2014

A Multiple COTS Receiver GNSS Spoof Detector -- Extensions

Peter F. Swaszek
University of Rhode Island, swaszek@uri.edu

Richard J. Hartnett

Citation/Publisher Attribution
Swaszek, Peter F. and Richard J. Hartnett. "A Multiple COTS Receiver GNSS Spoof Detector -- Extensions." *Proceedings of the 2014 International Technical Meeting of The Institute of Navigation, January 27-29, 2014, San Diego, CA.*

A Multiple COTS Receiver GNSS Spoof Detector – Extensions

Peter F. Swaszek
University of Rhode Island
Kingston, RI USA

Richard J. Hartnett
U.S. Coast Guard Academy
New London, CT USA

Biographies

Peter F. Swaszek is a Professor in the Department of Electrical, Computer, and Biomedical Engineering at the University of Rhode Island. His research interests are in statistical signal processing with a focus on digital communications and electronic navigation systems.

Richard J. Hartnett is a Professor of Electrical Engineering at the U.S. Coast Guard Academy, having retired from the USCG as a Captain in 2009. His research interests include efficient digital filtering methods, improved receiver signal processing techniques for electronic navigation systems, and autonomous vehicle design.

Abstract

Spoofing refers to the intentional (and considered malicious) interference to a GNSS user's inputs so as to distort the derived position information. A variety of approaches to detect spoofing have been proposed in the literature. Much of this prior work has focused on the conceptual level with limited analysis of the resulting detection performance, and/or has proposed fundamental redesign of the receiver itself. Little effort has been directed towards using existing, commercial-off-the-shelf (COTS) standalone receiver technology to perform spoof detection.

At ION ITM 2013 these authors proposed a simple spoof detection concept based on the use of multiple COTS receivers and analyzed the performance of several ad hoc detection algorithms from a Neyman-Pearson perspective assuming Gaussian statistics. At ION GNSS+ 2013, by restricting attention to a horizontal platform and assuming an independent measurement error model, we were able to develop the optimum Neyman-Pearson hypothesis test. That paper also included an analysis of performance, yielding closed form expressions for the false alarm and detection probabilities and an optimization of the performance over the locations of the receivers' antennae.

This current works extends the earlier results by considering more realistic statistical models, considers the processing of several sequential outputs from the receivers, and addresses 3-D receiver antennae patterns.

Introduction

GNSS systems are well known to be accurate providers of position and timing information across the globe. As such, they are commonly used to locate and navigate craft in various transportation modes (e.g. land vehicles, boats and ships, and aircraft). Because of high signal availabilities, capable/robust receivers, and well-populated satellite constellations, operators typically believe that the location information provided by their GNSS receiver is correct. Researchers (who are arguably a more skeptical group) often think more about the integrity of location information, and are interested in how a receiver might calculate its measure of integrity, or even how one or more receivers might be used to determine if position information is illegitimate, or spoofed. Here, spoofing refers to intentional (and considered malicious) interference to a GNSS user's inputs so as to distort the derived position information. Depending upon the cargo and/or mission of the transport, calculation of the user's position in the presence of one or more spoofers can provide hazardously misleading information, possibly resulting in disastrous consequences in safety critical applications. One way to detect a spoofing event is to use some sort of ground-based augmentation system; however, such detection methods inherently require additional ground-based infrastructure. A much more attractive approach is to have the user self check against spoofing; essentially a RAIM-like integrity test, but monitoring for spoofing and not satellite faults. A variety of approaches of this second type have been proposed in the literature. Some are based upon including additional signal processing capability in the receiver. While these techniques are valid, they are less appropriate for use in a cockpit where receiver certification is a concern or, perhaps, where the cost of new equipment is a significant issue.

