allowing TCP/IP based communications worldwide at speeds not lower than an ISDN connection.
    - The above does not mean that a dedicated broadband connection is necessary. Instead a ship-wide broadband connection can be used, if the connection is made available using wireless access. The global broadband gateway could be fitted with a 802.11 b/g router. The corresponding endpoint device, a WLAN endpoint dongle, is available for less than 10 EUR.
    - The maritime communications terminal is often bundled with a tariff plan. Depending on the tariff plan, the communications terminal itself is either "free" (in case of which voice communications are expensive) or costs 1000-3000 EUR per month (in case of which voice communications are considerably cheaper). Volume discounts can be negotiated, and costs can be shared between the vessel skipper and the authorities.
  - In case of **terrestrial applications**, an UMTS endpoint dongle is available for about 30 EUR. Commonly available UMTS service plans permit up to 5 GB per month of traffic for about 20 EUR per month.

- In any case, the tracker's data communications are fully based on TLS and TCP/IP, implying both encryption and reliable transmission over lossy media.

As projected above, total materials per-unit cost for a "low cost" hardened tracker would be about 170 EUR. Total materials per unit cost for a "high security" tracker would be similar, but such trackers require additional infrastructure set-up, as discussed in the chapter on the "hybrid Psiaki-Lo method".

As time passes, the adversary will gain additional capabilities. But with any luck, the defences would also become cheaper and more widely available. Therefore, a likely scenario is that in about a decade, the present "low cost" variant is no longer adequate, in case of which the "high security variant" may be used for consumer grade applications such as road tolling systems.

Lifecycle

Production site

The manufacturer should take care in securing the production and configuration site. The site would have the following features:

- **Access control**: only qualified staff should be allowed on the site where such trackers are produced and configured. The installation of CCTV cameras may be a useful complementary measure.
- **Hardware security modules**: The private keys that permit firmware updates must be secured, for instance using "hardware security modules", such as those used at "Certificate Authority" type Public Key Infrastructure sites.
- **EM shielding**: It also is necessary to shield the site, in order to defend against e.g. "Van Eck phreaking". To accomplish this, EM shielding with respect to EMSEC specifications such as NATO SDIP-27 Level B or C would be implemented, as appropriate given the site's location. It should be difficult for an outside adversary to direct a wireless communications link at the site, and/or to eavesdrop on communications inside of the site.

Configuration

At the production site of the tracker, the manufacturer configures the tracker as follows:

- Connect to the internal USB interface. The tracker reacts by wiping any key material.
- Ask the tracker to generate two key pairs. The private keys are stored on the secure controller.
- Load the tracker with a bootstrap server and Certificate Authority certificate.
- Ask the tracker to provide its public keys. Send the public keys to a Certificate Authority.
- The internal USB interface can then be disconnected. The volume is then closed, enabling the volume protection seal.
- From thereon, the *tracker is sealed*. Any opening of the volume should trigger an alarm condition.
- Connect the tracker to a power supply and ensure that an internal shielded wireless service is available.
- When the Certificate Authority responds with the signed certificate, perform a firmware upgrade of the tracker. An encrypted[72] firmware upgrade is then sent to the device using wireless communications, and is encrypted using both client and server side certificates.

At that point, the *tracker is ready to be installed*. It can then leave the manufacturer's production site.

[72] The manufacturer, the control agency, or both can encrypt the firmware upgrade. The reason for double encryption would be to mitigate the risk of the adversary stealing a private key.

Once the tracker has left the production site, it would be stored at a safe place, and should be attached to a power supply at least intermittently. Failure to charge the device would trigger an alarm condition, implying refurbishment. Ideally the devices are stored at a site controlled by a government control agency.

Each bolt securing the tracker is configured as follows.

- Ask a secure controller to produce a public/private key pair. Generate a certificate for the public key, and have a CA sign the certificate.
- Load the secure controller with the current time and date. Assemble the bolt. While it has not been fitted to an asset, the bolt should remain silent, meaning it would only use the battery to power its clock. This prolongs shelf life.

The manufacturer of the tracker need not be the same entity as the manufacturer of the bolt. The bolts are stored at a premise controlled by a manufacturer or by a government agency.

Installation

The tracker is installed by qualified service staff under the presence of an inspector, as follows:

- The control agency selects a spot on the asset to be tracked. For instance this could be a spot on top of a ship's cabin.
- The tracker is mounted in place. The power connection wires are attached to a power supply.
- The bolt(s) secure(s) the tracker to the asset. When a bolt is locked, it becomes active, meaning that it can be queried.
- The service staff will have the public certificate(s) of the bolt(s). The integrity of each certificate is verified by checking the certificate's chain of trust and the certificate revocation list. The bolt is queried, and will reply with a signed message including the time and date. The certificate is then used to verify the message signature.
- The certificate(s) are sent to the manufacturer of the tracker. The manufacturer will create an encrypted firmware update for the tracker, which it will send back to the tracker using wireless communication.
- The control agency sends a special configuration message to the tracker, which causes the tracker to update its firmware.

At that point, the tracker is installed, armed and operational.

The tracker therefore knows it(s) bolt(s), but the reverse is not true. Each bolt simply communicates with any reader that issues a request.

Operation

The device sends messages as described in the chapter on authenticated messaging.

If the communications private key is present, any communications is made using TCP/IP transport layer security (TLS), with client certificates, and optionally using FIPS 140-2 levels 3 or 4 approved key agreement schemes. If the secure controller has wiped the communications private key, then the position messages are no longer authenticated. This should trigger an inspection as soon as possible.

At a periodic time interval (e.g. every day) the tracker queries the bolt(s). If the bolt(s) fail(s) to respond, then this is flagged in a message to the control centre. In that case, the control centre communicates with the vessel, asking it to remove possible sources of interference. Afterwards, the control centre will ask the tracker to re-scan the bolt(s). The vessel should be inspected if the bolt(s) fail to respond repeatedly, or if the bolt(s) issues a message indicating that it has been breached.

Measures upon reaching any alarm condition

Alarms are triggered by integrity failures of the tracker itself, such as by physical breach of the tracker's housing, extreme temperature, or by depletion of all of the tracker's power sources. If such alarms happen, then the tracker wipes its key material. (Repeated non-responses of a bolt or breach of bolt do not constitute alarm conditions for the tracker to wipe its key material.)

The tracker has (at least) two private keys, one for communications with the control agency, and one for signing its log files. Upon alarm, the tracker's secure controller proceeds as follows:

1. It wipes the communications private key before any other action.
2. It signs the event log using its log private key (which is a private key different from the communications private key).
3. It wipes the log private key.
4. The tracker then sends its signed log file to the control agency.

In order to accomplish the above tasks, the secure controller uses any available energy, including the smaller CMOS battery and/or a backup capacitor if provided.

A bolt proceeds in a similar way, except that it attempts to communicate with its tracker, not the control agency.

Decommissioning and Refurbishment

The tracker is decommissioned by qualified staff as follows:

- The integrity of the bolt(s) and of the tracker is verified. They are queried and should respond with signed messages including date and time.
- Each bolt is removed. If the tracker is still armed, it will react by sending messages about the bolt(s) being disabled. The tracker thus documents its removal from the asset at the time of decommissioning.
- The tracker is shipped to the manufacturer for a post-mortem exam.
    - The manufacturer opens the device, and connects to it using the internal USB port.
    - The manufacturer downloads the log file, and performs diagnostics.
- If the tracker is not damaged beyond repair, it can afterwards be re-commissioned, starting at the configuration step.

Depending on their design, the individual bolts can also be re-furbished. To accomplish refurbishment, the bolt is opened, and the secure controller is extracted from the bolt. Then the manufacturer re-commissions the bolt by starting at the configuration step.

Use

Such hardened GNSS trackers can be used in a number of contexts.

In maritime domains,

- Merchant vessels shipping valuable and/or sensitive cargo can incorporate such a tracker for LRIT or AIS reporting purposes. If used for AIS, the tracker sends IMO-agreed VHF signals like any other AIS system. However the tracker would also send authenticated position messages via satellite communications.
- For fisheries, a hardened tracker could complement existing trackers for high-value fisheries or for larger vessels.

For terrestrial applications, a hardened GNSS tracker would use 3G services for data communication, and would incorporate an INS for dead reckoning abilities.

- The transportation of dangerous goods could presently use such a tracker, in particular in cases in which the use of the Galileo PRS is undesirable.
- Concerning the digital tachograph, a "low cost" hardened tracker variant may be a useful complement to, or replacement of, the present-day equipment that connects to a truck's gearbox.
- In about a decade, a "low cost" hardened GNSS tracker may be useful on the mass market, for pay-as-you-go car insurances, highway tolling charges, or for issuing urban congestion fees. The use of such a hardened tracker would however only be warranted if the adversary fabricates mass-market devices that defeat existing measures.

Concerning air traffic applications, most navigation devices are not "GNSS tracking systems" in the strict sense, and also make use of sensor fusion between an INS and a GNSS component. However, given that it may be possible to spoof an airliner's GPS components [B. Ledvina, personal communication], it may be warranted to insert safeguards against GNSS spoofing into airplane navigation equipment. This is the avenue taken e.g. by the "GPS Assimilator" product of Coherent Navigation. The crucial difference between the latter type of products and existing GNSS/INS products is that the former does not blindly trust the GNSS readings.

If at some point in time, GNSS spoofing becomes commonplace, and *such an attack may be carried out by "script kiddies" using software downloaded from the Internet* [Scott03], then a hardened GNSS tracker may be welcome in all domains that involve position-based GNSS services.

Annex: Dual antenna GPS receiver defence prototype

Authors: U. Kröner, J. Fortuny-Guasch, R. Giuliani, D. Shaw, D. Symeonidis

Abstract

The JRC IPSC action groups FISHREG and CORSA jointly built a prototype dual-antenna GPS receiver. The receiver is built on a Software Defined Radio platform using commodity hardware. The software is based on a modified version of an open source GPS software receiver. One then measures the difference in carrier phase between the antennae versus time, in two settings. The first setting is with view of the sky (outdoors), the second uses a GPS repeater (indoors), the latter to mimic a GPS spoofer. Although the measurements were not of the length or quality as described by Montgomery, Humphreys, and Ledvina [Montgomery09], the experiments show that research on spoof detection can be conducted using Open Source Software Defined Radio hardware and software.

