The Dissertation Committee for Kyle Douglas Wesson
certifies that this is the approved version of the following dissertation:

Secure Navigation and Timing Without Local Storage of Secret Keys

Committee:

Todd E. Humphreys, Supervisor

Brian L. Evans, Co-Supervisor

Ross Baldick

Lili Qiu

Ahmed H. Tewfik

Secure Navigation and Timing Without Local Storage of Secret Keys

by

Kyle Douglas Wesson, B.S.; M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2014

Dedicated to my parents.

Acknowledgments

I would like to express my sincere thanks and gratitude to a number of people who helped me complete my Ph.D.

I thank Dr. Todd Humphreys for his training and support through my journey. He has always maintained high standards that kept me pushing to learn and achieve over the past five years. I am most thankful for the editing and writing training that I received from him; these skills will serve me for the rest of my life.

I thank Dr. Brian Evans for his holistic and thoughtful mentorship from day one. He was the first person that I talked to at UT and has believed in me at times more than I believed in myself. I appreciate every one of our "office" hour chats.

I thank my committee members, Dr. Ross Baldick, Dr. Lili Qiu, and Dr. Ahmed Tewfik, for reviewing my Ph.D. research and serving on my committee.

I thank the members, past and present, of the Radionavigation Lab, the Embedded Signal Processing Lab, and the Wireless Networking and Communications Group.

I appreciate the generous financial support from numerous sources. In particular, The University of Texas at Austin and the National Defense Science and Engineering Graduate Fellowship supported me the most.

I appreciate the many friendships and good times spent training and racing with the members of the University of Texas Triathlon Team. I appreciate the advice from Coach Joanna Williamson who was supportive and enthusiastic every day.

I thank those mentors at the U.S. Army Cold Regions Research and Engineering Lab whose early mentorship encouraged me to pursue this degree: Dr. Sally Shoop, Dr. Joyce Nagle, Dr. Lindamae Peck, and Mr. Barry Coutermarsh.

I must also thank many friends who helped me along the way: Henri for the love of cycling; Karl for many Austin adventures; Gustavo and Paul for many, many laps in the pool; Kristin for pep talks; Marcus for reality checks; Zak for coffee breaks; Heidi for considerate and thoughtful conversations; Mr. Taupier for loud "hoorahs"; Ken for many hours studying; and Jahshan for fruitful discussions.

My biggest thanks goes to my Mother and Father. Their boundless love and support has been the biggest influence in my life. At every turn, they have supported my education and sacrificed on my behalf to allow me the privilege of studying at three fine institutions: Kimball Union Academy,

Cornell University, and The University of Texas at Austin. I could not possibly thank or love them enough for all that they have given me.

For all the others whose friendship and advice has assisted me on my journey, I thank you as well. It takes a village to graduate a Ph.D.

Secure Navigation and Timing Without Local Storage of Secret Keys

Publication No. _____

Kyle Douglas Wesson, Ph.D.
The University of Texas at Austin, 2014

Supervisors: Todd E. Humphreys
Brian L. Evans

Civil Global Navigation Satellite System (GNSS) signals are broadcast unencrypted worldwide according to an open-access standard. The virtues of open-access and global availability have made GNSS a huge success. Yet the transparency and predictability of these signals renders them easy to counterfeit, or spoof. During a spoofing attack, a malefactor broadcasts counterfeit GNSS signals that deceive a victim receiver into reporting the spoofer-controlled position or time. Given the extensive integration of civil GNSS into critical national infrastructure and safety-of-life applications, a successful spoofing attack could have serious and significant consequences.

Unlike civil GNSS signals, military GNSS signals employ symmetric-key encryption, which serves as a defense against spoofing attacks and as a barrier to unauthorized access. Despite the effectiveness of the symmetric-key approach, it has significant drawbacks and is impractical for civil applications.

First, symmetric-key encryption requires tamper-resistant receivers to protect the secret keys from unauthorized discovery and dissemination. Manufacturing a tamper-resistant receiver increases cost and limits manufacturing to trusted foundries. Second, key management is problematic and burdensome despite the recent introduction of over-the-air keying. Third, even symmetric-key encryption remains somewhat vulnerable to specialized spoofing attacks.

I propose an entirely new approach to navigation and timing security that avoids the shortcomings of the symmetric-key approach while maintaining a high resistance to spoofing. My first contribution is a probabilistic framework that develops necessary components of signal authentication.

Based on this framework, I develop my second and third contributions: an asymmetric-key cryptographic signal authentication technique and a non-cryptographic spoofing detection technique, both of which operate without a locally stored secret key. These techniques stand as viable spoofing defenses for civil users and could augment—or even replace—current and planned military anti-spoofing measures.

Finally, I offer an in-depth case study of the security vulnerabilities of a modern GNSS-based aviation surveillance technology. I then evaluate possible cryptographic enhancements to the system in the context of the technical and regulatory aviation environment.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Since its development in 1973, the Global Positioning System (GPS)—a type of Global Navigation Satellite System (GNSS)—has become the worldwide standard for globally accurate and precise position, navigation, and timing (PNT). As previous decades welcomed a wealth of low-cost, user-friendly equipment and modernized signal processing techniques into the GPS Control, Space, and User Segments, civil GPS—the family of GPS signals that are freely broadcast worldwide for all civilian uses—has become the default technology for PNT in today's critical national infrastructure (e.g., telecommunications, power, finance, and transportation), giving rise to GPS's moniker as the "invisible utility." Yet, civil GPS receivers are eminently vulnerable to counterfeiting-type attacks, commonly referred to as spoofing attacks.

1.1 The Civil GPS Spoofing Threat

The popularity of civil GPS is due, in part, because the GPS signal structure is defined in a freely-available and open-access Interface Specification (IS) [1]. (Although the discussion here focuses on GPS, other GNSS, such as the European Galileo, also have publicly-available signal definitions [2].)

The open-access nature of GPS signals coupled with a lack of embedded cryptographic safeguards (e.g., digital signatures) in the signal modulation or data bits means that civil GPS signals are highly predictable. This predictability renders civil GPS receivers vulnerable to spoofing attacks in which an attacker transmits matched-GPS-signal-structure interference in an attempt to commandeer the tracking loops of a victim receiver [3, 4]. If an attack is successful, the spoofer can manipulate the receiver's timing or navigation solution. Such attacks can be launched from a spoofer that is co-located with the victim receiver or even from a stand off distance of several kilometers. The vulnerability of civil GPS receivers to spoofing is a serious risk for GPS-dependent critical national infrastructure and safety-of-life applications [3, 5–10].

The spoofing threat has garnered significant attention from the U.S. Government, industry, and academia over the last decade [11–14]. A 2001 U.S. Department of Transportation report, commonly referred to as the "Volpe Report," highlighted the vulnerability of U.S. transportation infrastructure to civil GPS disruption [5]. The Volpe Report was the first publicly-available evaluation of the threats that spoofing poses to critical infrastructure. More recently in 2010, the U.S. Position, Navigation, and Timing National Executive Committee requested that the U.S. Department of Homeland Security (DHS) conduct a comprehensive risk assessment on the short- and long-term risks from civil GPS use in critical national infrastructure [15]. The DHS Homeland Infrastructure Threat and Risk Analysis Center published its findings in a 2011 report that remains classified. The bottom line, however, was that depen-

Figure 1.1: The University of Texas at Austin Radionavigation Laboratory spoofing test and development platforms. Shown here are a synchrophasor measurement unit (upper left), radio-frequency test chamber (upper right), civil GPS spoofer (lower left), cell phone base station oscillator (lower right), and several antennas and commercial receivers.

dency begets vulnerability—the growing dependency on GPS is increasing the risks and ramifications of a successful spoofing attack [16]. Additionally, the hardware and software necessary to develop and test spoofing techniques and defenses becoming cheaper and more readily accessible (e.g., see Figure 1.1).

Consider the implications of a successful spoofing attack against the following three sectors:

1.1.1 Spoofing Implications for the Telecommunication Sector

Modern wireless digital communication employs GPS for reliable time synchronization. In the last decade, the required timing accuracy in cellular

and wireless data standards has increased by a factor of 3 to 10 [17]. Cellular code division multiple access base stations must maintain ±3–10 μs timing accuracy over the air interface according to the CDMA2000 standard [18]. This ensures that base stations do not interfere with their neighbors and that calls are not disrupted during call hand-over. Time-Division Duplex (TDD) LTE, introduced in 2004, requires ±1.5 μs timing accuracy [19], and TDD WiMAX, introduced in 2005, requires ±1 μs timing accuracy [20, 21]. Given about 30 minutes, a single spoofer could effectively disable a cellular base station thereby preventing call handoff (i.e., islanding) [6]. Larger-scale attacks involving a coordinated network of spoofers could target multiple base stations throughout a dense urban population [3].

1.1.2 Spoofing Implications for Smart Power Grids

Like cellular networks, smart power grids demand accurate global synchronization, and the synchronization accuracy is often much smaller than the period of the main power frequency. Real-time voltage and current phasor measurement units installed throughout future power networks will offer engineers unprecedented visibility into power consumption and generation across the smart power grid. One possible outcome will be the ability of the grid to increase power distribution efficiency [22]. Over the coming decade, power engineers will be installing phasor measurement units (PMUs) as a critical component of the modernized smart power grid. PMUs rely on GPS to ensure timing accuracy to within 26.5 μs, which is far more demanding than the accu-

racy required by today's power monitoring equipment [23]. In a recent study, an in-lab attack against a model power grid in Mexico succeeded in disrupting their grid by targeting key nodes in the network [7, 8]. Other research has also considered timing and data forgery in the context of smart power grids [24–26].

1.1.3 Spoofing Implications for the Finance Sector

Global financial exchanges are now digital—their "brains" reside in large data centers connected by kilometers of cables and switches [27]. As a trade is executed, a time stamp is generated. Regulatory requirements state that these time stamps be accurate to within one second [28]; however, in practice, competition between exchanges for high-frequency traders, who are particularly concerned about measuring trading latency, have pushed the exchanges toward millisecond-accurate timing or better [29, 30]. Indeed, traders now even consider relativistic effects of their trades [31, 32]. Not only do traders depend on these time stamps, but facilitators also disseminate the national best bid and offer, which is offered as an "instantaneous" view of the best prices for financial instruments across all participating markets [33]. Manipulation of exchange and market participant timing via GPS spoofing could lead to confusion in the markets or illicit financial gains [34].

1.2 Shortcomings of Symmetric-Key Anti-Spoofing

Unlike civilian and commercial users, the military is afforded spoofing protection via the symmetric-key-encrypted security codes that modulate the

military GPS signals [35]. Symmetric-key encryption not only serves as an anti-spoofing technique but also ensures access to the signals are only available to authorized users. Civilians are not granted access to the cryptographic keys that are required to despread the military signals. Symmetric-key cryptographic techniques offer efficient (i.e., low computational cost and low latency) verification of cryptographic signatures and decryption of encrypted messages. Furthermore, the fact that the military GPS navigation data is encrypted means that signal authentication happens in real-time unlike non-symmetric-key approaches that can only offer authentication every few seconds or minutes [36–38].

Despite the advantages of the symmetric-key approach, its serious drawbacks make it inconvenient and costly for military users and unsuitable for civilians. First, symmetric-key encryption requires tamper-resistant hardware so that unauthorized access to the cryptographic keys stored in the device remain secret. Tamper-proof hardware is complicated and expensive and limits its manufacturing options to a handful of trusted foundries [39]. Second, key management is a burden on soldiers because of the elaborate protocols that must be followed to securely transfer the secret keys onto military receivers. These drawbacks make military receivers expensive and inconvenient. It is no surprise that warfighters prefer civilian GPS receivers 40-to-1 over military hardware in combat scenarios [40]. Finally, even symmetric-key encryption has limitations that could be exploited by a well-motivated and well-funded attacker [41].

The symmetric-key approach—today's state-of-the-art spoofing defense—is effective but has serious drawbacks. Future systems will need to overcome the drawbacks of this approach while simultaneously defending against increasingly sophisticated spoofing attacks [42].

Civil anti-spoofing techniques can be broadly categorized as either dependent on or independent of the GNSS space segment in the following sense: dependent techniques require GNSS signals to contain unpredictable but verifiable modulation structures derived from cryptographic techniques, whereas independent techniques require no special cryptographic modulation of the GNSS signals. This categorization can be broken down further into two groups per category as described in the next two subsections. The specific benefits and drawbacks of these anti-spoofing techniques will be discussed in Chapters 3 and 4 to motivate the cryptographic and non-cryptographic techniques presented in those chapters.

1.2.1 Space Segment Side Dependent Techniques

Space segment side dependent techniques are sometimes referred to as cryptographic techniques because they rely on unpredictable encryption or digital signatures to modulate the GPS signals. There are two types:

1. Civil GPS cryptographic techniques that rely on changes to the GPS IS to insert encryption or digital signatures in either in the data or spreading codes [14, 36, 37, 41, 43–47]. Such techniques can be made extremely

effective, but no civil GPS structure yet incorporates unpredictable elements in their signal definitions.

2. Civil signal authentication techniques that exploit the encrypted military signals without knowledge of the secret keys [38, 48–51]. These techniques can quickly detect an attack but require an always-on network connection between multiple receivers or additional hardware.

1.2.2 Space Segment Side Independent Techniques

Space segment side independent techniques are sometimes referred to as non-cryptographic techniques because they monitor statistical properties of the received signals to detect anomalies from the nominal broadcast that are indicative of a spoofing attack. There are two types:

1. Antenna-oriented non-cryptographic techniques employ multiple antennas or require antenna movement to differentiate between spoofed and authentic signal sets [52–57]. These antenna techniques often require large separation between antennas or additional hardware.

2. Receiver-autonomous signal-processing-type techniques that employ statistical measures to monitor specific signal properties, such as code rate, carrier-to-noise ratio, total in-band power, signal deformations, among others [58–64, 64–70]. These techniques can readily detect an attack, but need careful consideration to ensure that the probability of false alarm due to multipath is minimized.

1.3 Thesis Statement and Expected Contributions

In this dissertation, I defend the following thesis statement:

> *Both cryptographic and non-cryptographic anti-spoofing techniques can secure civil GPS and GNSS navigation and timing while avoiding the serious drawbacks of local storage of secret cryptographic keys that hinder military symmetric-key-based anti-spoofing.*

The drawbacks to symmetric key methods are discussed in Sec. 1.2. I propose an entirely new approach to navigation and timing security that avoids the shortcomings of the symmetric-key approach while maintaining an high resistance to spoofing. The following is a summary of the contributions of my dissertation:

1. **Probabilistic Framework**: I contribute a probabilistic framework that abstracts the particulars of GNSS anti-spoofing to establish necessary conditions for secure location and timing under a security-enhanced GNSS signal model. I illustrate the need for a probabilistic security model in the context of authenticating a timing signal as opposed to the traditionally non-probabilistic security models of message authentication and cryptography. The framework combines cryptography and statistical signal processing across multiple network layers while supporting combined cryptographic and non-cryptographic anti-spoofing techniques. See references: [37, 71, 72].

2. **Asymmetric Cryptographic Signal Authentication**: I develop an asymmetric cryptographic civil Global Positioning System (GPS) signal authentication strategy that is both practical and effective. The specific technique exploits the flexibility of the modernized GPS L2 or L5 civil navigation broadcast message and is packaged for immediate implementation. I further assess the effectiveness of the technique against a challenging spoofing attack scenario. See references: [37, 46, 47, 73].

3. **Non-Cryptographic Spoofing Detection**: I develop and evaluate a non-cryptographic GNSS anti-spoofing technique. The strategy relies on the difficulty of carrying out an effective spoofing attack that simultaneously maintains a low-enough counterfeit signal power to avoid alarms while minimizing tell-tale distortions of the received cross-correlation profile. I evaluate the technique against the Texas Spoofing Test Battery, which is the only publicly-available spoofing dataset. See references: [42, 58, 62, 63].

Finally, I offer an in-depth case study of the security and privacy concerns that face a GPS-based aviation surveillance technology known as ADS-B. My research considers practical cryptographic enhancements for this protocol in the context of the complex technical and regulatory practicalities that are inherent in aviation [74, 75].

1.4 Dissertation Organization

Chapter 2 illustrates the stark differences between data message authentication and navigation signal authentication, while demonstrating the additional challenges facing those who seek to achieve the latter. I formulate a probabilistic framework for timing assurance that combines cryptography and statistical signal processing across multiple network layers. The chapter discusses two specific attacks again security-enhanced GNSS signals: (1) record and playback (i.e., meaconing) and (2) security code estimation and replay.

Chapter 3 develops the asymmetric cryptographic signal authentication technique, known as navigation message authentication. I design and evaluate a practical and effective technique that can be implemented on the civil GPS L2 or L5 civil navigation broadcast message. The chapter concludes with a quantitative assessment of the technique's performance against a sophisticated spoofing attack and offers guidance on implementation.

Chapter 4 develops a non-cryptographic anti-spoofing technique that employs statistical analysis to detect anomalies in the received cross-correlation profile and the total in-band power. The chapter presents the power–distortion tradeoff that a spoofer faces when conducting a spoofing attack. The non-cryptographic technique is quantitatively evaluated against the Texas Spoofing Test Battery, which is the only publicly-available data set of spoofing recordings.

Chapter 5 examines the security of a critical aviation technology known as ADS-B, or Automatic Dependent Surveillance–Broadcast. By 2020, nearly all aircraft flying through U.S. airspace must carry ADS-B transponders to continuously transmit their precise real-time location and velocity to ground-based air traffic control and to other *en route* aircraft. Surprisingly, the ADS-B protocol has no built-in security mechanisms, which renders ADS-B systems vulnerable to a wide range of malicious attacks. In particular, I address the question "can cryptography secure ADS-B?"—in other words, is there a practical and effective cryptographic solution that can be retrofit to the existing ADS-B system and enhance the security of this critical aviation technology? The case study in Chapter 5 considers technical and regulatory challenges in the context of aviation security.

Chapter 6 concludes this dissertation with a summary of contributions and suggestions for future research.

1.5 Nomenclature

AADS	: Airplane Asset Distribution System
ADS-B	: Automatic Dependent Surveillance–Broadcast
AGC	: Automatic Gain Control
APNT	: Alternative Position Navigation and Timing
ARNS	: Aeronautical Radio Navigation Services
ASDI	: Aircraft Situation Display to Industry
ATC	: Air Traffic Control
CDMA	: Code Division Multiple Access
C/N_0	: Carrier-to-Noise Ratio
CA	: Certificate Authority
CNAV	: Civil Navigation

CFR : Code of Federal Regulations
CRC : Cyclic Redundancy Check
CRLB : Cramer–Rao Lower Bound
CSWAP : Cost, Size, Weight, and Power
DHS : Department of Homeland Security
DLL : Delay Locked Loop
DME : Distance Measuring Equipment
DSA : Digital Signature Algorithm
ECDSA : Elliptic Curve Digital Signature Algorithm
eLORAN : Enhanced Long Range Navigation (LORAN)
FAA : Federal Aviation Administration
FEC : Forward Error Correction
FEU : Front-End Units
FPE : Format Preserving Encryption
GMPlib : GNU Multiple Precision Arithmetic Library
GNSS : Global Navigation Satellite System
GPS : Global Positioning System
IEEE : Institute of Electrical and Electronics Engineers
ICAO : International Civil Aviation Organization
ICD : Interface Control Document
IMU : Inertial Measurement Unit
IP : Intellectual Property
IS : Interface Specification
J/N : Jamming-to-Noise
KDE : Kernel Density Estimation
LTE : Long Term Evolution
MAC : Message Authentication Code
NMA : Navigation Message Authentication
NSA : National Security Agency
NIST : National Institute of Standards and Technology
OCX : Operational Control Segment
OCXO : Oven-Controlled Crystal Oscillator
P_D : Probability of Detection
P_F : Probability of False Alarm
PKI : Public-Key Infrastructure
PMU : Phasor Measurement Unit
PNT : Position, Navigation, and Timing

PSR : Primary Surveillance Radar
PVT : Position, Velocity, and Time
RAIM : Receiver Autonomous Integrity Monitoring
RF : Radio Frequency
ROC : Receiver Operating Characteristic
RSA : Rivest, Shamir, and Adleman public-key cryptography
SDR : Software Defined Receiver
SCA : Spreading Code Authentication
SSSC : Spread Spectrum Security Codes
SCAP : Security Certification and Accreditation Procedures
SCER : Security Code Estimation and Replay
SHA : Secure Hash Algorithm
SNR : Signal-to-Noise Ratio
SQM : Signal Quality Monitoring
SSR : Secondary Surveillance Radar
TDD : Time-Division Duplex
TEXBAT : Texas Spoofing Test Battery
TESLA : Timed Efficient Stream Loss-Tolerant Authentication
TCXO : Temperature-Compensated Crystal Oscillator
UAT : Universal Access Transceiver
VSD : Vestigial Signal Defense
WAAS : Wide Area Augmentation System

Chapter 2

A Probabilistic Framework for Global Navigation Satellite System Signal Timing Assurance

2.1 Introduction

Signal authentication, the topic of this chapter, and message authentication, such as is used to sign data transmitted across the Internet, can be distinguished from one another by the models employed to describe their security. Message authentication security is predicated on the computational infeasibility of performing a brute-force search for the secret key used to sign the original message, or of reversing a so-called one-way function to discover the key [76]. While it is true that this assumed computational infeasibility can be couched in probabilistic terms (e.g., the probability that over the next 30 years a weakness will be found in a certain one-way hash function), such language is seldom used, either because the probabilities involved are too subjective or too small to be meaningful. In contrast to message authentication, the security of signal authentication is much weaker and demands a probabilistic model, as described in this chapter.

To defend against spoofing, GNSS receivers seek to authenticate GNSS signals—that is, to verify that the received signals (1) originated from the

declared satellite transmitter, and (2) arrived without delay [37, 77]. GNSS timing assurance, the topic of this chapter, and message authentication, which ensures data security [78], can be distinguished by their security models. Message authentication is predicated on the computational infeasibility of finding weaknesses in the underlying cryptographic functions or discovering the private signing key—tasks whose probability of success is vanishingly small [76]. By contrast, the intrinsic security of timing assurance is weaker and demands a probabilistic security model because the information of interest is conveyed through the signal timing in addition to the modulated data [37, 79]. Thus, even without reading or altering the modulated data, malefactors can manipulate the information content of a timing signal simply by delaying the signal itself.

GNSS anti-spoofing techniques are broadly categorized as either cryptographic methods that employ secure keys [37, 38, 41] or as non-cryptographic methods that are designed to be sensitive to certain GNSS signal statistics [58, 80]. To date, there is no encompassing framework that addresses the probabilistic nature of each technique or offers an expedient way to combine multiple techniques for a probabilistic security analysis.

The contribution of this chapter is to establish necessary conditions for timing authentication of security-enhanced (i.e., cryptographic) GNSS signals under a probabilistic framework that combines cryptographic and statistical signal processing. The chapter concludes by demonstrating how statistics

meeting these necessary conditions can be coupled with non-cryptographic statistics in a generalized probabilistic framework.

2.2 Data Message Authentication

Data message authentication is predicated on the computational infeasibility of (1) performing a brute-force search for the secret signing key, or of (2) reversing one-way hash functions. The probability of success of either task even under the most optimistic assumptions—the fastest supercomputers running the most advanced cryptanalysis techniques—is so vanishingly small that standards bodies assume near-absolute security of data authentication techniques over periods of years. The National Institute of Standards and Technology considers standardized data authentication techniques with an underlying cryptographic secret key strength of 112 bits secure through the year 2030 [81].

Public-key digital signature algorithms are often employed to achieve data message authentication (e.g., signing emails with the Digital Signature Algorithm). Here, a cryptographic signature algorithm $\mathbb{S}$ generates a message signature s based on the input message m and a secret cryptographic key k_{private}: $\mathbb{S}(k_{\text{private}}, m) = s$. Application of a cryptographic verification algorithm $\mathbb{V}$ to the message-signature pair $\{m, s\}$ with a corresponding cryptographic public key k_{public} derived from k_{private} results in a Boolean: $\mathbb{V}(k_{\text{public}}, \{m, s\}) = T$ or F. If true, the result confirms that the owner of k_{private} generated $\{m, s\}$ and

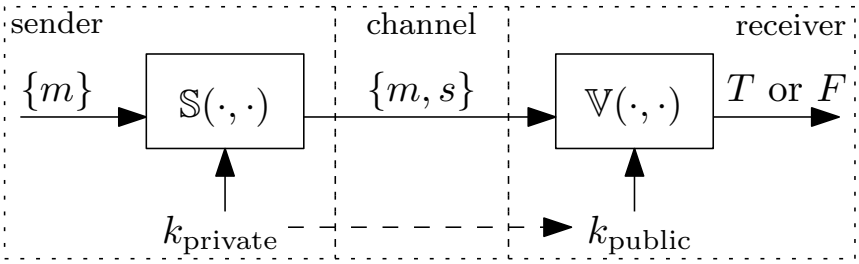


Figure 2.1: Diagram illustrating the public-key digital signature system. The verification algorithm $\mathbb{V}(k_{\text{public}}, \{m,s\}) = T$ iff the message-signature pair $\{m,s\}$ is authentic: the holder of k_{private} generated $\{m,s\}$ exactly.

that $\{m,s\}$ arrived without modification. The public-key digital signature model is illustrated in Fig. 2.1.

I assume that s is unpredictable prior to reception, because so far as it is known, the tasks of either (1) recovering k_{private} from any number of signed messages or from k_{public}, or (2) predicting s based on m or k_{public} are computationally infeasible. These tasks are difficult to talk about in probabilistic terms. Instead, the assumption is based on the mathematics of the underlying cryptographic functions and the scrutiny of security experts worldwide that has yet to reveal a weakness in the approach.

In data message authentication, the result of $\mathbb{V}$ is a sufficient statistic; no other metric is assumed to offer any additional information about the authenticity of the message-signature pair. By analogy with other detection tests described later, one can consider this statistic in the context of a hypothesis test: $\mathbb{V}$ is tested against a threshold to determine the difference between the null hypothesis H_0 (no spoofing) and the alternate hypothesis H_1 (spoofing).

The probability of detection $P_{D,\mathbb{V}}$ of an attack against a cryptographic message authentication system, either an attack that modifies $\{m, s\}$ or forges s, is effectively perfect (i.e., $P_{D,\mathbb{V}} = 1$). The probability of false alarm $P_{F,\mathbb{V}} = 0$.

Given the near certainty with which the technique guarantees data message authentication, it may be surprising that data message techniques alone are insufficient to authenticate timing signals. In the next section, two types of attacks against security-enhanced GNSS signals will illustrate why *signal* authentication requires both data message and timing authentication. Data message authentication is a necessary, but not sufficient, component of comprehensive signal authentication. The latter requires components that span the sub-physical to presentation layer.

2.3 Generalized Model for Security-Enhanced GNSS Signals

Current and proposed security-enhanced GNSS signals can be represented by a simple model from the perspective of a GNSS receiver. Let the signal exiting the radio frequency (RF) front-end of a GNSS receiver after having been downmixed and sampled be modeled as:

$$Y_k = w_k c_k \cos(2\pi f_{IF} t_k + \theta_k) + N_k \tag{2.1a}$$

$$= w_k s_k + N_k \tag{2.1b}$$

Here, at sample index k, w_k is a ± 1-valued security code with chip length T_w, c_k is a known ± 1-valued spreading (ranging) code with chip length T_c, f_{IF}

is the intermediate value of the downmixed carrier frequency, θ_k is the beat carrier phase, and N_k is a sequence of independent, identically distributed zero-mean Gaussian noise samples with variance σ^2 that models the effects of thermal noise in the RF front end. The signal and noise have been normalized so that the modeled signal amplitude is unity. For convenience, $s_k = c_k \cos(2\pi f_{IF} t_k + \theta_k)$ is used to represent the deterministic signal components. Also for convenience, and without loss of generality, the receiver time t_k is assumed to be equivalent to true time with a uniform sampling interval $T_s = t_k - t_{k-1}$.

The model's security code w_k is a generalization of a binary modulating sequence that is either fully encrypted or contains periodic authentication codes. The defining feature of w_k is that some or all of its symbols are unpredictable to a would-be spoofer prior to broadcast from a legitimate GNSS source. The unpredictable symbols of w_k serve two related functions: they enable verification of w_k as originating from a GNSS Control Segment (standard message authentication) and they make possible a hypothesis test for a security code estimation and replay attack [41]. Various security code implementations will be considered in a later section.

2.4 Attacks against Security-Enhanced GNSS Signals

GNSS spoofing is the transmission of counterfeit GNSS signals with the intent to manipulate the position, velocity, and timing (PVT) readout of a GNSS receiver. A spoofer matches its counterfeit signal structure to that of the

authentic signals, as modeled by (2.1). To circumvent the security afforded by the unpredictable security code w_k, the spoofer may attempt one of the following specialized spoofing attacks.

2.4.1 Record and Playback Attack

A record and playback attack is one in which the attacker records and plays back an entire block of RF spectrum containing an ensemble of GNSS signals [5]. It is also sometimes referred to as a "meaconing" attack. Constituent GNSS signals are not typically separated during record and playback, which implies that a meaconing attack cannot arbitrarily manipulate the PVT of target receivers; rather, target receivers will display the position and velocity of the meaconer and a time in arrears of true time. For a single GNSS signal corresponding to a particular satellite, the combined meaconed and authentic received signals can be modeled as

$$Y_k = \alpha w_{k-d} s_{k-d} + N_{m,k} + w_k s_k + N_k \tag{2.2}$$

where $N_{m,k}$ is the noise introduced by the meaconer's RF front end, N_k is the noise introduced by the target receiver's RF front end, and $d > 0$ is the number of samples of meaconing delay, such that the meaconed signal $\alpha w_{k-d} s_{k-d}$ arrives at the target receiver with a delay of d samples relative to the authentic signal $w_k s_k$. The coefficient α is the meaconed signal's amplitude advantage factor.

High performance digital signal processing hardware permits a meaconer located close to its intended target to drive the delay d to ever smaller

values. In the limit as d approaches zero the attack becomes a zero-delay meaconing attack with the meaconed signals code-phase-aligned with their authentic counterparts. Such alignment enables a seamless liftoff of the target receiver's tracking loops, following which a meaconer can increase d at a rate that is consistent with the target receiver clock drift and gradually impose a significant timing delay.

2.4.2 Security Code Estimation and Replay (SCER) Attack

A SCER attack allows greater flexibility than a meaconing attack in manipulating the target receiver's PVT solution. In a SCER attack, a spoofer receives and tracks individual authentic signals and attempts to estimate the values of each signal's unpredictable security code chips on-the-fly. It then reconstitutes a consistent ensemble of GNSS signals, with the security code chip estimates taking the place of the authentic codes, and re-broadcasts these with some delay. For a single GNSS signal corresponding to a particular satellite, the combined SCER-spoofed and authentic received signals can be modeled as

$$Y_k = \alpha \hat{w}_{k-d} s_{k-d} + w_k s_k + N_k \tag{2.3}$$

where $\hat{w}_{k-d}$ represents the security code estimate arriving with a delay of d samples relative to the authentic security code w_k and other quantities are as described previously. The delay d can be modeled as the sum $d = p + e$ of a processing and transmission delay p, which represents the required signal processing and propagation time and which does not contribute to better

estimates of the security code chips, and an estimation and control delay e, which represents an additional delay imposed by the spoofer to improve its estimate of the security code chip values and to control the relative phasing of the spoofed signals so as to impose spoofer-defined position and timing offsets on the target receiver. If the initial delay d exceeds the spreading code chip interval (i.e., if $dT_s > T_c$), then the spoofer will be unable to dislodge the target receiver's tracking loops without forcing re-acquisition. Thus, if the spoofer has an irreducible delay $dT_s > T_c$ then it must first jam or obstruct the incoming GNSS signals to force the target receiver to perform re-acquisition. Attacks in which the spoofer avoids this condition by transmitting the counterfeit signals at a power level such that the sidelobe power is sufficient to disrupt tracking at the victim receiver would trigger the J/N detector under typical received signal strength conditions and in cases where the attacker is unable to physically block the victim antenna. Therefore, such attacks are excluded from consideration.

The success of a SCER attack depends on the accuracy of the security code estimate. Let k_l be the index of the first sample within the lth authentic security code chip. Then for the received sample Y_{k+d}, with $k_l \leq k < k_{l+1}$, a maximum of $\min(e+k-k_l+1, \lfloor T_w/T_s \rfloor)$ security code samples will have been summed within the spoofer to produce the security code estimate $\hat{w}_{k+d-d} = \hat{w}_k$, where $\lfloor x \rfloor$ is the floor of x (the largest integer not greater than x). The accuracy of the chip estimates improves with increasing number of participating samples. For example, the probability of error for hard-decision chip estimates

is $p_e = \text{erfc}(\sqrt{mT_s(C/N_0)_s})/2$ where m is the number of participating samples at sampling interval T_s, $(C/N_0)_s$ is the spoofer's carrier-to-noise ratio, and erfc$(\cdot)$ is the complementary error function. Thus, because $m \leq \lfloor T_w/T_s \rfloor$, small T_w severely limits the accuracy of the security code estimates. Consider that a spoofer receiving the legacy Y-code GPS signal, for which $T_w \approx 2$ μs (i.e., W-code period) [82], at a nominal carrier-to-noise ratio of 48 dB-Hz, generates hard-decision chip estimates with a 30 percent probability of error. A detection strategy for short-delay SCER attacks is detailed in [41].

Long security code chips (e.g., $T_w = 20$ ms for data-bit or navigation message authentication as discussed in Chapter 3) allow the spoofer to increase e and thereby generate highly accurate chip estimates. A large delay $d = p+e$, however, is itself a liability for the spoofer. The signal denial prelude to a SCER attack must be made long enough that d is consistent with the target receiver's clock drift during the denial interval; otherwise, d will lead to a suspicious increment in the target receiver's pseudorange measurements. Thus, the spoofer finds itself vulnerable to detection at low d due to poor security code chip estimates and at high d due to timing anomalies. This is suggestive of the probabilistic nature of signal authentication, which is further elucidated in the following section.

2.4.3 Insufficiency of Data Message Authentication

Consider applying the data message authentication technique of Sec. 2.2 to the attack modeled in (2.2). For a very strong α (i.e., $\alpha \gg 1$), the spoofed

signals overpower the authentic signals. In turn, the GNSS receiver would authenticate the signal: w_{k-d} would pass $\mathbb{V}$ because it was generated from k_{private}. Note that $\mathbb{V}$ cannot identify d. The result is a successful attack that modifies the victim receiver's time estimate by d. The SCER attack proceeds similarly, but its success depends on the accuracy of $\hat{w}_{k-d}$. Clearly, signal authentication requires verifying the consistency of the incoming signal timing (i.e., timing of the spreading and security code) with the receiver's own time estimate. While $\mathbb{V}$ is effective for source authentication, it does not offer timing authentication.

The presence of noise N_k in (2.1) causes additional difficulties for $\mathbb{V}$. Strong noise can cause bit errors, despite application of error correction techniques [83], which results in $\mathbb{V} = 0$. Bit errors occur at a known rate under H_0 (see Sec. 2.5.1 for further details). The probability of a false alarm when verifying $\{m, s\}$ of length $N_{\{m,s\}}$ is $P_{F,\mathbb{V}} = 1 - (1 - p_e)^{N_{\{m,s\}}}$ where p_e is the probability that a single bit is decoded incorrectly.

To reduce the false alarm rate of message authentication in the presence of noise, it is appropriate to consider the statistic $B = \overline{\mathbb{V}} \wedge E$ where E represents the output of an error detection routine (i.e., $E = 1$ for no errors detected). If B asserts under H_1, then an attack is detected: $\mathbb{V} = 0$ and $E = 1$. If B remains low under H_0, then either verification passes or errors were detected in the bit stream. If B asserts under H_0, then there was a false alarm.

The probability of false alarm $P_{F,B}$ is the probability that the error detection routine failed to detect errors when errors were present. For mod-

ernized GNSS signals, this is a very low probability because both error correction and error detection are applied. Note that error correction and detection only applies to low-rate security codes (e.g., at the bit-level) and not high-rate security codes (e.g., embedded in the security code) [41]. For the latter, $\mathbb{V}$ is considered alone. Finally, note that cryptographic operations occur at the presentation layer as defined by the Open Systems Interconnection model [84].

The remaining sections describe the necessary elements for signal authentication, including timing consistency and SCER detectors at the physical layer, and illustrate why the intrinsic security model of signal authentication demands a probabilistic framework compared to pure data authentication.

2.5 Components of an Integrated Probabilistic GNSS Signal Authentication Strategy

In simplest terms, GNSS signal authentication means certifying that a received signal is not counterfeit, that it originates from a GNSS satellite and not a spoofer. As opposed to data authentication, however, GNSS signal authentication is far from absolute; rather, it involves a set of hypothesis tests each with a probability of false alarm. In the formulation adopted here, the tests are designed to detect a spoofing attack under the assumption that a spoofer will either (1) generate a falsified security code that does not match the authentic security code, (2) attempt a non-zero-delay meaconing attack, or (3) attempt a SCER attack. Framed by these assumptions, GNSS signal authentication can be interpreted as involving two authentication sub-types:

(1) code origin authentication, a certification that the security code originates with the GNSS Control Segment, and (2) code timing authentication, a certification that the security code arrives promptly and intact.