Technical discussions on spoof detection can vary widely depending upon the assumed capabilities and a priori knowledge of the spoofer. In 2003 Warner and Johnston suggested several possible methods to detect a spoofing event at a single GNSS receiver [1]: monitoring the power levels of the GNSS signals (absolute, relative, and across satellites), checking that the observed constellation is correct for the given time (e.g. number of and IDs of the

satellites), testing the accuracy of the clock component, and checking the computed position against that derived from some non-GNSS source (e.g. an IMU). Since then various authors have experimented with spoofing and suggested detectors including correlating the P(Y) code at the RF level [2], looking for vestigial peaks in the correlator outputs [3], comparing to trusted reference signals [4], and using an antennae array to spatially locate and identify signals [5]. Much of this prior work has focused on the conceptual level with limited analysis of the resulting detection performance, and/or has proposed fundamental redesign of the receiver itself. Unfortunately, little work has been directed towards using existing commercial off-the-shelf (COTS) stand-alone receiver technology to perform spoof detection.

Recently, at ION ITM 2013, we examined a simple spoof detection concept based upon the use of multiple COTS receivers and attempted to assess its performance under nominal assumptions on the signal environment [6]. Specifically, the detector monitors GNSS signals using not one, but two or more receivers with their antennae at known relative positions. With no spoofer present, each antenna would receive a unique RF signal consistent with its position in space. Under the assumption that the spoofer is present, and has only one broadcast antenna, these multiple receivers would receive nearly identical spoofer RF signals; the presence of spoofing is thus discernible from the near equivalence of the receivers' receptions. While one could compare these multiple receptions at the RF level, we proposed comparing the position solutions across receivers, declaring a spoofing event if the resulting position solutions are too close to each other as compared to the known relative locations of the antennae. The primary advantage of such an approach is that an implementation of the hypothesis test does not require receiver hardware modification (hence, no recertification is necessary) or even access to software GNSS methods; a separate processor could easily monitor the positions generated by each of the antennae/receivers and decide spoof or no spoof. Our January 2013 work proposed several ad hoc detection algorithms (resulting from different assumptions on the receivers' antennae locations; specifically, known positions, known relative positions with orientation information, and known relative positions without orientation information) and analyzed each detector from a Neyman-Pearson perspective assuming Gaussian statistics.

Since that time we have improved our approach to spoof detection, presenting the results at ION GNSS+ 2013 [7]. Specifically, by restricting the antennae location model to a horizontal platform and assuming that the measurement errors were independent Gaussian variables, we were able to develop the form of the optimum hypothesis test under the Neyman-Pearson criterion; this optimum test (actually a generalized likelihood ratio test to account for unknown platform rotation and position) was shown to examine the separation of and relative locations of the estimated positions. This work included an analysis of performance, yielding closed form expressions for the false alarm and detection probabilities. We were then able to optimize the performance over the locations of the receiver antennae. Of note, if the antennae are restricted to fall within a disk of radius r, we were able to show that the optimum configuration has all of the antenna on the edge of the disk in a symmetric arrangement. Further, we showed that using 3 or 4 antennae/receivers spaced 4σ or more apart (σ being the standard deviation of the measurement error in any direction, typically 1-2 meters for current receivers) yields excellent performance.

The current paper extends this prior work [7] in the following ways:

- Better statistical models – While the choice of independent East and North errors simplified the analysis, it does not match reality; the exact GNSS constellation being viewed results in correlation between the resulting East and North errors (this leads to the use of GDOP as a figure of merit in GNSS performance). This work extends the detection analysis, employing unequal variances and a non-zero correlation for the East and North errors for the case of a symmetric antennae configuration.
- Time sequential processing – Our original approach might be called a "snapshot" method in that it processes just one set of positions to decide spoof versus no spoof. Allowing multiple measurements, we develop a test that looks for similar/equivalent geometric relationships across time. In this paper we report on two ways to combine such data: a "coherent" approach which assumes that the platform does not rotate between measurements and a "non-coherent" approach which allows for rotation.
- Extension to 3D – We extend the problem and solution formulations to allow for antennae at different altitudes (which may be especially appropriate for some platforms).