Introduction

The Global Positioning System (GPS) is a U.S. space-based global navigation satellite system. It provides reliable positioning, navigation, and timing services [GPS-WP]

A GPS receiver calculates its position by precisely timing the signals sent by the GPS satellites high above the Earth. Each satellite continually transmits messages which include

- *the time the message was sent*
- *precise orbital information (the ephemeris)*
- *the general system health and rough orbits of all GPS satellites (the almanac).*

The receiver utilizes the messages it receives to determine the transit time of each message and computes the distances to each satellite. These distances along with the satellites' locations are used with the possible aid of trilateration to compute the position of the receiver. [GPS-WP]

Civilian GPS receivers are vulnerable to a threat known as GPS spoofing. In GPS spoofing, an adversary sends a fake GPS signal, which is then picked up by a victim GPS receiver.

In the InsideGNSS cover article of March/April 2009, the authors Montgomery, Humphreys, and Ledvina *demonstrate the use of a dual-antenna receiver that employs a receiver-autonomous angle-of-arrival spoof countermeasure. […] It is based on observation of L1 carrier differences between multiple antennae referenced to a common oscillator.* [Montgomery09]

Materials and Methods

A dual-antenna Software Defined Radio (SDR) was devised, by fitting a single "Universal Software Radio Peripheral" [EttusUSRP1] with two DBSRX daughterboards [EttusUSRP2], which together represent an open source hardware component able to sample the L1 frequency. Two low-budget active GPS antennas ("Jinchang GPS", type JCA002-GAACZ-D, www.jinchanggps.com) were fitted on a North-South axis at about 1.2m distance to each other. Visibility was hampered by tall trees, mostly pine, and some built-up areas, together obstructing an estimated 45% of the sky. Once the outdoor experiments concluded, identical antennas were used with a custom-built indoor GPS repeater for an indoor experiment.

The GPS-SDR software [Heckler-GPSSDR] was modified to be able to run two GPS processes on the same machine. Each process uses the data feeds from one antenna, in order to arrive at a Position, Time, and Velocity solution (PVT). Further the GPS-SDR software was changed to record the processes' raw measurements, which contain GPS L1 carrier phase observables. Care was taken to ensure that the carrier phase observables were measured at the same "epoch[73]" in both GPS processes. The carrier phase observables were fed in real time into a custom-written C++ programme. This programme is able to produce epoch-antenna differencing and triple-differencing observables [Rizos]. One then charts the epoch-antenna differencing observable.

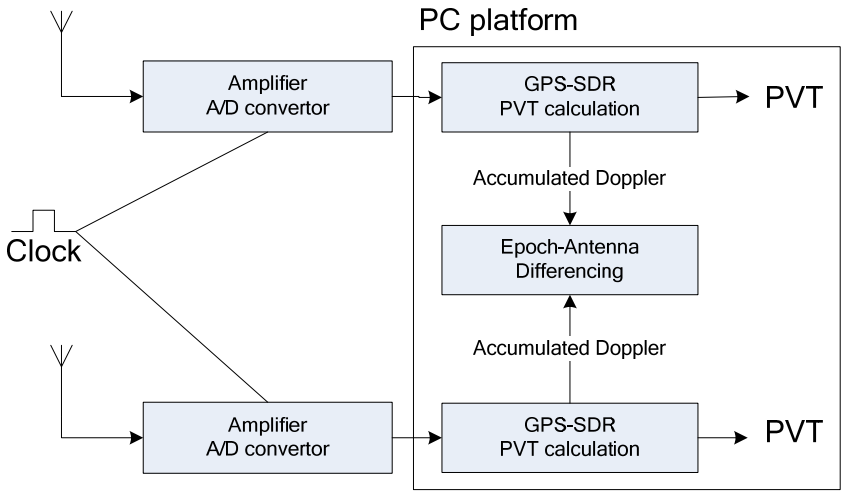


Figure 34: Dual antenna system with common clock

Note: Commodity GPS receivers cannot measure the carrier phase directly, and the software GPS-SDR receiver is no different in that aspect: After the USRP converts the signal from the L1 centre frequency to baseband, the GPS-SDR software receiver calculates the Doppler shift in the correlators. It then reconstructs the actual carrier phase by using the Doppler shift. This step introduces additional carrier phase measurement errors.

In practice, the "antenna-epoch differencing" observable was built as follows.

Algorithm 1: Build undifferentiated carrier phase

Given two software receivers RA and RB, operating each on an antenna A and B, calculate the undifferentiated carrier phase for RA and RB, at time t, for the satellites that antenna A and B have in common, as follows:

1. For each space vehicle (a.k.a. GPS satellite) SV^i that the receiver has acquired, obtain the integrated Doppler drift at a given epoch. Store N recent measurements in a cache of Meas(SV^i, antenna, epoch).
   Note: GPS-SDR was changed to output Measurement_M data; one file per antenna was used. When reading the results, it is only necessary to keep the last 3-4 measurements for each (SV, antenna) tuple.
2. Eliminate satellites from the measurement that are not acquired by both receivers, yielding the set of common observations for each SV, SVC^i.
3. Find an epoch (*t*) for which there exist measurements Meas(SVC^i, antenna, epoch) just before and just after the epoch.

[73] In GPS parlance, the epoch represents the GPS time synchronised millisecond.

4. Calculate Meas(SVC^i, antenna, *t*) by linearly interpolating between Meas(SVC^i, antenna, epoch1) and Meas(SVC^i, antenna, epoch2), where epoch1 is the most recent measurement right before *t*, and epoch2 is the most recent measurement right after *t*.
   Note: With the GPS-SDR, one receives measurements each 100 milliseconds, so linear interpolation of the accumulated Doppler come with a negligible numerical error.

Algorithm 2: Antennae-epoch differencing

Given algorithm 1 and the software receivers, differentiate between epochs and antennae as follows:

1. Let *t0* be some fixed point in time. Calculate Meas(SVC^i, antenna, *t0*) according to algorithm 1. Store *t0* and the above measurements, implying that Meas(SVC^i, antenna, *t0*) remain fixed.
2. Let *t1* be the current, running GPS epoch. Calculate Meas(SVC^i, antenna, *t1*).
3. Calculate delta_epoch(SVC^i, antenna, *t0, t1*)
   = Meas(SVC^i, antenna, *t1*) - Meas(SVC^i, antenna, *t0*)
4. Calculate delta_epoch_antenna(SVC^i, A, B, *t0, t1*)
   = delta_epoch(SVC^i, B, *t0, t1*) - delta_epoch(SVC^i, A, *t0, t1*), where A and B designate two antennae.

Interpretation

If the double difference is applied to the carrier phase, then the double difference observable has a direct and physical relationship with the changing geometry between antennae A and B, and the observed GPS satellites. As the satellite moves overhead in the ENU (East-North-Up) reference frame, the number of waves counted between antenna A and B changes.

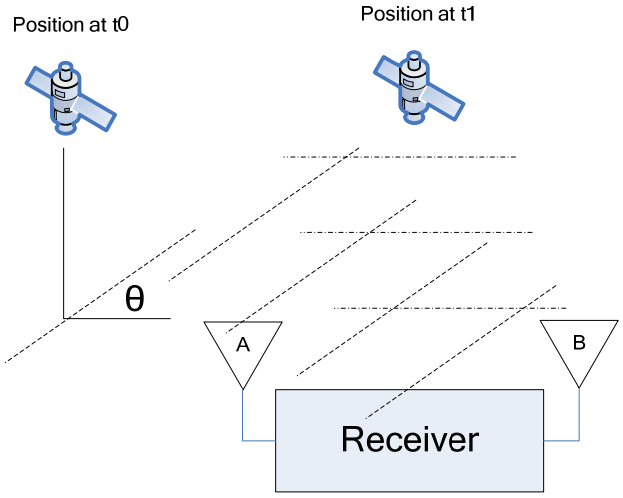


Figure 35: Satellite movements imply carrier phase changes

Considering just one satellite, assume that A-B are on a horizontal plane and lie in the same axis as the arriving satellite.

One has

$$d\varphi_i = \lambda\, |ab| \cos(\theta)$$

where

- φ designates the carrier phase, in units of L1 cycles
- $d\varphi_i$ is the difference in carrier phase between A and B
- λ is the apparent L1 frequency, including Doppler shifts
- |ab| is the distance of A to B
- θ is the azimuth or elevation in the sky of the satellite as seen from A or B. For practical intents and purposes, θ is the same whether seen from A or B.

Results

Two kinds of result sets are presented below, namely those obtained "outdoors" from two antennae spaced apart by 1.2m, and that obtained by two similar antennae, which received a signal radiated by an indoor GPS repeater.

First, three separate experiments were conducted "outdoors". As can be seen from the graphs representing the L1 phase differential versus time, it is apparent that the satellites are moving with respect to the baseline formed by antennae A and B. Even in the presence of significant phase noise, the effect of the satellites' movement can be spotted in about 60 seconds. This is because the phase noise has a different frequency and amplitude with respect to the effect created by the moving satellites.