In the sections that follow, the functional components that support code origin authentication and code timing authentication are described. As a guide to the discussion, the components and their interconnections are presented schematically in Fig. 2.2 for a security code based on Navigation Message Authentication (NMA), which is a cryptographic anti-spoofing technique that is discussed further in Chapter 3.

For simplicity of presentation, Fig. 2.2 represents the authentication process for a single GNSS signal, i.e., a signal identified by a unique combination of spreading code and carrier frequency. An entire ensemble of GNSS signals is assumed to be downmixed and sampled in the RF front end to produce the sampled signal output Y_k, which is routed to the signal tracking and navigation processor where the raw digital output of the RF front end is correlated against receiver-generated signal replicas to acquire and track multiple constituent GNSS signals. However, from the perspective of downstream components, which are associated with a single GNSS signal, Y_k can be modeled as in (2.1) for unspoofed signals and in (2.2) and (2.3) for meaconed and SCER-spoofed signals, respectively.

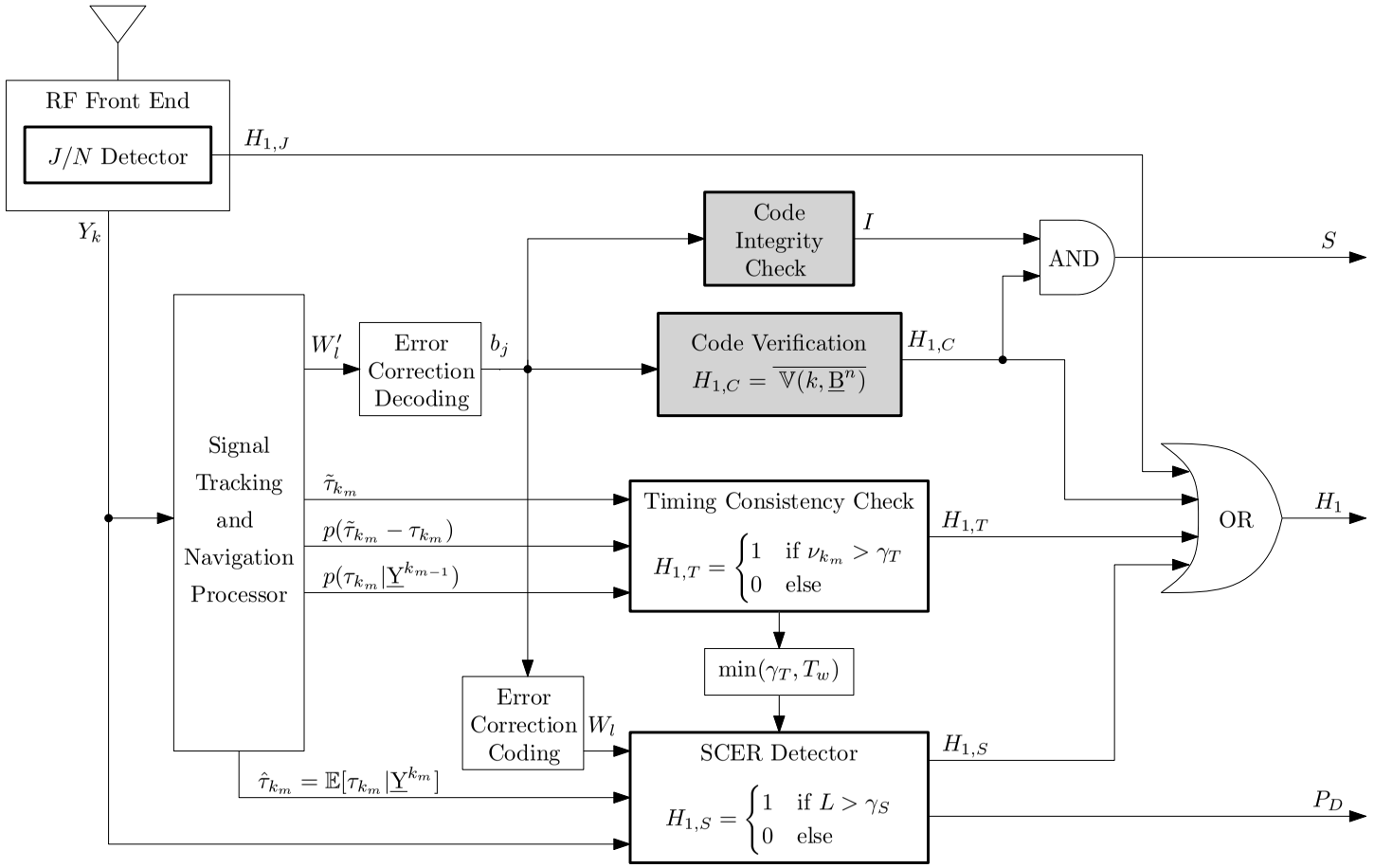


Figure 2.2: Schematic showing GNSS receiver components required for GNSS signal authentication. Components that support code origin authentication are outlined in bold and have a gray fill, whereas components that support code timing authentication are outlined in bold and have no fill. The schematic assumes a security code based on navigation message authentication.

2.5.1 Code Origin Authentication

In the case of a security code based on NMA (c.f., Chapter 3), the signal tracking and navigation processor produces a sequence W_l' of received navigation message symbol estimates. In most cases, these symbols are an error-correction-encoded version of the navigation message data (e.g., the GPS CNAV message is convolutionally encoded before transmission [1]). As the sequence W_l' passes through the error correction decoder, errors introduced by noise in the transmission channel are corrected and the navigation message symbols b_j are recovered. At low carrier-to-noise (C/N_0) ratios some errors may remain in b_j. The code integrity check exploits redundant symbols in b_j (e.g., cyclic redundancy check codes in the GPS CNAV message [1]) to determine whether errors remain. Upon success, the code integrity check sets its logical output I high. For practical purposes, a successful integrity check indicates that the navigation message is correct as received.

The nth block of N_b navigation message symbols $\underline{\mathrm{B}}^n \equiv [b_{j_n}, b_{j_n+1}, \ldots, b_{j_n+N_b-1}]^T$, which in an NMA scheme includes both navigation data and a digital signature, is passed to a code verification algorithm $\mathbb{V}(k, \underline{\mathrm{B}}^n)$ that verifies $\underline{\mathrm{B}}^n$ against a cryptographic key k. If the verification check passes, then $\underline{\mathrm{B}}^n$ can be safely assumed to originate with the GNSS Control Segment. In this case, the logical output signal $H_{1,C}$ remains low. Otherwise, if the verification fails, $H_{1,C}$ is asserted; however this does not necessarily indicate a spoofing attack.

Despite error correction, there may yet remain errors in the symbol stream b_j. A single error within the block $\underline{\mathrm{B}}^n$ would cause the code verification

to fail. Because of this possibility, and by analogy with other hypothesis tests to be introduced shortly, it is convenient to view the code verification as a statistical hypothesis test. The probability of false alarm for the nth verification is $P_{F,\mathbb{V}} \equiv P_{F,C} = 1 - (1 - p_{e,j})^{N_b}$, with $p_{e,j}$ being the probability that b_j is wrong, which depends on C/N_0 over the jth symbol, where $j_n \leq j < j_n + N_b$.

To get a sense for the size of $P_{F,C}$, consider a conservative scenario in which the satellites broadcast a block of N_b = 10,000 non-error-correction-encoded navigation message symbols $\underline{\mathrm{B}}^n$. In this case, the probability that b_j is wrong is $p_{e,j} = \mathrm{erfc}(\sqrt{T_w(C/N_0)_r})/2$. For $(C/N_0)_r \approx 29$ dB-Hz and T_w = 20 ms, $P_{F,C} \approx 0.0001$. If error correction were employed, $P_{F,C}$ would be smaller for a given $(C/N_0)_r$. To ensure that $P_{F,C}$ remains negligible relative to $P_{F,J}$, $P_{F,T}$ and $P_{F,S}$, a receiver can ignore signals whose $(C/N_0)_r < 30$ dB-Hz.

The output $H_{1,C}$ is combined in a logical 'OR' operation with outputs from other hypothesis tests to produce H_1. If the code verification fails ($H_{1,C}$ high) but the code integrity check passes (I high), then, with a very high likelihood, the code verification failure cannot be attributed to symbol errors caused by noise. In this case, the output S is asserted, indicating a nearly certain spoofing attack. As opposed to $H_{1,C}$, which goes high with false alarm rate probability $P_{F,C}$ even under normal unspoofed conditions, the infinitesimal probability of false alarm associated with output S suggests that S need not be viewed probabilistically.

One might ask why $H_{1,C}$ should be considered independently from S. The answer is that if only S is considered then a would-be spoofer could always maintain S low by injecting a symbol stream b_j that repeatedly fails the code integrity check. Thus, the outputs S and $H_{1,C}$ are monitored independently both to prevent this type of an attack and in recognition of the clear certainty of a spoofed condition when S goes high.

2.5.2 Code Timing Authentication

The following functional blocks are involved in code timing authentication: the timing consistency check, the SCER detector, and the jamming-to-noise (J/N) detector.

2.5.2.1 Timing Consistency Check

The timing consistency check is a hypothesis test on the timing of the received spreading code c_k. It amounts to a consistency check on the code phase measurement innovation, or the difference between the measured and predicted code phase, and is essentially a special case of so-called receiver autonomous integrity monitoring [85]. The check takes three inputs from the signal tracking and navigation processor:

$\tilde{\tau}_{k_m}$: the receiver's mth measurement of code phase, expressed as the arrival time of some feature of the incoming signal and defined at receiver time t_{k_m}.

$p(\tilde{\tau}_{k_m} - \tau_{k_m})$: the probability distribution of the code phase measurement noise error.

$p(\tau_{k_m}|\underline{Y}^{k_{m-1}})$: the *a priori* probability distribution of the code phase τ_{k_m} given all input data $\underline{Y}^{k_{m-1}} \equiv [Y_1, Y_2, \ldots, Y_{k_{m-1}}]^T$ up to the $(m-1)$th code phase measurement.

In the consistency check, the difference, or innovation, between the measured code phase $\tilde{\tau}_{k_m}$ and the predicted code phase $\bar{\tau}_{k_m} = \mathbb{E}[\tau_{k_m}|\underline{Y}^{k_{m-1}}]$ is compared against a threshold γ_T. Let $\nu_{k_m} = \tilde{\tau}_{k_m} - \bar{\tau}_{k_m}$ be the innovation. Then the output $H_{1,T}$ is asserted if $\nu_{k_m} > \gamma_T$; otherwise, $H_{1,T}$ remains low. The value of γ_T, which in general varies with time, depends on a pre-selected false alarm probability $P_{F,T}$ for the timing consistency check and on the innovation's conditional distribution, $p(\nu_{k_m}|\underline{Y}^{k_{m-1}})$, which is derived from $p(\tilde{\tau}_{k_m} - \tau_{k_m})$ and $p(\tau_{k_m}|\underline{Y}^{k_{m-1}})$. Commonly, the distributions involved can be modeled as Gaussian, in which case $p(\nu_{k_m}|\underline{Y}^{k_{m-1}})$ can be summarized by its mean $\mathbb{E}[\nu_{k_m}|\underline{Y}^{k_{m-1}}] = 0$ (assuming an unbiased estimator and unbiased measurements) and variance $\sigma_\nu^2 = \sigma_{\Delta\bar{\tau}}^2 + \sigma_{\Delta\tilde{\tau}}^2 + \sigma_m^2$, where $\sigma_{\Delta\bar{\tau}}^2 = \mathbb{E}[(\tau_{k_m} - \bar{\tau}_{k_m})^2|\underline{Y}^{k_{m-1}}]$, $\sigma_{\Delta\tilde{\tau}}^2 = \mathbb{E}[(\tilde{\tau}_{k_m} - \tau_{k_m})^2]$, and σ_m^2 is the pseudorange error due to multipath. The threshold γ_T is the value of γ for which

$$P_{F,T} = \int_\gamma^\infty p(\nu_{k_m}|\underline{Y}^{k_{m-1}}) d\nu_{k_m} \tag{2.4}$$

Note that by comparing ν_{k_m}, not $|\nu_{k_m}|$, against the threshold, the consistency check doubles its sensitivity by making the implicit assumption that the spoofer can only delay the code phase (increase τ_{k_m}).

Another interpretation of γ_T is as the "window of acceptance" referred to in [36]. Between code phase measurement updates, the innovation's conditional distribution $p(\nu_{k_m}|\underline{\mathrm{Y}}^{k_{m-1}})$ widens as receiver clock drift and position uncertainty cause the *a priori* code phase estimate $\bar{\tau}_k$ to become less certain. The distribution can become especially wide if the receiver has a poor clock and is subjected to prolonged jamming or signal blockage. If, after re-acquisition, the innovations remains below γ_T, then the timing of the re-acquired signal is within the window of acceptance; i.e., it is consistent with the assumed uncertainty in $\bar{\tau}_k$.

It should be noted that $p(\nu_{k_m}|\underline{\mathrm{Y}}^{k_{m-1}})$ depends on all signals being tracked by the receiver, not only on the individual signal whose code phase measurement is $\tilde{\tau}_{k_m}$. This is because the *a priori* distribution $p(\tau_{k_m}|\underline{\mathrm{Y}}^{k_{m-1}})$, from which $p(\nu_{k_m}|\underline{\mathrm{Y}}^{k_{m-1}})$ is derived, is a complete summary of what the receiver knows about τ_{k_m} based on all the raw samples in $\underline{\mathrm{Y}}^{k_{m-1}}$. When a particular signal is acquired or re-acquired, its authentication depends on the time aiding provided by other signals. Vector tracking algorithms [86] are particularly well suited for GNSS signal authentication because they combine timing information from all signals and can be designed to produce $p(\nu_{k_m}|\underline{\mathrm{Y}}^{k_{m-1}})$ as part of their routine processing.

To give a better understanding of factors that affect γ_T, two scenarios are considered. The top panel of Fig. 2.3 shows γ_T in a static scenario as a function of $(C/N_0)_r$ for $P_{F,T} = 0.0001$. Under H_0 (no spoofing), the analysis assumes that $p(\nu_{k_m}|H_0) = \mathcal{N}(0, \sigma^2_{\Delta\bar{\tau}} + \sigma^2_{\Delta\tilde{\tau}} + \sigma^2_m)$ where

- $\sigma^2_{\Delta\tilde{\tau}}$ is the predicted code phase measurement error variance—a function of satellite geometry and $(C/N_0)_r$, which, for the purposes of this analysis, corresponds to a particular, but fairly typical 8-satellite arraignment and assumes every satellite has the same $(C/N_0)_r$;

- $\sigma^2_{\Delta\tilde{\tau}} = dB_{\mathrm{DLL}}T_c^2/(4(C/N_0)_r)$ is the measured code phase measurement error with correlator spacing $d = 1/2$ chip, $T_c \approx 1$ μs, and phase-lock-loop-aided delay locked loop (DLL) bandwidth $B_{\mathrm{DLL}} = 0.05$ Hz; and,

- σ^2_m is a conservative estimate of the assumed multipath error variance within a receiver that implements a multipath mitigation scheme; it is calculated by multiplying by 3 the maximum root mean square pseudorange multipath error for a typical (Fig. 5 [24]).

The plot shows how the window of acceptance must widen as $(C/N_0)_r$ decreases to maintain $P_{F,T} = 0.0001$.

The bottom plot of Fig. 2.3 corresponds to a scenario in which a stationary receiver falls victim to a complete satellite signal outage (e.g., via jamming or blockage) when driven by a temperature-compensated crystal oscillator (TCXO) with short-term stability $\sigma_{\mathrm{TCXO}} = 10^{-8}$ or an oven-controlled crystal oscillator (OCXO) with short-term stability $\sigma_{\mathrm{OCXO}} = 10^{-11}$ [87]. The plot assumes that the final tracking $(C/N_0)_r$ before the outage was 40 dB-Hz and that the outage lasts for duration T_{outage}. Clearly, the longer the interval T_{outage}, the greater γ_T must be to maintain $P_{F,T} = 0.0001$. As one might expect, OCXO-driven receivers maintain a lower γ_T for a given T_{outage} than their

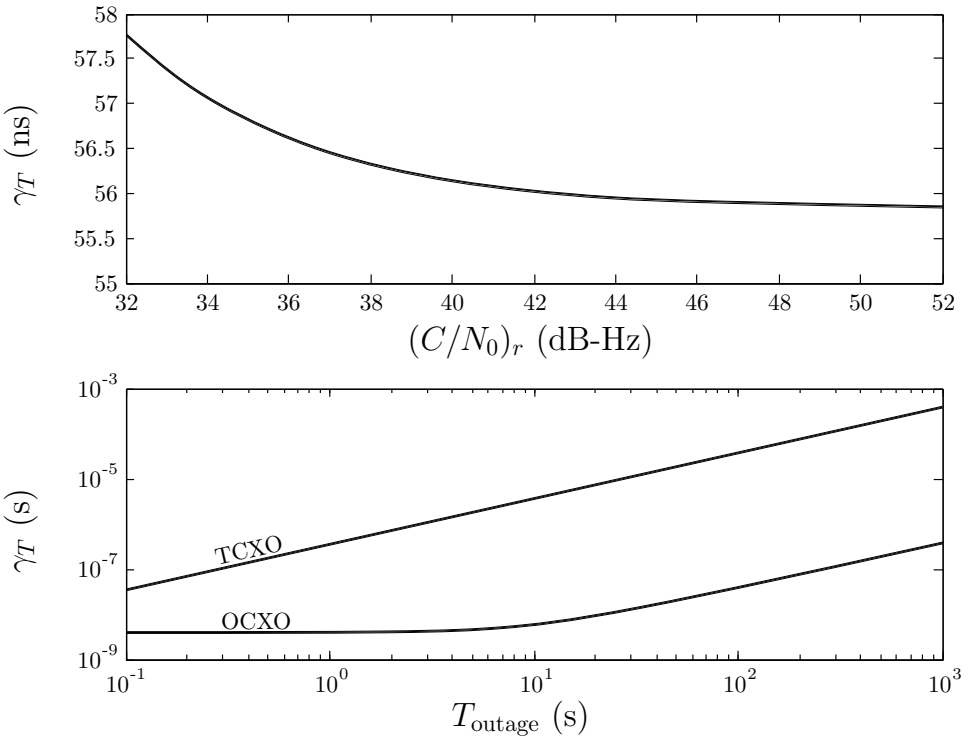


Figure 2.3: Sensitivity analysis for γ_T under two static scenarios with $P_{F,T} = 0.0001$. Top panel: γ_T versus $(C/N_0)_r$ for a particular 8 satellite constellation each with $(C/N_0)_r$. Bottom panel: γ_T versus T_{outage} for a TCXO- and an OCXO-driven receiver.

TCXO-driven counterparts. The bend in the OCXO plot marks a transition from an innovation distribution in which the measurement noise and initial timing uncertainty dominate to one in which the uncertainty contributed by the OCXO's frequency instability dominates.

The timing hypothesis test depends critically on the accuracy of the receiver's internal oscillator because the latter provides a reference for measuring the promptness of the incoming signal. Thus, somewhat counterintuitively, the receiver must already have an accurate estimate of time, and know its estimate to be accurate, if it is to validate the promptness of an incoming timing signal. Note that timing consistency alone cannot detect spoofing attacks in cases where the spoofed signal's delay remains below γ_T. Thus, timing consistency is necessary but not sufficient for timing authentication of security-enhanced GNSS signals; it must be combined with other tests to ensure a high probability of spoofing detection.

2.5.2.2 Security Code Estimation and Replay (SCER) Attack Detector

The SCER detector is a hypothesis test at the physical layer that decides whether the security code in the incoming samples Y_k arrives (1) intact and (2) near the *a posteriori* code phase estimate $\hat{\tau}_{k_m} = \mathbb{E}[\tau_{k_m} | \underline{\mathrm{Y}}^{k_m}]$ produced by the signal tracking and navigation processor [41]. At least one of these two conditions is violated if a SCER attack is underway. The SCER detector performs time-weighted correlations with Y_k over the lth unpredictable security

chip interval to produce a single-chip statistic S_l, which is derived in Sec. IV. of [41]. These correlations involve the error correction encoded symbols W_l, which are identical to the raw received symbols W_l' if no symbol errors are present in W_l, but, in general, include corrections to W_l made possible by the operation of error correction decoding and subsequent re-encoding.

The SCER detector combines a set of N single-chip correlations S_l into a detection statistic L, which it compares with a threshold γ_S that is set by a pre-selected probability of false alarm, $P_{F,S}$. If a SCER attack is underway, and if the estimation delay e is sufficiently small, then L will rise above γ_S, causing $H_{1,S}$ to assert. The SCER detector assumes that the spoofer's C/N_0 advantage over the target receiver's is limited to approximately 3 dB (i.e., $(C/N_0)_s \leq (C/N_0)_r + 3$ dB). This assumes the spoofer and defender are physically close and both use a commercially-available antenna with similar gain patterns. The at-most-3-dB advantage accounts for a scenario in which the spoofer's antenna may have a better noise figure or a better line-of-sight to the satellite, but not scenarios in which the spoofer employs a high-gain antenna array. The SCER detector further assumes that a J/N detector is monitoring the incoming in-band power so that the power advantage of the received spoofing signal ensemble is limited to approximately 4 dB above the authentic signal ensemble. Attacks in which the spoofer broadcasts its counterfeit signals with a power advantages greater than 4 dB fall outside the range of applicability of the SCER detector (Sec. VI.B. in [41]) and can be detected at a low false alarm rate by a properly configured J/N detector (c.f., Sec. 2.5.3 and [88]).

This is why a J/N detector is a necessary component of an integrated signal authentication strategy. The J/N detector threshold is governed by a pre-determined false alarm probability $P_{F,J}$ [88].

The distribution of L, $p_{L|H_j}(\xi|H_j)$ for $j = 0, 1$, is distributed as a non-central chi-square distribution with N degrees of freedom and non-centrality parameter λ_j. Given $p_{L|H_j}(\xi|H_j)$ for $j = 0, 1$, the threshold γ_L can be chosen to satisfy a pre-determined probability of false alarm $P_{F,L}$ by solving for γ_L in

$$P_{F,L} = \int_{\gamma_L}^{\infty} p_{L|H_0}(\xi|H_0)d\xi \tag{2.5}$$

A corresponding probability of detection $P_{D,L}$ is

$$P_{D,L} = \int_{\gamma_L}^{\infty} p_{L|H_1}(\xi|H_1)d\xi \tag{2.6}$$

In a typical application, the SCER detector performs a hypothesis test just after each code verification $\mathbb{V}(K, \underline{\mathrm{B}}^n)$. There is little point in performing the test more frequently, since the authenticity of the symbols b_j, and by extension the encoded symbols W_l used in the SCER detector correlations, cannot be guaranteed until the code verification has been performed.

The SCER detector outputs a probability of detection P_D that depends on the detector's model for the statistics of a SCER spoofing attack, which in turn depend on the possible estimation delay e (Sec VI.C. in [41]). In setting P_D, the SCER detector pessimistically assumes that the total estimation delay in seconds eT_s could be as large as γ_T, which means that at each security code chip transition the spoofer could already have an estimate based on as much

as $\min(\gamma_T, T_w)$ seconds into the upcoming chip. A degraded P_D reflects the penalty paid, in terms of ability to detect spoofing, for uncertainty in ν_{k_m}, which could be caused by an extended period of GNSS jamming or blockage. As $p(\nu_{k_m}|\underline{Y}^{k_{m-1}})$ widens and γ_T increases, the limitations on spoofing delay d become less stringent. Knowing this, a SCER-attack spoofer can increase the estimation time e, thereby improving the reliability of its security code chip estimates. When the spoofer's $(C/N_0)_s$ is high and γ_T is large (e.g., $(C/N_0)_s > 50$ dB-Hz and $\gamma_T > 300$ μs), then the null and spoof hypotheses become virtually indistinguishable within the SCER detector and P_D drops. Even though γ_T may subsequently contract and P_D increase, a low P_D creates a window of vulnerability after which signal authentication assurance is permanently degraded.

2.5.3 Total In-Band Power Monitor

During a spoofing attack against a security-enhanced GNSS signal, an admixture of authentic and spoofed signals are present [c.f., (2.3) and (2.2)], which will increase the measured in-band signal power P_T. The purpose of this J/N detector is to monitor the nominal in-band power levels and detect when additional power is present due to spoofed signals, thereby limiting the power advantage of the spoofer.

Consider the following hypothesis pair, which models P_T as measured by a defender's front-end:

$$H_0 : P_T = P_A + N_0 B, \tag{2.7a}$$

$$H_1 : P_T = P_A + P_S + N_0 B \tag{2.7b}$$

Here, $P_A = \sum_i P_{A,i}$ is the total received signal power from each authentic signal $P_{A,i}$, $P_S = \sum_i P_{S,i}$ is the total received signal power from each spoofed signal $P_{S,i}$, N_0 is the one-sided noise power density at the low-noise amplifier (LNA), and B is the one-sided LNA filter bandwidth.

A spoofer seeking to maximize the likelihood of a successful attack will set its power advantage factor $\eta \equiv P_S/P_A > 1$ since higher values of η reduce the defender's probability of detecting a spoofing attack (c.f., [41], Sec. IV.B). Applying this notation to the hypothesis pair in (2.7) yields

$$H_0 : P_T = P_A + N_0 B, \tag{2.8a}$$

$$H_1 : P_T = P_A(1 + \eta) + N_0 B \tag{2.8b}$$

Given the densities $p_{P_T|H_j}(\xi|H_j)$ for $j = 0, 1$, an optimal detection test exists:

$$P_T \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{P_T} \tag{2.9}$$

The threshold γ_{P_T} corresponding to a specific probability of false alarm P_{F,P_T} can be computed:

$$P_{F,P_T} = \int_{\gamma_{P_T}}^{\infty} p_{P_T|H_0}(\xi|H_0) d\xi \tag{2.10}$$

A corresponding probability of detection P_{D,P_T} is

$$P_{D,P_T} = \int_{\gamma_{P_T}}^{\infty} p_{P_T|H_1}(\xi|H_1)d\xi \tag{2.11}$$

In practice, computing analytical forms of $p_{P_T|H_j}(\xi|H_j)$ for $j = 0, 1$ for the detection test of (2.9) is intractable because η has no determinable distribution and N_0 can vary widely depending on the number and time-varying magnitudes of natural and man-made interference sources that contribute to T_I. Given these difficulties, a more modest goal for the in-band signal power test is sought.

Because the SCER detector assumes that $\eta \leq \eta_{\max}$, the modest goal of the operational in-band signal power detection test is to limit $\eta \leq \eta_{\max}$ so that values of $\eta > \eta_{\max}$ result in the measured P_T exceeding γ_{P_T} for an acceptable P_{F,P_T}. A value of γ_{P_T} that meets these goals can be derived based on historical atmospheric data from [89]. In addition, so-called personal privacy devices (i.e., jammers) are becoming increasingly prevalent. Statistics of these devices in [90] can further help set γ_{P_T}.

2.5.4 Other Security Code Implementations

The above components of a GNSS signal authentication system are specific to a security code based on NMA (c.f., Chapter 3). The components are also valid for the civil public spreading code authentication technique introduced in [36] except that in this case the symbols b_j are routed directly to the SCER detector where they are used to seed a pseudorandom spreading code

generator a segment of whose output gets inserted into the local spreading code replica.

For private spreading code authentication schemes such as the civil level-3 technique introduced in [36] and military GPS Y- and M-code security, the code verification block in Fig. 2.2 is unnecessary. The figure can be adapted to these cases by setting $H_{1,C}$ permanently low and by routing the symbols b_j directly to the SCER detector. These private-key techniques rely on storage of a secure "red key" in tamper-resistant hardware within the receiver. Segments of the symbol stream b_j are coupled with the red key in the SCER detector to produce a seed for a pseudorandom spreading code generator. Only segments of the generated code are used in the civil private-key technique of [36], whereas the continuous output of the generator constitutes the security code for GPS Y- and M-code security.

2.6 Operational Definition of GNSS Signal Authentication

With the authentication components and their interactions specified, an operational definition of GNSS signal authentication—in other words, how signals are declared authentic in practice—can now be formulated. A GNSS signal is declared authentic at a given moment if and only if, during the time elapsed since some initialization event at which the receiver was known to be tracking only genuine GNSS signals, (1) the logical output S has remained low,

(2) the logical output H_1 has remained low, and (3) the real-valued output P_D has remained above an acceptable threshold (e.g., 0.9).

Some comments about this operational definition are in order. First, although there may be reasonable alternatives to this definition, they cannot be substantially different. Aside from the variations that occur when implementing other security codes as discussed previously, the components of the proposed definition are each unique and necessary. Second, although a GNSS signal may be pronounced authentic by the above operational definition, it may in fact be counterfeit. Practical constraints of hypothesis testing prevent P_D from reaching unity. For example, for the NMA-based security codes discussed later on, nominal P_D may drop as low as 0.97. Moreover, jamming or signal blockage can temporarily reduce P_D. Inversely, even though a signal may be declared unauthentic, it may actually be authentic. In the case that S is asserted, the incoming signal is certainly unauthentic; on the other hand, H_1 will at times assert even under unspoofed conditions. It has a false alarm probability

$$P_F = 1 - (1 - P_{F,J})(1 - P_{F,C})(1 - P_{F,T})(1 - P_{F,S})$$

which is greater than any of the false alarm probabilities for the individual tests that can trigger H_1. Third, movement of P_D below the acceptable threshold does not necessarily indicate a SCER spoofing attack, it only indicates that the SCER detector's probability of detecting a SCER attack has been compromised, and thus the currently tracked signal cannot be considered authentic.

2.7 Probabilistic Framework

In the case of data message authentication, only the measurement $z = \mathbb{V}$ was necessary to determine the authenticity of $\{m, s\}$. In the case of signal authentication, the timing consistency, SCER, and in-band power detector and error correction are required to authenticate the GNSS signal. Under the probabilistic framework for cryptographic GNSS signal authentication, the measurement incorporates all of the statistics:

$$\boldsymbol{z} = [\overline{\mathbb{V}} \wedge E, \nu, L, P_T]^T \tag{2.12}$$

Given $\boldsymbol{z}$, one can consider the joint probability distribution $p_{\boldsymbol{z}|H_j}(\boldsymbol{\xi}|H_j)$ for $j = 0, 1$ and form the appropriate tests based on the density function. In this case, the system-wide probability of false alarm P_F is

$$P_F = \int_{\gamma}^{\infty} p_{\boldsymbol{z}|H_0}(\boldsymbol{\xi}|H_0) d\boldsymbol{\xi} \tag{2.13}$$

for a given γ. A corresponding system-wide probability of detection P_D is

$$P_D = \int_{\gamma}^{\infty} p_{\boldsymbol{z}|H_1}(\boldsymbol{\xi}|H_1) d\boldsymbol{\xi} \tag{2.14}$$

The probabilistic framework for signal authentication offered here illustrates how the intrinsic security of signal authentication is much weaker than that of data message authentication. The security depends on multiple detection tests at several network layers (i.e., sub-physical, physical, and presentation layers) each with their own probabilities of detection and false alarm. Furthermore, the system-wide P_D and P_F are set subject to a security risk assessment unique to individual users and scenarios.

2.7.1 Combination with Non-Cryptographic Techniques

The statistics that represent the necessary conditions for security-enhanced GNSS signal authentication can be readily coupled with other non-cryptographic statistics in a generalization probabilistic framework. Non-cryptographic techniques have been proposed that examine incoming signal statistics of Y_k for distortions that are present during a spoofing attack [58]. One example is the complex early-minus-late tap difference D. To combine this statistic with the cryptographic statistics in (2.12), D is simply appended to $\boldsymbol{z}$:

$$\boldsymbol{z} = [\overline{\mathbb{V}} \wedge E, \nu, L, P_T, D]^T \tag{2.15}$$

Then, a new characterization of $p_{\boldsymbol{z}|H_j}(\boldsymbol{\xi}|H_j)$ can be computed either analytically or empirically.

2.7.2 Characterizing the Joint Probability Distribution

The success of this probabilistic approach to GNSS signal authentication hinges on the correct characterization of $p_{\boldsymbol{z}|H_j}(\boldsymbol{\xi}|H_j)$. Thus far, only two hypotheses were considered: the null hypothesis of no spoofing, and the alternative hypothesis of spoofing. In practice, additional hypotheses need to be tested. For example, multipath causes statistical variations similar to spoofing [91]. If the spoofing and multipath hypothesis are indistinguishable then a high false alarm rate exists [58]; hence, a multipath hypothesis is necessary to reduce false alarm rates between spoofing and multipath. Thus, three hypothesis will each need to be characterized.

Characterizing $p_{\boldsymbol{z}|H_0}(\boldsymbol{\xi}|H_0)$ under the null hypothesis H_0 is amenable to an analytical solution assuming the thermal noise N_k takes on a Gaussian distribution. Characterizing $p_{\boldsymbol{z}|H_1}(\boldsymbol{\xi}|H_1)$ under the multipath hypothesis H_1 is suited to a combined analytical and empirical approach. Multipath can be modeled analytically [92] but the combinations of real-world recordings with a theoretical analysis will offer a better characterization of $p_{\boldsymbol{z}|H_1}(\boldsymbol{\xi}|H_1)$ than analysis alone. Finally, characterizing $p_{\boldsymbol{z}|H_2}(\boldsymbol{\xi}|H_2)$ under the spoofing hypothesis H_2 is only possible empirically, and even then, only partially. The number of spoofing attack vectors is enormous; only a subset can be considered. Empirical analysis will leverage the Texas Spoofing Test Battery [42]. This collection of recorded spoofing scenarios is available for evaluating civil Global Positioning System signal authentication techniques and offers a wide-range of potential spoofing attacks with which to generate $p_{\boldsymbol{z}|H_2}(\boldsymbol{\xi}|H_2)$.

2.8 Conclusion

This chapter has illustrated why data message authentication techniques alone are not sufficient for timing assurance in the context of a security-enhanced Global Navigation Satellite System (GNSS) signal. Instead, a probabilistic framework that combines cryptography and signal processing detection tests at multiple network layers is necessary to capture the subtleties and the weaker intrinsic security of signal authentication. The next chapters demonstrate how this theoretical framework can be applied to develop and evaluate cryptographic and non-cryptographic GPS spoofing defenses.

Chapter 3

Practical Cryptographic Civil GPS Signal Authentication

3.1 Introduction

It is convenient to distinguish cryptographic spoofing defenses, which rely on secret keys that encrypt or digitally sign components of the broadcast signals, from non-cryptographic defenses, which do not depend on encryption or digital signatures. Among non-cryptographic defenses, the multi-antenna defense [53, 93] appears to be one of the strongest, although it remains vulnerable to the coordinated spoofing attack explored in [3]. This defense requires two or more antennas spaced by an appreciable fraction of the approximately 20-cm GPS signal wavelength, which would tend to increase receiver cost, weight, and size. As a result, the multi-antenna defense is unlikely to be widely adopted by commercial GPS manufacturers. This is also true of other non-cryptographic defenses involving inertial measurement units or other hardware, which would exceed the cost, mass, or size constraints of a broad range of applications.

Cryptographic spoofing defenses are attractive because they offer significant protection against spoofing relative to the additional cost and bulk

required for implementation. While it must be conceded that no anti-spoofing technique is impervious to the most sophisticated attacks, a cryptographic defense significantly raises the bar for a successful attack and can be combined with non-cryptographic spoofing defenses for better security than either category could offer separately.

Several civil GPS cryptographic spoofing defenses have been proposed whose implementation would require fundamental changes to the legacy GPS signal structure (e.g., [14, 36, 43]). These defenses are unlikely to be implemented over the next decade given the static nature of GPS signal definitions [94].

A growing literature suggests navigation message authentication (NMA) is a practical basis for civil GPS signal authentication [13, 14, 36, 95]. In NMA, the low-rate navigation message is encrypted or digitally signed, allowing a receiver to verify that the GPS Control Segment generated the data. NMA could be implemented without fundamental changes to the GPS Interface Specification by exploiting the extensibility of the modern GPS civil navigation (CNAV) messaging format. Moreover, NMA has been proposed for implementation in the European Galileo GNSS [44, 96].

Previous papers have pointed out that signal authentication based on NMA may be vulnerable to replay-type spoofing attacks [14, 36]. Thus, whereas it is clear that NMA authenticates the origin of the navigation data, there has been uncertainty regarding whether NMA can be used to authenticate the underlying GPS signal, which demands resistance against replay-

type spoofing attacks. The combination of this work in this chapter and the statistical test recently developed in Ref. [41] clears up this uncertainty by demonstrating that NMA can in fact offer integrated civil GPS signal authentication—that is, combined data and signal authentication—if it is paired with timing authentication based on statistical hypothesis tests.

The present work offers contributions beyond those given in [13, 14, 36, 44, 95, 96]. First, it identifies sensible design criteria for civil GPS signal authentication and, second, applies this framework to evaluate several proposed candidate authentication strategies. Third, it proposes a specific cryptographic signal authentication implementation for civil GPS that meets the design criteria and is packaged for immediate adoption.