The Problem

Imagine a configuration of m GPS antennae/receivers, each of which provides a two dimensional position solution based upon its observed RF signals (while latitude and longitude are the nominal coordinates, we will assume that they are converted to East and North in a local reference frame). For simplicity of the resulting analysis, we will parameterize the position of each antenna as a point on the complex plane relative to some fixed origin. Specifically, the k^{th} antenna, $k = 1, \ldots m$, is at position

$d_k = d_{k,r} + jd_{k,i}$ (in this decomposition into real and imaginary components, we will think of the real part as the East component and the imaginary part as the North component of the position). Further, and without loss of generality, we will assume that the origin of our reference frame is such that the centroid of these antennae positions is zero, so that

$$\sum_{k=1}^{m} d_k = 0$$

Our interest is in mounting this array of antennae onto a moving platform; hence, relative to the location of the centroid, the array could have a random orientation with respect to true East/North. Keeping the array horizontal, we model this as an angular rotation by angle θ (in radians) on the complex plane. As such, the position of the k^{th} antenna is now $d_k e^{j\theta}$. For our spoof detector each antenna processes the RF signals it receives, yielding an estimate of its position; this position is a complex number in that same reference frame which we will denote by $x_k = x_{k,r} + jx_{k,i}$. We will assume that the error in this estimate is dominated by additive Gaussian noise, so will employ complex Gaussian distributions when describing the statistics of these positions. In [7] we assumed a white noise component; here we want to take into account the satellite geometry as described by HDOP (i.e. unequal variances on the components along with correlation) and will use a more general model.

We consider two situations, the null hypothesis, H_0, in which no spoofer is present and the alternative hypothesis, H_1, in which a spoofer is present:

H_0: With no spoofer present we assume that each individual antenna is giving an accurate estimate of its actual position. For notation, let b represent the true position of the centroid of the antennae array; including this position offset, the rotation for each antenna, and additive Gaussian noise terms (n_k), we have a model for the position observations of

$$x_k = b + d_k e^{j\theta} + n_k$$

for $k = 1, 2, \ldots m$.

H_1: With a spoofer present we assume that the antennae all receive identical RF signals; hence, all would provide noisy estimates of the same constant position. (We assume that the spoofer takes over all satellite signals; we would have to modify this characterization if only some satellites were spoofed. And with only one radiator, a spoofer can create only one possible position solution). Letting c represent this spoofed position, we have the observation model

$$x_k = c + n_k$$

for $k = 1, 2, \ldots m$.

In [7] we assumed white noise for n_k; here we will keep the Gaussian model and zero mean, but allow the individual real and imaginary parts to have different variances and to be correlated. As a complex random variable

$$n_k \sim \mathcal{CN}(0, \Gamma_k, C_k)$$

Note that the parameters of this model are dependent upon the satellites in view to antenna k. Since we assume that the antennae are nearly co-located, then we assume that the sky view is the same for each antennae, so $\Gamma_k = \Gamma$ and $C_k = C$; in other words, the noise at each antenna has the same statistics. We also assume independence of the noise for different k.

For those not familiar with complex Gaussian variates, consider a pair of jointly Gaussian variables x_r and x_i. Standard notation (listing the two means, two variances, and the correlation coefficient) that is used to describe their statistics is

$$(x_r, x_i) \sim \mathcal{N}\left(\mu_r, \mu_i, \sigma_r^2, \sigma_i^2, \rho\right)$$

If we construct the complex random variable x from these two random variables as

$$x = x_r + jx_i$$

then

$$x \sim \mathcal{CN}(\mu, \Gamma, C)$$

with

$$\mu = \mu_r + j\mu_i$$

$$\Gamma = \sigma_r^2 + \sigma_i^2$$

and

$$C = \sigma_r^2 - \sigma_i^2 + j2\rho\sigma_r\sigma_i$$

Obviously, one can also write expression for the parameters in the reverse direction.

The Test

Our hypothesis test from [7], which assumed white statistics on the measurement noise, is

$$T(x_1, \ldots, x_k) = -\left|\sum_{k=1}^{m} d_k^* x_k\right| \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

While this form is quite simple, its performance does vary as a function of antenna configuration, satellite constellation, and platform orientation. Our goal is to analyze these effects. Before presenting the analysis, we note that the test statistic is independent of translation of the data. This is particularly important is that it means that any common-mode bias in the GNSS positions is ignored by the test.