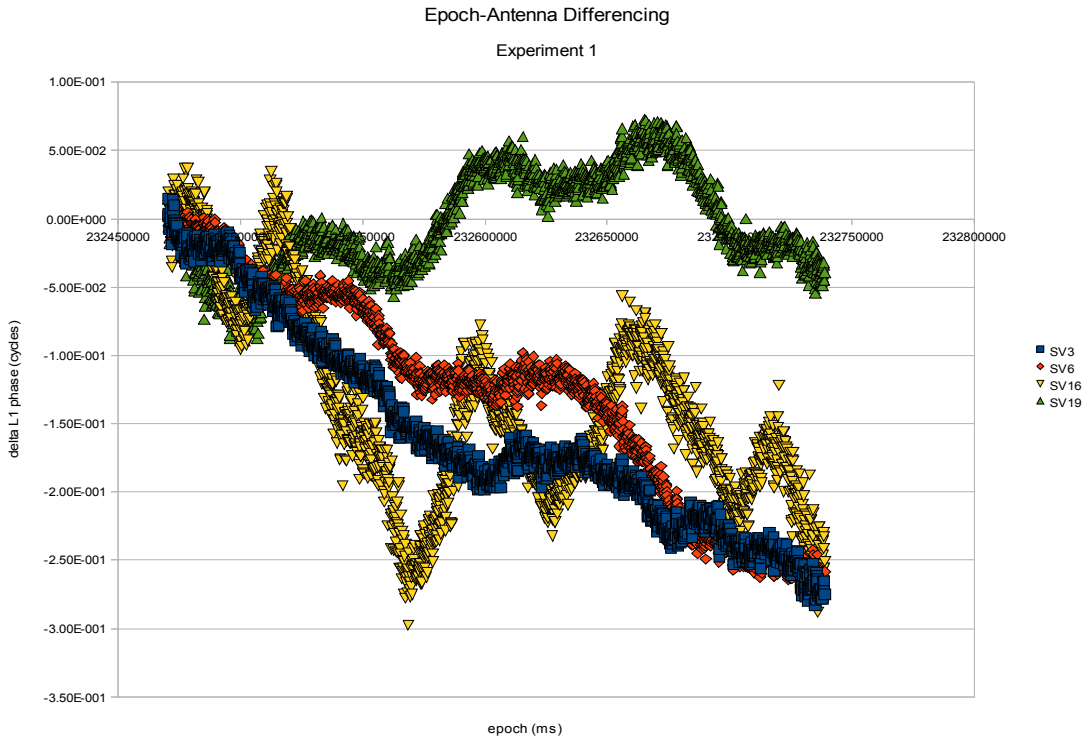


Figure 36: Experiment 1 on 29/SEP/2009, ended 16:38 GMT, duration 268 seconds

Concerning experiment 1, out of four observed satellites (SV 3, SV 6, SV 16, and SV 19), all but SV 19 showed phase shifts of magnitude of 0.25 L1 cycles towards the end of the experiment.

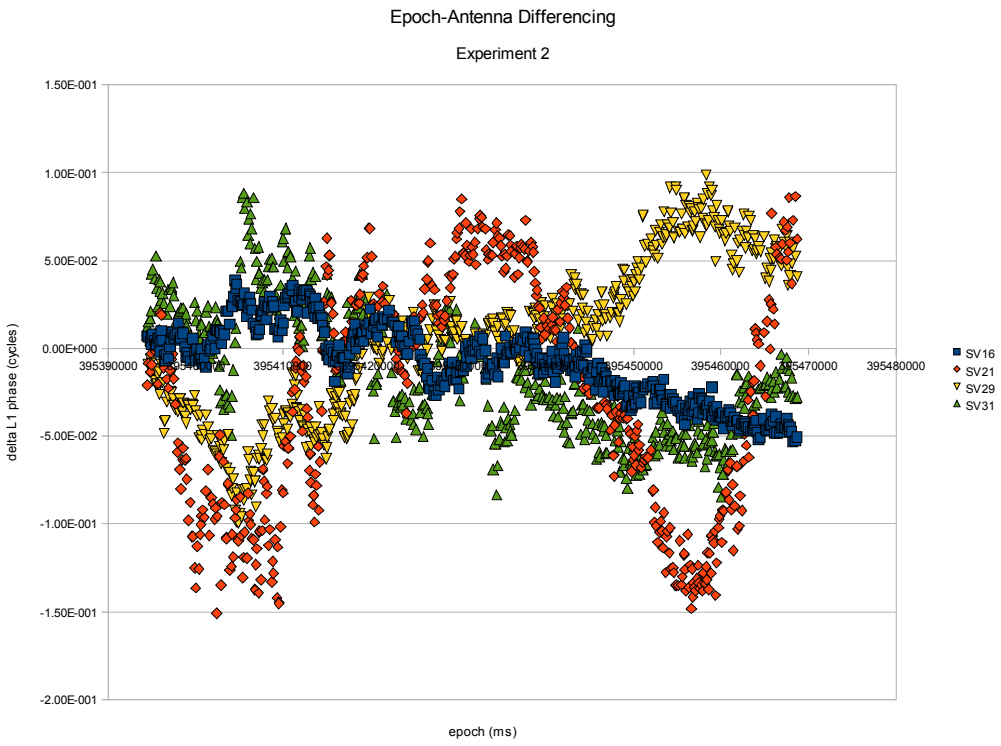


Figure 37: Experiment 2 on 01/OCT/2009, ended 13:53 GMT, duration 74 seconds

Concerning experiment 2, out of four observed satellites (SV 16, SV 21, SV 29, and SV 31), two (SV 16 and SV 29) showed clear trends towards phase shifts of 0.05 L1 cycles towards the end of this rather brief experiment.

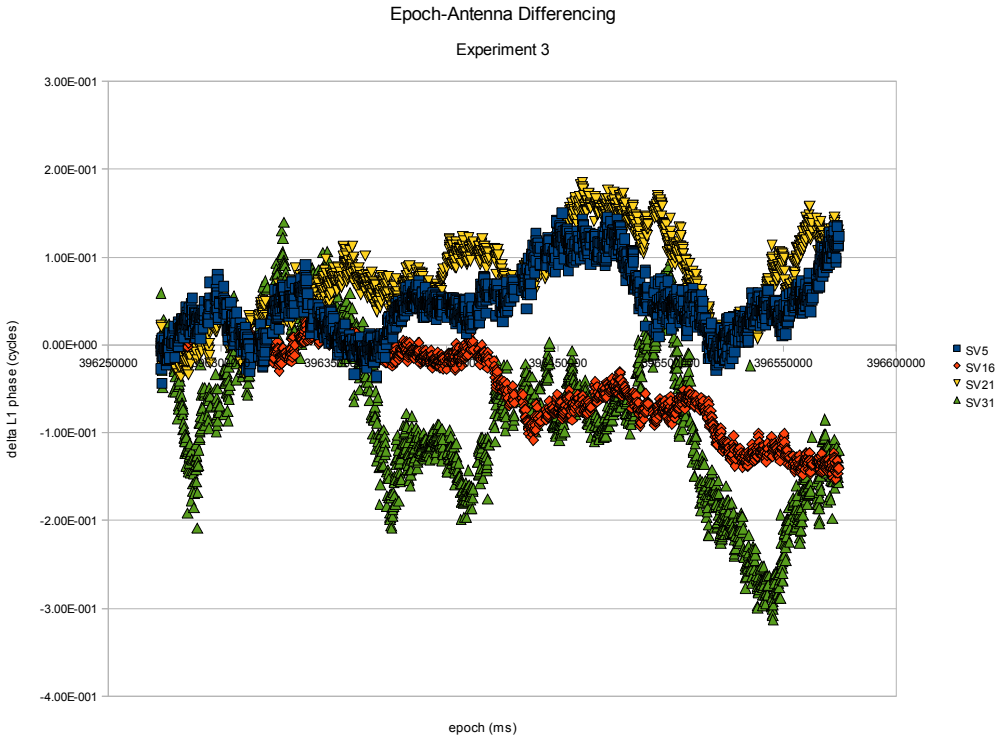


Figure 38: Experiment 3 on 01/OCT/2009, ended 14:09 GMT, duration 301 seconds

In the longest outdoor experiment #3, out of four observed satellites (SV 5, SV 16, SV 21, and SV 31), all of them showed phase shifts of 0.1 L1 cycles.

Summarising the outdoor experiments, the satellites phase shift is consistent with the earlier experiment by Montgomery, Humphreys, and Ledvina [Montgomery09], who for three out of four satellites, observed a phase shift of 0.25 L1 cycles in about 300 seconds. The authors' GPS antennae were spaced apart by 1.46 metres, in the present case they are spaced apart by a comparable 1.2 metres. In the document [Montgomery09], the reason for the observed phase shift is clearly laid out.

But is this in fact measuring the same phenomenon as described in the document [Montgomery09]? One way to confirm the measurement is to place the antennae indoors, within range of a GPS repeater, and then to perform another measurement.

The indoor measurement is indeed consistent with the phenomena described in the document [Montgomery09]:

Figure 39: Experiment 4 on 02/OCT/2009, ended 08:33, duration 341 seconds

In indoor experiment #4, out of four observed satellites (SV 9, SV 12, SV 26, and SV 27), none of them showed a clear tendency to shift their phase. Instead one seems to be observing some form of noise that has amplitude of 0.1 L1 cycles, without any trend, despite the experiment's comparatively long duration of 341 seconds. Noise is larger than the one in observed in an outdoor type situation. Indeed, there were noise outliers in the phase measurements of SV 12, of which the receiver subsequently lost track. The outliers appear to be distributed more or less evenly on both sides of a zero trend.

Note: In the indoor experiment, it does not matter how much the antennae are spaced apart. All GPS receivers within the area of influence of an indoor GPS repeater (or outdoor spoofer) ideally arrive at the same position and velocity solution. The GPS time of affected GPS receivers appears shifted into the past, by the time it takes for the signal to travel any additional path from the GPS repeater's

receiving antenna[74]. However, the aforementioned method differentiates between epochs, and the antennae do not move during the experiments. Therefore any constant time delay is eliminated.

The findings are summarised as follows:

- In an outdoor type situation, two GPS antennae spaced apart by a few multiples of the L1 carrier wave length, will invariably pick up trends in the L1 carrier wave phase shift, when differentiated between antennae and epochs. After 60 seconds, a first trend becomes apparent. After about 300 seconds, antennae spaced apart by 1.2 metres will pick up a trend in excess of 0.1 L1 cycles for a majority of satellites.
- In an indoor type situation, no matter the distance between two antennae, the same method will invariably fail to pick up any trend in the L1 carrier wave phase.
- The indoor GPS repeater introduces an additional phase noise. However a true GPS position spoofer may behave differently.