3.2 Design of NMA in Consideration of the Probabilistic Anti-Spoofing Framework

It is easy to appreciate the advantage of short over long security code chips given the authentication architecture proposed in Fig. 2.2. Short chips such as the $T_w \approx 2$ μs chip of the legacy GPS Y code keep $\min(\gamma_T, T_w)$ to less than a few microseconds and thereby prevent significant degradation in P_D (c.f., Chapter 2) even during a prolonged signal blackout, whereas long chips such as $T_w \approx 20$ ms for NMA allow significant degradation in P_D for the same outage. This weakness of NMA-based GNSS signal authentication has been noted—although not in these formal terms—in [36] and [14]. Practically, the weakness translates into the following additional requirements for NMA-based

GNSS security: For a static receiver in a known location, maintaining P_D high requires either continuous tracking of at least one strong GNSS signal or a clock that does not drift significantly during whatever complete signal outages occur. For a receiver mounted on a dynamic platform, either continuous tracking of at least 4 strong GNSS signals or a clock and inertial measurement unit (IMU) combination that does not drift significantly are required.

Given these requirements, one may question whether NMA-based GNSS security will be useful in practice. One should bear in mind that for many applications of interest the prolonged signal denial required to significantly degrade P_D would be highly suspicious. For example, consider a static receiver with a TCXO having short-term stability 10^{-8}. A spoofer would be forced to preface a spoofing attack with a 150-second complete signal denial interval in order to increase γ_T beyond 5 μs (assuming $P_{F,T} < 0.002$) and thereby cause a significant reduction in P_D [41]. If the complete signal denial is done via jamming, then the J/N detector will trigger; if done by obstructing the target receiver's antenna, this requires close physical access. In any case, the signal outage will appear suspicious.

Also, it is worth noting that security code alternatives to NMA are not foolproof and are likely to be less practical. Indeed, it appears that no exclusively cryptographic defense, no matter how short the security chip interval T_w, can detect a well-executed near-zero-delay meaconing attack. (This is why such an attack is excluded from the attack model in the discussion on components of signal authentication in Chapter 2.) Universal vulnerability

to near-zero meaconing suggests the need for a layered approach that combines cryptographic signal authentication with non-cryptographic techniques such as the vestigial signal defense [58]. It also suggests that expectations for GNSS signal authentication must be modest: the goal should not be preventing a successful attack at all cost, but making one difficult. Furthermore, a GNSS signal authentication scheme's potency must be weighed against its practicality. This tradeoff is the subject of the next section.

3.3 Design and Evaluation of Cryptographic Signal Authentication Strategies

The previous section considered general GNSS signal authentication, which relies in part on some or all of the security code w_k being unpredictable to a would-be spoofer. This section considers candidate signal authentication strategies (i.e., the design of w_k) specifically for civil GPS. These strategies are evaluated based on their:

effectiveness: how difficult they make it for a spoofer to carry off a successful spoofing attack; and their

practicality: how likely they are to be implemented.

In practice, a tradeoff emerges between effectiveness and practicality with the most effective approaches being impractical. This section elucidates this tradeoff and selects the most effective strategy from the set of practical ones.

3.3.1 Selecting T_w

The security code chip length T_w is fundamental to the design of a signal authentication strategy. To evaluate potential choices of T_w, the notions of effectiveness and practicality can be refined as follows. Effective strategies enable frequent signal authentication and offer receivers a high probability of detecting an attack. Such strategies significantly raise the bar for a successful spoofing attack but are not necessarily impervious to the most sophisticated attacks. Additionally, practical strategies (1) remain backward compatible, meaning legacy equipment will function correctly without modification if the approach is implemented (e.g., GPS L1 C/A remains unaltered) and (2) avoid fundamental modifications to the GPS Interface Specification (IS). The GPS Control Segment is less likely to support implementation of a civil GPS signal authentication strategy that fundamentally alters the GPS IS [94].

As noted earlier, a short T_w has the advantage that it prevents significant degradation of P_D due to timing uncertainty. Although one could define a new signal definition to support an arbitrarily small T_w, this approach is impractical because it fundamentally modifies the GPS IS. A more practical approach is to leverage one of the fundamental intervals defined for civil GPS signals in the GPS IS: the spreading code chip interval (approximately 100 ns for L5 and approximately 1 μs for L1 and L2) or the navigation data bit interval (20 ms for all civil frequencies).

In terms of effectiveness, setting T_w equal to the spreading code chip interval, a strategy known as spreading code authentication (SCA), is best. SCA

meets the first criteria for practicality: backward-compatible SCA strategies have been proposed for GPS L5 [14, 36]. However, SCA does not satisfy the second requirement for practicality: it requires modification of the civil spreading codes, which must be considered a fundamental—and therefore unlikely—alteration of the GPS IS.

Consider instead setting T_w equal to the navigation data bit interval. This is the navigation message authentication (NMA) approach. In other words, $w_k = d_k$ where d_k are samples from the ± 1-valued navigation message and $T_w = 20$ ms. One can either make all or part of the navigation message unpredictable to generate w_k. A possible approach encrypts all or nearly all of the navigation message with a cryptographic cipher (e.g., message recovery mode [13] or hybrid message recovery mode). This approach generates a high average rate of unpredictable navigation data bits, which reduces the required interval between signal authentication tests, but it is ultimately impractical since complete or nearly complete navigation message encryption would not be backward compatible and would require a fundamental alteration of the GPS IS.

The only practical strategy, then, is to form w_k by introducing periodic randomness into the navigation message. This NMA-based approach is assumed hereafter.

3.3.2 Generating Periodic Unpredictability

The previous discussion settled on a strategy of forming w_k by transmitting a periodically unpredictable navigation message. Unpredictable, however, does not mean unverifiable. A receiver can verify the origin of w_k—that is, who generated the security code—to prevent being spoofed with a forged w_k. Cryptographic digital signature protocols would enable receivers to verify the origin of signed messages. By their very nature, the signatures that enable this authentication are unpredictable prior to broadcast. The unpredictability of digital signatures allows receivers to treat the digital signature as a security code.

Before comparing various digital signature protocols for NMA, refined definitions of effectiveness and practicality with respect to digital signature protocols are offered to guide the selection process. A digital signature protocol is considered effective for signal authentication if it is standardized, is cryptographically secure, and offers a high P_D for a low P_F:

- Standardization indicates that the protocol has been well-studied by the cryptography community and is thought to be secure against even the strongest cryptographic attacks such as those described in [13] and [97]. Standardized protocols also facilitate adoption: verified open-source implementations often exist and certification programs can validate proper operation of cryptographic modules.

- The equivalent symmetric-key strength b_s in units of bits is a useful measure of the strength of a cryptographic protocol. The U.S. National Institute of Standards and Technology (NIST) considers $b_s \geq 112$ secure for the years 2011–2030 [81]. To meet NIST guidelines, a cryptographic civil GPS signal authentication strategy must therefore set $b_s \geq 112$.

- A high P_D means that the receiver has a high likelihood of detecting a spoofing attack. The number N of chip-level correlations S_l that are combined to generate each SCER-attack detection statistic L increases with the length of the digital signature. Since a larger N tends to increase SCER-attack P_D, a longer signature leads to a higher P_D for a fixed probability of false alarm $P_{F,S}$ and threat model [41]. It is reasonable to define strategies offering $P_D \geq 0.95$ for $P_{F,S} = 0.0001$ as effective. For NMA, a digital signature that produces a signature length of approximately 400 bits will exceed this requirement in typical scenarios [41].

A practical digital signature protocol is one that does not burden the limited resources of the Control, User, or Space Segment. This chapter considers a protocol to be practical if:

- its implementation does not adversely affect a standard receiver's ability to determine its position from the broadcast ephemeris;

- the percentage of the GPS navigation message required to transmit the digital signature is low (e.g., 10 percent or less);

- the computational resources of the receiver that are devoted to authentication are a small fraction of those devoted to standard GPS signal processing;
- it requires no additional receiver hardware, which would increase receiver cost, size, or weight; and,
- it allows feasible key management.

Protocols that have a short signature length for a given b_s, that have a low computational burden, and that can be implemented entirely in software are practical.

Given the foregoing definitions of the terms effective and practical as applied to digital signature protocols, the following discussion settles on a protocol appropriate for civil GPS signal authentication.

3.3.2.1 Public vs. Private Key Protocols

The primary categorization for digital signatures is their classification as either public key (i.e., asymmetric) or private key (i.e., symmetric) [76]. Private key algorithms are generally more computationally efficient and offer shorter signature length than public key protocols, but they require a shared and secure private key. This requirement makes private key digital signatures, however effective, impractical for civil GPS signal authentication because securely storing a private key requires tamper-proof receiver hardware [39]. Furthermore, key management for symmetric protocols would be complicated: if

any one of the private keys were disclosed, then every receiver would need to securely update the private key. Thus, private key protocols are impractical for civil GPS signal authentication.

On the other hand, public key protocols are practical because the public key k_{public} can be stored unsecured in receiver memory. Despite the fact that the cryptographic key may be widely known, public key protocols offer as much security as private key methods for a given b_s and are believed to be secure against even the strongest cryptographic attacks, such as those described in [13] and [97]. Although public key protocols generally have a higher computational burden than private key protocols, some public key digital signature protocols still have a low computational burden relative to the standard GPS signal processing, which makes them practical for this application. Moreover, public key techniques allow feasible key management in the form of a public key infrastructure.

3.3.2.2 Public Key Management

Key management for a civil GPS authentication scheme based on public key digital signatures would be fairly straightforward. The GPS Control Segment would publish a unique public key $k_{\text{public},i}$ for each satellite i (i.e., for each unique pseudorandom spreading code), hereafter referred to as k_{public} for convenience. A unique k_{public} for each satellite offers an additional layer of defense against cryptographic attacks. GPS receivers would then store k_{public} in local (potentially unsecured) memory. Although some proposals have sug-

gested transmitting k_{public} over the GPS navigation message, this creates a new spoofing attack possibility whereby a spoofer broadcasts a counterfeit key to the receiver. Instead, the Control Segment should leverage the key management techniques already developed to facilitate public key protocol implementation. In general, k_{public} should be distributed through a secure secondary channel, such as over the Internet, with the guarantee of a mutually trusted third party. One frequently employed framework is called public key infrastructure (PKI) [76, 98]. In this framework, trusted Certificate Authorities would certify (i.e., sign) the Control Segment public key thereby binding k_{public} to the identity of the Control Segment and preventing a spoofer from publishing a forged key. The certified Control Segment public key would then be stored on the receiver for signature verification. Because public keys can have a valid lifetime (i.e., cryptoperiod) of several years, a receiver's stored public key can be updated infrequently [81]. Thus, a receiver need not be continuously connected to the Internet to take advantage of public key signature methods. In the case of a security breach, PKI also offers key revocation techniques upon which the Control Segment can rely [99].

3.3.2.3 Public Key Digital Signature Generation and Validation

An overview of public key digital signatures will clarify the code verification block of Fig. 2.2. To digitally sign d_k and embed the signature in the navigation message thereby forming w_k, the GPS Control Segment would compute a private key k_{private} that remains secret and a public key k_{public} that

is distributed to users and stored in receiver memory. To sign a message m, the Control Segment would compute a digital signature s based on m and k_{private} with a signing algorithm $\mathbb{S}$:

$$\mathbb{S}\left(k_{\text{private}}, m\right) = s. \tag{3.1}$$

The Control Segment would then transmit the signed message $\{m, s\}$ over d_k, thereby forming w_k. Public key cryptography assures that even with precise knowledge of k_{public} and of m, there is no computationally feasible method for a spoofer to predict s; or, once s is known, to infer k_{private}. Once the receiver obtains an unauthenticated signed message $\underline{\mathrm{B}}^n = \{m', s'\}$, it runs a code verification protocol $\mathbb{V}$ as in Fig. 2.2 to validate the message origin:

$$\mathbb{V}\left(k_{\text{public}}, m', s'\right) = \{\text{true}, \text{false}\}. \tag{3.2}$$

If $\mathbb{V}$ asserts, then $H_{1,C}$ remains low and the receiver can trust that the Control Segment generated $\{m', s'\}$ (i.e., $\{m', s'\} = \{m, s\}$).

3.4 Evaluating Digital Signature Protocols

By focusing on high-level design criteria, the discussion of cryptographic signal authentication thus far has settled on a NMA technique whereby a public key digital signature is embedded in the navigation message. This section evaluates four potential digital signature protocols that could generate the signed navigation message: a delayed-disclosure symmetric-key protocol called TESLA and three public key protocols called RSA, DSA, and ECDSA.

The most effective and practical protocol for civil GPS signal authentication is sought.

3.4.1 TESLA

The Timed Efficient Stream Loss-Tolerant Authentication (TESLA) protocol, described in [100] and adapted for radionavigation authentication in [95] and [101], is similar to the S/KEY protocol from [97], in that it uses a one-way chain of symmetric keys k_n. A chain of intermediate keys is generated by applying a secure hash function H iteratively N times to a seed key k_0 to yield $N-1$ intermediate keys such that for $m \leq n$, $H^{n-m}(k_m) = k_n$, along with a base key k_N that can be used to authenticate any intermediate key [e.g., $H^2(k) = H(H(k))$]. Intermediate keys are broadcast in reverse order $\{k_N, k_{N-1}, k_{N-2}, \ldots\}$. Verification can be achieved by comparison to any previously-released key: if k_{n+m} has already been validated and $H^m(k_n) = k_{n+m}$, then k_n must also be valid. Intermediate keys are broadcast as part of the navigation message, and because they are generated using a one-way function, they are unpredictable in advance but verifiable afterward.

To authenticate the navigation message, an unreleased intermediate key k_i is used to compute a message authentication code (MAC) for part of the navigation message. MAC_i corresponding to k_i is then broadcast over the data bits. According to the key-release schedule, k_i is broadcast after MAC_i is broadcast. When k_i is received, MAC_i can be validated. Since MACs are based on private-key algorithms that do not provide data non-repudiation (i.e.,

a valid MAC can be generated by any user with knowledge of the private key), only received MACs corresponding to keys not yet broadcast can be considered suitable for authentication. When used for both timing and navigation message origin authentication, keys and MACs need verification; each of these tasks is independent and could be computationally intensive.

Although TESLA is a novel approach, it does not meet all of the design criteria discussed in the previous section. Foremost, TESLA is not standardized. The protocol was designed for broadcast authentication and has been tested and studied only in that context, including a trial implementation on an eLORAN system. A concrete suggestion for implementation is given in [44, 101].

In addition, TESLA may not be effective in the sense defined above because of its low equivalent symmetric key strength b_s. Various proposals suggest that sufficient cryptographic strength can be achieved with keys that are 160 bits or shorter, which implies that the output of the secure hash function that generates the keys is also 160 bits (i.e., $b_s = 80$). But, hash functions used in signal authentication cannot have an output less than 224 bits as this is the minimum length necessary to achieve $b_s = 112$. Becker *et al.* suggest that the short cryptoperiod of individual keys and frequency of key updates dispels this concern [101]. However, if the hash protocol were broken off-the-air because $b_s < 112$, then the short cryptoperiod may no longer assure their security: all keys could potentially be computed before their release.

If TESLA were designed for a $b_s \geq 112$, then the computational burden to support TESLA would likely increase.

Another concern for TESLA is the computational burden of key management. The public key k_N, distributed to receivers over a PKI scheme and then stored to the receiver, can authenticate any intermediate key. One proposal suggests intermediate keys are generated once per second and the public key k_N would be valid for several years [101]. If this were the case and a receiver obtained a one-year-old k_N from the PKI, then it would need approximately 2^{25} computations of H in order to generate the current intermediate key. This would impose a large computational burden on the receiver relative to standard GNSS signal processing. Although k_N could be published more often, frequent key updates discourage adoption.

3.4.2 RSA

The Rivest, Shamir, and Adleman (RSA) algorithm has become a *de facto* standard for data security [76, 97]. It was one of the first public key algorithms and can be applied for pure encryption and signature generation. It is believed that the only way to defeat RSA is to factor a number with large prime factors. As factoring has become faster, the length of RSA keys needed to preserve security has increased. RSA requires a 2048-bit modulus to achieve $b_s = 112$ and would therefore occupy a significant portion of the low-data-rate navigation message (i.e., the RSA digital signature is too long to be practically

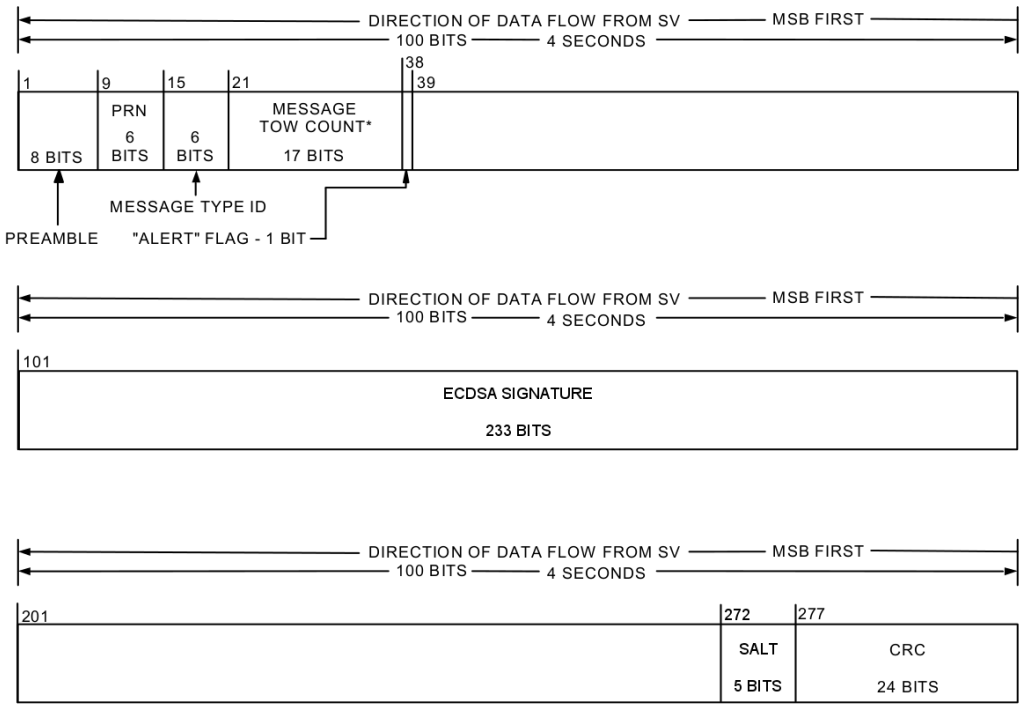


Figure 3.1: Diagram showing the format of the proposed CNAV ECDSA signature message, which delivers the first or second half of the 466-bit ECDSA signature and a 5-bit salt in the 238-bit payload field (figure adapted from [1]).

broadcast over the navigation message). Thus, RSA is impractical according to the earlier discussion of practical digital signatures.

3.4.3 DSA

The Digital Signature Algorithm (DSA) belongs to a class of algorithms that rely on the difficulty of finding logarithms in finite groups [76, 97]. It was developed by the U.S. National Security Agency (NSA) for NIST and adopted

for use in U.S. government applications in 1993. Its widespread use indicates that it is cryptographically valid and strong, as it has been implemented in a variety of critical applications.

DSA has two domain parameters that determine the strength of the algorithm. In order to achieve $b_s = 112$, it is necessary to use a 2048-bit prime p and a 224-bit prime q. Verification of digital signatures relies on p, making DSA comparable in computational complexity to RSA. Yet, DSA signatures are only twice as long as q (i.e., 448-bit signature for $b_s = 112$). Despite having signature length shorter than RSA, DSA is still not practical enough for cryptographic signal authentication because of its computational complexity.

3.4.4 ECDSA

Based on DSA, the Elliptic Curve Digital Signature Algorithm (ECDSA) operates on groups associated with an elliptic curve space [76]. For a given b_s, ECDSA signatures are the same length as DSA signatures. But by operating on a more complicated underlying elliptic curve space, ECDSA has smaller domain parameters and more efficient verification algorithms [102–104]. Furthermore, NSA recommends that systems built after 2010 implement ECDSA, which has been standardized by NIST [78].

3.4.5 Selecting the Appropriate Signature

With short signatures, efficient verification, and standardization, ECDSA appears to be both an effective and a practical digital signature protocol for NMA-based civil GPS signal authentication. Given the discussion above, ECDSA appears to be the best among current options although other signature schemes could be used if weaknesses in ECDSA were found. NIST offers several choices of standardized ECDSA domain parameters for a key strength $b_s \geq 112$ [78]. Among these, the standardized ECDSA 233-bit Koblitz curve (K-233) is attractive because it generates a short 466-bit signature amenable to optimized software-defined verification routines [105].

To sign messages, ECDSA first applies a secure hash function to generate a digital fingerprint of the message, which is typically shorter than the message itself, and then signs the fingerprint rather than the whole message. For proper implementation the following two conditions must be met: (1) the length of the signed navigation message must be at least as long as the output of the hash function (i.e., $2b_s$), and (2) each signed navigation message must vary in at least a single bit from previous messages to generate an unpredictable signature. These conditions are easily satisfied. The randomness introduced by the hash function along with the additional randomness introduced by the so-called salt, described in the next section, causes the signature to remain unpredictable even with knowledge of previous signed navigation messages.

In selecting the appropriate hash function for GPS signal authentication, NIST offers a standardized cryptographic hash family named SHA-2 [106]. Setting $b_s \geq 112$ implies implementing SHA-2 with at least a 224-bit key (i.e., SHA-224). Since there is no computational difference between SHA-224 and the stronger SHA-256, SHA-256 is proposed for implementation. Although the SHA-256 fingerprint is longer than the SHA-224 fingerprint, the digital signature length remains the length of the ECDSA signature, which is 466 bits long.

3.5 A Cryptographic Civil GPS Signal Authentication Proposal

This section proposes a concrete strategy for cryptographic civil GPS signal authentication. Consistent with the conclusions of the previous two sections, the strategy is based on public key elliptic curve cryptographic signatures inserted periodically into the flexible GPS civil navigation (CNAV) message. Specific details of the strategy, offered here, facilitate near-term adoption by the GPS Control Segment. The proposed strategy enables civil GPS signal authentication as described in the second section and diagrammed in Fig. 2.2 with the following properties: (1) a probability of detection of $P_D > 0.97$ for $P_{F,S} = 0.0001$, (2) a cryptographic strength of $b_s = 112$ bits, and (3) authentication every five minutes per channel.

3.5.1 Digital Signature Conveyance via CNAV

The flexible CNAV message format that modulates modernized GPS signals offers a convenient conveyance for a digital signature. The CNAV format was designed to be extensible so that new messages can be defined within the framework of the GPS IS. The CNAV message format is broadcast from Block IIR-M GPS spacecraft at the L2 frequency and Block IIF GPS spacecraft at the L2 and L5 frequencies [1]. Plans call for CNAV to be broadcast from Block IIIA GPS spacecraft at the L2 and L5 frequencies and additionally at L1. Thus, future single-frequency receivers can benefit from the extension to the CNAV message proposed in this section.

Every 12 seconds, a CNAV message delivers a 300-bit packet, which includes a 38-bit header, a 238-bit payload, and a 24-bit cyclic redundancy check (CRC). The flexibility of CNAV is due in part to the information broadcast over the header, which delivers a 6-bit message type identification field identifying up to 64 unique message types. The current GPS IS defines only 15 of these messages, reserving the others for future applications [1].

The following proposal defines two new CNAV messages to deliver an ECDSA signature. This is not a fundamental change to the GPS IS, but rather an extension to CNAV. Thus, this extension to CNAV can be considered practical in the sense defined earlier.

3.5.2 CNAV Message Signature Type Definition

Since the CNAV structure does not support payloads larger than 238 bits, the 466-bit ECDSA signature selected at the end of the last section must be broadcast across two CNAV messages. It is proposed to define two CNAV messages that deliver the 466-bit ECDSA signature, each message having the format shown in Fig. 3.1. The first ECDSA CNAV message type contains the first 233-bit half of the signature and the second message type contains the second half of the signature.

A 466-bit signature broadcast over two 238-bit payloads leaves 10 bits undefined. It is proposed to uniquely and randomly generate these bits for each instance of a signed message with a standardized pseudorandom number generator [107]. This technique is known as adding cryptographic "salt." Since the 10 salt bits are unpredictable prior to broadcast, they contribute to the total number of unpredictable w_k symbols available to a receiver to perform SCER detection tests. However, they do not increase b_s since they are not part of the digital signature. Like other components of the navigation message, they are digitally signed and can therefore be authenticated as originating from the Control Segment. Together, the two CNAV signature messages transmit 476 unpredictable bits.

3.5.3 Signing the CNAV Message

The frequency at which the CNAV navigation message can broadcast signatures requires consideration of several factors. First, although the CNAV

message format is flexible, it is not without constraints. Ephemeris message types 10 and 11 and a timing message of type 30–39 must be broadcast at least every 48 seconds to ensure accurate GPS receiver operation [1, 44]. Since a practical signal authentication strategy cannot adversely affect a receiver's position solution, the CNAV signature must respect these requirements. Given these constraints, the smallest block of data in which a complete signature can be embedded is the 96-second signature block such as the one shown in Fig. 3.2. In this structure, the two CNAV signature messages are interleaved between the ephemeris and clock data to meet the broadcast requirements.

A second consideration when signing the CNAV message is the duration between signature blocks. This choice involves a tradeoff between effectiveness (i.e., offering frequent authentication) and practicality (i.e., imposing a low computational burden relative to standard GPS signal processing and maintaining a low percentage of the CNAV message reserved for the digital signature). The maximum rate at which the CNAV message can be signed corresponds to a scenario in which the 96-second signature block in Fig. 3.2 is broadcast continuously back-to-back. However, this strategy is not practical: besides the high percentage of the navigation message reserved for the signature (i.e., 25 percent), this back-to-back configuration would eliminate the possibility of sending any other message types than 10, 11, 30–39, and the signature. Instead, a reasonable approach would be to sign every 336 seconds (about every five minutes). In this case, one signature block would authenticate every 28 CNAV messages as illustrated in Fig. 3.3. This means

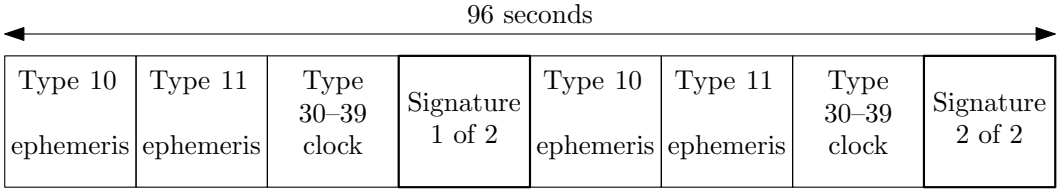


Figure 3.2: Schematic illustrating the shortest broadcast signature block that does not violate the CNAV ephemeris and timing broadcast requirements. To meet the required broadcast interval of 48 seconds for message types 10, 11, and one of 30–39, the ECDSA signature is broadcast over a 96-second signature block that is composed of eight CNAV messages.

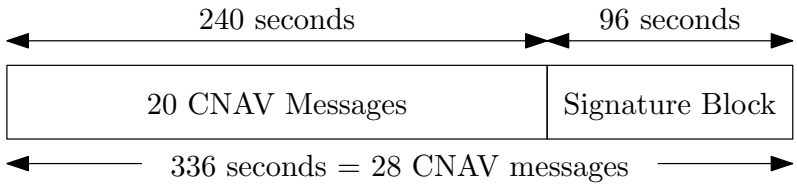


Figure 3.3: Schematic illustrating a signed 336 second broadcast. The proposed strategy signs every 28 CNAV messages with a signature broadcast over two CNAV messages on each broadcast channel.

the percentage of the navigation message devoted to the digital signature is a more practical 7.5 percent.

To broadcast a signature every five minutes, the Control Segment would first compute the next five minutes worth of CNAV navigation message including the salt. It would then concatenate signable navigation message bits in order—that is the first 23 CNAV messages (i.e., the 20 CNAV message in Fig. 3.3 and the first three in Fig. 3.2), the first signature header, the first five bits of the salt, the 5th through 7th CNAV messages from Fig. 3.2, the second signature header, and remaining five salt bits—and then generate the SHA-256 fingerprint. After generating the ECDSA signature from the fingerprint,

the Control Segment would break the signature into two parts and insert each part into a ECDSA signature message shown in Fig. 3.1. These two signature messages would then be transmitted at the appropriate times as part of the CNAV message signature block as seen in Fig. 3.2.

Note that the signature and corresponding CRC are not themselves signed. This is because neither is known until after signature generation. Unlike the signature field, which is entirely unpredictable, the CRC can be deterministically computed by a receiver immediately upon receiving the last unpredictable bit of any CNAV message. Thus, the CRC symbols cannot be used for SCER detection.

It is worth noting that a single uncorrected bit error would cause the verification algorithm to fail. CNAV has the option of being broadcast with forward error correction enabled. As described in the second section, FEC would enhance the robustness of NMA-based signal authentication. It is therefore recommended that FEC be enabled to support civil GPS signal authentication.

3.5.4 Constellation-Wide Signature Scheduling

Under the proposed strategy, each channel is authenticated every five minutes. However, the per-channel signature block could be offset from other channels (i.e., other satellites in the GPS constellation) such that a receiver tracking several satellites would see signatures more frequently. This offset strategy would substantially constrain the degrees-of-freedom that a spoofer could manipulate. An optimal offset strategy would minimize the maximum

time between authentications T_{ba} [i.e., $\min(\max(T_{ba}))$] that a receiver at any point on earth between a certain upper and lower latitude would observe based on the current constellation spatial arrangement. The optimal satellite offset assignment problem can be reduced to a directional graph coloring problem [108] that is likely best solved via a genetic algorithm similar to the one proposed for use in future optimization of the GPS constellation itself [109]. A sub-optimal solution computed through a greedy algorithm for the constellation in August 2011 computed that $\min(\max(T_{ba})) = 144$ seconds was possible between $\pm 70^{\circ}$ latitude. Thus, even with a simple sub-optimal signature offset assignment, a receiver could receive signatures with a T_{ba} of at most about two minutes and a T_{ba} on average of about one minute.

3.5.5 Authentication Performance

The proposed civil GPS signal authentication strategy broadcasts 476 unpredictable symbols approximately every five minutes. Given this, the P_D output in Fig. 2.2 can now be computed for a given threat model based on the statistical tests in [41]. To appreciate the effectiveness of the proposed authentication strategy, consider the following challenging scenario from the target receiver's perspective:

- the spoofer has a 3 dB carrier-to-noise ratio advantage over the receiver (i.e., $(C/N_0)_s = (C/N_0)_r + 3$ dB);
- the received spoofed signals are 1.1 times stronger than the received authentic signals;

- the spoofer has introduced a timing error of 1 μs in the receiver through jamming or other means and exploits this entire delay to improve its estimates of the security code chip values (i.e., the quantity e from the discussion of the SCER attack is equal to 1 μs); and,

- the false alarm probability for the SCER detector in Fig. 2.2 is $P_{F,S} = 0.0001$.

The statistics developed in [41] can be used to show that, under this scenario, the output P_D in Fig. 2.2 will be maintained above 0.97 over the range 34–51 dB-Hz of authentic signal carrier-to-noise ratio $(C/N_0)_r$ values as seen in Fig. 3.4. This indicates that the proposed NMA-based strategy enables effective anti-spoofing.

3.5.6 Implementation Details

The receiver modifications required to exploit the proposed civil GPS signal authentication strategy can be readily implemented on a software-defined receiver such as those presented in [110, 111] and [112]. A traditional receiver with application-specific correlation hardware would require some redesign to take advantage of the proposal. First, the correlation hardware would need to be modified to accommodate the new correlations needed for SCER detection [41]. Second, a traditional receiver would need to monitor J/N. This could be a natural extension of the GNSS spectrum monitoring that some GNSS receivers already offer [113, 114]. Third, the traditional receiver would need

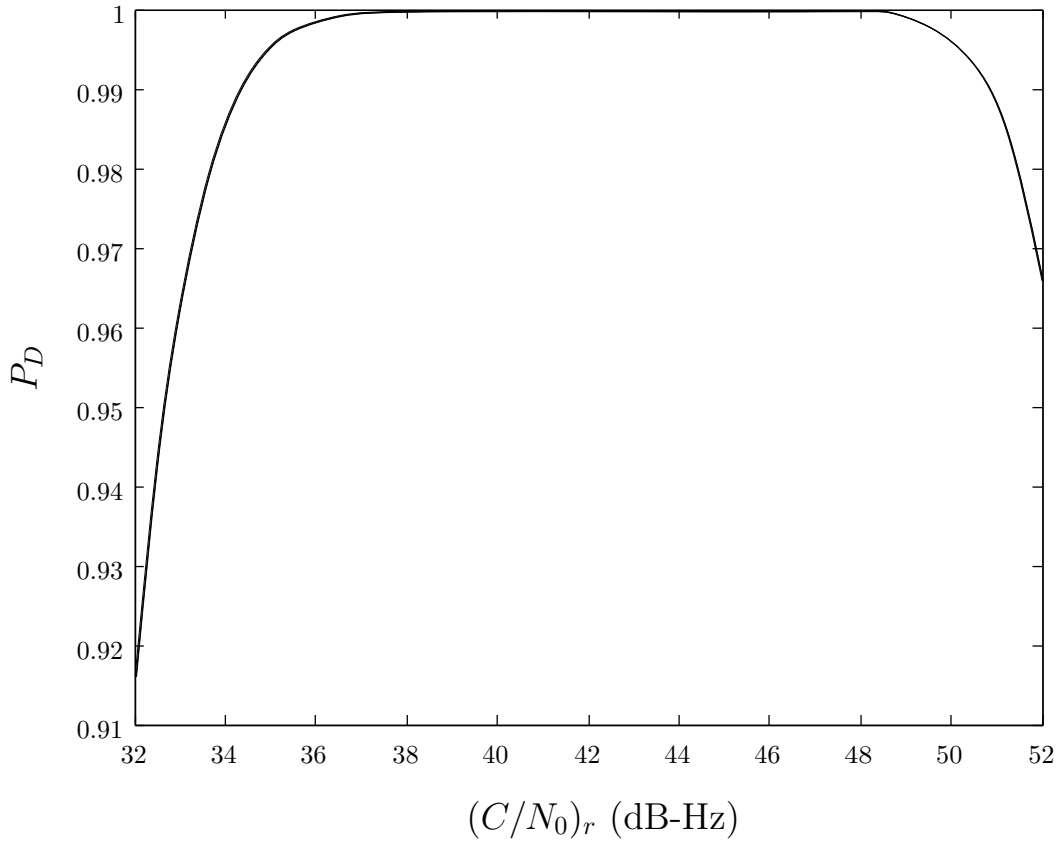


Figure 3.4: P_D as a function of $(C/N_0)_r$ for a challenging spoofing attack scenario. The proposed civil GPS signal authentication strategy maintains $P_D > 0.97$ for $P_{F,S} = 0.0001$ over 34–51 dB-Hz $(C/N_0)_r$ as shown.

to implement the remaining elements of Fig. 2.2 such as signature verification and the timing consistency check in its baseband processor, which is typically a general-purpose processor that is modifiable via firmware updates.

Although software receivers can be immediately modified to exploit the proposed authentication strategy and traditional receivers can be replaced as next-generation receivers are manufactured, there is a large number of receivers installed in critical applications that are not easily upgradeable. The GPS Assimilator introduced in [115] could be employed to protect such receivers by monitoring and sanitizing the incoming RF signals before they are ingested by the receiver.

The computational burden of verifying an ECDSA digital signature has been compared in a laboratory experiment to the computational burden of tracking GPS satellites. For this, P-256 ECDSA (i.e., a prime-curve-based ECDSA with a 256-bit key) was implemented in C++ with the GNU Multiple Precision Arithmetic Library (GMPlib). The code design was not optimized for implementation in a secure application. P-256 was implemented instead of K-233, the algorithm proposed earlier, because a reference design and test vectors were available to verify P-256. An actual ECDSA implementation of K-233 is likely even faster than P-256 because of optimizations that could be applied to Koblitz-based curve calculations [105, 116]; thus, if P-256 is shown to be computationally acceptable, then so will K-233 [117]. The computational expense of verifying a P-256 signature under this implementation was compared to the signal processing burden of the routine signal tracking in the

post-processing software-defined GPS receiver presented in [111]. Over a 336-second authentication segment on one channel, the CPU time spent on routine signal processing was approximately two seconds. By comparison, the CPU time spent verifying the ECDSA signature was approximately 10 milliseconds. Thus, the expected verification burden is roughly 0.5 percent of the overall signal processing burden per channel.

It should be noted that one drawback of ECDSA is the intellectual property landscape. A company called Certicom holds 130 elliptic-curve-related patents. Although NSA purchased a license to allow ECDSA use in national security applications, the license only covers prime-curve ECDSA signatures with key sizes of 256, 384, or 512 bits [118]. A civil GPS signal authentication strategy that implemented ECDSA signatures would likely be included under the purview of the NSA license. However, the smallest key size among NSA-licensed curves is 256 bits, which would generate a 512-bit signature requiring three CNAV messages for broadcast.