Let's begin an analysis. For convenience, we write the test as

$$T(x_1, \ldots, x_k) = -|y| \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda$$

with

$$y = \sum_{k=1}^{m} d_k^* x_k = \sum_{k=1}^{m} y_k$$

Consider the individual terms in the summation for y

$$y_k = d_k^* x_k$$

As separate linear functions of independent complex Gaussian random variables (the x_k), the y_k are jointly complex Gaussian and maintain their independence; the sum of these terms, y, is also complex Gaussian. Simplifying the algebra, the distributions are

$$y \sim \mathcal{CN}\left(e^{j\theta}\sum_{k=1}^{m}|d_k|^2, \Gamma\sum_{k=1}^{m}|d_k|^2, C\sum_{k=1}^{m}(d_k^*)^2\right)$$

under H_0 and

$$y \sim \mathcal{CN}\left(0, \Gamma\sum_{k=1}^{m}|d_k|^2, C\sum_{k=1}^{m}(d_k^*)^2\right)$$

under H_1. We note that these expressions depend upon two functions of the antennae locations

$$\sum_{k=1}^{m}|d_k|^2 \quad \text{and} \quad \sum_{k=1}^{m}(d_k^*)^2$$

In [7] the optimality of a circular symmetric antenna configuration was noted. Focusing our attention on such a pattern, let

$$d_k = re^{j\left(\frac{2\pi k}{m}+\phi\right)}$$

for $k = 1, 2, \ldots m$ with ϕ an arbitrary phase shift and r a known radius. After some algebra, the first constant can be seen to reduce to

$$\sum_{k=1}^{m}|d_k|^2 \quad = \quad mr^2$$

The second summation can be shown to equal zero if $m > 2$. The combined result is that for symmetric arrays of $m = 3$ or more antennae, the complex random variable y has distributions

$$y \sim \mathcal{CN}\left(e^{j\theta}md^2, mr^2\Gamma, 0\right)$$

and

$$y \sim \mathcal{CN}\left(0, mr^2\Gamma, 0\right)$$

under H_0 and H_1, respectively. We defer the $m = 2$ case to a separate section below.

To develop the performance expressions, it is convenient to convert the complex variable y back into two separate components, the real part y_r and the imaginary part y_i. The means for each part are just the real and imaginary parts of the complex mean, respectively. Under H_0

$$\mu_{y,r} = mr^2\cos\theta \quad \text{and} \quad \mu_{y,i} = mr^2\sin\theta$$

Since θ is an unknown angle, the center of the distribution on the (y_r, y_i) plane is somewhere on a circle of radius mr^2 about the origin. Under H_1 both means are zero. Since Γ is real and C is zero for both hypotheses, we have equal variances for the two components

$$\sigma_{y,r}^2 = \sigma_{y,i}^2 = \frac{mr^2\Gamma}{2}$$

Finally, since Γ is real and C is zero, we have zero correlation

$$\rho_y = 0$$

In other words, the bivariate distributions are

$$(y_r, y_i) \sim \mathcal{N}\left(mr^2\cos\theta, mr^2\sin\theta, \frac{mr^2\Gamma}{2}, \frac{mr^2\Gamma}{2}, 0\right)$$

and

$$(y_r, y_i) \sim \mathcal{N}\left(0, 0, \frac{mr^2\Gamma}{2}, \frac{mr^2\Gamma}{2}, 0\right)$$

under H_0 and H_1, respectively.

The test itself can also be written in terms of these components

$$\begin{aligned} T(x_1, \ldots, x_k) &= -|y| \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda \\ &= -\sqrt{y_r^2 + y_i^2} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \lambda \end{aligned}$$

Multiplying by -1 reverses the direction of the test

$$T(x_1, \ldots, x_k) = \sqrt{y_r^2 + y_i^2} \underset{\mathrm{H}_1}{\overset{\mathrm{H}_0}{\gtrless}} \lambda$$

In other words, if the observation when viewed on the (y_r, y_i) plane is outside a circle of radius λ, then we decide H_0, no spoofer; if inside the circle, we decide H_1.