Discussion

Hardware

The initial plan was to use a phase coherent antenna array, with two antennae spaced apart such as to obtain a compact spoofing defence, implying distances of less than 30 cm. This is about the same order of magnitude as the L1 wavelength of 19 cm. There is reason to believe such an approach may lead to results: With angular velocity of GPS satellites being about 15 degrees per hour, and phase coherent arrays being able to detect a 1 degree move, an observation of several minutes should lead to sufficient data to validate or invalidate the presence of a spoofer. The phase shift would then need to be filtered using statistical methods.

A compact spoofing defence, with two antennae located within a shielded environment, is vital in a practical setting.

- If the antenna elements are exposed, a spoofer can isolate each antenna element, and feed two separate spoofed GPS signals, via two separate transmit antennae. This was previously explored in references [Humphreys08] and [Montgomery09]. These two spoofed GPS signals would in principle need to be phase-coherent. However even with some phase noise it would still achieve the desired results, as the target spoof detection technology, based on antenna diversity, must invariably account for the phase noise introduced by its own antenna array.
- If on the contrary, the antenna elements cannot be exposed separately without physically opening a sealed multi-antenna device, an attack as detailed in reference [Humphreys08] and [Montgomery09] would be faced with the additional task of beamforming. Otherwise the victim antennae will not pick up the separately emitted spoofed signals, but some mix between the two signals, which is undesirable from the vantage point of the spoofer.
- Are there physical volume enclosing seals that render the job of physically breaking such a device quite expensive? This is indeed the case, and that this is widely used in the banking and military circles. Nearly all of them use a series of OEM kits from a single vendor, namely the Scottish division of W.L. Gore.

[74] From a purely logical point of view, any information streamed from the GPS repeater is the same as (or a distorted version of) the signal captured by the repeater's receiving antenna. This is so because that is the very purpose of a GPS repeater. GPS receivers affected by the repeater must use information encoded in the repeater's signal to arrive at a Position, Velocity, and Time solution, because that is the definition of "affected". Because the information essentially is the same between the affected GPS receivers and the repeater's GPS receiving antenna, position and velocity must match, save for GPS receiver and GPS repeater imperfections. As for the 3rd component of a PVT solution, there is a time delay introduced by the signal travelling an additional distance, so affected GPS receiver time appears shifted into the past.

In conclusion of the above, a sealed multi-antenna device with a small form factor would be quite interesting as a deterrent to spoofing. It hinges on proper sealing, which is available, and on a phase-coherent antenna array for detecting relatively small phase changes.

Initial hopes of using the DBSRX boards as a phase-coherent antenna array were dashed, as it is impossible to achieve such an approach with the above hardware. In an experiment of CORSA colleagues that attempted to determine the Angle of Arrival of FM signals using the USRP and the DBSRX daughterboards, the colleagues found that two DBSRX daughterboards are not phase coherent. For a 100 MHz signal fed directly from an oscilloscope, the signal's angle of arrival remains fixed for some time, but then jumps randomly. This may be due to the way in which the DBSRX boards upscale the 64MHz USRP clock to the L1 frequency of 1.57542 GHz. For the above reasons, one should encourage the makers of the USRP and/or other radio engineers to build daughterboards capable of being used in a phase coherent array setting.

Because of the lack of a phase-coherent hardware, it was not attempted to co-locate two antennae within one L1 cycle length. This reduced the approach to reproducing the results obtained by Montgomery, Humphreys, and Ledvina [Montgomery09], but using open source software and hardware.

In summary, in view of the results obtained, and in spite of the aforementioned nature of the DBSRX daughterboards, clearly the hardware can be used for experiments into antenna diversity.

Software

A possible candidate for open source software for GPS PVT determination, using an URSP, is the GPS-SDR package [Heckler-GPSSDR]. Apart from the USRP hardware, it requires GNU Radio version 3.1.x. [GR1] to run. The GPS-SDR uses UNIX pipes and threads to achieve functionality that in hardware would be implemented using circuit boards and microchips. Its minimalist approach makes it well suited to study the inner workings of GPS receivers. With some modifications, it is possible to extract the integrated carrier phase, reconstituted from the GPS receiver's Doppler measurements, from the GPS-SDR and to write it into a file. In plain English the "integrated carrier phase" stands for the number of waves that the GPS received from a particular satellite, since some arbitrary time in the past.

A limiting factor experienced during work with the GPS-SDR is that it occasionally loses the GPS fix. Compounding the problem, if one runs two separate programmes on the same computer, with data streams from individual antennae, the two GPS processes sometimes lose their GPS fix at different times. This could be due to the antennae, the USRP/DBSRX hardware, the host computer's USB interface, the GPS-SDR software itself, or insufficient available computing power, or of a combination thereof, to name some of the possible causes. Because of the above, it was not possible to record a longer time series of integrated carrier phase: When both programmes again achieve a GPS fix, the difference between both carrier phase measurements is commonly distorted by a large amount of cycle slips.

Binomial Bayesian inference of posterior probability of spoofing

Concerning the epoch-antenna carrier phase difference, the characteristics of the indoor versus the outdoor measurements suggest a Bayesian method for testing for the presence of a spoofer.

In a static scenario, consider what happens when one fits a linear model with a trend to the measurements.

- Outdoors, with antennas spaced 1.2 metres apart, one expects that a majority of satellites have a trend with a slope that according to preliminary tests, is between 1E-5 and 1E-4 in absolute terms.
- By contrast, indoors, the slope will typically be smaller than 1E-6 in absolute terms, for all satellites.

Algorithm 3: Binomial Bayesian inference of posterior probability of spoofing, based on slope spreads, for static scenarios (non-kinematic)

At present consider the following algorithm, for a static case:

1. Let TS denote the time series of measurements, in a matrix form, where the rows denote the time, and the columns denote the carrier phase differential measurements for a given common satellite SVC_i. Given TS, proceed with fitting linear models through the carrier phase difference for all valid satellites SVC_i. Let $TR(SVC_i)$ be the trend (slope) associated with these linear models.
2. Let the spread in trends be the magnitude of the difference between the maximal and the minimal trend value: $SPREAD = ABS(MAX(TR(SVC_i)) – MIN(TR(SVC_i)))$. If the two antennae can see at least 2 common satellites, then there will be exactly one SPREAD value for a given time series TS.
3. Design a Bayesian posterior estimation system, by
    a. Having a hypothesis H_0, which is that the device is located outdoors, and the complementary hypothesis denoted H_1, which is that the receiver is located outdoors. This implies that at all times, $P(H_1) = 1-P(H_0)$
    b. Finding non-overlapping ranges for the SPREAD value, typical of an indoor or outdoor scenario, SPREAD_INDOOR and SPREAD_OUTDOOR. These values depend on the distance between antennae A and B, and on the axis in which the antennae are configured.
    c. Determining four conditional probabilities, needed for Bayesian modelling:
    $P(SI|H_0)$ = P (SPREAD in SPREAD_INDOOR range | outdoor)
    $P(SO|H_0)$ = P (SPREAD in SPREAD_OUTDOOR range | outdoor)
    $P(SI|H_1)$ = P (SPREAD in SPREAD_INDOOR range | indoor)
    $P(SO|H_1)$ = P (SPREAD in SPREAD_OUTDOOR range | indoor)
4. At runtime, the receiver is switched on and initialises $P(H_0) = 0.95$, meaning it starts by believing with near certainty that it is located outdoors.
5. While the receiver is running, let it conduct a time series measurement TS for a given duration. With antennae in a North-South axis spaced at 1.2 metres, that duration would be approximately 90 seconds. Then fit linear models through each TS column, and calculate the SPREAD value.
6. At each measurement, use Bayesian calculus to update a posterior probability.
    a. In case less than two common satellites are tracked, no spread value will be available. In this case let $P'(H_0) = P(H_0)$.
    b. If the spread falls in the indoor range, calculate $P'(H_0) = P(H_0 | SI)$ as follows:
    $P(H_0|SI) = P(SI|H_0) * P(H_0) / (P(SI|H_0)* P(H_0) + P(SI|H_1) * P(H_1))$
    c. If the spread falls in the outdoor range, calculate $P'(H_0) = P(H_0 | SO)$ as follows:
    $P(H_0|SO) = P(SO|H_0) * P(H_0) / (P(SO|H_0) * P(H_0)+ P(SO|H_1) * P(H_1))$
    d. If there is a spread value, but it does not fall within either range, let $P'(H_0) = P(H_0)$. Such a spread value is obtained either when it is larger than SPREAD_OUTDOOR, meaning that cycle slips have occurred, or between SPREAD_INDOOR and SPREAD_OUTDOOR, in case of which one cannot make any statement.
    e. Whatever was calculated in steps a, b, c, or d, update $P(H_0) = P'(H_0)$.
    f. As soon as $P(H_0)$ drops beneath a threshold, such as 5%, assert that the system is not operating in an outdoor scenario. A GPS tracker should then either actively send a message, or set some information in the next periodic message to be sent.

7. Discard the time series TS and start over at point 5.

Note that "Bayesian inference techniques have been a fundamental part of computerized pattern recognition techniques since the late 1950s"[75]. Therefore it stands to reason that the above algorithm should work.

Algorithm 4: Extension of algorithm 3 to a kinematic environment, by dividing observations into slots

More speculatively, in a kinematic environment (dynamic case), instead of having a single measurement, one could adapt the approach by instead using a set of time series, where each time series corresponds to a range for the heading of the vehicle. Then the same algorithm as above can be used.

1. Divide the 360 degrees of possible headings into N slots, where N is chosen appropriately. (For example with N = 24 one obtains slots of a width of 15 degrees each).
2. Let TS_k denote the time series where the heading of the vehicle, as measured by the GPS PVT solution, falls between 15*k and 15*(k+1) degrees.
3. Whenever 90 seconds of data are available for any TS_k, perform model fitting as described in the static algorithm, disregarding any gaps in the data while fitting a linear model.
4. Perform housekeeping by wiping off stale TS_k series where no new data has been received for a longer time, the "staleness period". Housekeeping is necessary because the available GPS satellites change. The staleness period is discussed further below.
5. When the GPS PVT solution indicates that the vessel has performed a complete rotation, and if TS_k is not stale, adjust the phase difference by adding or subtracting a complete cycle from the measurements in TS_k as appropriate.