Finally, the cryptographic anti-spoofing techniques proposed here can be augmented with a software-defined non-cryptographic technique such as the vestigial signal defense [58] for additional protection during the initial stages of a code-phase-aligned spoofing attack when the SCER detector P_D can drop to around 0.5.

3.6 Conclusion

This chapter offers a practical technique to authenticate civil GPS signals. The proposed technique embeds digital signatures in the GPS civil navigation (CNAV) message and exploits a recently-developed statistical hypothesis test to secure civil GPS receivers against replay-type spoofing attacks. In a challenging example scenario, the technique was shown to detect a replay-type spoofing attack with probability of detection greater than 0.97 for a false alarm probability of 0.0001. The proposed strategy enables receivers to authenticate each individual civil GPS signal every five minutes.

Chapter 4

Non-Cryptographic GPS Spoofing Detection

4.1 Introduction

Despite the effectiveness of cryptographic ant-spoofing, no civil GPS signals yet incorporate cryptographic modulation due to financial and technical hurdles. Despite recent interest and engagement by U.S. and European satellite navigation agencies, a space-segment-dependent solution remains years away. GPS anti-spoofing techniques that can be implemented in the near term are those that operate independently of the space segment [119] or those that piggyback on encrypted military GPS signals. Recently proposed anti-spoofing techniques include networked receiver cross correlation of military GPS signals [38], multiple-antenna angle-of-arrival discrimination [80], multiple-antenna statistical monitoring [56], and antenna-motion-based carrier-phase detection [55]. The drawback of these approaches is their reliance on constant network connectivity, multiple antennas, or motion. Such requirements may prove impractical in applications with cost, size, weight, or power constraints.

The space-segment-independent technique proposed herein is also receiver-autonomous: it requires no network connection and no additional hardware.

Moreover, it can be readily implemented via a software or firmware update. My technique makes two reasonable assumptions: (1) an admixture of spoofing and authentic signals are incident on the victim receiver's antenna during an attack, and (2) the spoofer can neither null nor block the authentic signals. These assumptions are based on the difficulty of these tasks: the former is exquisitely challenging, and the latter is physically preventable.

This chapter is a significant extension to my work in [62]. A major contribution of this work is the development of an anti-spoofing technique that leverages kernel density estimation and other nonparametric techniques while maintaining low computational complexity. A rigorous discussion and evaluation of the combined monitoring of nominal in-band power and symmetric difference measurements demonstrate their effectiveness as an indicator of spoofing. This chapter also contributes a sophisticated and effective method to distinguish multipath and spoofing. A quantitative evaluation of the proposed technique against the Texas Spoofing Test Battery [42], the only set of publicly-available GPS spoofing field recordings, constitutes another significant contribution.

The key insight behind my proposed detection technique is that, when authentic and spoofed signals interact, a spoofer who wishes to conduct a successful spoofing attack faces a tradeoff between (a) maintaining a low-enough counterfeit signal power to avoid power monitoring alarms, and (b) minimizing distortions in the victim receiver's autocorrelation profile that are hallmarks of spoofing. My proposed technique exploits the unavoidable difficulties facing a

would-be spoofer by monitoring for anomalous autocorrelation-profile distortions and anomalous total in-band received power measurements. The combination of these measurements and application of nonparametric anomaly-detection-type methods pose a formidable defense. Nonparametric techniques make no *a priori* assumptions about the underlying data and instead compute statistics directly from current or training data. They have proven successful in network and facial recognition anomaly detection [120, 121].

Previous work has, in fact, proposed anti-spoofing techniques that monitor the total in-band power [69] and autocorrelation profile [64, 65] for anomalies. However, the combination of these two techniques is not evaluated despite its potency. Still, the author of [69] recognizes that an automatic gain control (AGC) approach to monitoring total in-band power is insufficient for two reasons. First, a receiver that ignores correlation distortions allows a spoofer to transmit a weak spoofed signal that does not trigger the user-selected upper AGC limit. Second, so-called personal privacy devices cause false alarms in an AGC-only approach. Similarly, the approaches in [64, 65] that only monitor the correlation profile for distortions can be readily fooled by a spoofer broadcasting with a significant power advantage over the authentic signals (c.f., Sec. 4.2). Finally, the author of [122] recognizes the potential of monitoring for anomalous quantities, including power levels, noise levels, correlation peak height, and Doppler/code rate, but the research does not identify the potency and immediacy of the two-pronged approach presented herein.

This chapter is organized as follows. Section II develops the measurement model based on the power–distortion tradeoff. Section III presents the nonparametric GPS spoofing detection algorithm. Section IV qualitatively evaluates the proposed defense against real-world recordings of spoofing, jamming, and multipath. Section V summarizes the chapter.

4.2 Measurement Model

4.2.1 Power–Distortion Tradeoff

During a spoofing attack, an admixture of the spoofed and authentic signals are incident on a victim receiver's antenna, and their interaction causes distortions in the correlation profiles [58]. A spoofer can eliminate the hallmark distortions of a spoofing attack by generating an antipodal, or nulling, signal or by preventing reception of the authentic signal (e.g., emplacing an obstruction). To generate a nulling signal, a spoofer requires both (a) centimeter-accurate knowledge of the relative three-dimensional position vector from the phase center of its antenna to the phase center of the victim receiver's antenna, and (b) 100-picosecond-accurate knowledge of its processing and transmission delay. Blocking reception of the authentic signals necessitates near-physical-proximity access to the victim receiver. Assuming that generating nulling signals is impractical, as laboratory and field experiments have indicated [6, 10], and that physical access near the receiving antenna is controlled to prevent signal blockage, an admixture of authentic and spoofed signals will be present during a spoofing attack.

Instead of trying to annihilate or block the authentic signals, the spoofer can attempt to overpower them. Before automatic gain control (AGC) and quantization, the received signal $r(t)$ at time t is

$$r(t) = a(t) + I(t) + n(t). \tag{4.1}$$

Here, $a(t)$ is the authentic signal, $I(t)$ is interference (e.g., spoofing), and $n(t)$ is thermal noise due primarily to low-noise amplifiers and subsequent in-line amplifiers. At this stage, the signal-to-noise ratio (SNR) $a(t)/n(t)$ is constant regardless of the instantaneous interference power $I^2(t)$. After the operation of the AGC, the gain-controlled signal $r_\alpha(t)$ becomes

$$r_\alpha(t) = \beta(t)[a(t) + I(t) + n(t)] \tag{4.2}$$

where $\beta(t)$ is the AGC scaling factor that varies in time as the AGC attempts to maintain time average of $\langle r_\alpha(t)\rangle$ at an approximately constant level. Note that the SNR of $r_\alpha(t)$ still remains independent to increases in $I^2(t)$. Finally, consider the quantized signal $r_Q(t)$ under a multi-bit quantization scheme:

$$r_Q(t) = Q[r_\alpha(t)] = \beta(t)[a(t) + I(t) + n_{\text{eff}}(t)]. \tag{4.3}$$

Here, the effective noise $n_{\text{eff}}(t)$ is a mixture of thermal noise and quantization noise. As $\langle I^2(t)\rangle/\langle a^2(t)\rangle$ gets large, $\langle n^2_{\text{eff}}(t)\rangle$ incorporates $\langle a^2(t)\rangle$. The result is that the actions of the AGC and quantization push the authentic signal down into the noise floor set by thermal and quantization noise.

In other words, in the limit as the spoofed signal-ensemble power P_s greatly exceeds the nominal authentic signal-ensemble power P_a, the high-power spoofed signals will push the despread authentic signals into the thermal

noise floor and thereby eliminate the hallmark distortions of a spoofing attack. However, if the target receiver raises an alarm when the received power in the radio frequency (RF) band containing the authentic signal exceeds some threshold $\eta_{\max}$, then the spoofer is strictly limited in the power advantage $\eta \triangleq 10 \log_{10}(P_s/P_a)$ that it can covertly apply. By upper bounding η, the spoofer is unable to fully eliminate distortion in the correlation function by increasing its power advantage.

The spoofer can also attempt to eliminate correlation function distortions by selecting a small η. However, as shown in [6], reliable capture of the target receiver's tracking loops requires $\eta \geq 0.4$ dB $= \eta_{\min}$. Thus for reliable spoofing η is lower-bounded by $\eta_{\min}$ and for covert spoofing η is upper-bounded by $\eta_{\max}$. Therefore, so long as $\eta_{\max}$ can be made sufficiently low while maintaining a tolerable rate of false alarm in the in-band power monitor, then a spoofing attack that respects this upper bound yet successfully captures the target receiver's tracking loops will be guaranteed to significantly distort the correlation profiles, and this distortion is detectable. This power–distortion tradeoff is the fundamental premise of this spoofing detection technique.

The following subsections present a model of the autocorrelation profile and explain how the symmetric difference and the total-in band power measurements are formed. This section concludes with the measurement model applied in the rest of the chapter and offers comments on the difficulties of nonparametric techniques.

4.2.2 Autocorrelation Model

Let $R^i(\tau)$ be the autocorrelation function that results from correlating the incoming filtered pseudorandom spreading code corresponding to satellite i with the unfiltered receiver-generated local code replica at offset τ. For each i at sample index k with uniform sample period T_s (i.e., $t_k = kT_s$), the receiver-computed autocorrelation function $\xi_k^i(\tau)$ can be modeled as the following extension of the model in [91]:

$$\xi_k^i(\tau) = a_k^i(\tau) + m_k^i(\tau) + s_k^i(\tau) + n_k^i(\tau). \tag{4.4}$$

The quantities superimposed in $\xi_k^i(\tau)$ are now enumerated. The quantity $a_k^i(\tau)$ represents the authentic signal:

$$a_k^i(\tau) = \alpha_{k,a}^i R^i(\tau - \tau_{k,a}^i) e^{j\theta_{k,a}^i}. \tag{4.5}$$

Here, $\alpha_{k,a}^i$ is a real-valued amplitude scaling factor, $\tau_{k,a}^i$ is an offset, and $\theta_{k,a}^i$ is a phase delay. The latter two quantities are both measured relative to the receiver-generated local code replica. The subscript a denotes the authentic signal (i.e., direct-path signal).

The quantity $m_k^i(\tau)$ represents multipath components of the authentic signal. Multipath can be modeled as a superposition of N_m amplitude-scaled, offset-shifted, phase-modified replicas of $R^i(\tau)$ [123]:

$$m_k^i(\tau) = \sum_{n=1}^{N_m} \alpha_{k,n}^i R^i(\tau - \tau_{k,n}^i) e^{j\theta_{k,n}^i}. \tag{4.6}$$

Since multipath signals are delayed replicas of the authentic signals, $\tau^i_{k,n} > \tau^i_{k,a}$ for all n, i. The model assumes that reflections from satellite $\ell \neq i$ contributes nothing to the multipath model. This is a reasonable assumption because of the good (i.e., low) cross correlation between the pseudorandom spreading codes of different satellites. Also, let $m^i_{k,n}(\tau)$ represent the nth multipath component.

The quantity $s^i_k(\tau)$ models correlation with a received spoofing signal [58]:

$$s^i_k(\tau) = \left[\alpha^i_{k,s} R^i(\tau - \tau^i_{k,s}) e^{j\theta^i_{k,s}}\right] \times \mathbb{1}_s. \tag{4.7}$$

The indicator function $\mathbb{1}_s$ indicates the presence (i.e., $\mathbb{1}_s = 1$) or absence (i.e. $\mathbb{1}_s = 0$) of a spoofing attack. The model of a spoofing signal is similar to the model of a single multipath reflection except that $\tau^i_{k,s}$ is unconstrained. Note that spoofed signal multipath is not modeled but may be present. It can be safely omitted from (4.7) because it increases the spoofer-induced distortions of $\xi^i_k(\tau)$ and thus would only make detection easier than application of (4.7) would predict. Multiple simultaneous spoofing attacks are also not modeled for the same reason.

The quantity $n^i_k(\tau)$ in (4.4) represents thermal noise from the RF front end that has been spread by the receivers early $E = R^i(\tau_p - \tau_c)$, prompt $P = R^i(\tau_p)$, and late $L = R^i(\tau_p + \tau_c)$ code replicas, where τ_p is the center tap value and τ_c is the tracking correlator offset. In this case, the inphase and quadrature components of $n^i_k(\tau)$ are independent and can be modeled as

zero-mean Gaussian with variance σ^2_{IQ}, where

$$E[\mathbb{R}\{n^i_k(\tau)\}\mathbb{I}\{n^i_k(\nu)\}] = 0 \quad \forall \tau, \nu. \tag{4.8}$$

When $2\tau_c \leq 1$ chip, the early and late noise samples are correlated [124]:

$$E[n^i_k(\tau)\{n^i_k(\nu)\}^\star] = 2\sigma^2_{IQ}(1 - |\tau - \nu|) \qquad |\tau - \nu| \leq 2\tau_c \tag{4.9}$$

$$E[n^i_k(\tau)\{n^i_k(\nu)\}^\star] = 0 \qquad 2\tau_c < |\tau - \nu| \tag{4.10}$$

The presence of of unintentional interference (e.g., solar flares [89]) or intentional interference (e.g., personal privacy devices or jammers [90]) can cause $\alpha^i_{k,a}$, $\alpha^i_{k,n}$, and $\alpha^i_{k,s}$ to vary significantly. Assuming a properly operating AGC, σ^2_{IQ} will remain fairly stable even during the presence of interference. Fig. 4.1 illustrates a potential $\xi^i_k(\tau)$ that is composed of authentic, multipath, and spoofing components.

4.2.3 Symmetric Difference Measurements

The symmetric difference measures distortions in $\xi^i_k(\tau)$ that are indicative of a spoofing attack. Although it is just one of a variety of signal quality monitoring (SQM) metrics that have been applied to detect anomalous signals [125, 126], it has substantial benefits for spoofing detection that will be explained shortly. Other SQM metrics include measures of ratios [64, 65], deltas [64, 66], early-late phases [67], and signs [68]. When applied to spoofing detection independently from other measurements, SQM metrics are generally unreliable because they have difficulty distinguishing between multipath

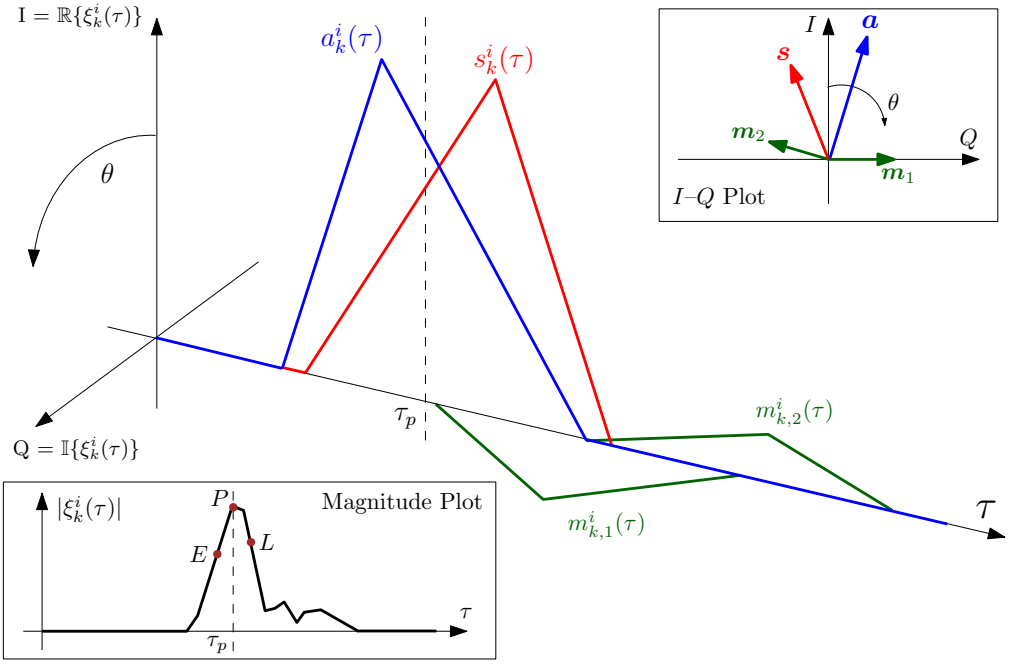


Figure 4.1: Illustration of a noise-free $\xi_k^i(\tau)$ composed of authentic $a_k^i(\tau)$, multipath $m_k^i(\tau)$, and spoofing $s_k^i(\tau)$ components. The center illustration shows each component of $\xi_k^i(\tau)$ in three dimensions. The upper right I–Q plot shows the maximum magnitude and angle of authentic $\boldsymbol{a}$, multipath $\boldsymbol{m}_n$, and spoofing $\boldsymbol{s}$ phasors. The lower left magnitude plot shows the resulting distortions in $|\xi_k^i(\tau)|$.

and spoofing [58]. However, the combination of the symmetric difference measurement with the total in-band power monitor, proposed herein, enhances multipath and spoofing discrimination.

For signal i at time $t_k = kT_s$, the complex-valued symmetric difference is

$$D_k^i(\tau_d) \triangleq |\xi_k^i(\tau_p - \tau_d) - \xi_k^i(\tau_p + \tau_d)|. \tag{4.11}$$

Here, τ_p is the prompt, or center, tap and τ_d is the symmetric difference tap offset, both in units of chips. $D_k^i(\tau_d)$ is measured in front-end units (FEUs). The function $|\cdot|$ is the absolute value. In an ideal noise-, multipath-, and spoofing-free scenario, $\xi_k^i(\tau_p + \tau_d)$ is even in τ_d and $D_k^i(\tau_d) = 0$ for all τ_d. In practice, $D_k^i(\tau_d)$ deviates from zero, with large deviations possibly indicating a spoofing attack.

$|D_k^i(\tau_d)|$ is a powerful test statistic for two main reasons. First, it is simple to implement. Second, it is sensitive to the distortions caused by a matched-structured spoofing signal that fails to maintain perfect code-phase alignment with the authentic signal that it is trying to replicate. In any successful spoofing attack, the spoofing signal must necessarily violate code-phase-alignment to commandeer, or "pull off," the tracking points from the authentic correlation peak. $D_k^i(\tau_d)$ measures the resulting distortions in the autocorrelation profile.

It is important to note that a weakness of $D_k^i(\tau_d)$ is its insensitivity at the onset of a spoofing attack when code-phase-alignment exists. At this stage

of the attack, however, the spoofer has yet to manipulate the victim receiver's navigation solution. Fig. 4.2 illustrates distortions in $\xi_k^i(\tau)$ under nominal and spoofed conditions.

A normalized symmetric difference metric, called the delta test, has also been proposed [66, 126]. $D_k^i(\tau_d)$ is un-normalized, because the noise statistics of $D_k^i(\tau_d)$ under thermal noise conditions are independent of the receiver's carrier-to-noise ratio if $D_k^i(\tau_d)$ remains un-normalized. In addition, $D_k^i(\tau_d)$ is independent of any nonlinear distortions in $R^i(\tau)$ that are due to a finite precorrelation bandwidth. $D_k^i(\tau_d)$ is also insensitive to differences in the slope of $R^i(\tau)$ caused by peak-flush and peak-adjacent sidelobes dependent on the pseudorandom spreading code properties of signal i [66]. Thus, $D_k^i(\tau_d)$ is insensitive to the specific function $R^i(\tau)$ or receiver front-end properties.

The maximum distortion of $D_k^i(\tau_d)$ is a function of τ_d. Consider the scenario with a single authentic and single spoofing signal, assuming that (a) $\alpha_{k,a} < \alpha_{k,s}$, (b) $\tau_{k,a} < \tau_{k,s} < \tau_c$, and (c) $\theta_{k,a} = \theta_{k,s} = 0$. In this case,

$$\tau_{\max} = \arg\max_{\tau_d} D_k(\tau_d) = \frac{\tau_{k,s}}{\alpha_{k,s}^2 + 1} \tag{4.12}$$

Here, $\tau_{\max}$ is parametrized by $\tau_{k,s}$ and $\alpha_{k,s}$. As the spoofer increases its $\eta = \alpha_{k,s}^2$, $\tau_{\max}$ moves closer to the peak. To appreciate the variability in $\max D_k^i(\tau_d)$ and $\tau_{\max}$, consider Fig. 4.3, which shows simulated spoofing attacks over a range of $\{\alpha_{k,s}^i, \tau_{k,s}^i, \theta_{k,s}^i\}$, assuming (a) and (b) above. The top plot shows that the greatest distortions occur when the spoofer is 180° out-of-phase with

the authentic signals, and that $\tau_{\max}$ varies between 0.20 and 0.01 chips for a $\tau_c = 0.25$ chips.

4.2.4 In-Band Power Measurements

The total in-band power measured by a GPS receiver is an essential component of interference monitoring [127]. The total in-band power at time $t_k = kT_s$ is given by P_k in Watts. A high P_k relative to nominal measurements indicates when additional power is present, possibly due to the presence of a spoofer. Recall that a spoofer must transmit counterfeit signals with enough power to commandeer the tracking loops of the victim receiver. Counterfeit signals will increase P_k provided the authentic signal remains. In controlled laboratory experiments where the spoofed signals were transmitted to the victim receivers via coaxial cable, $\eta \geq 0.4$ dB led to successful capture for every civil GPS receiver tested [6]. During a field test where the counterfeit signals were broadcast over-the-air to an unmanned aerial vehicle, successful capture necessitated $\eta \gg 0.4$ dB to overcome spoofed signal multipath and commandeer the craft with fine-grained control (c.f., [10], Sec. 3.2.1). These experiments demonstrate that the power monitor is an essential component of any spoofing defense [69, 119].

Note that a power advantage is only required if the spoofer seeks fine-grained control of the navigation solution of the victim receiver. If $\eta < 0.4$ dB, then the spoofer could still increase the error of the navigation solution or disrupt tracking of individual signals. In this sense, the spoofer can be thought

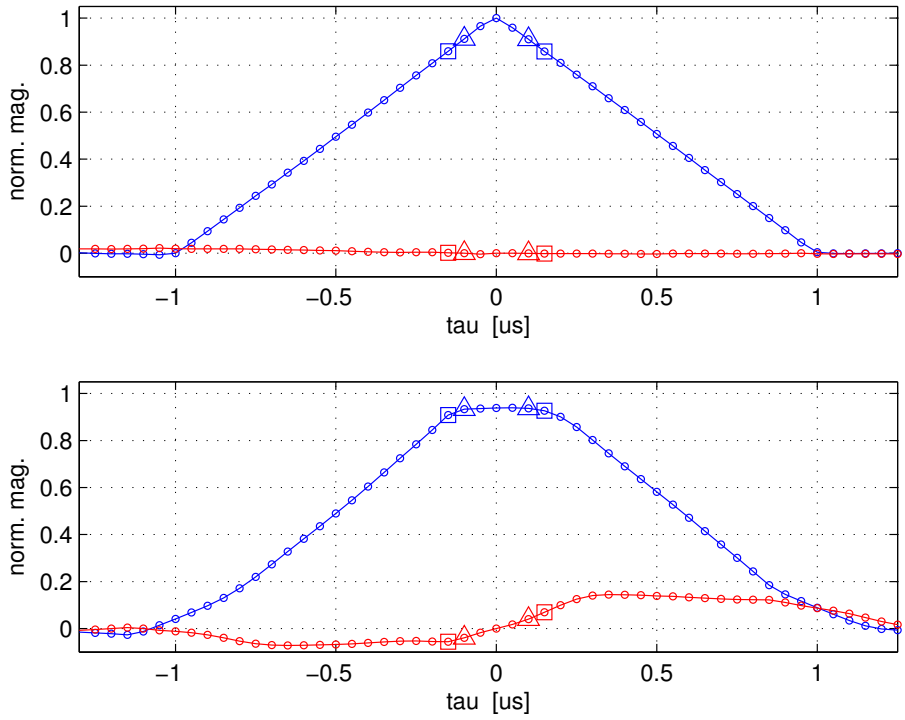


Figure 4.2: Plot showing the measured autocorrelation function $\xi_k^i(\tau)$ along with the early–late tracking taps marked by a square and $\pm\tau_d$ marked by a triangle. The in-phase components $\mathbb{R}\{\xi_k^i(\tau)\}$ are shown in blue, and the quadrature components $\mathbb{I}\{\xi_k^i(\tau)\}$ are shown in red. The top plot was generated from data recorded during nominal conditions, and the bottom plot was generated during a static matched-power time push spoofing attack.

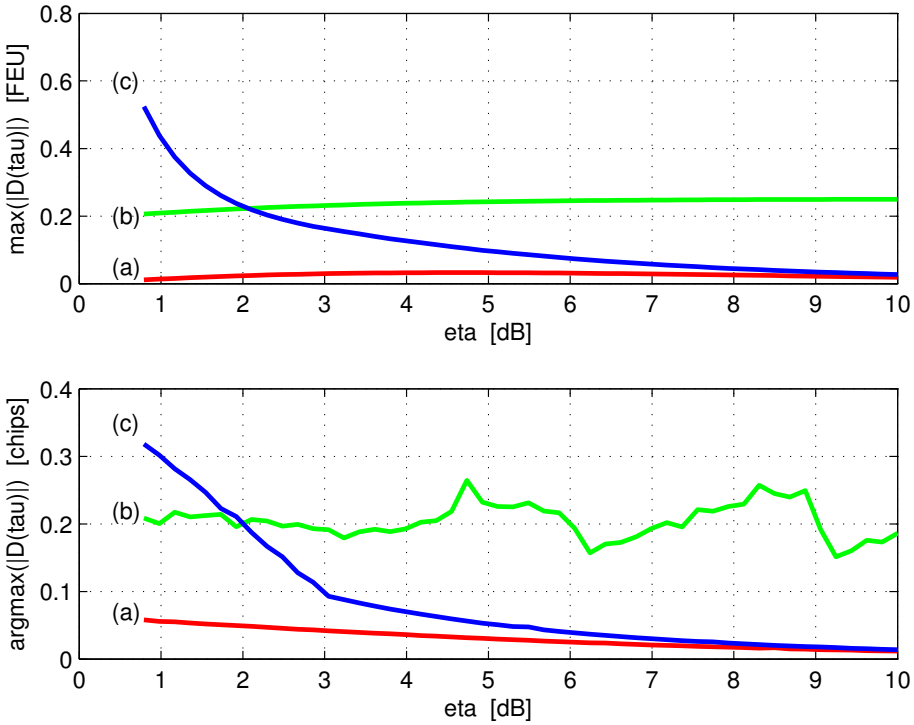


Figure 4.3: Plot showing $\max |D_k^i(\tau_d)|$ in front-end units [FEU] and $\tau_{\max} = \arg\max_{\tau_d} |D_k^i(\tau_d)|$ in chips versus $\eta = \alpha_{k,s}^2/\alpha_{k,a}^2 = \alpha_{k,s}^2$ for simulated steady-state tracking with an infinite bandwidth coherent delay-locked loop when the spoofed and authentic signals are (a) in phase, (b) 90° out-of-phase, and (c) 180° out-of-phase. The lines are averages of $\tau_{k,a} < \tau_{k,s} < \tau_c$, where the early–late offset τ_c was 0.25 chips.

of as acting like severe multipath. This chapter assumes that the spoofer's goal is complete capture.

Fig. 4.4 shows the power spectrum in the Global Positioning System (GPS) L1 C/A band with vertical lines indicating 2 and 10 MHz bandwidths during a nominal operation in the top plot and a spoofing attack in the bottom plot. Spoofers may inadvertently generate modulation distortions such as mixing, image, and jamming signals that manifest as additional power outside of the 2 MHz main lobe of the GPS L1 C/A signal. The lower plot of Fig. 4.4 indicates the presence of these artifacts.

Fig. 4.5 shows a time history of P_k measured with bandwidths of 2 MHz and 10 MHz during the same nominal and spoofed scenarios as Fig. 4.4. The normalized P_k measured with 2 MHz bandwidth shows a greater increase than measured with the 10 MHz bandwidth, because the spoofed power fills a greater proportion of the 2 MHz bandwidth than the 10 MHz bandwidth even with the presence of the mixing, image, and jamming signals. Because P_k is more sensitive to power increases inside of a narrower band, P_k measurements are made about the 2 MHz GPS L1 C/A band. This choice also favors implementation in standard civil GPS receivers with typical front-end bandwidth of 2 MHz.

4.2.5 Measurement Model Formation

In a probabilistic global navigation satellite system (GNSS) anti-spoofing framework [71], each measurement of $D_k^i(\tau_d)$ for every $i = 1, 2, \ldots, N_i$ is com-

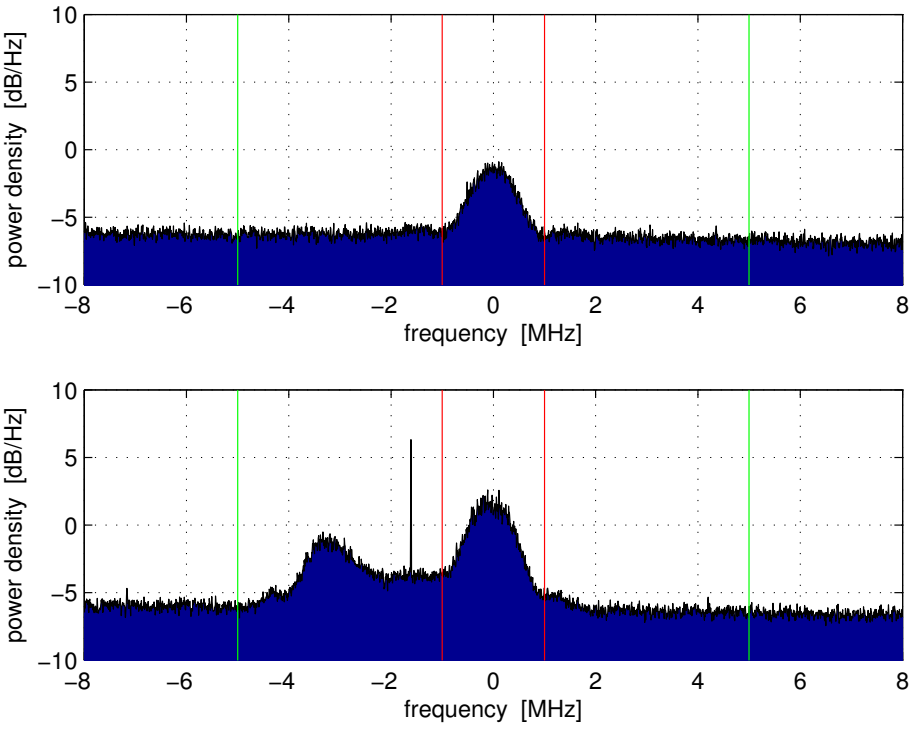


Figure 4.4: Plot showing the power spectral density in dB/Hz about the GPS L1 C/A center frequency of 1575.42 MHz for a static-receiver-platform during (top) nominal conditions and (bottom) a matched-power time push spoofing attack. The vertical lines represent the 2 MHz bandwidth (red) and 10 MHz bandwidth (green). In addition to power in the GPS L1 C/A main lobe, the spoofer introduces mixing and image distortions that manifest as additional power outside of the 2 MHz main lobe.

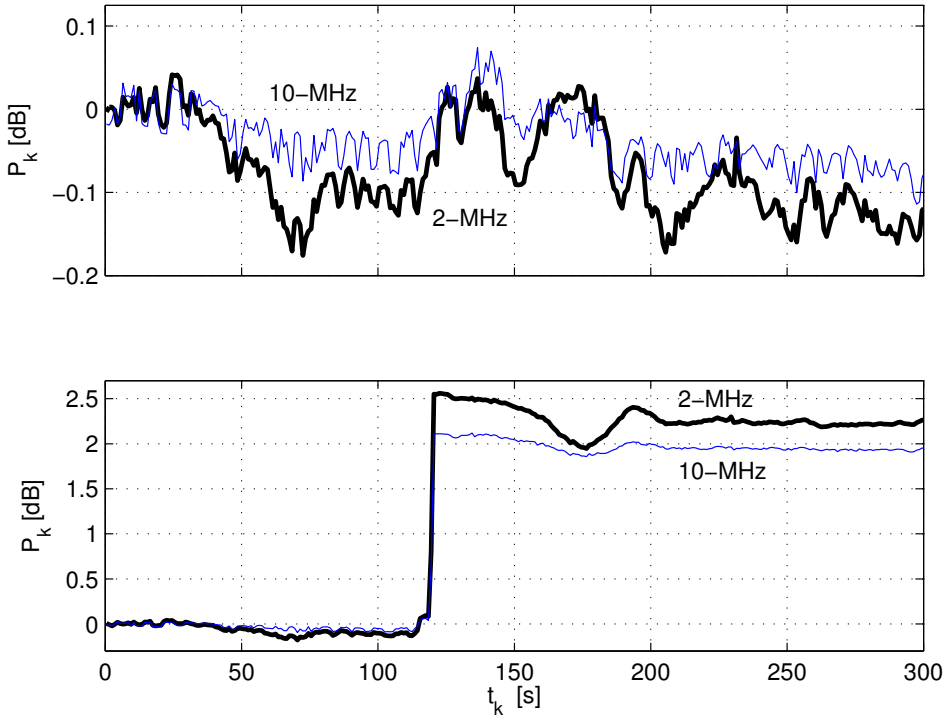


Figure 4.5: Plot showing the time history of the normalized in-band power measurements P_k for a static-receiver-platform during (top) nominal conditions and (bottom) a static matched-power time push spoofing attack scenario. The black, bold line represents the 2 MHz bandwidth and the slim, blue line represents the 10 MHz bandwidth.

bined with P_k in a single measurement vector $\boldsymbol{z}_k$:

$$\boldsymbol{z}_k = [D_k^1(\tau_d), D_k^2(\tau_d), \ldots, D_k^{N_i}(\tau_d), P_k]^{\mathsf{T}}. \tag{4.13}$$

But for theoretical and computational simplicity, I will only analyze the proposed defense based on individual signal measurements:

$$\boldsymbol{z}_k^i = [D_k^i(\tau_d), P_k]^{\mathsf{T}}. \tag{4.14}$$

Extensions of the per-channel test in (4.13) to the full test in (4.14) improves hypothesis detection performance (i.e., a lower probability of false alarm and a higher probability of detection) and builds straightfowardly on the principles of the per channel test.

In a general attack versus no-attack hypothesis test, the null hypothesis H_0 of no attack is distributed as $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ and the hypothesis of an attack H_1 is distributed as $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$. A receiver monitors $\boldsymbol{z}_k^i$ and decides at each decision time $k = K$ if a spoofing attack is or has been initiated at some time $\lambda \leq K$. If a spoofing attack initiates, $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$, the pre-change distribution, becomes $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$, the post-change distribution:

$$\begin{aligned} H_0 &: \boldsymbol{z}_k^i \sim p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0) \quad k = 1, \ldots, K \\ H_1 &: \boldsymbol{z}_k^i \sim \begin{cases} p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0) & k = 1, \ldots, \lambda - 1 \\ p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1) & k = \lambda, \ldots, K \end{cases} \end{aligned} \tag{4.15}$$

Detection techniques seek to minimize the time-to-alarm $\mathbb{E}[K - \lambda | K \geq \lambda]$, the probability of false alarm $P_F \equiv P(H_1|H_0)$, and the probability of detection P_D [128].

The success of an anti-spoofing technique hinges on its ability to characterize and differentiate between $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ and $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$. A number of complications make this problem particularly challenging [129]. A primary complication is the limited amount of available training data: reasonable bounds on $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ can be derived from training data, but the infinite variety of attack vectors make characterizing $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ particularly challenging. A secondary complication is the similarity of multipath and spoofing [c.f., Sec. 4.2.2, (4.7)], which has been demonstrated to limit the effectiveness of distortion-metric-based anti-spoofing [58]. A tertiary complication is that nominal conditions vary with the varying radio-frequency and physical environments.

As a final consideration, note that spoofing can be thought of as "intentional interference." In this sense, jamming and spoofing are the same. However, their varied statistics mean that $\boldsymbol{z}_k^i$ is sensitive to their variations. The remainder of the chapter considers H_1 to be either spoofing or jamming, but differentiates the two with statistical methods applied to $\boldsymbol{z}_k^i$ described in the next section.

Given these complications, spoofing defenses cannot offer foolproof security. Instead, the limited goals of anti-spoofing are to (a) constrain the spoofer to mimic multipath, thereby reducing the attack's effects, and to (b) decrease the appeal of spoofing by increasing the cost to conduct a successful attack.

4.3 Nonparametric GPS Spoofing Detection

Note: for the purposes of the remainder of this chapter, assume $z_k^i = [\mathbb{R}\{D_k^i(\tau_d)\}, \mathbb{I}\{D_k^i(\tau_d)\}, P_k]^\intercal$. *Also assume,* H_0 *corresponds to thermal and/or multipath, while* H_1 *corresponds to spoofing and/or jamming.*

This section introduces the nonparametric techniques that together form the proposed GPS spoofing defense. Nonparametric statistical techniques make no *a priori* assumptions about the underlying data; rather, they form statistical or probabilistic estimates directly from current or historical data. Nonparametric techniques excel in contexts where the data is poorly modeled by closed-form densities [130].