Test Performance for $m > 2$

The probability of detection is the probability under H_1 that the test statistic is smaller than the threshold

$$\mathrm{P_d} = \mathrm{Prob}_{\mathrm{H}_1}\left(\sqrt{y_r^2 + y_i^2} < \lambda\right)$$

Since (y_r, y_i) is bivariate Gaussian under H_1 then

$$\mathrm{P_d} = \iint_{\Omega} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(y_r^2 + y_i^2\right)} dy_r dy_i$$

in which, for simplicity, we have introduced the notation

$$\sigma^2 = \frac{mr^2\Gamma}{2}$$

and Ω is the disk about the origin of radius λ. Changing variables to polar coordinates of magnitude s (chosen to avoid using r with two definitions) and phase angle ϕ yields

$$\mathrm{P_d} = \int_0^{2\pi}\int_0^{\lambda} \frac{s}{2\pi\sigma^2} e^{-\frac{s^2}{2\sigma^2}}\, ds d\phi$$

in which we have explicitly described the limits of integration of Ω. Integrating first over ϕ, then over s yields

$$\mathrm{P_d} = 1 - e^{-\frac{\lambda^2}{2\sigma^2}} = 1 - e^{-\frac{\lambda^2}{mr^2\Gamma}}$$

The false alarm of the test is the probability under $\mathrm{H_0}$ that the test statistic is smaller than the threshold

$$\mathrm{P_{fa}} = \mathrm{Prob_{H_0}}\left(\sqrt{y_r^2 + y_i^2} < \lambda\right)$$

Again, (y_r, y_i) is bivariate Gaussian, but with non-zero means, so

$$\mathrm{P_{fa}} = \iint\limits_{\Omega} \frac{1}{2\pi\sigma^2} e^{-\frac{\left(y_r - mr^2\cos\theta\right)^2 + \left(y_i - mr^2\sin\theta\right)^2}{2\sigma^2}} dy_r dy_i$$

(again using the notation σ^2). Changing to polar coordinates yields

$$\mathrm{P_{fa}} = \int_0^{\lambda} \frac{s}{\sigma^2} e^{-\frac{s^2+m^2r^4}{2\sigma^2}} \left[\int_0^{2\pi} \frac{1}{2\pi} e^{\frac{smr^2}{\sigma^2}\cos(\phi-\theta)} d\phi\right] ds$$

Now, the inner integral in brackets can be manipulated by changing variables to $\zeta = \phi - \theta$, using the periodicity of the cosine function to shift the integration limits, and recognizing the definition of the modified Bessel function of the first kind. The result for the false alarm probability is then

$$\mathrm{P_{fa}} = \int_0^{\lambda} \frac{s}{\sigma^2} e^{-\frac{s^2+m^2r^4}{2\sigma^2}} I_0\left(\frac{smr^2}{\sigma^2}\right) ds$$

To simplify this expression, we first change variables to

$$z = \frac{s}{\sigma}$$

so

$$\mathrm{P_{fa}} = \int_0^{\frac{\lambda}{\sigma}} z e^{-\frac{1}{2}\left[z^2+\gamma^2\right]} I_0\left(z\gamma\right) dz$$

with

$$\gamma = \frac{mr^2}{\sigma}$$

This final form can be written in terms of Marcum's Q function [8, pp.344-346]

$$\mathrm{P_{fa}} = 1 - Q\left(\gamma, \frac{\lambda}{\sigma}\right)$$

Substituting for γ and σ

$$\mathrm{P_{fa}} = 1 - Q\left(\sqrt{\frac{2mr^2}{\Gamma}}, \lambda\sqrt{\frac{2}{mr^2\Gamma}}\right)$$

At this point we have expressions for $\mathrm{P_{fa}}$ and $\mathrm{P_d}$ in terms of the system parameters of number of antennae, m, spacing of the antennae, r, and the variance of the position error, Γ. We can invert the $\mathrm{P_d}$ expression for the threshold λ

$$\lambda = \sqrt{-mr^2\Gamma\ln\left(1 - \mathrm{P_d}\right)}$$

Inserting this result into the expression for $\mathrm{P_{fa}}$, we have

$$\mathrm{P_{fa}} = 1 - Q\left(\sqrt{\frac{2mr^2}{\Gamma}}, \sqrt{-2\ln\left(1 - \mathrm{P_d}\right)}\right)$$

We acknowledge that some might think that this expression is backwards, that it is more usual in hypothesis testing to write the detection probability as a function of the false alarm probability. However, the utility of this closed-form expression is that for a fixed $\mathrm{P_d}$ and noise variance parameter Γ, the known monotonically of Marcum's Q function in its arguments implies that our test's performance improves with increasing r and m.