If the vehicle is truly operating in an outdoor scenario, and the antennae can see some common satellites, but no measurements can be calculated for some time τ, then that implies that the vehicle is manoeuvring in a circle at an angular speed bigger than 10 degrees per minute. The angular speed is then also slow enough not to accomplish an entire rotation in the given time τ. In fisheries, that could happen when a purse seiner is performing a seine set.

On the other hand, care must be taken to design the system in such a way as to detect a spoofer antenna, mounted on a beam turning in the horizontal plane, with the beam slowly turning over the two victim antennae. Such a setup, when done correctly, could prevent the above algorithm from collecting enough data for any TS_k over 90 seconds. With N being the number of slots, one must therefore set the "staleness period" to a time strictly longer than 90 * N seconds. With N = 24, the staleness term can be set to a value greater than 36 minutes. This ensures that either the beam turns too slow, giving enough data for performing the Bayesian measurement, or turns too fast, meaning it comes back to its point of origin before the measurements become stale.

Alternate algorithm for kinematic environments

There are other algorithms for dynamic cases. A particularly simple algorithm was mentioned by Humphreys [personal communication]. It is based on the observation that for a fixed spoofer-antennae geometry, the carrier phase between the antennae and satellites will have certain constant properties.

1. Execute algorithm 2 for *t0* and t1, both of which are some points in the past. This yields delta_epoch_antenna(SVC^i, A, B, *t0, t1)*

[75] Wikipedia, http://en.wikipedia.org/wiki/Bayesian_inference#Computer_applications

2. For any two satellites SVC^i and SVC^j, calculate
   delta_epoch_antenna_satellite(SVC^i, SVC^j,A, B, *t0, t1)* =
   delta_epoch_antenna(SVC^j, A, B, *t0, t1)*
   - delta_epoch_antenna(SVC^i, A, B, *t0, t1)*
3. If the following inequality is satisfied
   abs (delta_epoch_antenna_satellite(SVC^i, SVC^j,A, B, *t0, t1)*) < *eps*
   then this counts towards the hypothesis H_1, which is that the equipment is located indoors. At system calibration time, a value for *eps* must be chosen. It is a small positive value close to zero.
4. If the above inequality does not hold, then this counts towards the hypothesis H_0, which is that the equipment is located outdoors.
5. Feed above hypothesis assessments into a hypothesis test, for instance a or a "Binomial Bayesian inference of posterior probability"
6. Set *t0* = *t1*, wait some time, pick a new *t1* value, and start over at step 1.

One can visualise this algorithm on a time-versus-phase plot. First, the following figure depicts an indoor type scenario with the GPS wave front coming from a single source:

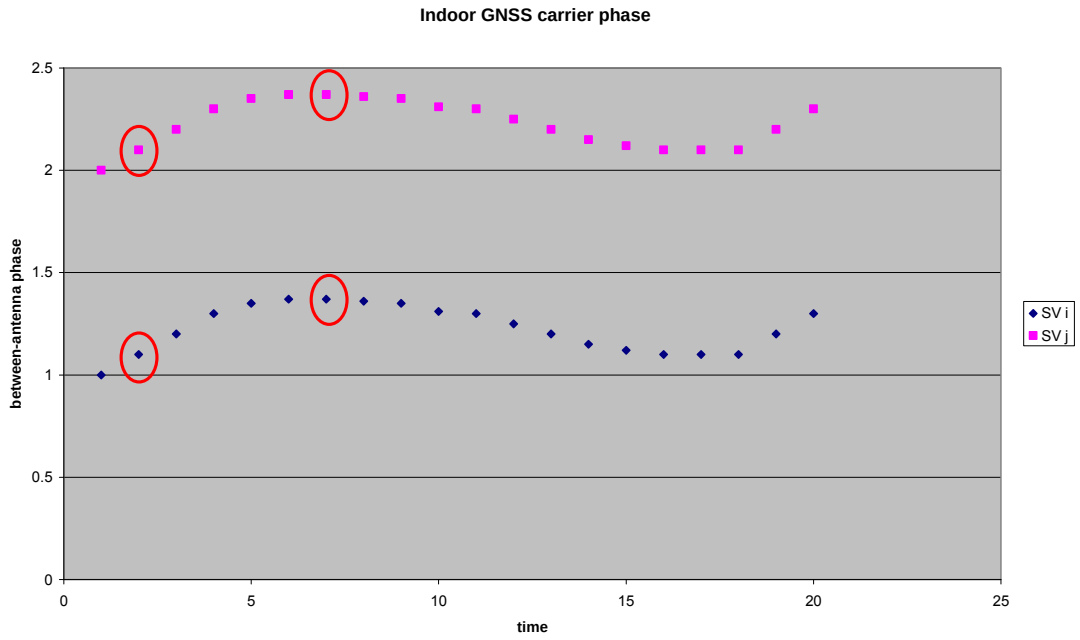


Figure 40: Visualisation of an indoor type scenario. The phase difference is constant between antennae ("delta antenna") for two Space Vehicles with indices *i* and *j*, at two different time values. The algorithm takes differences with respect to time and Space Vehicles (delta_epoch_antenna_satellite), which will yield a near-zero value for an indoor scenario.

Contrast the above with the scenario that would arise from an outdoor antenna GPS wavefront, depicted below:

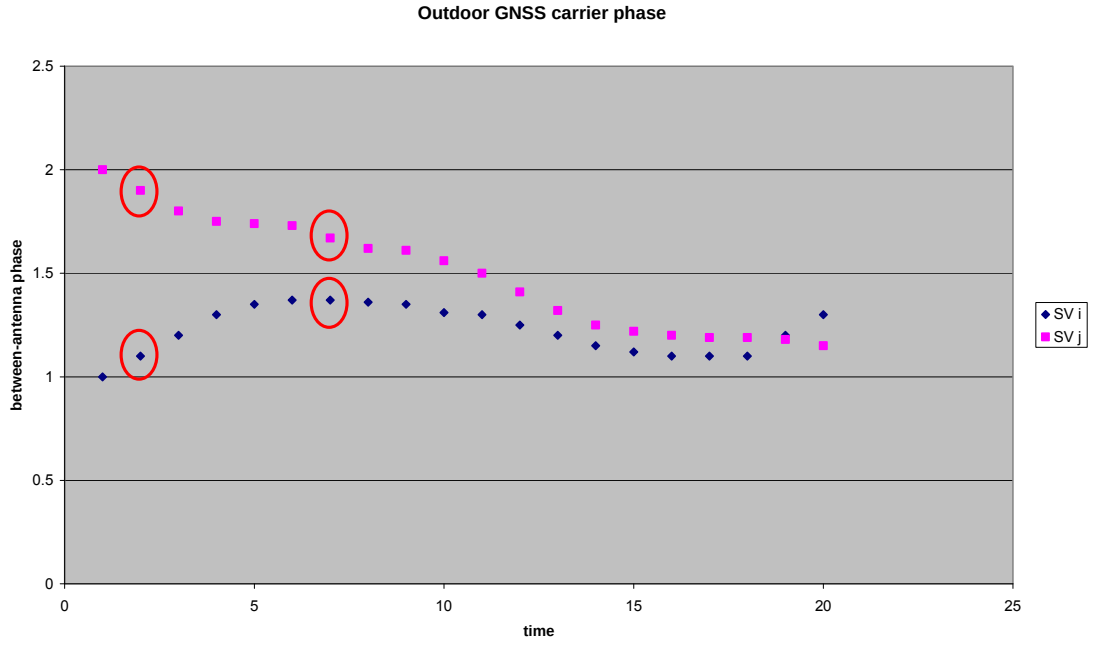


Figure 41: Visualisation of an outdoor scenario. The phase difference is varies between antennae ("delta antenna") for two Space Vehicles with indices *i* and *j*, at two different time values. The term delta_epoch_antenna_satellite would be much larger in absolute than in the indoor scenario.

The simplicity of the algorithm and its use for a kinematic environment implies that it is not able to compare the evolution of the carrier phase against the satellites' almanac data. Such a comparison was performed in the reference [Montgomery09], but that comparison referred to a static scenario.

Glossary[76]

ADSL: Asymmetric Digital Subscriber Link. A data communications technology that is widely used to enable broadband internet access over copper telephone lines.

AES: Advanced Encryption Standard. *In cryptography, the Advanced Encryption Standard (AES) is an encryption standard adopted by the U.S. government.* It is based on the Rijndael cipher. *The Rijndael cipher was developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen.*

ALOHA: A pioneering and presently defunct networking communications system, which broadcasted data packets using radio frequency, in order to achieve point to point data communication. One of its principles is used today in the Ethernet 802.3 protocol, namely the Carrier Sense Multiple Access. An improvement on the original ALOHA protocol was the "slotted ALOHA" concept: to divide time into slots, and start sending data packets only at the boundaries of each slot.

AGC: Automatic Gain Control. *An adaptive system found in many electronic devices, effectively reducing the amplitude if the signal is strong and raising it when it is weaker.* The principle of Automated Gain Control was also discovered in biology by natural selection, as is the case e.g. for eye pupil dilation.

AIS: Automatic Identification System. A maritime short range tracking and anti-collision system. *Used on ships and by Vessel Traffic Services (VTS) for identifying and locating vessels by electronically exchanging data with other nearby ships and VTS stations. On-board AIS transponders automatically broadcast information, such as their position, speed, and navigational status, at regular intervals via a VHF transmitter built into the transponder.*

BGAN: Broadband Global Area Network. A term coined by Inmarsat to designate a "global" satellite network used for voice (telephony) and data traffic at "broadband" rates of up to 492 kbit/s. The service covers most of the globe, except for polar latitudes beyond 75 degrees north or south. For use in maritime, BGAN is marketed as FleetBroadband.