To appreciate the complexity of $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ and $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$, consider Fig. 4.6a. It shows the contour surfaces $R_p = \{\boldsymbol{z} : \hat{p}_{\boldsymbol{z}_{1:K}^i}(\boldsymbol{z}; \boldsymbol{B}) \geq p\}$ of $\boldsymbol{z}_k^i$ during nominal, spoofed, and jamming conditions at three probability levels decreasing in probability density shown in colors green, red, and blue, respectively. The quantity $\hat{p}$ will be described shortly. The cluster of contours with mean power $\bar{P}_k^i$ about 0 dB represents nominal data, the large blue contour with $\bar{P}_k^i \approx 2$ dB represents spoofing data, and the cluster with $\bar{P}_k^i \approx 7$ dB represents jamming data. The marginals are plotted in Fig. 4.6b as a probability density for P_k and a scatter plot for $D_k^i(\tau_d)$.

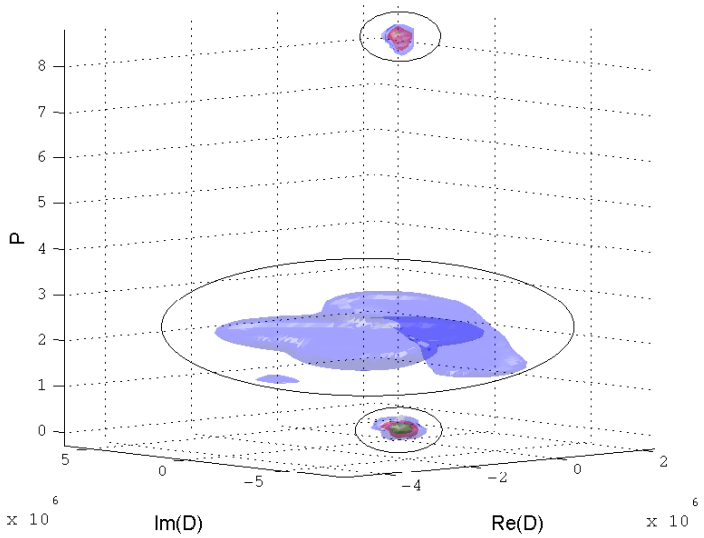
Clearly, $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ has no closed-form distribution; still, Fig. 4.6 motivates the nonparametric techniques of this proposed detection technique. Because the nominal data is confined to a volume about $\bar{P}_k^i = 0$ dB, it sug-

gests that H_0 is a reasonable assumption in a small volume about $\bar{P}_k^i = 0$ dB, provided that the training data supports this hypothesis. Theory also supports the assumption: the probability distribution $p(D_k^i(\tau_d))$ is distributed as a zero-mean, complex Gaussian with variance σ_D^2. A fixed volume, however, cannot adapt to changing nominal conditions that may well occur in practice. Therefore, real-time probability distribution estimates attempt to determine when to increase the volume in response to variations in $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$. Finally, in an attempt to distinguish spoofing and jamming, a windowed statistical estimate further differentiates $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ into a detection of spoofing and jamming. A side benefit of the windowed statistical estimate is the identification of multipath.

The remainder of this section describes the following nonparametric nonparametric techniques: (a) a volume subset to define an *a priori* nominal region, (b) a windowed kernel density estimator to adapt to changing conditions, and (c) a windowed statistics monitor that attempts to more finely identify multipath, spoofing, and jamming. The receiver runs each of these techniques simultaneously and independently for each satellite (i.e. channel) tracked. Algorithm 1 provides an overview of the technique.

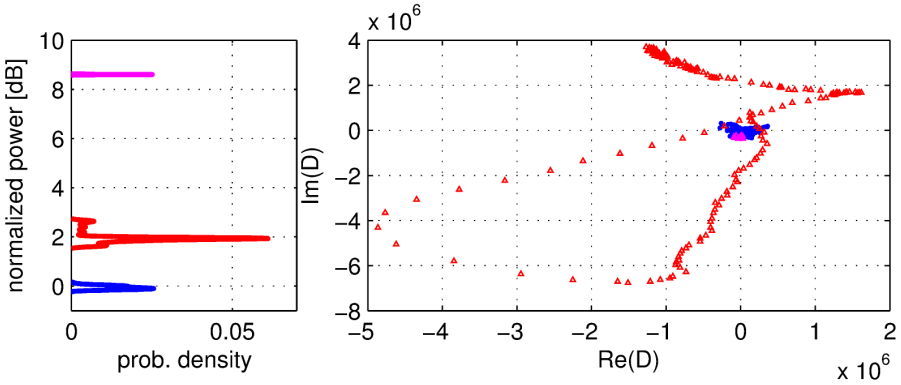
4.3.1 Volume Subset

A volume subset is used to define bounds in which $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ is always declared [130, 131]. The appeal of a volume-subset-type technique is that it establishes an *a priori* acceptable region defined by training data where



(a) Plot of contour surfaces $R_p = \{\boldsymbol{z} : \hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) \geq p\}$ during nominal conditions (lowest group), a static matched-power time push spoofing attack (middle group), and a jamming attack (highest group). The highest to lowest probability density is represented by green, red, and blue, respectively.

.



(b) Plot of $\boldsymbol{z}^i_k$ statistics during nominal conditions (blue), a static matched-power time push spoofing attack (red), and a jamming attack (magenta). The left plot shows the probability density of P_k and the right plot shows $\mathbb{R}\{D^i_k(\tau_d)\}$ versus $\mathbb{I}\{D^i_k(\tau_d)\}$.

Figure 4.6: Visual comparison of $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ and $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ during nominal conditions, a static matched-power time push spoofing attack, and a jamming attack.

$P_F = 0$. A drawback is that $P_D = 0$ within the volume and, if a spoofer can operate within this volume, it does so undetected. In this chapter, the volume subset is defined as the region where a squared distance function $D^2(\boldsymbol{z}, \boldsymbol{z}_k^i)$ remains below an upper bound $\bar{\boldsymbol{z}}$ (i.e., $D^2(\boldsymbol{z}, \boldsymbol{z}_k^i) \leq \bar{\boldsymbol{z}}$).

A volume subset raises three important questions. First, what distance metric is suitable? The Mahalanobis distance is appropriate because it scales each dimension of $\boldsymbol{z}_k^i$ so that the scaled changes in any one dimension are comparable to scaled changes in any other dimension. The Mahalanobis distance between $\boldsymbol{z}$ and $\boldsymbol{z}_k^i$ is

$$D^2(\boldsymbol{z}, \boldsymbol{z}_k^i) = D^2(\boldsymbol{z}, \boldsymbol{z}_k^i; \boldsymbol{P}) = (\boldsymbol{z} - \boldsymbol{z}_k^i)^T \boldsymbol{P}^{-1} (\boldsymbol{z} - \boldsymbol{z}_k^i) \tag{4.16}$$

where $\boldsymbol{P}$ is the covariance matrix. Since $E[\boldsymbol{z}_k^i] = \boldsymbol{0}$, set $\boldsymbol{z} = \boldsymbol{0}$. Now, $D^2(\boldsymbol{0}, \boldsymbol{z}_k^i) \leq \bar{\boldsymbol{z}}$ defines an ellipsoid with volume $V = 4\pi(\det[\boldsymbol{P}])^{-1/2}\bar{\boldsymbol{z}}^3/3$.

Second, how are $\boldsymbol{P}$ and $\bar{\boldsymbol{z}}$ set? Here, the importance of sufficient training data is again prominent. The sequel sets $\bar{\boldsymbol{z}} = \max[D^2(\boldsymbol{0}, \boldsymbol{z}_k^i)]$ and $\boldsymbol{P}$ equal to a covariance matrix with diagonal entries of $\text{var}[\mathbb{R}\{D_k^i(\tau_d)\}], \text{var}[\mathbb{I}\{D_k^i(\tau_d)\}]$ and $\text{var}[P_k]$. Here, $\text{var}[x]$ indicates the variance estimate of the quantity x. Both $\bar{\boldsymbol{z}}$ and $\boldsymbol{P}$ employ all training data where $\boldsymbol{z}_k^i \sim p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$, $\forall k, i$.

A $\boldsymbol{z}_k^i$ distributed as $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ that was not observed in training may cause $D^2(\boldsymbol{0}, \boldsymbol{z}_k^i) \geq \bar{\boldsymbol{z}}$. The capability of the defense to update $\bar{\boldsymbol{z}}$ during operation ensures that nominal conditions that exceed $\bar{\boldsymbol{z}}$ avoid declaring false alarm. Updates to $\bar{\boldsymbol{z}}$ will be based on a probability estimate $\hat{p}$ that $\boldsymbol{z}_k^i \sim p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$; the estimate $\hat{p}$ is described next. The downside of updating

$\bar{\boldsymbol{z}}$ is that slowly moving spoofing attacks may avoid triggering an alarm. Still, this is a known risk that is proportional to V. To reduce the missed detection rate, $\bar{\boldsymbol{z}} = \min\max[D^2(\mathbf{0}, \boldsymbol{z}_k^i)]$ for all training data where $\boldsymbol{z}_k^i \sim p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ $\forall k, i$.

It is a security risk to set $P_D = 0$ within the volume subset. To address this concern, consider the following. The distributions $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ and $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ are wildly different: in only two of the 14 spoofing data sets evaluated did $D^2(\mathbf{0}, \boldsymbol{z}_k^i) \leq \bar{\boldsymbol{z}}$ for a spoofing attack (c.f. Sec. 4.4, Table 4.3), and even then, $\max[D^2(\mathbf{0}, \boldsymbol{z}_k^i)] \gg \bar{\boldsymbol{z}}$. In addition, real-time estimation of $\hat{p}$ is running simultaneously to detect spoofing. Even if the spoofer *could* operate entirely within the defined volume, the defense still meets the goals presented earlier: constraining the spoofer and decreasing the appeal of spoofing. Results in Sec. 4.4 demonstrate the detection technique's efficacy despite allowing $P_D = 0$ within V. For applications with low risk tolerance, $\bar{\boldsymbol{z}}$ could be fixed or could be forced to remain below some user-defined upper bound $\bar{\boldsymbol{z}}_{\max}$.

4.3.2 Windowed Kernel Density Estimation

Kernel density estimation (KDE) is a nonparametric technique that generates empirical probability density estimates that converge to the true probability in the mean square sense [132, 133]. The KDE process does not induce artificial artifacts like those that exist in histogram-based techniques due to histogram bin size and location [62].

ID#	abbr.	qualitative description	TEXBAT?	platform mobility	interference	duration [s]	number channels	bandwidth [MHz]
1	cs	clean static rooftop	✓	static	–	370	8	20
2	ds1	static switch	✓	static	spoofing	375	8	20
3	ds2	overpowered time push	✓	static	spoofing	375	8	20
4	ds3	matched-power time push	✓	static	spoofing	375	8	20
5	ds4	matched-power pos. push	✓	static	spoofing	375	8	20
6	cd	clean dynamic	✓	dynamic	–	374	4	20
7	ds5	overpower time push	✓	dynamic	spoofing	220	4	20
8	ds6	matched-power pos. push	✓	dynamic	spoofing	228	4	20
9	cf	near dormitory		static	multipath	1680	5	10
10	wm	near trees and clock tower		static	multipath	120	7	10
11	fb	deep urban downtown		static	multipath	572	7	10
12	jam55	low jamming power		static	jamming	210	8	10
13	jam45	average jamming power		static	jamming	210	8	10
14	jam35	significant jamming power		static	jamming	210	8	10

Table 4.1: Summary of data used to evaluate the proposed nonparametric GPS spoofing detection technique. The Texas Spoofing Test Battery (TEXBAT), which is the only publicly-available data set of spoofing recordings, is available online [42].

Let $\boldsymbol{Z}^i_{1:K} = \{\boldsymbol{z}^i_1, \boldsymbol{z}^i_2, \ldots, \boldsymbol{z}^i_K\}$. The kernel density estimate $\hat{p}_{\boldsymbol{Z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B})$ at point $\boldsymbol{z}$ is defined as [134]:

$$\hat{p}_{\boldsymbol{Z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) = \frac{1}{K(\det[\boldsymbol{B}])^{1/2}} \sum_{k=1}^{K} \mathbb{K}\left(\frac{\boldsymbol{z} - \boldsymbol{z}^i_k}{(\det[\boldsymbol{B}])^{1/2}}\right). \tag{4.17}$$

Here, the kernel function $\mathbb{K}(\boldsymbol{z})$ is a symmetric, d-variate probability density function. The kernel bandwidth $\boldsymbol{B}$ is a $d \times d$ symmetric, positive-definite matrix. In this chapter, the distribution of $\mathbb{K}(\boldsymbol{z})$ is a three-dimensional Gaussian kernel: $\mathbb{K}(\boldsymbol{z}) = (2\pi)^{-3/2} e^{-\frac{1}{2}\boldsymbol{z}^T\boldsymbol{z}}$. The corresponding kernel density estimate is

$$\hat{p}_{\boldsymbol{Z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) = \frac{1}{K\sqrt[3]{2\pi}(\det[\boldsymbol{B}])^{1/2}} \sum_{k=1}^{K} e^{-\frac{1}{2}D^2(\boldsymbol{z}, \boldsymbol{z}^i_k; \boldsymbol{B})} \tag{4.18}$$

where $D^2(\boldsymbol{z}, \boldsymbol{z}^i_k; \boldsymbol{B})$ is the Mahalanobis distance. For any point $\boldsymbol{z}$, $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B})$ is an average of K, three-dimensional Gaussian distributions with mean $\boldsymbol{z}^i_k$ and covariance $\boldsymbol{B}$.

The KDE presented in (4.18) is windowed to limit its computational complexity and allow the nominal distribution to vary. At every $k = K$, $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B})$ is compared to a fixed threshold γ_p:

$$\gamma_p \underset{H_0}{\overset{H_1}{\gtrless}} \hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) \tag{4.19}$$

Typically, $\boldsymbol{z} = \boldsymbol{z}^i_{K+1}$. If $\hat{p} \geq \gamma_p$, then at the next sample $k = K + 1$, the windowed KDE computes the probability of the next point: $\boldsymbol{Z}^i_{1:K} \leftarrow \boldsymbol{Z}^i_{2:K+1}$ and $\boldsymbol{z} \leftarrow \boldsymbol{z}^i_{K+2}$. If, however, $\hat{p} \leq \gamma_p$, then only $\boldsymbol{z} \leftarrow \boldsymbol{z}^i_{K+2}$ occurs; the window remains. If the window were to include a below-threshold $\boldsymbol{z} = \boldsymbol{z}^i_\lambda$, possibly the first in a sequence $\boldsymbol{z}^i_{\lambda+1}, \boldsymbol{z}^i_{\lambda+2}, \ldots$ all distributed as $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$, then future deviations would likely produce $\hat{p} \geq \gamma_p$.

Algorithm 1 Nonparametric GPS Spoofing Detection

Given: K, $\boldsymbol{B}$, γ_p, $\boldsymbol{P}$, $\bar{\boldsymbol{z}}$, γ_σ
for $k = 1, 2, \ldots$ **and** $i = 1, 2, \ldots$ **do**
  Compute: $\hat{p}_{\boldsymbol{Z}^i_{1:K}}(\boldsymbol{z}^i_k; \boldsymbol{B})$
  if $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) > \gamma_p$ **then**
    $\boldsymbol{Z}^i_{1:K} \leftarrow \boldsymbol{Z}^i_{2:K+1}$
  else
    $\boldsymbol{Z}^i_{1:K} \leftarrow \boldsymbol{Z}^i_{1:K}$
  Compute: $D^2(\boldsymbol{0}, \boldsymbol{z}, \boldsymbol{P})$
  if $D^2(\boldsymbol{0}, \boldsymbol{z}, \boldsymbol{P}) > \bar{\boldsymbol{z}}$ **and** $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) > \gamma_p$ **then**
    $\bar{\boldsymbol{z}} \leftarrow D^2(\boldsymbol{0}, \boldsymbol{z}, \boldsymbol{P})$
  Compute: $\sigma_{\boldsymbol{Z}^i_{1:K}}$
  if $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z}; \boldsymbol{B}) < \gamma_p$ **and** $D^2(\boldsymbol{0}, \boldsymbol{z}, \boldsymbol{P}) > \bar{\boldsymbol{z}}$ **then**
    if $\sigma_{\boldsymbol{Z}^i_{1:K}} > \gamma_\sigma$ **then**
      Declare H_1**: spoofing**
    else
      Declare H_1**: jamming**
  else if $\sigma_{\boldsymbol{Z}^i_{1:K}} > \gamma_\sigma$ **then**
    Declare H_0

The windowed KDE approach immediately raises several questions. First, what is an appropriate window length K? The choice of K reflects expected receiver platform dynamics. A long K increases the computational burden, but retains a longer memory of historical $\boldsymbol{z}^i_k$ variations. A shorter K is more suitable if the nominal conditions are known to be varying rapidly but could be taken advantage of by a slowly varying spoofing attack. A 60 s window was applied with success in Sec. 4.4.

Second, what is an appropriate γ_p? In a one-dimensional Gaussian probability density function, a three sigma bound sets a confidence level greater than 99.7% that data will fall with that upper and lower bound. This corre-

sponds to a probability of false alarm $P_F \leq 0.003$. To approximate the three sigma bound in the 3–D Gaussian kernel function, compute

$$\gamma_p = \frac{\hat{p}_{Z=\mathbf{0}}(\mathbf{0}, \boldsymbol{B})}{\mathcal{N}(0;0,1)/\mathcal{N}(3;0,1)} \tag{4.20}$$

where $\mathcal{N}(x;\mu,\sigma^2)$ is the one-dimensional Gaussian density with mean μ and variance σ^2 at point x.

Third, how is $\boldsymbol{B}$ set? In theory, it can take the form of any positive definite matrix [134]. In practice, a diagonal matrix suffices, that is $\text{diag}(h_1, h_2, h_3)$ with $h_i \geq 0 \ \forall i$. A diagonal matrix with diagonal entries proportional to the standard deviation of each dimension of $\boldsymbol{z}_k^i$ is selected [132].

Earlier, $\bar{\boldsymbol{z}}$ was defined. If $D^2(\mathbf{0}, \boldsymbol{z}_k^i) \geq \bar{\boldsymbol{z}}$ while $\hat{p} \geq \gamma_p$, then $\bar{\boldsymbol{z}} \leftarrow D^2(\mathbf{0}, \boldsymbol{z}_k^i)$ to adapt to naturally varying $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$.

4.3.3 Windowed Statistics

Windowed statistical techniques have been previously applied to GPS interference detection [70]. The final nonparametric technique is the monitoring of the windowed standard deviation $\sigma_{\boldsymbol{Z}_{1:K}^i}$ of the magnitude $|D_k^i(\tau_d)|$:

$$\sigma_{\boldsymbol{Z}_{1:K}^i} = \left\{ \text{var}\left[|D_{1:K}^i(\tau_d)| \right] \right\}^{1/2}. \tag{4.21}$$

Based the observations in Fig. 4.6, this technique applies a threshold γ_σ to differentiate multipath, spoofing, and jamming. The value of γ_σ is set empirically (c.f. Table 4.2, Column 3).

4.4 Evaluation

This section introduces the training data used to evaluate the defense, describes the hardware- and software-based data processing and defense implementation, and quantitatively evaluates the proposed nonparametric GPS spoofing detection technique.

4.4.1 Training Data Descriptions

A description of the training data that were used to evaluate the defense follows; Table 4.1 provides a summary.

4.4.1.1 TEXBAT

The Texas Spoofing Test Battery (TEXBAT) is a set of six high-fidelity digital recordings of spoofing attacks against the civil GPS L1 C/A signals [42]. Both stationary- and dynamic-receiver-platform scenarios are provided along with their corresponding un-spoofed recording. Each 16-bit quantized recording was centered at the GPS L1 C/A center frequency of 1575.42 MHz with a bandwidth of 20 MHz and at a complex sampling rate of 25 MSps.

Each TEXBAT spoofing scenario makes use of the most advanced civil GPS spoofer publicly disclosed [3]. The spoofer can generate code-phase-aligned counterfeit signals (n.b., carrier-phase-alignment is only possible in controlled laboratory conditions), align counterfeit navigation data bits with authentic bits, and control η.

4.4.1.2 Multipath-Dense Recordings

To augment TEXBAT, data was collected where the receiver was statically positioned in multipath environments. The recording sites were located on a university campus near large buildings (e.g., clocktower) and in a dense urban environments (e.g., downtown metroplex). Each 16-bit quantized recording was centered at the GPS L1 C/A center frequency with a bandwidth of 10 MHz and at a complex sampling rate of 12.5 MSps.

4.4.1.3 Jamming Recordings

Jamming noise was recorded from a "cigarette lighter jammer" with a sweep range of 1550.02–1606.72 MHz and sweep period of 26 μs (c.f., [90], Table 1, Row #1 and Fig. 8.). This device and its performance typifies low-cost jammers that can be readily purchased online and easily operated, albeit illegally, from within cars. The recorded jamming noise was combined with clean, static-receiver data from a rooftop antenna and re-recorded. Each 16-bit quantized recording was centered at the GPS L1 C/A center frequency with a bandwidth of 10 MHz and at a complex sampling rate of 12.5 MSps.

4.4.2 Data Processing and Defense Implementation

The data processing and defense implementation was completed by hybridizing the functionality of high-end radio-frequency recording devices, a software-defined receiver (SDR), and MATLAB computational routines.

To measure $D_k^i(\tau_d)$, the GPS SDR, known as the Generalized Radio Interfusion Device (GRID) [112], processed the recorded data. In addition to signal acquisition and tracking, the receiver logs summary files with measured $\xi_k^i(\tau)$ in FEUs at 71 τ-offsets with 0.05 chip spacing.

A MATLAB routine imported this log file and output the appropriate $D_k^i(\tau_d)$. To compute P_k, MATLAB's power spectral density estimate with Welch's overlapping segment averaging estimator was applied at a rate of 1 Hz with a Hann window and discrete Fourier transform of length 4096 samples. P_k was measured centered about the GPS L1 C/A center frequency.

The combination of these two measurements as a single $\boldsymbol{z}_k^i$ was completed in MATLAB, where subsequent development and analysis of the nonparametric detection technique continued. Table 4.2 provides a summary of $|D_k^i(\tau_d)|$ and P_k statistics, and Table 4.3 provides a summary of $D^2(\mathbf{0}, \boldsymbol{z}_k^i; \boldsymbol{P})$ statistics. All statistics are computed for interference events. The parameters that generated the results are: $K = 60$ s, $\boldsymbol{B} = \text{diag}[1.4{\times}10^5, 1.6{\times}10^5, 0.21]$, $\gamma_p = 8.29{\times}10^{-12}$, $\boldsymbol{P}^{-1} = \text{diag}[4.6{\times}10^{-11}, 3.7{\times}10^{-11}, 23]$, $\bar{\boldsymbol{z}} = 73$, and $\gamma_\sigma = 12 \times 10^4$.

4.4.3 Quantitative Evaluation and Comparison

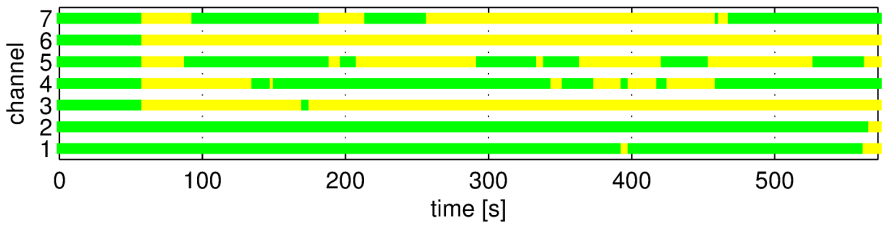
The defense was evaluated against all of the training data in Table 4.1. Overall, $E[K - \lambda | K \geq \lambda] = 1.6$ s with a worst-case delay of three seconds. Fig. 4.7a–c show the channel-by-channel decision between nominal (green), multipath (yellow), spoofing (red), and jamming (black) for data sets ID# 11,

	$\lvert D_k^i(\tau_d)\rvert$ [FEU$\times 10^{-4}$]				P_k [dB]			
ID#	μ	σ	min	max	μ	σ	min	max
1	18	10	0.3	70	-0.1	0.06	-0.2	0.1
2	104	40	36	170	-17	0.03	-17.2	-17
3	221	81	38	430	8	0.1	8	8.5
4	104	150	0.8	890	2.3	0.1	2	2.5
5	170	160	1.5	790	1.9	0.3	1.5	2.7
6	20	10	1	60	-0.2	0.1	-0.5	0.6
7	210	80	1.4	410	7.6	0.1	7	8
8	170	110	1.7	670	2.1	0.3	1.6	3.3
9	19	13	0.1	120	0	0.2	-0.3	0.8
10	15	9	0.3	50	0.1	0.2	-0.7	0.1
11	28	18	0.2	130	-0.4	0.1	-0.7	0.1
12	12	7	0.4	36	2.1	0.04	2	2.2
13	10	6	0.1	38	8.5	0.1	8.3	8.6
14	7	4	0.3	20	18	0.1	17	18

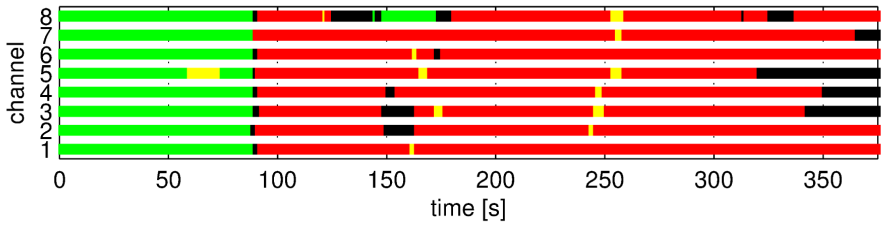
Table 4.2: Summary of statistics for $\lvert D_k^i(\tau_d)\rvert$ and P_k during $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ for spoofing and jamming files and for all data $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ files.

	$D^2(\mathbf{0}, \boldsymbol{z}_k^i; \boldsymbol{P})$			
ID#	μ	σ	min	max
1	4.4	3	0.1	28
2	41×10^4	1,280	41×10^4	42×10^4
3	5,200	150	5,120	5,900
4	490	44	360	610
5	53	16	30	115
6	9	6.2	0.04	50
7	4,220	130	3,490	4,650
8	55	16	30	124
9	3.6	5	0.01	73
10	4.7	5.6	0.01	55
11	9.8	6	0.2	38
12	2,450	100	2,240	2,650
13	7,480	170	7,140	7,750
14	25,700	320	25,000	26,200

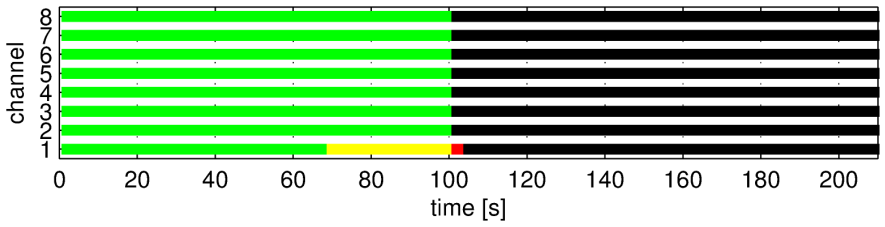
Table 4.3: Summary statistics for $D^2(\mathbf{0}, \boldsymbol{z}_k^i; \boldsymbol{P})$ during $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ for spoofing and jamming files and for all data $p_{\boldsymbol{z}|H_0}(\boldsymbol{\psi}|H_0)$ files.



(a) Results during multipath conditions (ID# 11).

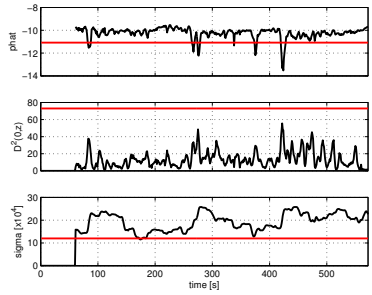


(b) Results during a spoofing attack (ID# 5).

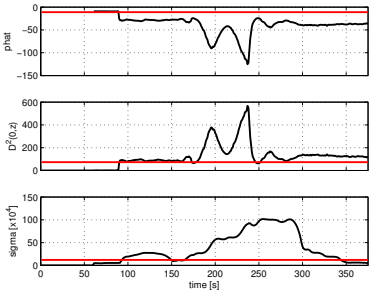


(c) Results during jamming (ID# 13).

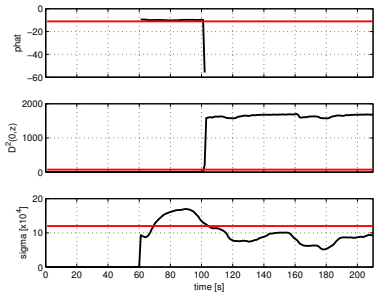
Figure 4.7: Plots showing the channel-by-channel decision between nominal (green), multipath (yellow), spoofing (red), and jamming (black). Three scenarios are shown (ID# 11, 5, and 13).



(a) Results during multipath conditions (ID# 11) on channel 3.



(b) Results during a spoofing attack (ID# 5) on channel 3.



(c) Results during jamming (ID# 13) on channel 1.

Figure 4.8: Plots showing $\log_{10}[\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z};\boldsymbol{B})]$, $D^2(\mathbf{0},\boldsymbol{z}^i_k)$, and $\sigma_{\boldsymbol{Z}^i_{1:K}} \times 10^{-4}$ (black) with their corresponding thresholds $\log_{10}[\gamma_{\hat{p}}]$, $\bar{\boldsymbol{z}}$, and $\gamma_\sigma \times 10^{-4}$ (red) versus time for three scenarios (ID# 11, 5, and 13).

5, and 13. Fig. 4.8a–c shows the time history of $\log_{10}[\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z};\boldsymbol{B})]$, $D^2(\mathbf{0},\boldsymbol{z}^i_k)$, and $\sigma_{\boldsymbol{Z}^i_{1:K}}$ and corresponding thresholds for a specific channel.

Fig. 4.8a corresponds to channel 3 of Fig. 4.7a. Here, multipath affects the recording throughout, but no spoofing is declared. Note that in Fig. 4.8a, $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z};\boldsymbol{B}) \leq \gamma_p$, but because $D^2(\mathbf{0},\boldsymbol{z}^i_k) \leq \bar{\boldsymbol{z}}$, H_0 remains declared.

Fig. 4.8b corresponds to channel 3 of Fig. 4.7b. Here, a spoofing attack initiates around 90 s. All channels declare H_1 within three seconds. Initially the detection method declares jamming, which is an artifact of the windowed statistics in $\sigma_{\boldsymbol{Z}^i_{1:K}}$. At attack onset, the window still contains samples of of the nominal data with a small standard deviation. However, the large deviation of $p_{\boldsymbol{z}|H_1}(\boldsymbol{\psi}|H_1)$ quickly raise $\sigma_{\boldsymbol{Z}^i_{1:K}}$, and spoofing is declared. Notice how in Fig. 4.8b the scale of $\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z};\boldsymbol{B})$ and $D^2(\mathbf{0},\boldsymbol{z}^i_k)$ varies when compared to clean and multipath data in Fig. 4.8a.

Fig. 4.8c corresponds to channel 1 of Fig. 4.7c. Here, a jamming attack initiates at 100 s. All channels initially declare jamming correctly with the exception of channel 1 where the initial classification is spoofing, likely as a result of multipath that affects the channel just before attack onset. Notice that $\log_{10}[\hat{p}_{\boldsymbol{z}^i_{1:K}}(\boldsymbol{z};\boldsymbol{B})] \to -\infty$ at onset.

The sensitivity analysis to $\bar{\boldsymbol{z}}$ is shown in Fig. 4.9. In the top plot, empirical worst-case P_D for spoofing and jamming and empirical worst-case P_F is plotted versus $\bar{\boldsymbol{z}}$. The lower plot shows a receiver-operating characteristic (ROC). A sensitivity analysis and ROC curve for γ_p is shown in Fig. 4.10. With

	$D_k^i(\tau_d)$	P_k	$[D_k^i(\tau_d), P_k]^\mathsf{T}$
worst-case P_F	0.0045	0.0027	0.00025
worst-case P_D spoof	0.173	0.994	0.969
worst-case P_D jam	0	~ 1	0.991
diff. jam–spoof	✗	✗	✓
diff. multipath–spoof	✗	✗	✓
multicorrelator taps	✓	✗	✓
reference	[58, 64, 66]	[69, 127]	[62]

Table 4.4: Comparison of the individual metrics $\boldsymbol{z} = D_k^i(\tau_d)$ and $\boldsymbol{z} = P_k$ against the combined measurement $\boldsymbol{z} = [D_k^i(\tau_d), P_k]^\mathsf{T}$.

the parameters listed in the previous section, an empirical worst-case P_D equal to 0.969 for spoofing and 0.991 for jamming results in a worst-case empirical $P_F = 0.00025$.

Table 4.4 quantitatively and qualitatively compares the proposed combined statistic $\boldsymbol{z} = [D_k^i(\tau_d), P_k]^\mathsf{T}$ against the two statistics individually, that is $\boldsymbol{z} = D_k^i(\tau_d)$ and $\boldsymbol{z} = P_k$. The combined statistic offer a lower probability of false alarm than either single metric. It further has the ability to differentiate multipath, spoofing, and jamming.

4.5 Conclusion

The nonparametric Global Positioning System (GPS) anti-spoofing technique proposed herein detects spoofing by monitoring real-time measurements of autocorrelation profile distortions and total in-band power. The defense was evaluated against the only publicly-available spoofing data set and detected

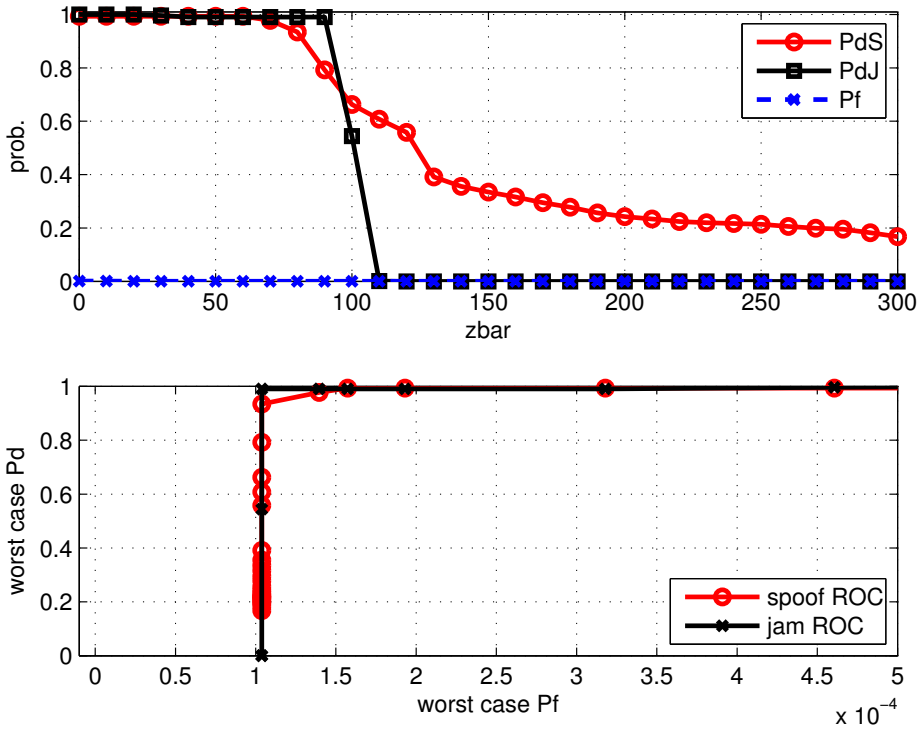


Figure 4.9: Sensitivity analysis to $\bar{\boldsymbol{z}}$ with $\gamma_p = 8.29 \times 10^{-12}$. Top: empirical worst-case P_D for spoofing and jamming along with empirical worst-case P_F versus $\bar{\boldsymbol{z}}$. Bottom: ROC curve varying $\bar{\boldsymbol{z}}$.