As examples, Figure 1 shows a typical ROC for $m = 4$ and various spacings of the antennae (r) while Figure 2 shows a ROC for $r = 1.5$ and various numbers of antennae (m). In both of these examples we have, for convenience, set Γ to unity. These figures demonstrate the monotonicity of the performance with increasing m and/or r.

Including HDOP

The expressions for the false alarm and detection probabilities developed above depend upon the parameters of the antennae array (through m and r) and the underlying position inaccuracies through Γ. The parameter Γ can also be written as

$$\Gamma = \sigma^2_{\mathrm{UERE}}\mathrm{HDOP}^2$$

in which σ_{UERE} is the user equivalent range error and HDOP is the horizontal dilution of precision. With this relationship, the performance of our spoof detector is

$$\mathrm{P_{fa}} = 1 - Q\left(\sqrt{\frac{2mr^2}{\sigma^2_{\mathrm{UERE}}\mathrm{HDOP}^2}}, \sqrt{-2\ln\left(1 - \mathrm{P_d}\right)}\right)$$

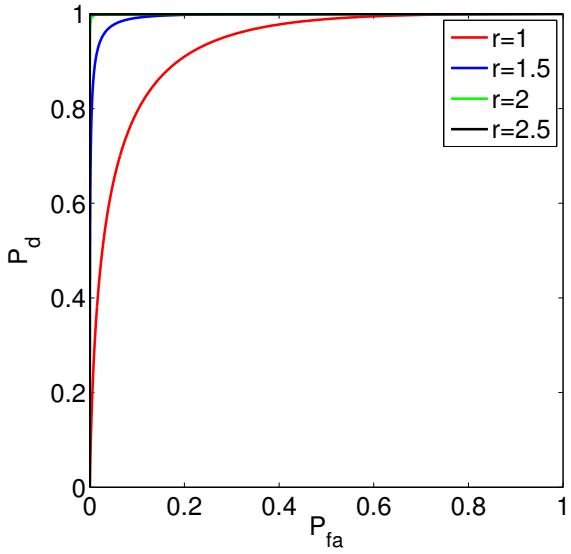

Figure 1: ROC for $m = 4$ and $\Gamma = 1$ with different values for r (in meters).

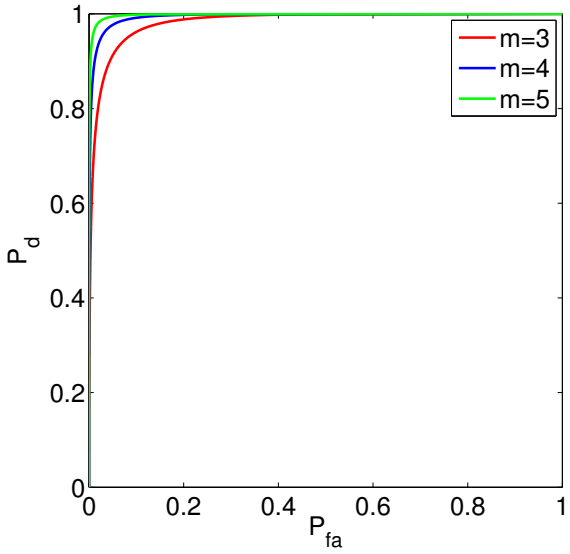

Figure 2: ROC for $d = 1.5$ meters and $\Gamma = 1$ with various m.