C/A: Coarse Acquisition (code). Signals emitted from GPS satellites at the L1 frequency comprise both the C/A code and the P(Y) code. The C/A code is well-documented and open for civilian use, therefore the term C/A by extension often refers to the civilian GPS signal.

dbW: decibel Watts. *Unit for the measurement of signal power, expressed in decibels relative to one watt.*

DLL: Delay Locked Loop. In GNSS, once the receiver identifies a signal, *it establishes two tracking loops. One is a phase-locked loop (*PLL) *in the frequency domain and the other is a time-locked loop* (same as delay locked loop) *in the 1023 bit code space domain with the goal of tracking both the code and carrier phases for that signal.* [Clark99]

DSP: Digital Signal Processor, *a specialized microprocessor with an optimized architecture for the fast operational needs of digital signal processing.* The latter implies *a large number of mathematical operations to be performed quickly and repetitively on a set of data. […] Many DSP applications have constraints on latency.* For portable devices, a DSP offers a better trade-off between power

[76] Words in italics are from Wikipedia unless attributed otherwise.

consumption and processing speed than a general-purpose CPU. A DSP is performing the calculations in the GPS signal simulator developed at Cornell University by Humphreys, Ledvina and others.

DT: Digital Tachograph. *A speed-logging device used in trucks throughout the European Union for the purpose of improving road safety.* Composed of two units, the Vehicle Unit (VU) and the Motion Sensor (MS).

DNID: Data Network Identifier. A term employed by the company Inmarsat to identify a logical participant in its communications satellite network. One logical participant can have up to 256 devices, each device being assigned one member number. If an actual legal entity needs to have more than 256 devices, it will be attributed more than one DNID.

ECDSA: *The Elliptic Curve Digital Signature Algorithm (ECDSA) is a variant of the Digital Signature Algorithm (DSA) which uses Elliptic curve cryptography. The Digital Signature Algorithm (DSA) is a United States Federal Government standard or FIPS for digital signatures. A digital signature or digital signature scheme is a mathematical scheme for demonstrating the authenticity of a digital message or document.*

ECIES: *Elliptic Curve Integrated Encryption Scheme. A public key encryption scheme where single messages are sent between two parties without a prologue key agreement scheme (as customary in many other encryption schemes).* Instead of using a key agreement scheme, party A sends a message to party B where the message contains its own shared secret key.

EGNOS: *The European Geostationary Navigation Overlay Service is a satellite based augmentation system (SBAS) under development by the European Space Agency, the European Commission and EUROCONTROL.* It is intended to supplement the GPS, GLONASS and Galileo systems by reporting on the reliability and accuracy of the signals.

FAO: Food and Agriculture Organization of the United Nations. The agency of the UN whose mission is food security.

FDE: Fault detection and Exclusion, a GNSS term. *An enhanced version of RAIM employed in some receivers is known as Fault Detection and Exclusion (FDE). It uses a minimum of 6 satellites to not only detect a possible faulty satellite, but to exclude it from the navigation solution so the navigation function can continue without interruption. The use of satellites from multiple GNSS constellations or the use of SBAS satellites as additional ranging sources can improve the availability of RAIM and FDE.*

FIPS: *Federal Information Processing Standard, a publicly announced standards developed by the U.S. Federal government.*

FMC: Fisheries Monitoring Centre. An installation operated by a flag state that participates in a Vessel Monitoring System (VMS). This installation collects the VMS position messages from the fishery vessels using computerised systems. The installation also forwards the messages to other states on a need-to-know basis, for instance if the vessel is fishing in the waters of another state. By extension the term FMC is often employed to mean the infrastructure and manpower that a state uses to monitor fishing activities of the vessels flying its flag.

Galileo: A fledging Global Navigation Satellite System (GNSS) owned by the European Union.

GLONASS: *"GLObal NAvigation Satellite System". A satellite navigation system (GNSS) developed by the former Soviet Union and operated for the Russian government by the Russian Space Forces.*

GNSS: *Global Navigation Satellite System. Generic term for satellite navigation systems (Sat Nav) that provide autonomous geo-spatial positioning with global coverage. GNSS allows small electronic receivers to determine precise time and location.*

GPGPU: General Purpose Graphics Processing Unit, a specialised type of microprocessor. *A graphics processing unit or GPU (also occasionally called visual processing unit or VPU) is a specialized processor that offloads 3D or 2D graphics rendering from the microprocessor. Their highly parallel structure makes them more effective than general-purpose CPUs for a range of complex algorithms. Because* typical GPU *computations involve matrix and vector operations, engineers and scientists have increasingly studied the use of GPUs for non-graphical calculations. A new concept is to use a general purpose graphics processing unit as a modified form of stream processor,* itself a *limited form of parallel processing.* Coincidentally, the GPS simulator developed at Cornell University by Humphreys, Ledvina and others requires parallel processing, therefore it may benefit from the additional processing power of a GPGPU.

GPS: Global Positioning System. A Global Navigation Satellite System (GNSS) owned by the United States of America.

IMO: International Maritime Organization.

IMU: Inertial Measurement Unit: *the main component of inertial guidance systems, an electronic device that measures and reports on a craft's velocity, orientation, and gravitational forces, using a combination of accelerometers and gyroscopes.* An IMU together with an embedded computing system constitute an INS.

INS: Inertial Navigation System: *A navigation aid that uses a computer, motion sensors (accelerometers) and rotation sensors (gyroscopes) to continuously calculate via dead reckoning the position, orientation, and velocity (direction and speed of movement) of a moving object without the need for external references.*

L1: The frequency of 1575.42 MHz, used as a centre frequency for some GNSS, notably GPS.

L2: The frequency of 1227.60 MHz, used as a centre frequency for some GNSS, notably GPS.

MAC: *Message Authentication Code. In cryptography, a MAC is a short piece of information used to authenticate a message, such that the receiver can be certain that the sender is who he claims to be.*

Mbps: Megabits (million bits) per second. A data transmission rate.

Mcps: Megachips (million chips) per second. A chipping rate. A “chip” is an information encoding element modulated into a radiofrequency (spread spectrum) signal. In Binary Phase Shift Keying, a single bit corresponds to a single chip and vice versa. Generally, if one modulates one or more bits onto a radiofrequency signal, one obtains one or more chips.

MEO: Medium Earth Orbit. A satellite orbit located between 2000 and 35786 km in altitude. Orbital periods range from two to 24 hours.

MS: Member State (of the European Union).

MS: Motion Sensor. The part of a Digital Tachograph (DT) that is located in the gearbox of the truck.

MSPS: Mega (million) samples per second. A sampling rate.

NIST: National Institute for Standards and Technology. *NIST is a non-regulatory federal agency within the U.S. Department of Commerce. NIST's mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve quality of life.*

OEM: Original Equipment Manufacturer. *A manufacturer that builds hardware that is subsequently used as a component in a different manufacturer's device, and rebranded according to the latter manufacturer.* For instance, many GPS navigation systems make use of independently built GPS receiver components. The company manufacturing and supplying the GPS receiver component would the OEM.

P(Y): The precision encrypted code emitted from GPS satellites for defence (military) purposes. The known P code is modulated with the secret W code, yielding the encrypted P(Y) code. This encryption is used in order to defeat GPS spoofing in the military domain. The W code is known only to a tightly controlled of devices available to the USA and some allies.

PEM format: PEM originally stands for Privacy Enhanced Mail. Digital certificates are mostly encoded in two formats, PEM and DER. The PEM format refers to a *Base64 encoded DER certificate, enclosed between "-----BEGIN CERTIFICATE-----" and "-----END CERTIFICATE-----"*.

PLL: Phase Locked Loop. In GNSS, once the receiver identifies a signal, *it establishes two tracking loops. One is a phase-locked loop (*PLL) *in the frequency domain and the other is a time-locked loop* (same as delay locked loop) *in the 1023 bit code space domain with the goal of tracking both the code and carrier phases for that signal.* [Clark99]

PKI: *Public Key Infrastructure. Public Key Infrastructure (PKI) is a set of hardware, software, people, policies, and procedures needed to create, manage, distribute, use, store, and revoke digital certificates. In cryptography, a PKI is an arrangement that binds public keys with respective user identities by means of a certificate authority (CA).*

PRN: Pseudo Random Number. In GNSS the term PRN refers to a pseudo random bit sequence. Every GNSS satellite modulates its signal using its own PRNs; it uses one PRN for the (civilian) Coarse Acquisition code and one PRN for the (military) P(Y) code. By extension the term PRN is often used to refer to identify a GNSS satellite.

PNT: Position, Navigation, and Timing. Same as PVT.

PVT: Position, Velocity and Timing. Provided by GNSS user equipment when they are fully operational.

RAIM: Receiver Autonomous Integrity Monitoring, a GNSS term. *RAIM detects faults with redundant GPS pseudorange measurements. That is, when more satellites are available than needed to produce a position fix, the extra pseudoranges should all be consistent with the computed position. A pseudorange that differs significantly from the expected value (i.e., an outlier) may indicate a fault. The use of satellites from multiple GNSS constellations or the use of SBAS satellites as additional ranging sources can improve the availability of RAIM and FDE.*

RFC: *Request for Comments. In computer network engineering, a Request for Comments (RFC) is a memorandum published by the Internet Engineering Task Force (IETF) describing methods,*

behaviors, research, or innovations applicable to the working of the Internet and Internet-connected systems.

RSA: Rivest, Shamir and Adleman. The term RSA stands for an algorithm for asymmetric key cryptography systems that uses large prime numbers and modulus operations. The term RSA also stands for the company "RSA Security" that was founded by Rivest, Shamir and Adleman, in order to commercialise the RSA algorithm. RSA Security no longer exists independently, as it was taken over in 2006.

SOLAS: *International Convention for the Safety of Life at Sea.*

SBAS: Satellite Based Augmentation System. Generic term for a system that improves GNSS accuracy, reliability and availability through additional satellite broadcast messages.

SDR: Software Defined Radio. A radio receiver or emitter system that is mostly using software routines, instead of using hardware processors.