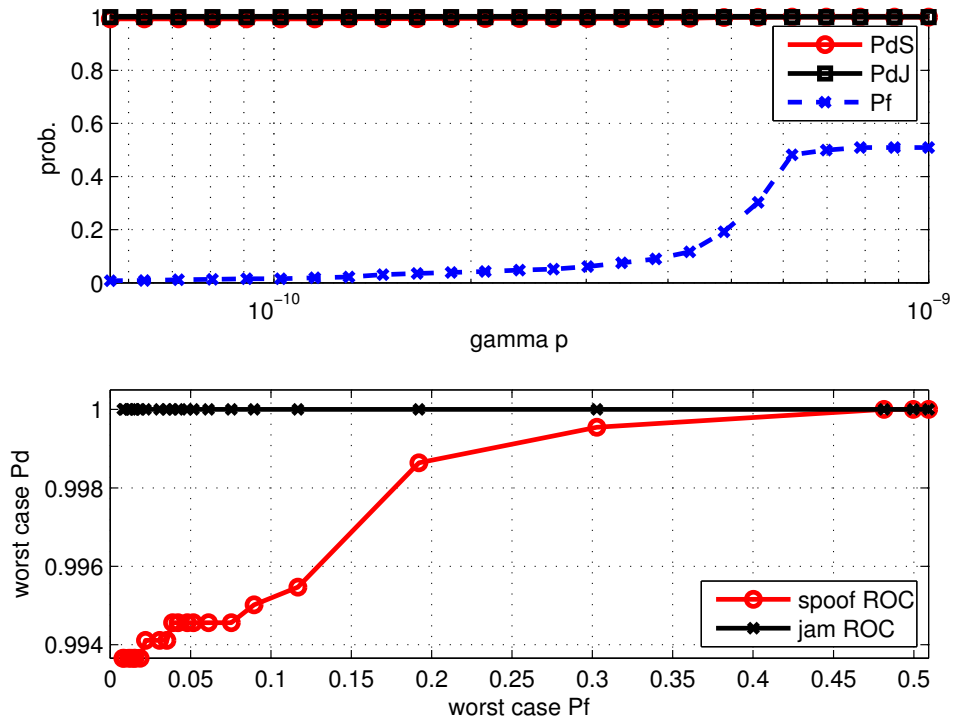


Figure 4.10: Sensitivity analysis to γ_p with $\bar{\boldsymbol{z}} = 0$. Top: empirical worst-case P_D for spoofing and jamming along with empirical worst-case P_F versus γ_p. Bottom: ROC curve varying γ_p.

spoofing attacks within three seconds of attack onset with a probability of detection $P_D \geq 0.969$ with corresponding false alarm probability $P_F = 0.00025$.

Chapter 5

Can Cryptography Secure Next Generation Air Traffic Surveillance?

5.1 Introduction

The year 2020 marks the dawn of aviation modernization. By that year, nearly all aircraft flying through U.S. airspace must carry Automatic Dependent Surveillance Broadcast (ADS-B) equipment, according to the Federal Aviation Administration's timeline to implement the Next Generation Air Transportation System (NextGen). ADS-B is central to NextGen, which shifts the burden of surveillance from antiquated ground-based radar to modern satellite-navigation-based aircraft transponders. Benefits of ADS-B include increased situational awareness, extended surveillance coverage, enhanced conflict detection, reduced operational costs, and improved routing efficiency [135].

Unfortunately, ADS-B as currently designed is riddled with security vulnerabilities [136–138]. ADS-B messages are broadcast in-the-clear according to an open protocol without cryptographic security mechanisms such as encryption or digital signatures that could protect and authenticate them. An open-access protocol has merits for international interoperability but renders ADS-B vulnerable to problems stemming from a lack of confidentiality, such

as aircraft targeting for electronic or kinetic attack, and malicious injection attacks, such as displaying ghost aircraft on cockpit displays.

Proposed cryptographic solutions attempt to mitigate these vulnerabilities [139–141]. These proposals merit evaluation in the context of the technologically-complex, cost-averse, and interoperability-focused aviation community. To this end, I address the question "can cryptography secure ADS-B within the constraints of the proposed NextGen system?" My holistic evaluation considers the historical design and policy constraints that shaped ADS-B, discusses the effectiveness of various candidate cryptographic solutions, and analyzes their implementation burden. I conclude with a quantitative assessment of the technological burden required to implement the most feasible cryptographic solution.

5.2 The Shift from Independent to Dependent Surveillance

Radar, developed in the 1940s, is the current state-of-the-art air traffic surveillance system. Primary surveillance radar (PSR) is considered an independent and non-cooperative surveillance system—independent because the radar on its own is sufficient to determine the necessary surveillance data (i.e., range and azimuth to target), and non-cooperative because the aircraft provides no assistance besides offering its cross-sectional area as a radar-reflective surface. Drawbacks of PSR include its need to perform sev-

eral radar sweeps of each target, measurement accuracy that degrades with increased target range, and susceptibility to so-called clutter interference.

Secondary surveillance radar (SSR) is independent and cooperative. Like PSR, SSR determines range and azimuth from radar sweeps, but SSR additionally interrogates aircraft equipped with Mode-S(elect) beacons at 1030 MHz. The cooperative responses from an aircraft's beacon transponder at 1090 MHz augment radar-derived surveillance with the aircraft's altitude and identity from Mode-C(ontract) and Mode-A(ddress) transmissions, respectively. Not all aircraft carry Mode-S transponders; for those that do not, non-cooperative radar and voice communication are the primary surveillance technologies.

The combined U.S. PSR–SSR network provides aircraft position accuracy of 1–2 nmi with updates every 5–10 s, which leads to a 3 nmi or greater separation requirement between aircraft in most U.S. airspace under FAA Order 7110.65. The current system has been sufficient to handle past and present air traffic densities, but it cannot support the high aircraft densities that are predicted. The combination of this fact with the high operating costs of PSR–SSR systems motivate the transition from radar to the modernized ADS-B system, which will provide an accurate, real-time view of air traffic purportedly at a lower cost than radar.

The acronym ADS-B conveys how the protocol operates. ADS-B transponders *automatically* broadcast without external interrogation or pilot input. The navigation data and its quality are *dependent* on the sensors installed on board the aircraft. The message contains *surveillance* data that is *broadcast*

so that anyone may receive it and no reply is sought. ADS-B offers position accuracy of 92.6 m (0.05 nmi), velocity accuracy of 10 m/s (19.4 nmi/h), and updates every second. These performance standards are designed to support (a) reduction in aircraft lateral separation from 90 nmi to 20 nmi and reduction in aircraft longitudinal separation from 80 nmi to 5 nmi in airspace that is outside of radar range, and (b) expansion of the 3 nmi aircraft separation requirements to airspace that currently sets a minimum 5 nmi separation [142].

When ADS-B was developed as an extension to the Mode-S beacon radar surveillance system in the 1980–90s, performance concerns focused on reliability, accuracy, range, operational capacity, and channel occupancy [143]. Note the omission of security—a topic that has received scant coverage in publicly-available reports from the FAA and other stakeholders. In response to concerns about ADS-B vulnerabilities, the FAA conducted a Security Certification and Accreditation Procedures (SCAP) study that, to date, remains protected from public disclosure because of its status as "Sensitive Security Information." Unable to discuss their test procedures or results, the FAA instead stated in 2009 that "using ADS-B data does not subject an aircraft to any increased risk compared to the risk that is experienced today" [142].

The FAA has committed to an annual review of its security study to evaluate new and evolving threats against ADS-B. One evolving threat targets the Global Positioning System (GPS) and other Global Navigation Satellite Systems (GNSS), from which ADS-B derives its surveillance data. GPS is vulnerable to denial-of-service and signal counterfeiting attacks known as

jamming and spoofing, respectively. GPS security has recently been the focus of vigorous research [15]. In 2012, the FAA tasked the "GNSS Intentional Interference and Spoofing Study Team" to evaluate the threat. Like the ADS-B SCAP study, their findings have yet to be released to the public.

At first glance, the FAA's claim of no increased risk seems implausible given the ease with which ADS-B can be spoofed and jammed in comparison to radar. Consider the difficulty facing an attacker who wishes to fool, or spoof, SSR. For one, the highly-directional SSR beam pattern makes it difficult for the attacker to inject a false target with an arbitrary bearing or altitude. The commonplace ASR-11 surveillance radar has a 5° elevation and 1.4° azimuth beamwidth. An attacker would either need to be within this narrow beam or would have to resort to injecting its signals through the antenna's side lobes, which would require high power or close proximity. For example, an attacker outside the main SSR radar antenna beam at a standoff distance of 1 km would need to transmit an 80 W signal, assuming a minimum 34 dBi sidelobe suppression, to match the received signal power of a 200 W Mode-S transponder at a range of 80 km. Furthermore, because radar is triggered, an attacker would need to detect when a radar pulse is sent and respond with an appropriately-timed response. Although these technical hurdles can be cleared, they increase the cost of an attack and limit its scale. Unsurprisingly, radar spoofing and jamming attacks "very rarely occur" [142].

By way of comparison, consider an attack against ADS-B. Omnidirectional ADS-B antennas afford attackers flexibility in orientation and proxim-

ity. The power from a 125 W ADS-B transceiver 80 km away is matched by a 20 mW transmitter 1 km away. Forged ADS-B message broadcasts can initiate anytime and can continue at 1 Hz, commensurate with the ADS-B transmission rate. Couple this relative physical flexibility with the lack of built-in security mechanisms, and it becomes clear just how vulnerable ADS-B is: a single, fraudulent, properly-formatted ADS-B transmission that passes parity is indistinguishable from an authentic message from the point-of-view of an ADS-B receiver.

Even so, the FAA's original claim regarding risk may not be inaccurate. In response to concerns about spoofing and jamming attacks against ADS-B or GPS, the FAA plans to retain near legacy levels of radar as a backup for ADS-B surveillance. The agency will continue to operate 100% of the 150 *en route* SSRs and will retain 40 legacy SSRs, or approximately 50%, in some high-density areas [142]. Class B airports—those with the highest air traffic density in the U.S.—will retain legacy-level coverage. By maintaining these radar systems, the FAA will not reap the cost-savings originally predicted from NextGen until after 2035, but air traffic control (ATC) will retain the ability to cross validate ADS-B broadcasts with radar, thereby providing near-legacy-level surveillance security.

There are good reasons, however, to demand *better* than legacy security. As with ADS-B, worrisome weaknesses also exist in the legacy air traffic surveillance system: Mode-S, A, and C have no cryptographic safeguards, and voice communication over radio between ATC and pilots is unencrypted.

Legacy surveillance systems also operate with aircraft separation requirements that NextGen will reduce in some airspace. If ADS-B is working as intended, the tighter spacing is likely no less safe than legacy spacing, but if an attack occurs, tighter spacing will increase the chance of a mishap. Under attack, legacy-level security cannot maintain legacy-level risk.

Besides, legacy-level security appears oddly out-of-date in a post-9/11 world. After the 9/11 attacks, the FAA oversaw the installation of reinforced cockpit doors, and air-bound passengers continue to endure enhanced screening procedures administered by the Transportation Safety Administration. Why then should NextGen be content with legacy-level security? The modern aviation risk landscape has also been altered by new technology. Whatever security concerns may have arisen during ADS-B development in the 1990s were likely assuaged by the high costs of acquiring ADS-B hardware and mounting a successful attack. Four decades later, a do-it-yourself ADS-B transponder that can produce counterfeit ADS-B messages can be made for just $1,000 [138]. Greater risk calls for greater security. Thus, even if the FAA's claim of no increased risk is accurate, there remain good reasons to pursue a cryptographic fix for ADS-B.

5.3 The Technical Ins and Outs of ADS-B

The following technical details will aid understanding of the security problems and the constraints of the ADS-B protocol. ADS-B Out messages are broadcast every second at a data rate of 1 Mbps over either 1090 MHz

Mode-S Extended Squitter (ES) or 978 MHz Universal Access Transceiver (UAT) [135]. This dual-link strategy is a compromise that the FAA made to satisfy international standards that require 1090 MHz Mode-S ES and those general aviation pilots who have already purchased UAT transceivers. Despite its name, UAT is a U.S.-only protocol for general aviation aircraft flying below Class A airspace, which begins at 18,000 ft, and outside of other controlled airspace, such as Class B airspace.

To support aircraft equipped with an ADS-B transponder that only operates at one frequency, the FAA will install ADS-R(ebroadcast) capabilities in ADS-B ground stations to rebroadcast Mode-S ES messages in UAT format and *vice versa* [142]. Each ADS-R system will have a range of 150–200 nmi, and the costs of installing and running the network will be borne by the FAA. To ensure ADS-R stations can receive ADS-B messages with sufficient power, the FAA has set the minimum transmission power of ADS-B at 125 W for 1090 MHz Mode-S ES broadcasts.

ADS-B Out messages are modulated with pulse position modulation (PPM), which is a type of pulse amplitude modulation (PAM). Differential phase shift keying (DPSK) was also considered. DPSK has a lower bit error rate than PAM for a given signal-to-noise ratio but had a higher hardware cost. Designers selected PPM to minimize costs and maintain interoperability—that is, the compatibility of ADS-B with existing protocols and equipped hardware.

ADS-B Out messages are 112-bits long. The first 8 bits indicate the data format, the next 24 bits indicate the aircraft's unique and fixed Interna-

tional Civil Aviation Organization (ICAO) address, the next 56 bits transmit the ADS-B surveillance data, and the final 24 bits are a cyclic redundancy check block. During flight, an aircraft's 112-bit ADS-B Out Data Format 17 messages contain the time and the aircraft's latitude, longitude, and altitude. Other 112-bit message formats are broadcast to communicate other operational events when the aircraft is on the tarmac.

The FAA only requires equipage of ADS-B Out by 2020; ADS-B In remains optional because of concerns regarding its implementation cost, equipment performance standards, and cockpit display requirements. Nonetheless, complete ADS-B In/Out systems will be popular because of the additional situational awareness, more efficient oceanic routing, and enhanced aircraft interval management that ADS-B In/Out offers over ADS-B Out alone. Figure 5.1 illustrates a basic operational ADS-B system.

No part of the ADS-B Out messages is encrypted or cryptographically signed. The lack of cryptographic safeguards is likely explained by the original designers' focus on interoperability, a principle that is evident throughout the design of ADS-B. Its frequencies, 1030 MHz interrogations and 1090 MHz responses, allow Mode-S and ATC to communicate over the same channel; its modulation scheme, PPM, was supported by existing, low-cost hardware in the 1990s; and its short message length, 112 bits, was an attempt to minimize communication interference with existing protocols. Interoperability facilitates adoption and keeps cost low, whereas cryptographic techniques limit

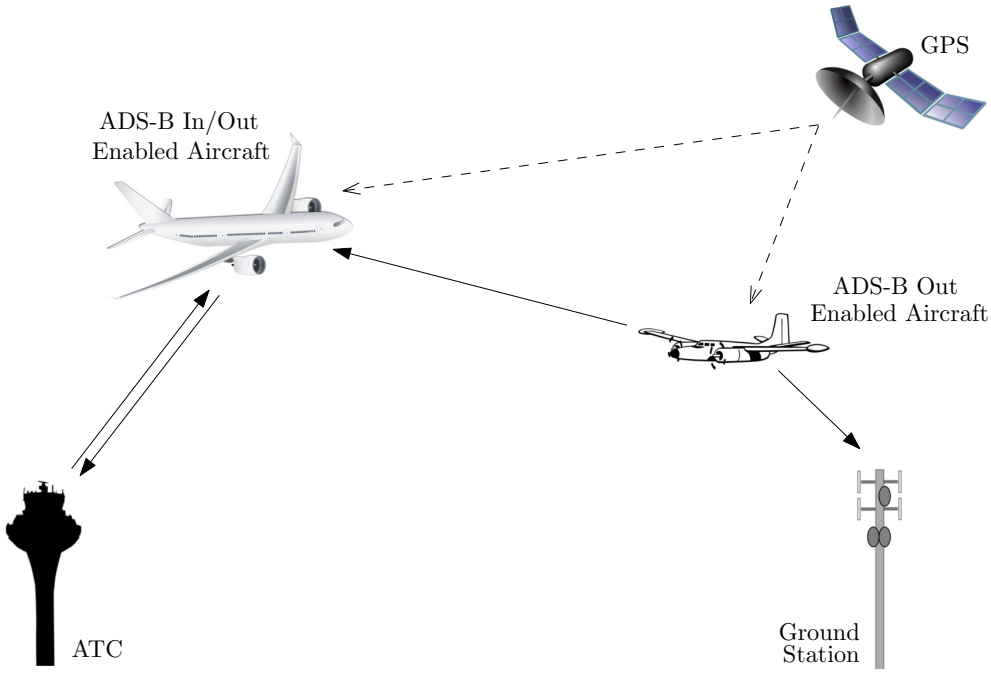


Figure 5.1: An overview of the ADS-B system, adapted from [135]. Aircraft are only mandated to broadcast ADS-B Out messages; receipt of ADS-B In messages is optional. Radar and other aviation broadcast messages are not shown.

international adoption and increase costs. When viewed in the context of interoperability, ADS-B is a well-designed open-access protocol.

5.4 Concerning Scenarios

Consider the following scenario: *Suppose a pilot wishes to fly in secret. During flight, the ADS-B transponder continuously broadcasts ADS-B messages that contain the aircraft's unique identifying number and real-time position. A network of ADS-B receivers operated by aviation enthusiasts through-*

out the country tracks all aircraft, including his, in real-time, and publishes the data online.

In response to privacy concerns voiced by the aviation community, the FAA stated that "there is no right to privacy when operating in the [National Air Space]" [142]. Aircraft flying through Class A, B, C, D, and E airspace must identify themselves to ATC during flight under 14 CFR § 91.215 regulations. However, the FAA does suggest a way to fly anonymously: pilots who choose to employ a UAT-equipped transceiver operating in pseudo-anonymity mode under visual flight rules can maintain anonymity if they do not file a flight plan and make no use of ATC services. In the U.S., this scenario is possible only in Class G airspace. Thus, anonymity remains elusive for aircraft equipped with 1090 MHz Mode-S ES transponders or for aircraft that fly through ATC controlled airspace.

While it is true that aircraft using public airports cannot expect privacy—a pair of binoculars will fare just as well as an ADS-B tracking system—the automation of ADS-B offers a far easier and more persistent way to track an aircraft than does manual surveillance. Such an automatic tracking capability presents an array of concerns similar to those that the U.S. Supreme Court faced in its 2012 ruling on GPS monitoring under the Fourth Amendment in *United States v. Jones*. The Court's 2012 ruling notes the striking difference between conventional and automatic surveillance, of which ADS-B is another example.

Beyond the concerns over the persistence of ADS-B tracking are concerns about its immediacy. Many flight tracking websites display information obtained from the Aircraft Situational Display to Industry (ASDI) that the FAA has offered to a variety of clients since 1998. While ASDI data is offered in real-time to commercial airline companies and flight management companies, most others receive ASDI with at least a five minute delay. The delay was implemented in response to the attacks of 9/11. With ADS-B, however, precise positions and velocities transmitted in real-time are accessible to anyone with an ADS-B receiver. A worrisome possibility of which the FAA is aware is one where real-time, in-the-clear ADS-B broadcasts are used to target passenger aircraft for kinetic or electronic attack [142].

Leaving privacy aside, consider the following scenario: *A rogue hobbyist living near a major airport decides to build a software-defined ADS-B transponder capable of broadcasting forged ADS-B messages. She programs the transponder to broadcast the positions of hundreds of counterfeit aircraft surrounding the airport. Some of these counterfeit positions are close enough to the actual aircraft that other surveillance techniques such as multilateration, angle-of-arrival discrimination, or radar scans cannot distinguish between the legitimate and forged aircraft. ATC and pilots respond by reverting to radar and voice, thereby vitiating the efficiency gains of ADS-B. A plane crash-lands when false aircraft trajectories and low-visibility conditions cause confusion in the cockpit.*

Attacks against ADS-B, such as the one in the preceding scenario as well as those listed in Table 5.1, can confuse pilots and ATC. Confusion is not deadly on its own, but when it is coupled with a stressful situation, such as takeoff or landing, or with compounding conditions, such as snow or wind, the results can be lethal. Recent events including the 2013 crash of Asiana Airlines Flight 214 have indicated a decline in airmanship in favor of technological reliance. How will pilots who have become increasingly reliant on an autopilot and GPS fare when faced with spoofed but plausible ADS-B messages?

5.5 Cryptography for ADS-B

In this section, my goal is to address the question "can cryptography secure ADS-B within the constraints of the current system?" In the discussion that follows, I evaluate proposed ADS-B cryptographic strategies based on their practicality and effectiveness in the technologically-complex, cost-averse, and interoperability-focused aviation community. Each proposal falls into one of four categories: symmetric-key encryption, message authentication codes, asymmetric-key encryption, or digital signatures.

Retrofitting a cryptographic technique to the existing ADS-B protocol faces many difficulties:

- ADS-B is an international protocol. A cryptographic solution must harmonize with existing policy, such as export control laws, and technological capabilities.

Attack	Description	Potential Ramifications
Interception	ADS-B Out messages can be decoded by any ADS-B receiver within range	Loss of privacy; persistent monitoring; targeting for kinetic or electronic attack
Jamming	A jammer can disrupt legitimate ADS-B message reception	Denial of service; fallback to older, less efficient technologies
False Injection	ADS-B messages can be forged and broadcast with intent to deceive air traffic control and aircraft	Falsely indicate a collision appears imminent; confuse pilots or ATC; interfere with legitimate message reception
Navigation	Satellite navigation systems (e.g., GPS) can be spoofed or jammed	False ADS-B position or velocity information; fallback to radar or voice communications

Table 5.1: There are a variety of attacks that can target ADS-B and the services from which it derives its surveillance data. Some of these attacks can be found in [15, 136–138].

- ADS-B is bandwidth constrained. Additional spectrum for ADS-B is scarce, and existing spectrum allocations may actually shrink (c.f., [135], Appendix F).

- ADS-B is interference constrained—that is, the number of aircraft that the ADS-B system can support is limited by interference in the Mode-S ES and UAT frequency bands. Extending the ADS-B message length will increase interference and reduce operational capacity.

- ADS-B operates in a cryptographically untrusted environment. Whatever cryptographic hardware, software, and keys are ultimately employed will be accessible to malicious parties.

The following discussion focuses on the ADS-B 1090 MHz Mode-S ES because of the limited operational scope of UAT. I outline a variety of proposed cryptographic enhancements to ADS-B, postponing until the next section a determination and discussion of the most feasible option.

5.5.1 Symmetric-Key Cryptography

Symmetric-key techniques are known to be computationally efficient. The premise of these techniques is that the sender and recipient share a secret cryptographic key. Without knowledge of the shared secret key, the encrypted messages and message authentication codes (MACs) generated via symmetric-key algorithms are computationally infeasible to forge or predict. In addition,

the secret key cannot be derived from the encrypted messages known as the ciphertext.

5.5.1.1 Symmetric-Key Encryption

Encrypting ADS-B messages via symmetric-key methods means (a) selecting an appropriate symmetric-key encryption algorithm (e.g., Advanced Encryption Standard [AES] or Triple Data Encryption Algorithm), (b) computing and disseminating a cryptographic secret key, and (c) broadcasting the encrypted ADS-B messages in place of the unencrypted, or plaintext, ADS-B messages. A byproduct of symmetric-key encryption is confidentiality: the encrypted message is unintelligible to those without knowledge of the secret key.

In the spectrum- and interference-constrained ADS-B system, a standout symmetric-key encryption protocol is format-preserving encryption (FPE), because the plaintext and resulting ciphertext are the same length. FPE also allows certain ADS-B message parameters to remain unencrypted, such as the data format field, which would facilitate interpretation [144]. Still, FPE remains under review at the U.S. National Institute of Standards and Technology (NIST), and despite favorable early reviews, FPE is not standardized. Other standardized, length-preserving alternatives are feasible, such as AES running output feedback mode with an 8-bit block size. My subsequent analysis, however, finds that no matter how appealing format-preserving protocols may be, symmetric-key encryption is impractical.

5.5.1.2 Symmetric-Key Message Authentication Codes

MACs are typically short messages that are derived from a longer message based on specific MAC-generating algorithms (e.g., keyed-hash message authentication code or parallelizable MAC). The MAC is generally appended to the longer message and the message–MAC pair is broadcast together to allow for immediate validation. A successful verification of the message–MAC pair ensures the recipient that the message–MAC pair were not manipulated after the MAC was generated. However, a MAC approach does not provide confidentiality, because the plaintext is still broadcast.

MACs would increase the message length and would thereby increase the potential of ADS-B message interference, or overlap, during broadcast. Supporting MAC-induced interference on the 1090 MHz channel could vitiate the gains of ADS-B by reducing the system's operational capacity. A potential alternative broadcast scheme is a "lightweight" approach: instead of broadcasting the message–MAC pair together, one transmits only portions of the MAC with every message [137]. The portioned MAC bits could be appended to regular ADS-B messages or broadcast over spare bits in alternate message formats [139]. The downside of the lightweight approach is that it introduces a delay between transmission of the original ADS-B message and the message's eventual MAC-based verification. The next section quantitatively discusses this interference tradeoff.

5.5.1.3 Symmetric Key Management

Symmetric-key techniques suffer from a serious drawback. Any party with knowledge of the secret key can generate a message that will pass cryptographic validation. This means that a single secret key leak compromises the entire system. The security of a symmetric-key system, therefore, depends crucially on the security of the secret key which is required for both encryption and decryption operations as well as MAC generation and validation. To support ADS-B, the secret key must be accessible to every ADS-B transceiver. Secret keys have a short lifetime when they are distributed among potentially untrustworthy groups. Consider that the Sony PlayStation 3 secret key was discovered only two years after its retail debut despite the intentions of system engineers to prevent a key leak.

Three secret key distribution strategies have been proposed: (1) distribute keys to all aircraft in tamper-proof hardware, (2) distribute keys only to select aircraft in tamper-proof hardware, or (3) distribute keys on a per-flight basis via air traffic control during preflight operations. The first approach remains vulnerable to the single-key disclosure leak problem and hinges on the security of the tamper-proof equipment. The feasibility of the second approach, while favored in [144] for civil and military applications, is questionable. How will these "secured" users interact with the "unsecured" users? Is a private-key-holding aircraft supposed to ignore unverifiable messages? What happens if valid yet unverifiable messages are ignored?

The third proposed approach is to distribute a unique secret key for every aircraft on a per-flight basis [139, 141]. During preflight, air traffic control could assign keys that are valid for only that flight and enter those keys into an international database to assist in interactions with other aircraft. The drawback of this approach is that the symmetric key must be securely distributed to every other agent who needs to validate the messages, and those users could, in turn, impersonate the intended user or leak the key. The approach is also vulnerable to a leak of the entire active key database.

5.5.2 Asymmetric-Key Cryptography

Asymmetric-key cryptographic techniques, while less computationally efficient and less length efficient than symmetric-key techniques, can be as secure as their symmetric-key counterparts. Asymmetric-key approaches distribute public–private key pairs via a public-key infrastructure (PKI) where every user has a public–private key pair bound to their identity by a Certificate Authority (CA). The FAA or ICAO could assume the role of CA.

Asymmetric-key techniques have an important advantage over symmetric-key techniques: Alice cannot forge Bob's asymmetric-key encrypted or signed message with her own private–public key pair. So, if a private key is compromised, then only a single key pair needs to be revoked. This stands in contrast to the symmetric-key approach where a single key leak renders the entire system compromised. A PKI has provisions for revoking compromised keys.

5.5.2.1 Asymmetric-Key Encryption

In an asymmetric-key encryption paradigm, users would encrypt the ADS-B message with the intended recipient's public key according to a specific public-key encryption technique (e.g., elliptic curve cryptography [ECC]). The recipient could then decrypt the message with his or her own private key. Confidentiality is also a byproduct of asymmetric-key encryption because only the sender's intended recipient can decrypt the transmission.

Asymmetric-key ADS-B message encryption has two significant drawbacks. First, asymmetric-key block or stream ciphers would increase the transmitted ADS-B message length, much like MACs. Second, and more problematically, unique encrypted ADS-B messages would be required for each recipient [141]. To maintain a fully-connected network of n aircraft would necessitate (n^2-n) unique broadcasts rather than n in the current system.

5.5.2.2 Digital Signatures

Digital signatures are similar to MACs in the sense that they are appended to the original in-the-clear ADS-B message. Digital signature algorithms (e.g., *the* digital signature algorithm [DSA] or elliptic curve DSA [ECDSA]) take a message and a user's private key as input and return a digital signature unique to the input. Upon reception of the message–signature pair, or signed message, the recipient can apply a verification algorithm that authenticates the signed message with the sender's public key. A successful authentication means that the signed message originated with the sender and

was not modified *en route*. Digital signatures could be transmitted in the same ways discussed earlier for MACs.

Within the family of digital signature algorithms, ECDSA generates the shortest digital signatures for a given equivalent symmetric-key security level, which makes ECDSA enticing for ADS-B when coupled with a PKI standard such as the International Telecommunications Union (ITU) X.509 standard [145, 146]. For a symmetric-key equivalent strength of 112 bits, which NIST claims is cryptographically secure until 2030, the ECDSA signature length is 448 bits. Note that this signature length is four times greater than the length of an ADS-B message.

5.5.2.3 Key Management

Public keys are public, like the name suggests, whereas private keys must remain secret to protect the security of the system. Asymmetric techniques can leverage a PKI to generate, disseminate, and revoke keys [146]. Before flight, a complete list of all known public keys or a list of those that had changed since the last flight could be uploaded to the aircraft. Real-time key creation and revocation could be communicated over satellite or ground data links that are available on most commercial flights.

5.6 Can Cryptography Secure ADS-B?

The previous section outlined four cryptographic ADS-B enhancements that were proposed to secure ADS-B. Yet a host of real-world considerations and practicalities mean that only one of these techniques is remotely practical.

First, consider encryption. One of the FAA's goals is to ensure international operation of ADS-B. While the FAA appears to have no policy that explicitly prohibits encryption on civil aviation protocols, the agency states that requiring encrypted ADS-B messages would "unnecessarily limit [ADS-B] use internationally" [142]. Even if the problem of international interoperability could be overcome, one suspects that the FAA and ICAO would reject ADS-B encryption because it undermines traditional safety: Legitimate but encrypted ADS-B messages may at times not be decryptable either due to a technical failure or human error, increasing the risk of aircraft collisions. It is extremely unlikely that the FAA or ICAO would trade this obvious increased risk for a reduction of the hypothetical risks associated with open-access real-time ADS-B broadcasts. Thus, I believe, ADS-B encryption is not viable.

It is worth pausing to consider the implications of this claim. Without ADS-B encryption, pilots of ADS-B–equipped aircraft who do not wish their aircraft's real-time precise position and velocity to be broadcast publicly to the curious and to the malign will have only one option in U.S. airspace: don't fly.

Next, consider symmetric-key techniques. Contrary to [137] and [144], I believe that the threat of symmetric-key leaks and the burden of key management renders symmetric-key encryption and MACs entirely impractical. It is unlikely that the FAA or ICAO would be willing to accept the risk of a symmetric key leak and the subsequent burden of securely re-keying every aircraft worldwide.

Therefore, of the four options discussed previously, asymmetric-key digital signatures are the only viable cryptographic enhancement for ADS-B within the constraints of NextGen. Among the possible digital signature algorithms, ECDSA generates the shortest digital signatures for a given key strength, making it the most appropriate choice in a bandwidth- and interference-constrained communication channel. To further investigate the practicality of an ECDSA-based ADS-B solution, I analyze the PKI and interference burden of its implementation.

5.6.1 Public Key Infrastructure Burden

To enable digital signatures, the aviation community would need to embrace a PKI infrastructure to handle public–private key creation, assignment, and revocation. The ITU X.509 standard, already implemented in non-aviation applications, specifies certificate formats, attributes, and algorithms to facilitate PKI. The authors of [145] and [146] propose X.509 to support cryptographic enhancements to ADS-B. A possible conduit for ground-to-plane data transfer of key certificates and revocation lists is the Airplane Asset Dis-

tribution System (AADS), which provides a framework and a nomenclature for aviation security. The authors of [146] propose AADS to support aviation security.

While feasible, PKI would be a significant financial and technical burden on the aviation community. This burden includes distributing public keys to aircraft and ground control, securing private keys during transmission and operation, and implementing real-time key revocation. A Verisign-like entity with experience in global PKI management is likely better suited for the task than either the FAA or ICAO.

According to FAA, there were approximately 225,000 general aviation aircraft and 7,500 commercial aircraft in the U.S. in 2011. Each wishing to use ADS-B would need a public–private key and would need to securely store the private key. To verify signatures, each plane would also need a list of all other public keys. Assuming the maximum size of a X.509 certificate is about 5 kB, then the size of the full U.S. public–private key database would be about 1.2 GB.

Real-time revocation remains a significant challenge as voice channels are not designed to support revocation. AADS as described in [146] is proposed for communication with aircraft on the ground and would need to be adapted to communicate with aircraft in flight. Another possibility would be to revoke keys over the Flight Information Services Bulletin (FIS-B), which is designed to communicate temporary flight restrictions and airspace information. However, FIS-B is broadcast over UAT frequencies, meaning that aircraft equipped

with 1090 MHz Mode-S ES transponders cannot receive FIS-B without additional hardware. General aviation is unlikely to equip even more technology to support cryptographic enhancements to ADS-B alone, and the FAA is sensitive to its own costs as well as those costs borne by the aviation community. Recall that costs were a driving factor for the dual-link ADS-B strategy.

5.6.2 Interference Burden

If the ECDSA signature is broadcast over 1090 MHz Mode-S ES, it will increase interference and reduce the number of aircraft that ATC can support. Here, I estimate the resulting reduction in operational capacity based on the operational scenarios presented in earlier ADS-B capacity analysis [143, 147].

The ECDSA signature length is 448 bits for a symmetric-key equivalent strength of 112 bits. Two possible broadcast scenarios were analyzed: (A) the broadcast of a 560-bit signed message consisting of a 112-bit ADS-B message and its 448 bit signature, and (B) the broadcast of a sequence of nine 112 bit messages where the first is the standard ADS-B message and the subsequent eight are 56-bit segments of the ECDSA signature packaged in the ADS-B framing structure. The former scenario assumes, optimistically, that the ADS-B message format could be altered, while the latter scenario assumes that the signature can be inserted into the 112 bit ADS-B message format in place of surveillance data but that the 112-bit ADS-B message structure is unchangeable.

The estimate of air traffic operational capacity is based on several assumptions from [147]. The model assumes that the probability distribution of message receipt times over the 1090 MHz channel is Poisson with rates proportional to the "moderately high" interference scenario in [143]. The model further assumes that only one interference message overlap can be tolerated per received message. Lastly, it assumes that aircraft employ a single bottom mounted 125 W antenna to transmit ADS-B messages [142]. Although reducing the transmit power would address the interference problem by reducing the range of receipt of ADS-B messages, the 125 W minimum was selected to ensure that the 150–200 nmi ADS-R separation could still support the dual-link ADS-B strategy as discussed in Sec. 5.3.

The result in Figure 5.2 shows the reduction in operational capacity for scenarios (A) and (B) with 6-sector ground-based receive antenna at a 150–200 nmi spacing. The capacity estimate is based on receiving a message with 99.5% probability of success [143]. The total number of supported aircraft in this range is reduced from 350 aircraft in the unauthenticated case to 80 and 190 aircraft for scenarios (A) and (B), respectively. Also, for scenario (B), the authentication delay is at least nine seconds from broadcast of the original signed ADS-B message.

These estimates are somewhat pessimistic because recent advances in antenna design (e.g., a 12-sector ground receive antenna) and processing techniques can decrease interference. Still, the results are troubling. Given the predicted increase in air traffic—and the estimated 10,000 unmanned aerial

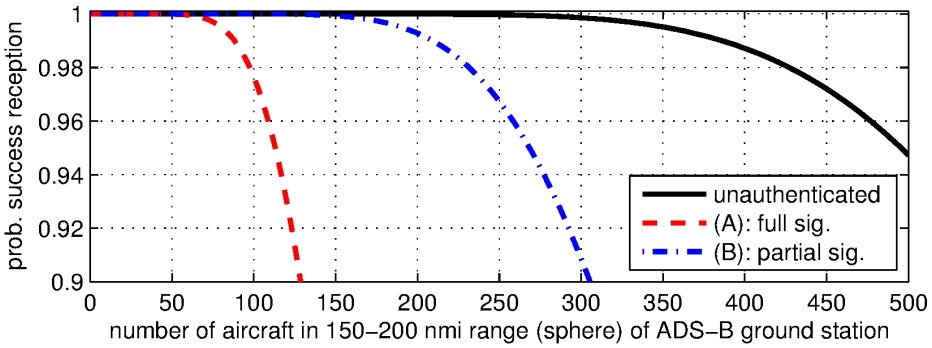


Figure 5.2: Plot showing air traffic operational capacity within a 150–200 nmi range (sphere) of an ADS-B ground station with the addition of ECDSA signatures as compared to unauthenticated broadcasts in the 1090 MHz Mode-S ES band. The red dashed line corresponds to scenario (A): a 560 bit signed message consisting of a 112 bit ADS-B message and its 448 bit signature. The blue dot-dashed line corresponds to scenario (B): a sequence of nine 112 bit messages where the first is the standard ADS-B message and the rest are 56-bit segments of the ECDSA signature packaged in the ADS-B framing structure.

vehicles operating throughout the national air space by 2030—this decrease in operational capacity may simply outweigh the benefits of digital signature broadcasts over the 1090 MHz channel.

One option would be to mitigate the interference with a multi-user modulation format that schedules transmissions in time, frequency, or code to limit interference [139]. A change of this magnitude to a nearly-operational protocol, however, is unlikely because of large signal definition inertia. Another option, which is potentially more practical and effective, would be to broadcast the authenticated messages in an alternate channel.

5.6.3 Alternative Authentication Channels

Instead of trying to retrofit digital signatures to the ADS-B protocol, would it be possible to transmit signed ADS-B messages over alternative channels? Imagine an alternative authentication channel over which signed ADS-B messages could be broadcast at the same rate as ADS-B messages at 1090 MHz or 978 MHz. Such an approach avoids the unpalatable reduction in operational capacity described in the previous section. The signed messages could take the structure suggested earlier, which consists of a 112-bit ADS-B message and its 448-bit ECDSA signature.