For example, HDOP $= 1$ and $\sigma_{\text{UERE}} = 1$ would yield the $\Gamma = 1$ employed for Figures 1 and 2. Realistic values for these parameters are HDOP of unity and σ_{UERE} of approximately 4 [9]. An obvious metric for performance (essentially a signal to noise ratio as a function of the parameters that we control, m and r) is the square of the first argument of the Marcum Q-function

$$\text{metric}(m, r) = \frac{2mr^2}{\sigma^2_{\text{UERE}}\text{HDOP}^2}$$

Clearly we are interested in very small P_{fa} and large P_{d}, so need to look more clearly at the performance expression to understand the roles of m and r. Toward this end, imagine that $m = 4$ and $\Gamma = 16$ (using the realistic HDOP and σ_{UERE} values), and that the threshold, λ, is set so that P_{d} is either 0.9, 0.99, or 0.999. Figure 3 shows the corresponding P_{fa} for a range of r values. For example, we observe that for $r = 10$ meters (4 antenna on a square with side lengths of $10\sqrt{2}$ meters) we achieve $\text{P}_{\text{fa}} \approx 10^{-5}$ and $\text{P}_{\text{d}} \approx 0.99$! The P_{fa} and P_{d} expressions depend upon the antenna array through the product mr^2 only; in other words, an array of 4 antenna on a circle of radius 10 meters has a score of $mr^2 = 4 \cdot 10^2 = 400$, the same as a 5 antenna array array with r just below 9 or 3 antennae with $d \approx 11.5$. To understand the scale of these values for platforms of interest, Figure 4 shows the 3 and 4 antennae locations on a Boeing 757 airplane.

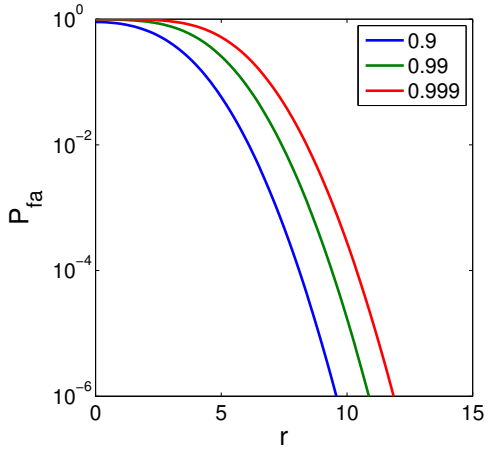

Figure 3: P_{fa} versus r for $m = 4$, $\Gamma = 16$, and selected values of P_{d}.

The Case of 2 Antennae

Some long and narrow platforms, track trailers or ships, might be better suited to a two antenna solution; for example, $r = 7$ as shown in Figure 5. Assuming that $d_1 = re^{j\phi}$ and $d_2 = -re^{j\phi}$, we can drop the common multiplication by $re^{j\phi}$ and simplify the test to

$$T(x_1, x_2) = -|x_1 - x_2| \underset{\text{H}_0}{\overset{\text{H}_1}{\gtrless}} \lambda$$

Concentrating on the term within the absolute value symbols

$$y = x_1 - x_2$$

we have the following characterizations:

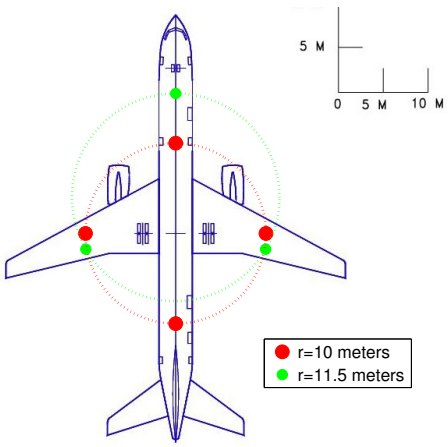

Figure 4: Antennae layout on a 757 aircraft.

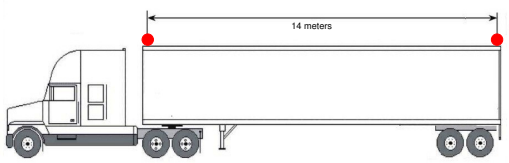

Figure 5: Antennae layout on tracker trailer.