SHA: *Secure Hash Algorithm. The Secure Hash Algorithm is one of a number of cryptographic hash functions published by the National Institute of Standards and Technology (NIST) as a U.S. Federal Information Processing Standard. A cryptographic hash function is a deterministic procedure that takes an arbitrary block of data and returns a fixed-size bit string, the (cryptographic) hash value, such that any change to the data will change the hash value.*

TCP/IP: Transmission Control Protocol over Internet Protocol: *Whereas IP handles lower-level transmissions from computer to computer as a message makes its way across the Internet, TCP operates at a higher level, concerned only with the two end systems, for example a Web browser and a Web server. In particular, TCP provides reliable, ordered delivery of a stream of bytes from a program on one computer to another program on another computer.*

TCXO: Temperature compensated crystal oscillator. A crystal oscillator is an electronics component used in electronic clocks. The oscillator generates a frequency for purposes of timekeeping. However an oscillator's frequency typically varies with temperature. Hence temperature-compensated oscillators have been devised; these integrate a device that it uses to compensate for the temperature-induced frequency variations. This typically leads to short term stability of less than 1 ppm (parts per million).

TLS: *Transport Layer Security and its predecessor, Secure Sockets Layer (SSL), are cryptographic protocols that provide security for communications over networks such as the Internet. TLS and SSL encrypt the segments of network connections at the Transport Layer end-to-end.* For data transport, TLS and SSL traffic are almost always sending payloads over TCP/IP.

VU: Vehicle Unit. The part of a digital tachograph (DT) that is located in the truck's cabin.

VMS: Vessel Monitoring System, a system that allows the Fisheries Monitoring Centre (FMC) to track the location of commercial fishing vessels in real time, in order to implement Monitoring, Control and Surveillance (MCS) programs regarding fisheries. The VMS is typically composed of an on-board GNSS tracker (the "VMS device") that sends position messages, a satellite communications infrastructure that relays messages, a Land Earth Station (LES) to bundle messages from the satellites, and a computerised system at the FMC that processes the messages from all vessels flying its flag.

W bits / W code: A pseudo-random number sequence that serves to encrypt the P code.

WAAS: Wide Area Augmentation System, an SBAS implementation. *WAAS is an air navigation aid developed by the Federal Aviation Administration to augment the Global Positioning System (GPS), with the goal of improving its accuracy, integrity, and availability.*

References

[AIS-WP] Wikipedia, “Automatic Identification System”, http://en.wikipedia.org/wiki/Automatic_Identification_System

[AIS-USCG] US Coast Guard, “AIS Messages”, http://www.navcen.uscg.gov/?pageName=AISMessages

[Aviation09] Michael A. Taverna, “Europe Cuts Galileo Sats Order”, Aviation Week, Oct 26th 2009, http://www.aviationweek.com/aw/generic/story_generic.jsp?channel=space&id=news/Gal102609.xml&headline=Europe%20Cuts%20Galileo%20Sats%20Order

[Backen07] Staffan Backén, “Towards Dynamic Array Processing for GNSS Software Receivers”, ISSN 1402-1757, Luleå University of Technology, 2007, http://epubl.luth.se/1402-1757/2007/65/LTU-LIC-0765-SE.pdf

[BBC05] BBC News, “Senegal suffers from fishing crisis”, http://news.bbc.co.uk/2/hi/business/4182972.stm

[BBC05-2] BBC News, “'Black fish' duo to pay back £1m”, http://news.bbc.co.uk/2/hi/uk_news/scotland/4349474.stm

[Bernstein05] D. J. Bernstein, “Cache-timing attacks on AES”, http://cr.yp.to/antiforgery/cachetiming-20050414.pdf

[BIS04] Bank for International Settlements, “Statistics on payment and settlement systems in selected countries, Figures for 2004”, http://www.bis.org/publ/cpss74.pdf

[BSI10] Bundesamt für Sicherheit in der Informationstechnik, “German Zoned Products List, BSI TL 03305 English Version”, 2010, https://www.bsi.bund.de/cae/servlet/contentblob/471342/publicationFile/56525/TL_03305eng_pdf.pdf

[Borre07] Kai Borre, Dennis M. Akos, Nicolaj Bertelsen, Peter Rindner, Søren Holdt Jensen, “A Software-Defined GPS and Galileo Receiver”, Applied and Numerical Harmonic Analysis, Birkhäuser Boston, 2007, ISBN 10 0-8176-4390-7

[Clark99] T. Clark, “How a GPS Receiver Gets a Lock”, http://gpsinformation.net/main/gpslock.htm

[DFT06]: United Kingdom Department for Transport, “Vehicle Licensing Statistics 2006”, http://webarchive.nationalarchives.gov.uk/+/http://www.dft.gov.uk/adobepdf/162469/221412/221552/228052/252186/vehiclelicensing2006.pdf

[ESA-GAL] European Space Agency, “Galileo Specifications”, http://www.esa.int/esaNA/SEMTHVXEM4E_galileo_0.html

[ETSC06] ETSC, “Traffic Law Enforcement across the EU”, ISBN 90-76024-24-3, http://www.etsc.eu/documents/ETS%20May%202006.pdf

[EttusUSRP1] Ettus Research LLC, “USRP Motherboard Datasheet”, http://www.ettus.com/download

[EttusUSRP2] Ettus Research LLC, “Datasheet for the BasicRX, BasicTX, LFTX, TVRX, and DBSRX daughterboards”, http://www.ettus.com/download

[FHMQV] A. P. Sarr, P. Elbaz–Vincent, J.–C. Bajard “A Secure and Efficient Authenticated Diffie–Hellman Protocol”, Cryptology ePrint Archive: Report 2009/408, http://eprint.iacr.org/2009/408.pdf

[FIPS140-2] Federal Information Processing Standards Publication, “FIPS PUB 140-2, Security Requirements For Cryptographic Modules”, National Institute of Standards and Technology, Gaithersburg, MD 20899-8900, May 25, 2001

[Gal02] Hein et al, “Status of Galileo Frequency and Signal Design”, September 2002, available at the ESA website.

[Gal05] ESA, “Galileo: The European Programme for Global Navigation Services”, ISBN 92-9092-738-0, ISSN 0250-1589.

[Gerg05] I. Gerg, “An Overview and Example of the Buffer-Overflow Exploit”, IAnewsletter Volume 7 Number 4 • Spring 2005, http://iac.dtic.mil/iatac

[GIGABEAM] Gigabeam Corp, “Gi-Linx line”, http://www.gigabeam.com/Products/GiLINX46GHz.aspx and “Gi-FLEX line”, http://www.gigabeam.com/Products/GiFLEX640GHz.aspx.aspx

[GNSS-WP] Wikipedia, “Global navigation satellite system”, http://en.wikipedia.org/wiki/Gnss

[Gore09] W. L. Gore, “Anti-Tamper Physical Security for Electronic Hardware”, http://www.gore.com/en_xx/products/electronic/anti-tamper/anti-tamper-respondent.html

[GPS-SIG-WP] Wikipedia, “GPS Signals”, http://en.wikipedia.org/wiki/GPS_signal

[GPS-WP] Wikipedia, “Global Positioning System”, http://en.wikipedia.org/wiki/Global_Positioning_System

[GPS-REF] B. Parkinson, J. Spilker et al, “Global Positioning System: Theory and Applications (Volume I)”, ISBN 1-56347-106-X, and “Global Positioning System: Theory and Applications (Volume II)”, ISBN 1-56347-107-8, both published by American Institute of Aeronautics and Astronautics.

[GPSPOWER] GPSInformation.net, “GPS Satellite power output”, http://gpsinformation.net/main/gpspower.htm

[GR1] “GNU Radio”, http://www.gnu.org/software/gnuradio/

[Grant09] A. Grant and P. Williams, “GNSS Solutions: What is the effect of GPS jamming on maritime safety?”, InsideGNSS, January/February 2009, http://www.insidegnss.com/node/1122

[Heckler-GPSSDR]: G. W. Heckler, “GPS Software defined radio”, http://www.gps-sdr.com/

[Hein07] G. Hein et al, “Authenticating GNSS: Proofs against Spoofs, Part 2”, Inside GNSS, Sept/Oct 2007

[Humphreys08] Todd E. Humphreys, Brent M. Ledvina, Mark L. Psiaki, Brady W. O'Hanlon, and Paul M. Kintner, Jr., "Assessing the Spoofng Threat: Development of a Portable GPS Civilian Spoofer", 2008 ION GNSS Conference, Savanna, GA, September 16-19, 2008

[Humphreys09] Todd E. Humphreys, Mark L. Psiaki, Joanna C. Hinks, Brady O'Hanlon, and Paul M. Kintner, Jr, "Simulating Ionosphere-Induced Scintillation for Testing GPS Receiver Phase Tracking Loops", IEEE Journal Of Selected Topics In Signal Processing, Vol. 3, No. 4, August 2009

[Infineon09] Infineon, "SLE 88 family: High End 32-bit Security Controller", http://www.infineon.com/cms/us/product/channel.html?channel=ff80808112ab681d0112ab693a350166

[IUU-WP] Wikipedia, "Illegal, unreported and unregulated fishing", http://en.wikipedia.org/wiki/IUU

[Jacobson07] Len Jacobson, "Expert Advice: Why we need eLoran", InsideGNSS, http://tl.gpsworld.com/gpstl/Expert+Advice+%26+Leadership+Talks/Expert-Advice-mdash-Why-We-Need-eLoran/ArticleStandard/Article/detail/486538

[Johnston06] R. G. Johnston, "Tamper-Indicating Seals", American Scientist, 2006, November-December issue, pp 515-523

[Johnston06-2] R. G. Johnston, "Tags and seals, Electronic Anti-evidence seals (2006)", LAUR-06-1312, available on request from rogerj@lanl.gov

[Johnston07] R. G. Johnston, M. J. Timmons, and J. S. Warner, "Protecting Nuclear Safeguards Monitoring Data from Tampering", Taylor & Francis Group LLC, DOI: 10.1080/08929880701715076, pp 185-204.