A variety of channels are worth considering to support signed ADS-B messages. Possibilities include the channels over which in-flight entertainment or internet connectivity are provided. Such high-bandwidth low-latency con-

nections could transmit a signed ADS-B message to a ground network, which would then relay it to a central ATC database.

Another channel to consider is the protected Aeronautical Navigation Radio Service (ARNS) L-band at 960–1215 MHz where distance measuring equipment (DME) broadcast. The DME band consists of 252 1-MHz-wide channels where DME synchronization pulses and replies are transmitted. The transponder-based position-measurement DME system transmits in this 252 MHz of spectrum with exceptions for UAT transmissions at 978 MHz, Mode-S ES transmissions at 1030 and 1090 MHz, and Global Positioning System transmissions at 1176.45 MHz (L5 frequency).

Employing L-band for ADS-B authentication is enticing for several reasons. First, both Mode-S ES and UAT hardware already operate in the L-band, meaning that additional hardware and additional "holes in the airframe" to support more antennas are unnecessary. The result is a cost savings for commercial and general aviation. Second, the band is already ARNS-protected and allocated for aviation operations. Third, the frequencies allocated to UAT, Mode-S, and GPS L5 were actually re-purposed DME channels. This suggests that one or more 1-MHz-wide DME channels could similarly be allocated to support ADS-B authentication. Finally, the L-band is enticing because the FAA's alternative position navigation and timing (APNT) efforts has already considered this band to transmit additional data and navigation services with bit rates as high as 1000 bps [148].

A drawback of the L-band alternative is that the necessary spectrum redistribution would take significant, collaborative political and technical discussions involving major agencies, such as the FAA and FCC as well as international aviation agencies such as ICAO and EUROCONTROL. Furthermore, DME receivers would need to be replaced, unless they could be updated as part of a software upgrade. Still, if APNT and signed ADS-B message broadcasts could be packaged and implemented together, then only a single operational change could address two problems at once.

5.7 Conclusion

NextGen's ADS-B air traffic surveillance protocol is unacceptably insure, but implementing a cryptographic enhancement would face significant regulatory and technical complexities. The most practical and effective cryptographic approach is one in which ADS-B broadcasts are signed with an asymmetric-key elliptic curve digital signature algorithm. Still, the burden of public-key management and the reduction in operational capacity over the 1090 MHz Mode-S ES channel would likely prove unacceptable to regulatory agencies, commercial airline companies, and general aviation enthusiasts. To avoid these difficulties, a possible alternative would be to broadcast signed ADS-B messages over a side channel such as the aviation-protected L-band at 960–1215 MHz. Meanwhile, ADS-B will continue to rely on radar for authentication—ironically, the very technology it was designed to replace.

Chapter 6

Conclusion

GPS spoofing has become an increasing concern as civil GPS receivers have become enmeshed in critical national infrastructure and safety-of-life applications. To protect civil GPS receivers from spoofing attacks, entirely new anti-spoofing techniques are required than those that protect the military GPS signals. In this dissertation, I defend the following thesis statement:

> *Both cryptographic and non-cryptographic anti-spoofing techniques can secure civil GPS and GNSS navigation and timing while avoiding the serious drawbacks of local storage of secret cryptographic keys that hinder military symmetric-key-based anti-spoofing.*

My contributions toward proving this thesis statement are described in the following section.

6.1 Summary

- Chapter 2 contributes a probabilistic framework that abstracts the particulars of GNSS anti-spoofing to establish necessary conditions for secure location and timing under a security-enhanced GNSS signal model.

- Chapter 3 contributes an asymmetric cryptographic civil Global Positioning System (GPS) signal authentication strategy that is both practical and effective for the GPS L2 and L5 civil navigation message.

- Chapter 4 contributes a GPS anti-spoofing technique that exploits the dilemma facing a spoofer who wishes to simultaneously maintain a low-enough counterfeit signal power to avoid alarms while minimizing telltale distortions of the received cross-correlation profile.

- Chapter 5 offers an in-depth case study of the security and privacy concerns that face the GPS-based ADS-B surveillance technology that is soon to be employed worldwide in aviation.

6.2 Future Work

In this section, I discuss possible future research areas based on my dissertation work:

6.2.1 Hybrid ECDSA–TESLA Implementation for GNSS NMA

The work in [37, 41, 46] goes a long way toward offering a practical and effective signal authentication technique specific to GPS L5 CNAV signals. Future work in this area involves coordinating with the U.S. Air Force GPS Directorate to assess and implement the technique. A variety of practicalities and logistical constraints exist in the operational control segment (OCX) that are not readily accessible to the public. Additionally, Ratheyon is building a

next generation OCX that may more readily be able to accommodate digital signatures into the GPS messages, but again many of the system capabilities are classified. Current work is underway between the UT Radionavigation Laboratory with the Aerospace Corporation and the GPS Directorate to implement civil GPS navigation message authentication in OCX.

The GPS Directorate and others have also become interested in a hybrid ECDSA–TESLA scheme. Despite the fact that TESLA is not a standardized cryptographic technique and that it requires a loose time synchronization, the digital signatures that result are very low overhead [95, 101]. While a TESLA-only solution is ineffective, a hybrid ECDSA–TESLA approach could result in reduced times to authentication for receivers with an alternative timing source (e.g., a networked time source). Some details of the hybrid approach are offered in [47, 73].

6.2.2 Composite Hypothesis Testing

In simple binary hypothesis test, a random variable realizes a specific value from one of two probability distributions. The parameters of each distribution are precisely known and fixed (e.g., a standard normal random variable). For such problems, Neyman–Pearson analysis yields the most powerful test for a specific probability of false alarm.

Interfering signals are not bound to follow distributions, and a GNSS receiver will therefore be subject to interfering signals with unknown parameters. For example, the characteristics of a spoofing signal are unpredictable

and will vary with attack meaning that the precise probability distributions cannot be derived. Instead, parameter ranges can be specified, which changes the hypothesis test from "simple" to "composite."

The techniques in 4 will be extended to a composite-type test that makes full use of the power–distortion tradeoff described earlier. This technique will be referred to as the "pincer" defense (for more details, see Appendix B).

6.2.3 Developing and Testing Against More Sophisticated Spoofing Attacks

Both cryptographic and non-cryptographic anti-spoofing defenses will benefit from testing against additional training scenarios [63]. The Texas Spoofing Test Battery, or TEXBAT, described in Chapter 4 offers a wealth of recorded spoofing scenarios that can be replayed through a signal generator to test the operational capabilities of a defense [42]. Yet, as attacks evolve, so must TEXBAT. Future scenarios are being developed to include SCER-type attacks against a cryptographically-secured navigation message, higher dynamic data, and non-phase-locked tests [88]. The enhanced test set will also be available online and could one day be part of a receiver program to certify "spoof resistant" hardware and software [149]. The result of this work will be another publicly-available dataset: TEXBATv2.

It is also worth considering subtle spoofing attacks that play with the boundaries of detection test. For example, spoofing signals designed to look

like multipath might induce errors without capture. Analysis of these edge-case-type attacks would be beneficial to demonstrate practical limitations of defenses.

6.2.4 Implementation in Operational Conditions

The technique in Chapter 4 reveals the power–distortion tradeoff facing a spoofer and develops a statistical spoofing detection method [58, 63]. While TEXBAT data was employed, further testing in an operational scenario is necessary to ensure its performance. Specifically, the probability of false alarm P_F stands to be evaluated over longer time periods under varying operational conditions and receiver dynamics [150]. For many applications, a high false alarm rate would reduce the effectiveness of the defense because users would start to ignore the alert. A key area to study is the signal distortions present during aircraft flight along with takeoff and landing.

6.2.5 Coupled Frameworks and Evaluation Tools

The probabilistic framework in Chapter 2 demonstrates why signal timing authentication demands a probabilistic model as opposed to the traditionally non-probabilistic security models of message authentication and cryptography. Future work must characterize the joint distribution $p_{\boldsymbol{z}|H_j}(\boldsymbol{\xi}|H_j)$ under M-ary hypothesis testing for a combined cryptographic and non-cryptographic anti-spoofing approach.

The probabilistic framework is readily coupled with additional security assessments. For example, a general position, navigation, and timing (PNT) sensor gathers data from a variety of physical sources (e.g., gravity for compasses and atmospheric pressure for altitude). Such physical characteristics and readings are often more difficult to counterfeit. A physics-based security evaluation, coupled with the understanding of probabilistic anti-spoofing, could lead to even more secure receivers.

6.2.6 Wide Area Augmentation System Authentication

The Wide Area Augmentation System (WAAS) is an aviation broadcast that increases the accuracy of GPS receivers. It was developed specifically to assist aircraft navigation. Like GPS, WAAS contains no security provisions or cryptographic signatures to verify the authenticity of broadcasts. A wide range of GNSS anti-spoofing techniques readily apply to the WAAS signals [95]. With some modification to specific WAAS practicalities, the techniques in Chapters 3 and 4 could potentially bolster the security of WAAS.

Appendices

Appendix A

Challenges of Securely Integrating Unmanned Aircraft into the National Airspace

On August 2, 2010, a Navy helicopter entered the highly restricted airspace above Washington, D.C. without permission [151]. The event might have passed as unremarkable but for the fact that no-one was piloting the helicopter: as an unmanned aircraft, it carried no humans onboard, and—somehow—the vital communications link to its ground operators had been lost. The 1,429-kilogram MQ-8B Fire Scout flew entirely on its own for 30 minutes, blithely drifting through the airspace near nation's capital [152, 153].

Ground operators at Naval Air Station Patuxent River in Maryland eventually regained control of the craft and ordered it to return to base, later diagnosing the cause of the unintended excursion as a "software issue." But in fact more than one error had occurred: not only did the Fire Scout lose its communications link, it failed to execute its "return-to-base" lost-link protocol. So even as one Navy official put a good face on the incident by praising the reliability of the unmanned aircraft's autopilot system [153], most saw it as a disconcerting example of the unresolved safety and security issues surrounding unmanned aircraft.

The cost advantages of unmanned aircraft are compelling and will almost surely make these craft a component of everyday life in years to come. For the price of renting a human-piloted aircraft for a single power line inspection flight, a utility company could buy an entire unmanned aerial vehicle system to do the same job repeatedly. FedEx's CEO and founder, Fredrick W. Smith, has talked about using drones to replace the company's fleet of package-delivery aircraft [154]. For search and rescue, agriculture, infrastructure monitoring, research, and myriad other applications, unmanned aircraft—or drones in the common vernacular—provide convenience and economy. Recognizing this, the U.S. Congress passed the FAA Modernization and Reform Act in February 2012. The Act directs the FAA to draw up a "comprehensive plan to safely accelerate the integration of civil unmanned aircraft systems into the national airspace system" by 2015 setting the stage for broad drone use throughout the U.S.

But there is a growing public backlash. Having witnessed drones employed primarily for surveillance and missile strikes in conflict areas outside the U.S., many see no good reason to welcome them into the U.S. national airspace. Who will be piloting these craft anyhow? Where and why? And with no human pilot onboard looking out the window, won't they be more vulnerable to hijacking or hacking?

Echoing the concerns of their constituents, lawmakers in over 42 states have proposed drone legislation imposing limits on unmanned aircraft use. For example, Texas House Bill No. 912 would make it a misdemeanor for a drone

operator to capture images of private property from an unmanned aircraft without the property owner's "express consent"except under a set of narrow circumstances (e.g., law enforcement in pursuit of a suspected felon). At the federal level, the Preserving American Privacy Act of 2013 would prohibit law enforcement from conducting drone-based surveillance without a warrant and would outlaw armed drones by law enforcement or private citizens over the U.S.

A.1 A Sober Look at the FAA's Task

It is hard to imagine the FAA completing the task of drawing up a comprehensive unmanned aircraft integration plan by 2015 as required by the 2012 Modernization Act. Behind the FAA's standout safety record (witness the absence of fatal domestic aircraft incidents since 2010 [155]) is a slow-moving organization that reflexively associates innovation with risk. The FAA is already in the midst of a broad modernization called the Next Generation Air Transportation System, or NextGen, that will see satellite navigation replace radar as the primary sensor for air traffic control; the additional congressional demand to incorporate unmanned aircraft was no doubt unwelcome. In its 2012 report on the FAA's progress to-date on the Modernization Act, the Government Accountability Office concedes that the FAA has been handed a "daunting task" with an "aggressive time frame." [156].

Beyond the mundane logistical hurdles, integrating unmanned aircraft into the national airspace will also require the FAA to grapple with new secu-

rity and privacy issues. The FAA's primary task is to ensure the safety of our public airways. Historically, this task was limited to preventing accidents due to human error or adverse natural conditions. After 9-11, it became obvious that a safe aircraft must also be secure against an attack by a scheming adversary; consequently, the FAA saw its role expand to include overseeing the installation of reinforced cockpit doors and crafting new security-conscious crew training procedures. From the FAA's point of view, aircraft security is now an integral part of airworthiness. This thinking logically extends to unmanned aircraft, bringing their security squarely within the FAA's purview.

As with security, the FAA has historically not been expected to grapple with issues of privacy related to aviation: it was left to the courts to decide whether someone in an aircraft had invaded someone else's privacy. But after the passage of the Modernization Act, the public is understandably concerned about the prospect of pervasive unmanned aircraft with high-definition cameras. Privacy advocates and members of Congress are now calling on the FAA to employ its regulatory authority to prevent breaches of privacy.

One might expect the Transportation Security Administration and its parent agency, the Department of Homeland Security (DHS), to take the lead in addressing unmanned aircraft security and privacy concerns. On paper, the Department of Transportation agrees, noting in its 2010 annual performance report that "[DHS] has primary responsibility for the security of the transportation system" [157]. But in practice, the FAA is unlikely to get much help from the DHS. In July 2012, Chairman Michael McCaul, speaking be-

fore a subcommittee hearing on "Using Unmanned Aerial Systems Within the Homeland: Security Game Changer?", complained that "[DHS] officials repeatedly stated the Department does not see this function (domestic use of drones) as part of their mission and has no role in domestic unmanned aerial systems" [158].

The FAA appears resigned to shouldering the burden alone. Under questioning about drone security and privacy in a February 2013 House Committee on Science and Technology hearing, FAA representative Dr. Karlin Toner revealed that the Administration had formed a study group to examine security threats against drones and had taken the lead in soliciting advice from the public on questions of privacy. The DHS was conspicuous by its absence at the hearing and in the commentary.

In short, whether the FAA welcomes the changes or not, its regulatory role has expanded over the last decade to cover issues of aircraft security and can be expected to expand further over the next decade to cover issues of privacy.

A.2 Security Concerns

Leaving privacy matters to privacy experts, I offer here a clear-eyed assessment of the security challenges that the FAA will confront as it integrates unmanned aircraft into the national airspace.

Whereas traditional pilots control their aircraft from within, with hands on the yoke and eyes in the sky, unmanned aircraft pilots control their craft remotely, sometimes allowing them to fly autonomously (whether by accident or intent). Autonomous operation leaves drones uniquely dependent on their various radio links: the receive-only links to Global Positioning System (GPS) satellites, the two-way command-and-control link to the aircraft's remote pilot, and one or more links to other aircraft. Disruption or corruption of any one of these links can have serious consequences.

A.2.1 Navigation

Almost all unmanned systems in the coming years will depend on civil satellite navigation systems like GPS for navigation. To be sure, the navigation sensor suite of a typical civil unmanned aircraft also includes inertial sensors (accelerometers and rate sensors), magnetometers, altimeters, and in some cases a camera. Even so, a GPS receiver is fundamental to the sensor suite because, unlike the other navigation sensors, it works in all weather conditions and does not drift.

Military GPS signals have long been encrypted to prevent counterfeiting and unauthorized use. Civil GPS signals, on the other hand, were designed as an open standard, unencrypted and freely-accessible to all [1]. These virtues have made civil GPS enormously popular, but the transparency and predictability of its signals give rise to a dangerous weakness: they can be

counterfeited, or spoofed. In fact, civil GPS is the most popular unauthenticated protocol in the world.

The vulnerability of civil GPS to spoofing has serious implications for unmanned aircraft, as was illustrated by a dramatic remote drone hijacking at White Sands Missile Range in June 2012. The University of Texas at Austin Radionavigation Laboratory conducted the demonstration at the behest of the DHS. From a standoff range of half a mile, our spoofing device commandeered an 80 thousand dollar drone and forced it to plummet toward the desert floor [10].

How was this possible? Spoofing signals can be near-perfect forgeries of authentic GPS signals because (1) the civil GPS signal definition is publicly available, and (2) there are no security provisions, such as digital watermarking or encryption, to thwart counterfeiters [37]. In the White Sands experiment, the drone, unable to distinguish between the authentic GPS signals and the forged signals we were transmitting, ultimately decided to believe the forged signals. Once fooled, it began taking its position cues from our spoofing device. When these signals indicated that the drone was rising vertically upward, the drone's autopilot system reacted by descending to "maintain altitude." The craft was only saved from crashing by a safety pilot who forced a manual override.

The spoofing threat is not new; it was well documented in a 2001 Department of Transportation report, known as the "Volpe Report" [5]. But policymakers and GPS manufactures largely ignored the report's warnings un-

til very recently, perhaps reasoning that a spoofing attack was so unlikely as to not warrant attention. And while GPS researchers have proposed a variety of fixes since 2001, stubborn challenges remain. Techniques that harden GPS signals with cryptographic watermarking are years away from implementation, and non-cryptographic defenses that could be implemented sooner must first prove their reliability in the dynamic signal environment in which drones operate.

Jamming is another concern for GPS-reliant drones. Near the earth's surface GPS signals are extraordinarily weak: they have no more flux density than light received from a 50 Watt bulb 22,000 kilometers away. As a result, their reception is easily disrupted, or jammed, by non-GPS radio-frequency noise in the GPS spectrum. In fact, it is harder not to degrade GPS signals than otherwise: almost any modern electronic system (e.g., a laptop) will dump substantial noise power into a GPS receiver at close range.

Not surprisingly, intentional jamming can be much more targeted and powerful than unintentional jamming, with serious consequences for drones. In May 2012, operators lost control of a 150-kg South Korean Schiebel S-100 Camcopter, which finally crashed into its ground control station, killing an engineer and wounding two remote pilots [159]. A follow-up investigation revealed that North Korean GPS jamming directed into South Korea had precipitated a sequence of events (including erroneous pilot actions) that led to the crash.

As this jamming incident and the University of Texas spoofing demonstration make clear, secure navigation systems are vital for the safe integration of unmanned aircraft into our skies. These systems will need to be spoof- and jam-resistant, detecting and artfully adapting to a disruption of the fragile GPS signals. In case of prolonged interference, they will need a safe "GPS denied" protocol, such as landing nearby or returning to base.

A.2.2 Sense and Avoid

By the FAA's own estimate, more than 10,000 unmanned aircraft will fly the U.S. skyways by 2030. Needless to say, interaction between unmanned aircraft, and between manned and unmanned aircraft, had better be collision-free. Just as traditional pilots use visual and radar cues to sense the presence of other aircraft and avoid collisions, so unmanned systems must also have a sense-and-avoid capability. But the Government Accountability Office notes that, so far, "no suitable technology has been deployed that would provide unmanned aircraft with the capability to sense and avoid other aircraft and airborne objects" while also complying with current FAA regulations [156].

Sense-and-avoid is especially challenging for small drones because these cannot accommodate existing airborne radar systems, which are prohibitively bulky and power hungry. Visible-light and infrared cameras offer an attractive alternative: modern cameras are high resolution, inexpensive, low-power, and compact. Unfortunately, cameras can't be trusted to see through clouds.

Several experts have come to conclude that the only viable primary sense-and-avoid solution for small drones is Automatic Dependent Surveillance-Broadcast, or ADS-B, a critical piece of technology from the FAA's NextGen air traffic system [135]. An ADS-B transponder broadcasts an aircraft's position and velocity every second and receives similar reports from nearby aircraft. By 2020, the FAA will require almost all aircraft to operate ADS-B transponders [142]. So long as all aircraft in a given neighborhood—manned and unmanned—dutifully broadcast their positions and velocities through their ADS-B transponders, the sense-and-avoid problem becomes a multi-agent path planning exercise for which there are many safe protocols.

However, like civil GPS, ADS-B has a serious Achilles's heel: its transmissions are unauthenticated and can thus be counterfeited. This omission stems from the fact that development of ADS-B took place in a time when security was a minor concern. No-one was expected to broadcast fake ADS-B signals because this had never happened before and it was hard to imagine what would motivate someone to spend the time and effort do so. Needless to say, such a naive assumption is out of place in the 21st century. The cost and effort required to mount an ADS-B attack are now alarmingly low; researchers from the Air Force Institute of Technology showed in 2012 that a variety of "false injection" attacks can be readily coded on a commercial software-defined radio platform and launched from the ground or air with a cheap antenna [136]. Attacks could cause aircraft to believe a collision is im-

minent, flood the airspace with hundreds of false transmissions, or prevent reception of legitimate messages.

False ADS-B messages would be problematic for small drones. Whereas a pilot in a snowstorm may quickly verify with onboard radar that a false aircraft is not, in fact, sitting 100 yards ahead in the flight path, a small drone may have no effective secondary sense-and-avoid capability with which to make such a determination.

The FAA is working to address the problem of false ADS-B messages through multilateration, a technique for locating the source of a transmission by measuring its relative arrival time at multiple ground receivers. But reliable multilateration depends on a robust and precise time alternative to GPS, a cost-effective embodiment of which remains elusive [160]. The FAA remains nonetheless, reporting in a 2010 assessment that "using ADS-B data does not subject an aircraft to any increased risk compared to the risk that is experienced today" [142]. To the dismay of security researchers, the Administration declined to explain how it had arrived at this summary dismissal of the problem, citing the sensitivity of its study.

A.2.3 Command and Control

Unmanned aircraft are controlled by a wireless tether, the so-called command and control radio link between the operator and the craft. This link enjoys much better intrinsic security than the GPS and ADS-B signals because it fits in the mold of standard wireless communications signals, for which secure

protocols have been developed. Thus, while the command and control link is in theory vulnerable to eavesdropping or counterfeiting, industry-standard encryption, if employed, should prevent this.

Nonetheless, as for any radio-frequency link, jamming is a concern. Loss of the command-and-control link is referred to as a "lost link" event. Much like with the loss or corruption of GPS signals, no satisfactory solution to the lost link problem has emerged. Operators typically configure their drones with a lost link protocol (e.g., return to base if link lost for more than 30 seconds), but these protocols invariably assume an absolutely reliable navigation system, which, as has been argued, may be an unreasonable expectation. If GPS signals are, for whatever reason, unavailable, and the command and control link is suddenly lost, what should a drone be programmed to do?

Another acute challenge related to the command and control link is the scarcity of protected radio spectrum. Owing to this scacity, many drone manufacturers currently resort to transmitting command and control signals in unprotected radio bands (e.g., the so-called industrial, scientific and medical bands), rendering unmanned aircraft susceptible to unintentional interference from the many electronic systems that already legally occupy these bands.

A.3 Discussion

The extent to which an attacker could exploit the vulnerabilities of unmanned aircraft depends somewhat on the regulations that will govern their operation. In crafting regulations, the FAA will be continually confronted

with a safety/utility tradeoff. A requirement that licensed unmanned aircraft always be maintained within line-of-sight of their (not so remote) operators would be good for safety, but would render drones utterly useless for a great number of legitimate applications. Likewise, requiring continuous active piloting of unmanned aircraft via the command-and-control link, and not allowing a remote operator to command more than one aircraft at a time, may increase resilience in the face of unforeseen events, but would put "dull" back in the "dull, dirty, and dangerous" missions that drones promise to eliminate. Remote control begs for autonomy, and autonomy is the future of unmanned systems.

Perspective is important when considering the security of unmanned aircraft, as their vulnerabilities have either exact parallels or close analogs in the world of manned aircraft. Planes can be hijacked, pilots coerced, communications interrupted, luggage compromised. Yet we continue to fly, not because we're unaware of the risks, but because convenience trumps them. Drones will seek from us the same concession.

Appendix B

Outline of "Pincer" Defense

This appendix offers an outline of the so-called "pincer" defense that exploits the same power–distortion tradeoff and measurements of total in-band power and correlation distortion described in Chapter 4. Recall the power–distortion tradeoff: as the spoofed signal power increases relative to the authentic signal power, the automatic gain control pushes the authentic signals into the thermal noise floor.

An admixture of authentic and spoofed signals causes distortion in the correlation function. Assuming that the spoofer cannot null or block the authentic signals, then the spoofer's only recourse is to broadcast signals with a significantly high power advantage so as to eliminate distortions in the correlation function.

Fig. B.1 illustrates the reduction in distortion brought about by a large power advantage. As the total in-band power (which can be thought of as a proxy for the power advantage) increases, then the total distortion decreases. By limiting the total allowable in-band power before declaring an alarm, the defending receiver can ensure that a spoofing attack will cause detectable distortion in the correlation function.

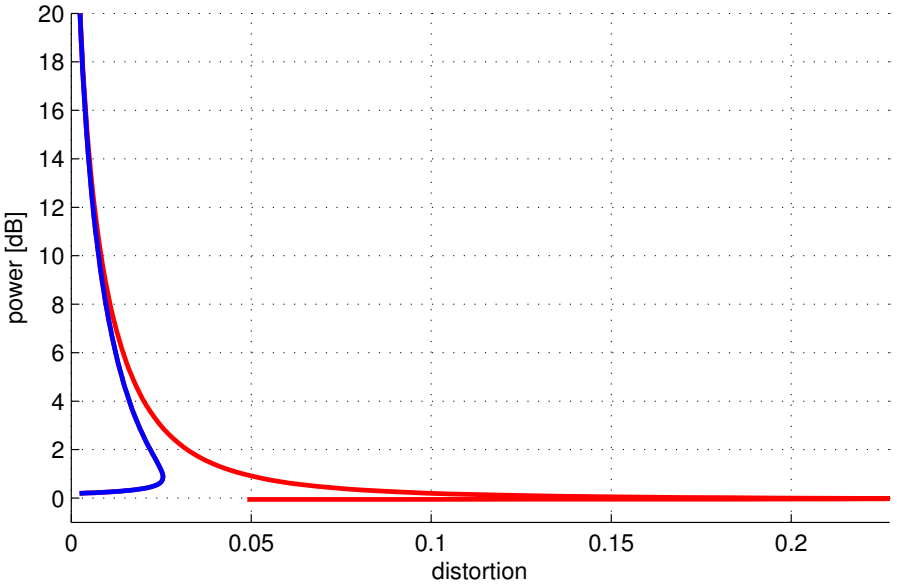


Figure B.1: Plot showing the amount of distortion caused as the total in-band power level increases. An increase in total in-band power corresponds to a higher spoofer power advantage. The blue line shows the distortion caused when the spoofed and authentic signal are in phase, while the red line shows the case where the two are out-of-phase. These two lines define an envelope within which the spoofer can operate.

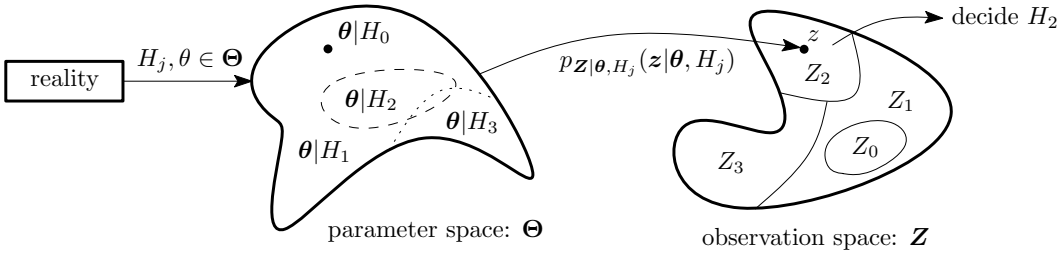


Figure B.2: An illustration of the composite hypothesis testing framework.

Two difficulties remain after measuring distortion and power: $z_k^i = [D_k^i, P_k]^\mathsf{T}$. First, how do we decide between multiple hypothesis (e.g., multipath, jamming, spoofing, etc.)? Second, how can we represent uncertainty in the interference model? A defender can never be exactly sure how an attack will proceed.

the composite hypothesis testing framework illustrated in Fig. B.2 addresses these questions. Here, reality selects a specific hypothesis and parameter $\boldsymbol{\theta}$ in the parameter space $\boldsymbol{\Theta}$. Possible $\boldsymbol{\theta}$ can take on a range of values and may or may not have a readily defined probability density function. A probabilistic transition mechanism $p_{\boldsymbol{Z}|\boldsymbol{\theta},H_j}(\boldsymbol{z}|\boldsymbol{\theta}, H_j)$ maps the parameters to the observation space $\boldsymbol{Z}$. The measurement space is divided into regions where the various hypotheses are declared. In Fig. B.2, $\boldsymbol{z}$ falls into the H_2 region.

Parameter space models agreed with the following assumptions:

- Multipath signals had Rayleigh distributed amplitudes, exponentially distributed time delays (consistent with a Poisson process), and uniformly distributed phases.

- Spoofing signals had power advantages greater than 0.4 dB (assuming the spoofer wanted complete capture), time delays greater than the authentic signal time delay, and phases of zero. Recall from Fig. B.1 that an in-phase spoofing signal caused the least amount of distortion; hence, this assumption makes the detection test more powerful.

- Jamming signals had significantly high power advantages and uniform phase offsets. Time delay is not applicable in the jamming scenario.

A simulation of the observation space for random realizations of the parameters under each hypothesis is shown in Fig. B.3.

Consider the difficulty of differentiating these hypotheses based only on the single measurement of distortion of power alone as some other literature has suggested [58, 69]. Fig. B.4 shows the marginals of a full observation space simulation. Clearly the power of any detection test in this space is severely limited.

The simulated data was compared to three real-world experimental data sets. The first set was collected during a "wardriving" experiment in downtown Austin, TX. The data contains recordings in deep urban environment with static and dynamic receiver platforms. The goal of this data set was to record GPS signals in a severe multipath environment. The second data set was one that recorded 18 so-called personal-privacy devices (i.e., jammers) [90]. The final data set was the Texas Spoofing Test Battery (TEXBAT) [42]. This is the only publicly-available dataset of spoofing attacks. It contains recordings

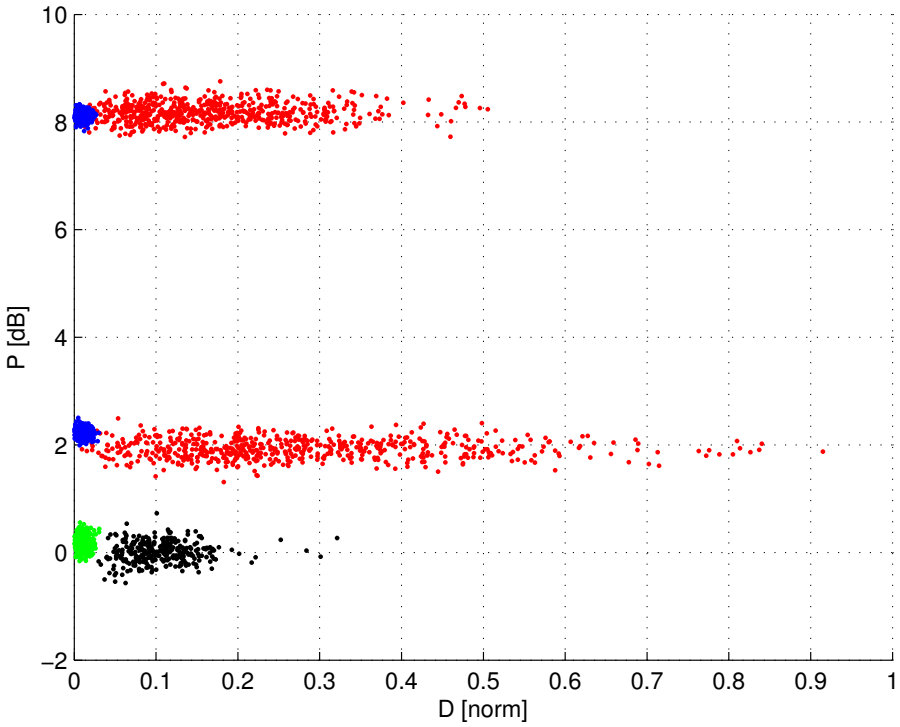


Figure B.3: Plot of the simulated observation space showing four hypotheses: clean in green, multipath in black, spoofing in red (two simulations with various power advantages), and jamming in blue (two simulations with various power advantages).

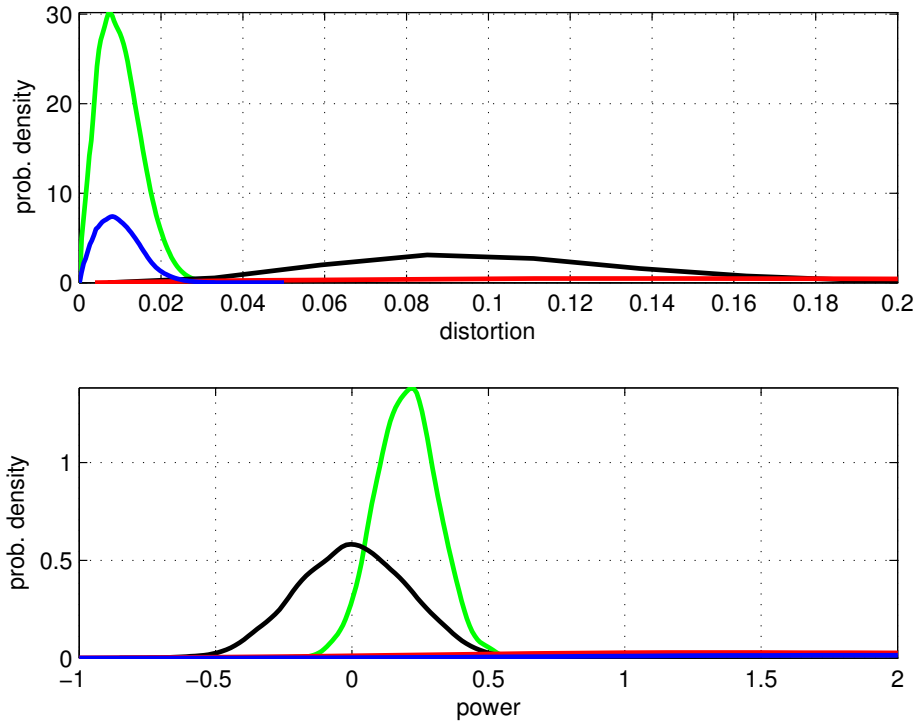


Figure B.4: The marginals of a simulated probability space. Clean is shown in green, multipath in black, spoofing in red, and jamming in blue. Note the difficulty facing a detection test based solely on one of these measurements.

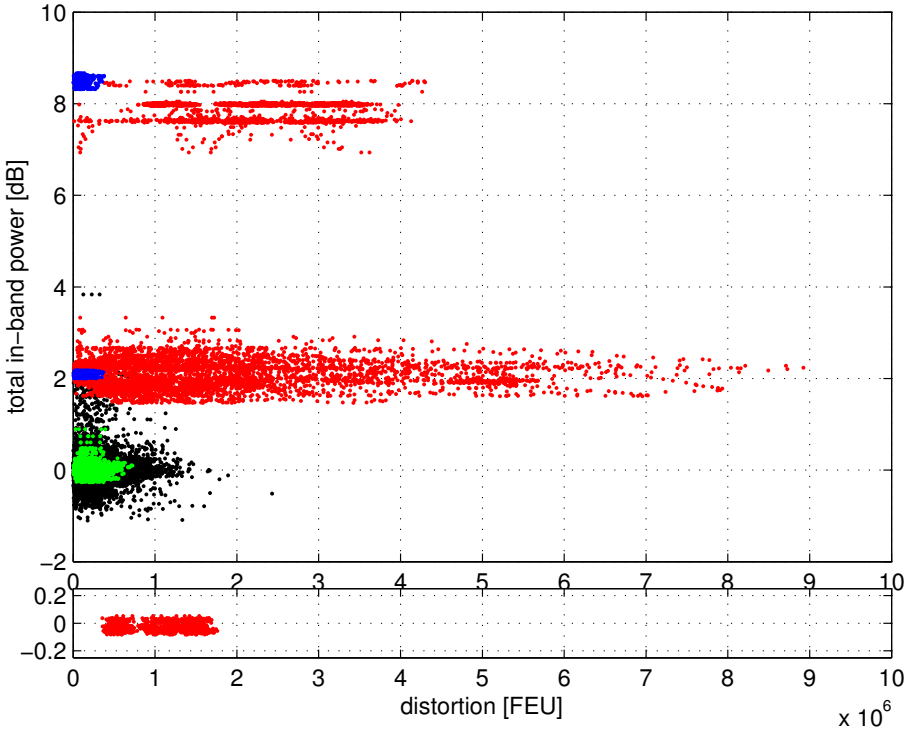


Figure B.5: Plot showing experimental data in the observation space. Clean data is shown in green, multipath in black, spoofing in red, and jamming in blue. Note that there are five spoofing experiments shown with similar power advantages.

of six high-fidelity recordings of static and dynamic receiver platforms under sophisticated spoofing attacks.