[Jorgensen00] Terje H Jorgensen, "Loran-C/Eurofix in Europe - A NELS Status Report", http://www.loran.org/Meetings/Meeting2000/pdffiles/papers/04.pdf

[Kowoma] Kowoma.de, "Composition of the Data signal", http://www.kowoma.de/en/gps/data_composition.htm

[Kroen09] U. Kröner et al, "Report on Authentication in Fisheries Monitoring", ISBN 978-92-79-11095-5, available via the European Bookshop, http://bookshop.europa.eu/

[Langley08] Richard Langley, "Innovation: Interference Heads-Up", GPSWorld, http://sidt.gpsworld.com/gpssidt/article/articleDetail.jsp?id=523633

[Ledvina10] Brent M. Ledvina, William J. Bencze, Bryan Galusha, and Isaac Miller, "An In-Line Anti-Spoofing Device for Legacy Civil GPS Receivers", Proceedings of the ION ITM 2010, San Diego, CA, January 25-27, 2010.

[Lo07] Sherman Lo, "What is eLoran and how is it different from Loran-C?", InsideGNSS Sept-Oct 2007, http://www.insidegnss.com/node/173

[Lo09] Sherman C. Lo, Benjamin B. Peterson, Per K Enge, "Assessing the Security of a Navigation System: A Case Study using Enhanced Loran", Stanford University, 2009, http://waas.stanford.edu/~wwu/papers/gps/PDF/LoENCGNSS09.pdf

[Lo09-2] Sherman C. Lo et al, "Signal Authentication: A secure civil GNSS for today", InsideGNSS Sept/Oct 2009, http://www.insidegnss.com/node/1633

[Loran10] InsideGNSS, "USCG Publishes Loran-C Termination; DHS Says Not Needed for GPS Backup", January 2010, http://www.insidegnss.com/node/1806

[Maxim08] Dallas Semiconductor / Maxim, TCXO Oscillator Modules, DS4026, http://para.maxim-ic.com/en/search.mvp?fam=osc_mod&980=TCXO&hs=1

[MDA10] Mobile Entertainment and MDA, "UK sends 11 million text messages an hour", http://www.mobile-ent.biz/news/35839/UK-sends-11-million-text-messages-an-hour

[Montgomery09] P. Montgomery, T. Humphreys, and B. Ledvina, "A multi-antenna defense, Receiver-Autonomous GPS Spoofing Detection" InsideGNSS Mar/Apr 2009, http://www.insidegnss.com/node/1370

[Mrag02] Marine Resources Assessment Group (MRAG), "Evidential Value of VMS Position Reports", DG FISH/2002/11 Draft Final report, European Commission

[Nav01] Navigs, "Data Risk Final Report", QLAM-2001-00092, European Commission internal paper

[Nav05] Navigs, "Fishing Vessel Monitoring Systems: Past, Present and Future", Prepared for: The High Seas Task Force, OECD, Paris, 17 October 2005

[NIST07] National Institute of Standards and Technology, "NIST Special Publication 800-56A", March 2007, http://csrc.nist.gov/publications/nistpubs/800-56A/SP800-56A_Revision1_Mar08-2007.pdf

[Nordvik07] J.-P. Nordvik, D. Landat, J. Bishop, A. Poucet, "Report on the attacks to security of the Digital Tachograph and on the risk associated with the introduction of adaptors to be fitted into light vehicles", Limited Distribution JRC Technical Note, 2007

[Pacific07] PACIFIC consortium, "Say 'Hello' to Galileo's PRS", Inside GNSS Fall 2007, http://www.insidegnss.com/node/364

[Perrig02] A. Perrig, R. Canetti, J.D. Tygar, D. Song, "The TESLA Broadcast Authentication Protocol", CryptoBytes, 5:2, Summer/Fall 2002, pp. 2-13

[Petovello08] Petovello, Lapachelle, Borio, and Lo Presti, "What are the main classes of interference that can degrade the GNSS signals? What are the possible countermeasures?", InsideGNSS, March/April 2008, pages 25-27, http://www.insidegnss.com/auto/igm_022-027.pdf

[Popescu04] C. Popescu, "A Secure Authenticated Key Agreement Protocol", University of Oradea, Department of Mathematics, Oradea, Romania, 2004, http://www.q2s.ntnu.no/publications/open/2004/Paper_rev/2004_popescu_SAK.pdf

[PSEC-KEM] NTT Corporation, "PSEC-KEM Specification version 2.2", April 14, 2008, http://info.isl.ntt.co.jp/crypt/eng/psec/dl/iso/psec-kem_v2.2_20080414e.pdf

[Psiaki03] Hee Jung, Mark L. Psiaki, Steven P. Powell: "Kalman-Filter-Based Semi-Codeless Tracking of Weak Dual-Frequency GPS Signals", The Institute of Navigation (ION), GPS 2003 meeting

[Psiaki09] M. L. Psiaki, Spoofing detection for civilian GNSS signals via aiding from encrypted signals," in *Accepted as an alternate paper at ION/GNSS 2009.* Savannah, GA: Institute of Navigation, sep 22-25 2009.

[Rizos] C. Rizos, "Modelling GPS Observations", http://www.gmat.unsw.edu.au/snap/gps/gps_survey/chap6/chap6.htm

[RFID-WP] Wikipedia, "Radio-frequency identification", http://en.wikipedia.org/wiki/Rfid

[RFID08] Project RFID-AP, "Projet rfid-ap - Context and state of the art", http://www.rfid-ap.fr/Public/pages_web/context_and_state_of754.html

[SCA-WP] Wikipedia, "Side channel attack", http://en.wikipedia.org/wiki/Side_channel_attack

[Scott03] L. Scott, "Anti-Spoofing & Authenticated Signal Architectures for Civil Navigation Systems", The Institute of Navigation (ION), GPS 2003 meeting

[Scott07] L. Scott, "Location Assurance", Expert Advice in GPS World, July 2007.

[Septentrio] Septentrio NV, "AsteRx1i OEM, MU enhanced GPS/GALILEO Single-frequency OEM Receiver", http://www.septentrio.com/products/receivers/asterx1i-oem

[Sionpower] Sionpower , "Lithium-Sulfur Rechargeable Batteries: Characteristics, State of Development, and Applicability to Powering Portable Electronics", http://www.sionpower.com/pdf/articles/PowerSources2004.pdf

[Silberman04] Peter Silberman, Richard Johnson, "A Comparison of Buffer Overflow Prevention Implementations and Weaknesses", IDefense, 1875 Campus Commons Dr. Suite 210 Reston, VA 20191, 2004

[Sluiman10] Dr. Ir. Sluiman, "Risk of Information Theft on Inmarsat C", March 31, 2010, available through "Maritime Professional", http://www.maritimeprofessional.com/News/333820.aspx

[Spiegel09] Spiegel Online International, "Dutch Road Toll System Gets Surprising Green Light", 18/02/2009, http://www.spiegel.de/international/europe/0,1518,608406,00.html

[TEMPEST-WP] Wikipedia, "TEMPEST", http://en.wikipedia.org/wiki/TEMPEST

[TPM] Trusted Computing Group, "Trusted Platform Module (TPM) Summary", http://www.trustedcomputinggroup.org/resources/trusted_platform_module_tpm_summary

[TT1] Thrane & Thrane, "VMS Capsat® Transceiver TT-3020C, TT-3022C, TT-3022D and TT-3028CM Configuration Manual", available on request from Thrane & Thrane Support

[TT2] Thrane & Thrane, "Capsat® Transceiver TT-3020C, TT-3022C, TT-3022D, TT-3022E and TT-3028CM Software Interface Reference Manual", available on request from Thrane & Thrane Support

[USCG10] US Coast Guard, "Loran-C General Information", http://www.navcen.uscg.gov/loran/default.htm

[Volpe01] “Vulnerability Assessment of the Transportation Infrastructure Relying on the Global Positioning System”, Final report, August 29, 2001, John A. Volpe National Transportation Center.

[Wang02] J. Wang, “Pseudolite Applications in Positioning and Navigation: Progress and Problems”, Journal of Global Positioning Systems (2002), Vol. 1, No. 1: 48-56

[Ward07] P. Ward, “What’s going on? RFI situational awareness in GNSS receivers”, InsideGNSS Sept-Oct 2007, http://www.insidegnss.com/node/168

[Warfield08] European GNSS Supervising Authority, “Secure Vehicle Communications: Results and Challenges Ahead”, http://icapeople.epfl.ch/panos/SVCWCR/presentations/Luasanne%20Presentation2-neil.pdf

[Weston99] Ed Weston, “GPS Navigation Satellite message format and protocol details”, http://gpsinformation.net/gpssignal.htm

[Woo00] Woo, K.T., "Optimum Semicodeless Carrier-Phase Tracking of L2," Institute of Navigation, 47(2), 2000, pp. 82-99.

[Xsens] XSens, “MTi: Miniature Attitude and Heading Reference System”, http://www.xsens.com/en/general/mti

{End of document}

European Commission

EUR 24390 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen
Title: Hardening of GNSS based trackers - Final Report
Author(s): U. Kröner , C. Bergonzi, J. Fortuny-Guasch, R. Giuliani, F. Littmann, D. Shaw, D. Symeonidis
Luxembourg: Publications Office of the European Union
2010 – 149 pp. – 21 x 29.7 cm
EUR – Scientific and Technical Research series – ISSN 1018-5593
ISBN 978-92-79-15878-0
doi:10.2788/97633

Abstract
Civilian GNSS based real-time tracking systems are presently used in a number of fields, such as the fisheries Vessel Monitoring System (VMS), the maritime Automatic Identification System (AIS), and the transportation of dangerous goods. Such trackers are commonly composed of a GNSS receiver module and a communications module for transmission of positions. These GNSS-based trackers are vulnerable to tampering. Modelling such trackers, one identifies multiple ways that an adversary could use to introduce false tracking information. This document summarises an array of vulnerabilities and options for hardening such trackers, such as against fake GNSS signals, physical tampering, side channel attacks, and the substitution of position reports.

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

LB-NA-24390-EN-N