Selected data from these data sets is shown in Fig. B.5. Note the agreement with the simulated observation space. Also note the bottom of the figure shows a spoofing scenario in which the authentic signals were not present. This reveals the amount of "natural" distortion in the spoofed signals.

Simulated data of the four hypotheses was then assumed as a first pass to be a Gaussian distribution. The mean vector and covariance matrix of the two-dimensional Gaussian distributions were estimated directly from the simulated data. Then, the observation space was divided into regions where the likelihood function of the various hypotheses were greatest. Fig. B.6 shows the decision region based on these estimates. Costs can also be incorporated to modify the boundary regions depending on the particular user's sensitivity to false alarms or missed detection.

The regions make intuitive sense. The null hypothesis (clean data) is tightly constrained to about the origin with small deviations. The multipath data has more distortion but not significant power increases. The spoofing region lies on a region with significant distortion and high power. The jamming region has high power but little distortion.

Applying these fixed decision regions to three experimental data sets yielded the results in Fig. B.7. The overall empirical probability of detecting an attack (either spoofing or jamming) was 0.999 while the overall empirical probability of false alarm was 0.004. While the overall process will be refined in future work, these initial results demonstrate that the pincer defense is indeed a powerful spoofing defense.

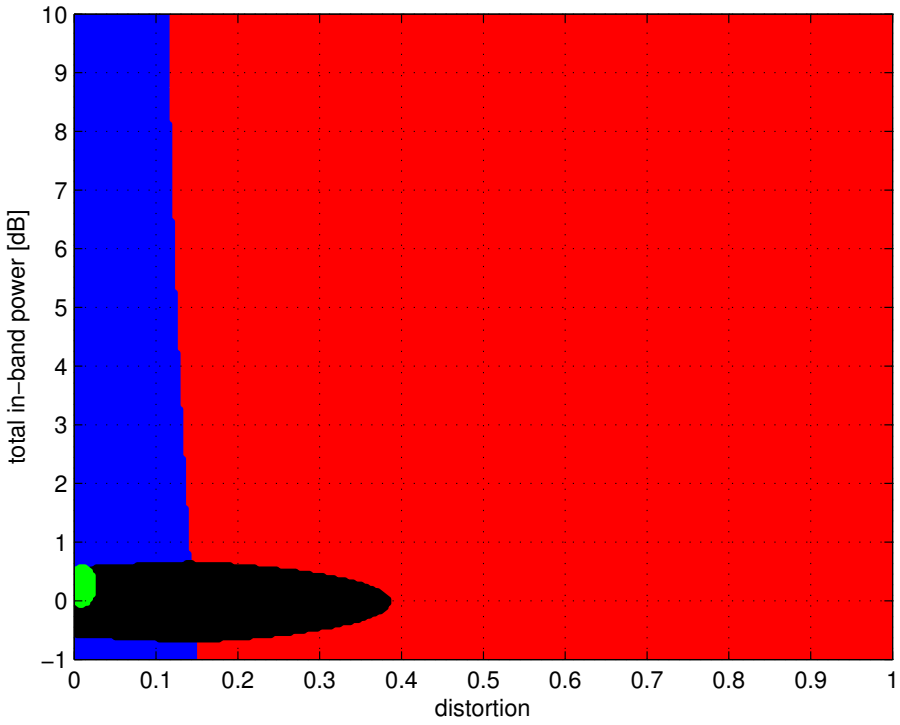


Figure B.6: Plot showing the decision regions based on the likelihood functions.

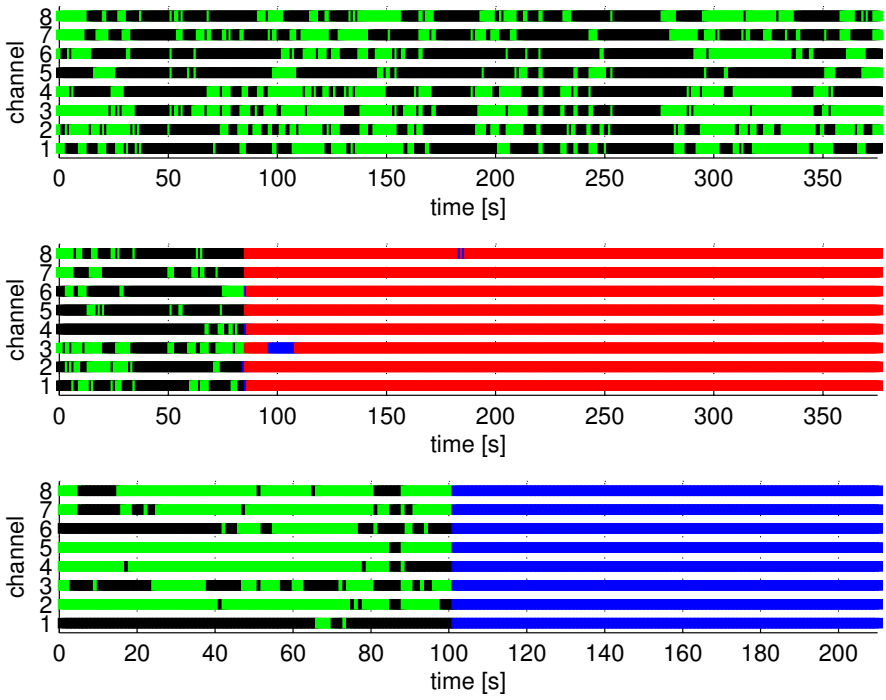


Figure B.7: Plot showing decisions for three experimental data sets. The top plot shows clean data; the middle shows a spoofing attack that initiates at about 80 seconds; and the bottom shows a jamming attack that initiates at 100 seconds.

Bibliography

[1] GPS Directorate. Systems engineering and integration Interface Specification IS-GPS-200G, 2012. http://www.gps.gov/technical/icwg/.

[2] European Union. European GNSS (Galileo) open service signal in space interface control document, 2010. http://ec.europa.eu/enterprise/policies/satnav/galileo/open-service/.

[3] Todd E Humphreys, Brent M Ledvina, Mark L Psiaki, Brady W. O'Hanlon, and Paul M Kintner, Jr. Assessing the spoofing threat: development of a portable GPS civilian spoofer. In *Proceedings of the ION GNSS Meeting*, Savannah, GA, 2008. Institute of Navigation.

[4] Jon S. Warner and Roger G. Johnston. A simple demonstration that the Global Positioning System (GPS) is vulnerable to spoofing. *Journal of Security Administration*, 2003.

[5] John A. Volpe National Transportation Systems Center. Vulnerability assessment of the transportation infrastructure relying on the Global Positioning System, 2001.

[6] Daniel Shepard and Todd E Humphreys. Characterization of receiver response to a spoofing attack. In *Proceedings of the ION GNSS Meeting*, Portland, Oregon, 2011. Institute of Navigation.

[7] Daniel P. Shepard, Jahshan A. Bhatti, Todd E. Humphreys, and Aaron A. Fansler. Evaluation of smart grid and civilian UAV vulnerability to GPS spoofing attacks. In *Proceedings of the ION GNSS Meeting*, 2012.

[8] D. P. Shepard, T. E. Humphreys, and A. A. Fansler. Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks. *International Journal of Critical Infrastructure Protection*, 5(3-4):146–153, 2012.

[9] Kyle D. Wesson, Todd E. Humphreys, and Brian L. Evans. Position paper: Secure time transfer for CPS. In *NSF/NSA National Workshop on The New Clockwork for Time-Critical Systems*, 2012.

[10] Andrew J. Kerns, Daniel P. Shepard, Jahshan A. Bhatti, and Todd E. Humphreys. Unmanned aircraft capture and control via GPS spoofing. *Journal of Field Robotics*, 2014. to be published.

[11] Edwin L. Key. Techniques to counter GPS spoofing. Internal memorandum, MITRE Corporation, Feb. 1995.

[12] M.U. Iqbal and S. Lim. Legal and Ethical Implications of GPS Vulnerabilities. *Journal of International Commercial Law and Technology (JICLT)*, 3(3), 2008.

[13] Guenter Hein, Felix Kneissl, Jose-Angel Avila-Rodriguez, and Stefan Wallner. Authenticating GNSS: Proofs against spoofs, Part 1. *Inside GNSS*, pages 58–63, July/August 2007.

[14] Guenter Hein, Felix Kneissl, Jose-Angel Avila-Rodriguez, and Stefan Wallner. Authenticating GNSS: Proofs against spoofs, Part 2. *Inside GNSS*, pages 71–78, September/October 2007.

[15] Kyle D. Wesson, Daniel P. Shepard, and Todd E. Humphreys. Straight talk on anti-spoofing: Securing the future of PNT. *GPS World*, Jan. 2012.

[16] Department of Homeland Security. National risk estimate: Risks to U.S. critical infrastructure from Global Positioning System disruptions, November 2012. FOUO: No Public Version Available.

[17] V. Chandrasekhar, J. Andrews, and A. Gatherer. Femtocell networks: a survey. *Communications Magazine, IEEE*, 46(9):59–67, 2008.

[18] 3GPP2. Recommended minimum performance standards for CDMA2000 spread spectrum base stations, C.S0010-B. Technical report, 3rd Generation Partnership Project 2, Feb. 2004.

[19] Kenneth M. Pesyna, Jr., Kyle D. Wesson, Robert W. Heath, Jr., and Todd E. Humphreys. Extending the reach of GPS-assisted femtocell synchronization and localization through tightly-coupled opportunistic navigation. In *IEEE GLOBECOM Workshop*, 2011.

[20] I. Hwang, Y. Kang, H. Kim, and S. Kim. Synchronization issue for mobile WiMAX femtocell. In *Information and Communication Technology Convergence (ICTC), 2010 International Conference on*, pages 563–564. IEEE, 2010.

[21] R.Y. Kim, J.S. Kwak, and K. Etemad. WiMAX femtocell: requirements, challenges, and solutions. *IEEE Communications Magazine*, 47(9):84–91, 2009.

[22] J. Giri, D. Sun, and R. Avila-Rosales. Wanted: A more intelligent grid. *IEEE Power & Energy*, pages 34–40, April 2009.

[23] K. E. Martin and et al. Exploring the IEEE standard C37.118–2005 synchrophasors for power systems. *IEEE Transactions on Power Delivery*, 23(4):1805–1811, Oct. 2008.

[24] Zhenghao Zhang, Shuping Gong, Husheng Li, Changxing Pei, Qi Zeng, and Ming Jin. Time stamp attack on wide area monitoring system in smart grid. *Computing Research Repository*, Feb. 2011.

[25] Le Xie, Yilin Mo, and Bruno Sinopoli. False data injection attacks in electricity markets. In *Proceedings of IEEE Smart Grid Communications (SmartGridComm) Conference*, pages 226–231, 2010.

[26] André Teixeira, Saurabh Amin, Henrik Sandberg, Karl H Johansson, and Shankar S Sastry. Cyber security analysis of state estimators in

electric power systems. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 5991–5998. IEEE, 2010.

[27] Sanjeev Dewan and Haim Mendelson. Information technology and time-based competition in financial markets. *Management Science*, 44(5):595–609, May 1998.

[28] Financial Industry Regulation Authority. OATS reporting technical specifications. Online., July 2010. `http://www.finra.org/`.

[29] Jonathan A. Brogaard. High frequency trading, information, and profits. The Future of Computer Trading in Financial Markets. Foresight Driver Review–DR 10, March 2011.

[30] David Schneider. The microsecond market. In *IEEE Spectrum*, pages 67–71,80–81, June 2012.

[31] Alex D. Wissner-Gross and C. E. Freer. Relativistic statistical arbirage. *Physical Review E*, 82(5):056104–1–7, 2010.

[32] James J. Angel. Impact of special relativity on securities regulation. The Future of Computer Trading in Financial Markets. Foresight Driver Review–DR 15, April 2011.

[33] Tara Bhupathi. Technology's latest market manipulator? high frequency trading: the strategies, tools, risks, and responses. *North Carolina Journal of Law and Technology*, 11(2):377–400, Spring 2010.

[34] Olivia Solon. GPS 'spoofers' could be used for high-frequency financial trading fraud. Wired.co.uk, Feb. 2012.

[35] B.C. Barker, J.W. Betz, J.E. Clark, J.T. Correia, J.T. Gillis, S. Lazar, K.A. Rehborn, and J.R. Straton III. Overview of the GPS M code signal. Technical report, DTIC Document, 2006.

[36] Logan Scott. Anti-spoofing and authenticated signal architectures for civil navigation systems. In *Proceedings of the ION GNSS Meeting*, pages 1542–1552, 2003.

[37] Kyle D. Wesson, Mark P. Rothlisberger, and Todd E Humphreys. Practical cryptographic civil GPS signal authentication. *Navigation, Journal of the Institute of Navigation*, 59(3):177–193, 2012.

[38] M.L. Psiaki, B.W. O'Hanlon, J.A. Bhatti, D.P. Shepard, and T.E. Humphreys. GPS spoofing detection via dual-receiver correlation of military signals. *IEEE Transactions on Aerospace and Electronic Systems*, 49(4):2250–2267, 2013.

[39] Oscar Pozzobon, Chris Wullems, and Marco Detratti. Tamper resistance: Security considerations for GNSS receivers. *GPS World*, pages 37–41, April 2011. to appear.

[40] Don Jewell. PHGPST resurrected: Seeking the perfect device. *GPS World*, Dec. 2012.

[41] Todd E Humphreys. Detection strategy for cryptographic GNSS anti-spoofing. *IEEE Transactions on Aerospace and Electronic Systems*, 49(2):1073–1090, 2013.

[42] Todd E. Humphreys, Jahshan A. Bhatti, Daniel P. Shepard, and Kyle D. Wesson. The Texas Spoofing Test Battery: Toward a standard for evaluating GNSS signal authentication techniques. In *Proceedings of the ION GNSS Meeting*, 2012. `http://radionavlab.ae.utexas.edu/texbat`.

[43] Markus Kuhn. An asymmetric security mechanism for navigation signals. In *Proc. of the 6th Int. Information Hiding Workshop*, pages 239–252. Springer, May 2004.

[44] Chris Wullems, Oscar Pozzobon, and Kurt Kubik. Signal authentication and integrity schemes for next generation global navigation satellite systems. In *Proc. European Navigation Conference GNSS*, Munich, July 2005.

[45] Oscar Pozzobon. Keeping the spoofs out: Signal authentication services for future GNSS. *Inside GNSS*, 6(3):48–55, May/June 2011.

[46] Kyle D. Wesson, Mark P. Rothlisberger, and Todd E. Humphreys. A proposed navigation message authentication implementation for civil GPS anti-spoofing. In *Proceedings of the ION GNSS Meeting*, Portland, Oregon, 2011. Institute of Navigation.

[47] Andrew J. Kerns, Kyle D. Wesson, and Todd E. Humphreys. A blueprint for civil GPS navigation message authentication. In *Proceedings of the IEEE/ION PLANS Meeting*, May 2014. submitted for review.

[48] Brady W. OHanlon, Mark L. Psiaki, Todd E. Humphreys, Jahshan A. Bhatti, and Daniel P. Shepard. Real-time GPS spoofing detection via correlation of encrypted signals. *Navigation, Journal of the Institute of Navigation*, 60(4):267–278, 2013.

[49] Sherman Lo, David DeLorenzo, Per Enge, Dennis Akos, and Paul Bradley. Signal authentication. *Inside GNSS*, 0(0):30–39, Sept. 2009.

[50] Mark L. Psiaki, Brady W. O'Hanlon, Jahshan A. Bhatti, and Todd E. Humphreys. Civilian GPS spoofing detection based on dual-receiver correlation of military signals. In *Proceedings of the ION GNSS Meeting*, Portland, Oregon, 2011. Institute of Navigation.

[51] B.W. O'Hanlon, M.L. Psiaki, J.A. Bhatti, and T.E. Humphreys. Real-time spoofing detection using correlation between two civil GPS receiver. In *Proceedings of the ION GNSS Meeting*, Nashville, Tennessee, 2012. Institute of Navigation.

[52] David S. De Lorenzo, Jennifer Gautier, Jason Rife, Per Enge, and Dennis Akos. Adaptive array processing for GPS interference rejection. In *Proceedings of the ION GNSS Meeting*, Long Beach, CA, Sept. 2005. Institute of Navigation.

[53] Paul Y. Montgomery, Todd E. Humphreys, and Brent M. Ledvina. A multi-antenna defense: Receiver-autonomous GPS spoofing detection. *Inside GNSS*, 4(2):40–46, April 2009.

[54] A. Broumandan, A. Jafarnia-Jahromi, V. Dehgahanian, J. Nielsen, and G. Lachapelle. GNSS spoofing detection in handheld receivers based on signal spatial correlation. In *Proceedings of the IEEE/ION PLANS Meeting*, Myrtle Beach, SC, April 2012. Institute of Navigation.

[55] Mark Psiaki, Steven P. Powell, and Brady W. O'Hanlon. GNSS spoofing detection using high-frequency antenna motion and carrier-phase data. In *Proceedings of the ION GNSS+ Meeting*, pages 2949–2991, 2013.

[56] Daniele Borio. PANOVA tests and their application to GNSS spoofing detection. *IEEE Transactions on Aerospace and Electronic Systems*, 49(1):381–394, Jan. 2013.

[57] Peter F. Swaszek and Richard J. Hartnett. Spoof detection using multiple COTS receivers in safety critical applications. In *Proceedings of the ION GNSS+ Meeting*, 2013.

[58] Kyle D. Wesson, Daniel P. Shepard, Jahshan A. Bhatti, and Todd E. Humphreys. An evaluation of the vestigial signal defense for civil GPS anti-spoofing. In *Proceedings of the ION GNSS Meeting*, Portland, OR, 2011.

[59] V. Dehghanian, J. Nielsen, and G. Lachapelle. GNSS spoofing detection based on receiver C/No estimates. In *Proceedings of the ION GNSS Meeting*, Nashville, Tennessee, 2012. Institute of Navigation.

[60] J. Nielsen, A. Broumandan, and G. LaChapelle. Method and system for detecting GNSS spoofing signals, May 2011. US Patent 7,952,519.

[61] John Nielsen, Ali Broumandan, and Gerard Lachapelle. GNSS spoofing detection for single antenna handheld receivers. *Navigation, Journal of the Institute of Navigation*, 58(4):335–344, 2011.

[62] Kyle D. Wesson, Brian L. Evans, and Todd E. Humphreys. A combined symmetric difference and power monitoring GNSS anti-spoong technique. In *IEEE Global Conference on Signal and Information Processing*, 2013.

[63] Kyle D. Wesson, Todd E. Humphreys, and Brian L. Evans. Nonparametric GPS spoofing detection. *IEEE Transactions on Information Security and Privacy*, 2014. in preparation.

[64] Brent M. Ledvina, William J. Bencze, Bryan Galusha, and Isaac Miller. An in-line anti-spoofing module for legacy civil GPS receivers. In *Proceedings of the ION ITM*, San Diego, CA, Jan. 2010.

[65] Antonio Cavaleri, Beatrice Motella, Marco Pini, and Maurizio Fantino. Detection of spoofed GPS signals at code and carrier tracking level. In

5th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing, Dec. 2010.

[66] Robert Eric Phelts. *Multicorrelator Techniques for Robust Mitigation of Threats to GPS Signal Quality*. Stanford University, 2001.

[67] Omer Mohsin Mubarak and Andrew G. Dempster. Analysis of early late phase in single- and dual frequency GPS receivers for multipath detection. *GPS Solut*, 14:381–388, Feb. 2010.

[68] Micaela Troglia Gamba, Beatrice Motella, and Marco Pini. Statistical test applied to detect distortions of GNSS signals. In *International Conference on Localization and GNSS*, 2013.

[69] Dennis M Akos. Who's afraid of the spoofer? GPS/GNSS spoofing detection via automatic gain control (AGC). *Navigation, Journal of the Institute of Navigation*, 59(4):281–290, 2012.

[70] Asghar Tabatabaei Balaei and Andrew G. Dempster. A stasistical inference technique for GPS interference detection. *IEEE Transactions on Aerospace and Electronic Systems*, 45(4):1499–1511, 2009.

[71] Kyle D. Wesson, Brian L. Evans, and Todd E. Humphreys. A probabilistic framework for Global Navigation Satellite System signal timing assurance. In *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2013.

[72] Todd E. Humphreys, Jashan A. Bhatti, Daniel P. Shepard, Kyle D. Wesson, and Andrew J. Kerns. Toward a framework for evaluating pnt security. In *International Symposium on Certification of GNSS Systems and Services (CERGAL)*, Dresden, Germany, 2014. in preparation.

[73] Andrew J. Kerns, Kyle D. Wesson, and Todd E. Humphreys. Efficient navigation message authentication for civil GPS. *Navigation, Journal of the Institute of Navigation*, 2014. in preparation.

[74] Kyle D. Wesson, Todd E. Humphreys, and Brian L. Evans. Can cryptography secure next generation air traffic surveillance? *IEEE Security and Privacy Magazine*, 2014. submitted for review.

[75] Kyle D. Wesson and Todd E. Humphreys. Hacking drones. *Scientific American*, 309(5):54–59, 2013.

[76] Christof Paar and Jan Pelzl. *Understanding Cryptography: A Textbook for Students and Practitioners*. Springer, 2010.

[77] Mark Luk, Adrian Perrig, and Bram Whillock. Seven cardinal properties of sensor network broadcast authentication. In *Proceedings of the ACM workshop on security of ad hoc and sensor networks*, pages 147–156, Alexandria, VA, Oct. 2006.

[78] NIST. Digital signature standard. FIPS PUB 186-3, National Institute of Standards and Technology, June 2009.

[79] P. Papadimitratos and A. Jovanovic. Protection and fundamental vulnerability of GNSS. In *IEEE Int. Workshop on Satellite and Space Communications*, pages 167–171, 2008.

[80] Paul Y. Montgomergy, Todd E. Humphreys, and Brent M. Ledvina. Receiver-autonomous spoofing detection: Experimental results of a multi-antenna receiver defense against a portable civil GPS spoofer. In *Proceedings of the ION ITM*, Anaheim, CA, Jan. 2009.

[81] NIST. Recommendation for key management—Part I: General (revised). SP 800-57, National Institute of Standards and Technology, March 2007.

[82] K. T. Woo. Optimum semi-codeless carrier phase tracking of L2. *Navigation, Journal of the Institute of Navigation*, 47(2):82 – 99, 2000.

[83] Todd K. Moon. *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley, 2005.

[84] Hubert Zimmerman. OSI reference model—the ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications*, COM-28(4):425–432, April 1980.

[85] R. Grover Brown. *Global Positioning System: Theory and Applications*, volume 2, chapter 5: Receiver Autonomous Integrity Monitoring, pages 143–168. American Institute of Aeronautics and Astronautics, Washington, D.C., 1996.

[86] Matthew Lashley. *Modeling and Performance Analysis of GPS Vector Tracking Algorithms*. PhD thesis, Auburn University, Auburn, Alabama, December 2009.

[87] Michael A. Lombardi. NIST frequency measurement and analysis system: Operator's manual. Technical Report NISTIR 6610, National Institute of Standards and Technology (NIST), Aug. 2001.

[88] Todd E. Humphreys, Daniel P. Shepard, Jahshan A. Bhatti, and Kyle D. Wesson. A testbed for developing and evaluating GNSS signal authentication techniques. 2013. in preparation; available at `http://radionavlab.ae.utexas.edu/testbed`.

[89] G.M. Nita, D.E. Gary, LJ Lanzerotti, and DJ Thomson. The peak flux distribution of solar radio bursts. *The Astrophysical Journal*, 570:423, 2002.

[90] R.H. Mitch, R.C. Dougherty, M.L. Psiaki, S.P. Powell, B.W. O'Hanlon, J.A. Bhatti, and T.E. Humphreys. Signal characteristics of civil GPS jammers. In *Proceedings of the ION GNSS Meeting*, 2011.

[91] Richard D. J. van Nee. Spread-spectrum code and carrier synchronization errors caused by multipath and interference. *IEEE Transactions on Aerospace and Electronic Systems*, 29(4):1359–1365, 1993.

[92] Pau Closas, Carles Fernandez-Prades, and Juan A. Fernandez-Rubino. A Bayesian approach to multipath mitigation in GNSS receivers. *IEEE Journal of Selected Topics in Signal Processing*, 3(4):695–706, Aug. 2009.

[93] Randolph G. Hartman. Spoofing detection system for a satellite positioning system. US Patent 5557284, Aug. 1996.

[94] T.A. Stansell. Location assurance commentary. *GPS World*, 18(7):19, 2007.

[95] Sherman C. Lo and Per K. Enge. Authenticating aviation augmentation system broadcasts. In *Proceedings of the IEEE/ION PLANS Meeting*, pages 708–717, Palm Springs, California, 2010. Institute of Navigation.

[96] O. Pozzobon, C. Wullems, and K. Kubic. Secure tracking using trusted GNSS receivers and Galileo authentication services. *Journal of Global Positioning Systems*, 3(1-2):200–207, 2004.

[97] Bruce Schneier. *Applied Cryptography*. John Wiley & Sons, 2nd edition edition, 1996.

[98] Niels Ferguson and Bruce Schneier. *Practical Cryptography*. Wiley, 2003.

[99] Shimshon Berkovits, Santosh Chokhani, Judith A. Furlong, Jisoo A. Geiter, and Jonathan C. Guild. Public key infrastructure study final report. In *Produced by the MITRE Corporation for NIST*, April 1994.

[100] A. Perrig, R. Canetti, J.D. Tygar, and D. Song. The TESLA broadcast authentication protocol. *RSA CryptoBytes*, 5(2):2–13, 2002.

[101] G.T. Becker, S. Lo, D. De Lorenzo, D. Qiu, C. Paar, and P. Enge. Efficient authentication mechanisms for navigation systems—a radio-navigation case study. In *Proceedings of the ION GNSS Meeting*, Savannah, Georgia, 2009. Institute of Navigation.

[102] K. Lauter. The advantages of elliptic curve cryptography for wireless security. *IEEE Wireless Communications*, 11(1):62–67, 2004.

[103] D.R. Hankerson, S.A. Vanstone, and A.J. Menezes. *Guide to elliptic curve cryptography*. Springer-Verlag, 2004.

[104] A.H. Koblitz, N. Koblitz, and A. Menezes. Elliptic curve cryptography: the serpentine course of a paradigm shift. *Journal of Number Theory*, 2009.

[105] Jerome A. Solinas. Efficient arithmetic on Koblitz curves. *Designs, Codes, and Cryptography*, pages 195–249, 2000.

[106] Anon. Secure hash standard. FIPS PUB 180-3, National Institute of Standards and Technology, Oct. 2008.

[107] Elaine Barker and John Kelsey. Recommendation for random number generation using deterministic random bit generators (revised). NIST special publication 800-90, National Institute of Standards and Technology, Mar. 2007.

[108] Tommy R. Jensen and Bjarne Toft. *Graph Coloring Problems*. Wiley Series in Discrete Mathematics and Optimization. Wiley-Interscience, 1994.

[109] Erick Lansard, Eric Frayssinhes, and Jean-Luc Palmade. Global design of satellite constellations: a multi-criteria performance comparison of classical Walker patterns and new design patterns. *Acta Astronautica*, 42(9):555–564, 1998.

[110] Todd E Humphreys, Brent M Ledvina, Mark L Psiaki, and Paul M Kintner, Jr. GNSS receiver implementation on a DSP: Status, challenges, and prospects. In *Proceedings of the ION GNSS Meeting*, pages 2370–2382, Fort Worth, TX, 2006. Institute of Navigation.

[111] Todd E Humphreys, Jahshan Bhatti, Thomas Pany, Brent Ledvina, and Brady O'Hanlon. Exploiting multicore technology in software-defined GNSS receivers. In *Proceedings of the ION GNSS Meeting*, pages 326–338, Savannah, GA, 2009. Institute of Navigation.

[112] B.W. O'Hanlon, M.L. Psiaki, S.P. Powell, J.A. Bhatti, Todd E. Humphreys, G. Crowley, and G.S. Bust. CASES: A smart, compact GPS software receiver for space weather monitoring. In *Proceedings of the ION GNSS Meeting*, pages 2745–2753, Portland, Oregon, 2011. Institute of Navigation.

[113] Carl Fenger. u-blox 6 GPS receivers enhanced with many new features. Technical Report GPS-X-11012, ublox, July 2011.

[114] Javad. *TRIUMPH-VS Datasheet*, rev. 2.6 edition, June 2011.

[115] Todd E. Humphreys, J. Bhatti, and B. Ledvina. The GPS Assimilator: a method for upgrading existing GPS user equipment to improve accuracy, robustness, and resistance to spoofing. In *Proceedings of the ION GNSS Meeting*, Portland, Oregon, 2010. Institute of Navigation.

[116] Eric W. Smith. The implementation and analysis of the ECDSA on the Motorola StarCore SC140 DSP primarily targeting portable devices. Master's thesis, University of Waterloo, Ontario, Canada, 2002.

[117] Billy Bob Brumley and Kimmo U. Jarvinen. Conversion algorithms and implementations for Koblitz curve cryptography. *IEEE Trans. on Computers*, 59(1):81–92, Jan. 2010.

[118] National Security Agency. The case for Elliptic Curve Cryptography, January 2009. `http://www.nsa.gov/business/programs/elliptic\_curve.shtml`.

[119] Ali Jafarnia-Jahromi, Ali Broumandan, John Nielsen, and Gerard Lachapelle. Review article: GPS vulnerability to spoofing threats and review of antispoofing techniques. *International Journal of Naivgation and Observation*, pages 1–16, 2012.

[120] Mehmet Celenk, Thomas Conley, John Willis, and James Graham. Predictive network anomaly detection and visualization. *IEEE Transactions on Information Forensics and Security*, 5(2):288–299, 2010.

[121] Hong Chang, Yi Yao, Andreas Koschan, Besma Abidi, and Mongi Abidi. Improving face recognition via narrowband spectral range selection using Jeffery Divergence. *IEEE Transactions on Information Forensics and Security*, 4(1):111–122, 2009.

[122] Ali Jahromi Jafarnia. *GNSS Signal Authenticity Verification in the Presence of Structural Interference*. University of Calgary, 2013.

[123] Michael S. Braasch. Autocorrelation sidelobe considerations in the characterization of multipath errors. *IEEE Transactions on Aerospace and Electronic Systems*, 33(1):290–295, Jan. 1997.

[124] A. J. Van Dierendonck, Pat Fenton, and Tom Ford. Theory and performance of narrow correlator spacing in a GPS receiver. *Navigation, Journal of the Institute of Navigation*, 39(3):265–283, Fall 1992.

[125] Sanjeev Gunawardena, Zhen Zhu, Maarten Uijt de Haag, and Frank van Graas. Remote-controlled, continuously operating GPS anomalous event monitor. *Navigation, Journal of the Institute of Navigation*, 56(2):97–113, 2009.

[126] Markus Irsigler and G Hein. Development of a real time multipath monitor based on multi-correlator observations. In *Proceedings of the ION GNSS Meeting*, 2005.

[127] Phillip W. Ward. GPS receiver RF interference monitoring, mitigation, and analysis techniques. *Navigation, Journal of the Institute of Navigation*, 41(4):367–391, 1994.

[128] Alexander G. Tartakovsky, Boris L. Rozovskii, Rudolf B. Blazek, and Hongjoon Kim. A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods. *IEEE Transactions on Signal Processing*, 54(9):3372–3382, 2006.

[129] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: a survey. *ACM Computing Surveys*, 41(3):1–72, 2009.

[130] C.D. Scott and E.D. Kolaczyk. Annotated minimum volume sets for nonparametric anomaly discovery. In *IEEE Workshop on Statistical Signal Processing*, pages 234–238, 2007.

[131] Sutharshan Rajasegarar, Christopher Leckie, James C. Bezdek, and Marimuthu Palaniswami. Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks. *IEEE Transactions on Information Forensics and Security*, 5(3):518–533, 2010.

[132] B. W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.

[133] David W. Scott. *Multivariate density estimation: theory, practice, and visualization*. Wiley, 1992.

[134] Tarn Duong. `ks`: Kernal density estimation and kernal discriminant analysis for multivariate data in `R`. *Journal of Statistical Software*, 21(7), Oct. 2007.

[135] Special Committee 186. Minimum aviation system performance standards for Automatic Dependent Surveillance Broadcast (ADS-B), 2002. RTCA DO-242A.

[136] Donald L. McCallie. Exploring potential ADS-B vulnerabilities in the FAA's NextGen air transportation system. Master's thesis, Air Force Institute of Technology, 2011.

[137] Andrei Costin and Aurelien Francillon. Ghost in the Air(Traffic): on insecurity of ADS-B protocol and practical attacks on ADS-B devices. In *Blackhat*, July 2012.

[138] Domenic Magazu, III, Robert F. Mills, Jonathan W. Butts, and David J. Robinson. Exploiting the Automatic Dependent Surveillance-Broadcast system via false target injection. *Journal of Aviation and Aerospace Perspectives*, 2(2):5–19, 2012.

[139] Ken Samuelson, Ed Valovage, and Dana Hall. Enhanced ADS-B research. In *IEEE Aerospace Conference*, 2006.

[140] Krishna Sampigethaya, Radha Poovendran, Sudhakar Shetty, Terry Davis, and Chuck Royalty. Future e-enabled aircraft communications and security: The next 20 years and beyond. *Proceedings of the IEEE*, 99(11):2040–2055, Nov. 2011.

[141] Cindy Finke, Jonathan Butts, and Robert Mills. ADS-B encryption: confidentiality in the friendly skies. In *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*, Oak Ridge, TN, 2012.

[142] Federal Aviation Administration. 14 CFR Part 91: Automatic Dependent Surveillance–Broadcast (ADS–B) Out Performance Requirements To Support Air Traffic Control (ATC) Service; Final Rule. Federal Register, May 28, 2010.

[143] Robert E. Boisvert and Vincent A. Orlando. ADS-Mode S system overview. In *IEEE Digital Avionics Systems Conference*, Oct. 1993.

[144] Cindy Finke, Jonathan Butts, Robert Mills, and Michael Grimaila. Enhancing the security of aircraft surveillance in the next generation air traffic control system. *International Journal of Critical Infrastructure Protection*, 2013.

[145] Wei-Jun Pan, Zi-Liang Feng, and Yang Wang. ADS-B data authentication based on ECC and X.509 certificate. *Jounal of Electronic Science and Technology*, 10(1):51–55, Mar. 2012.

[146] Richard V. Robinson, Mingyan Li, Scott A. Lintelman, Krishna Sampigethaya, Radha Poovendran, David von Oheimb, and Jens-Uwe Busser. Impact of public key enabled applications on the operation and maintenance of commercial airplanes. In *AIAA Aviation Technology Integration and Operations Conference*, 2007.

[147] Vincent A. Orlando and W. H. Harman. Project Report ATC-214: GPS–Squitter capacity analysis. In *MIT Lincoln Laboratory*, May 1994.

[148] Sherman C. Lo, Benjamin Peterson, Dennis Akos, Mitch Narins, Robert Loh, and Per Enge. Alternative position navigation and timing (APNT) based on existing DME and UAT ground signals. In *Proceedings of the ION GNSS Meeting*, Portland, OR, 2011.

[149] Logan Scott. Spoofs, proofs, and jamming: Towards a sound national policy for civil location and time assurance. *Inside GNSS*, pages 42–53, Sept./Oct. 2012.

[150] R. Benjamin Harris. Evaluation, refinement and fusion of software-based pseudorange multipath mitigation techniques. In *Proceedings of the ION GNSS Meeting*, Portland, OR, Sept. 2002.

[151] Elisabeth Bumiller. Navy drone violated washington airspace. *The New York Times*, August 26, 2010.

[152] Northrop Grumman. MQ-8B Fire Scout: Vertical takeoff and landing tactical unmanned aerial vehicle system, 2012. http://www.northropgrumman.com/firescout.

[153] Kristin Quinn. Fire Scout incident called 'learning experience'. *Defense-News*, August 27, 2010.

[154] Chris Anderson. Fred Smith: FedEx wants UAVs. *DIY Drones*, February 12, 2009. http://diydrones.com/profiles/blogs/fred-smith-fedex-wants-uavs.

[155] Jad Mouawad and Christopher Drew. Airline industry at its safest since the dawn of the jet age. *The New York Times*, February 11, 2013.

[156] U.S. Government Accountability Office. Unmanned Aircraft Systems: Measuring progress and addressing potential privacy concerns would facilitate integration into the national airspace system. GAO-12-981, September 18, 2012. http://www.gao.gov/products/GAO-12-981.

[157] U.S. Department of Transportation. FY2010 DOT annual performance report, January 2010. http://www.dot.gov/mission/budget/fy2010-dot-annual-performance-report.

[158] Michael McCaul. DHS abandons oversight of unmanned aerial drones inside us. Press Release, July 18, 2012.

[159] Gary Mortimer. Schiebel S-100 crash kills engineer in South Korea. *sUAS News*, 11 May 2012.

[160] Mitch Narins, Michael Lombardi, Per Enge, Ben Peterson, Sherman Lo, Yu Hsuan Chen, and Dennis Akos. The need for a robust precise time and frequency alternative to global navigation satellite systems. *Journal of Air Traffic Control*, 55(1), 2012.