- Under H_0, this variable is

$$\begin{aligned} y &= x_1 - x_2 \\ &= 2re^{j\theta} + n_1 - n_2 \end{aligned}$$

- Under H_1, it is

$$\begin{aligned} y &= x_1 - x_2 \\ &= n_1 - n_2 \end{aligned}$$

Since both n_1 and $-n_2$ (the negative of n_2) are complex Gaussian then so is their sum

$$n_1 - n_2 \sim \mathcal{CN}(0, 2\Gamma, 2C)$$

To facilitate a performance analysis in this case, it is convenient to again return to a bivariate description of y under the two hypotheses. Under H_0 we have means

$$\mu_{y,r} = 2r\cos\theta \quad \text{and} \quad \mu_{y,i} = 2d\sin\theta$$

while under H_1 both means are zero. As Γ is real and C is complex, we have variances

$$\sigma^2_{y,r} = 2\sigma^2_r \quad \text{and} \quad \sigma^2_{y,i} = 2\sigma^2_i$$

under both hypotheses. As n_1 and $-n_2$ are independent and identically distributed, their sum has the same correlation coefficient of ρ under both hypotheses. Summarizing, under H_0

$$(y_r, y_i) \sim \mathcal{N}\left(2d\cos\theta, 2d\sin\theta, 2\sigma^2_r, 2\sigma^2_i, \rho\right)$$

and under H_1

$$(y_r, y_i) \sim \mathcal{N}\left(0, 0, 2\sigma^2_r, 2\sigma^2_i, \rho\right)$$

At this point we want to calculate the probabilities of false alarm and detection which correspond to integrals within a disk of radius λ on the complex plane corresponding to y. Figure 6 shows such a situation with $\lambda = 1$ (the red, dashed circle is the decision boundary), the elliptical contours of the pdf under H_1 are shown in green, the possible location of the mean under H_0 is shown as a blue, dashed circle with radius of $\sqrt{5}$, and three typical contours of the pdf under H_0 are shown in blue. Under both hypotheses, this figure assumes that $\sigma_r = 2\sigma_i$ and that the correlation coefficient is $\rho = 0.5$. It is very clear from this figure that while the probability of detection is fixed (integrating the green pdf inside the red circle is merely mathematically ugly), the probability of false alarm (integrating a blue pdf inside the red circle) depends upon the relative location of the center of the ellipse (equivalently, the rotation angle θ).

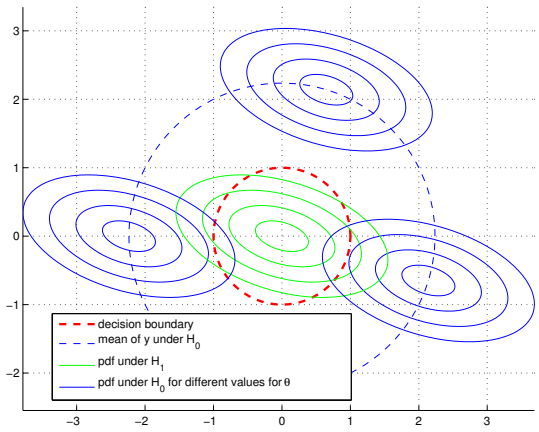

Figure 6: Distributions of y under H_0 and H_1 and the integration boundary.

To facilitate doing these computations, it is convenient to rotate the data by angle

$$\zeta = -\frac{1}{2}\tan^{-1}\frac{2\rho\sigma_r\sigma_i}{\sigma^2_r - \sigma^2_i}$$

so that the major axis of the ellipse is parallel to the horizontal axis. Let z represent this new pair of variables.

With this rotation, the probability density function of z under the two hypotheses is still bivariate Gaussian. Specifically, under H_0

$$(z_r, z_i) \sim \mathcal{N}\left(2r\cos\psi, 2r\sin\psi, \sigma_1^2, \sigma_1^2, 0\right)$$

with ψ an arbitrary angle and

$$\sigma_1^2 = \sigma_r^2 + \sigma_i^2 + \sqrt{(\sigma_r^2 - \sigma_i^2)^2 + 4\rho^2\sigma_r^2\sigma_i^2}$$

$$\sigma_1^2 = \sigma_r^2 + \sigma_i^2 - \sqrt{(\sigma_r^2 - \sigma_i^2)^2 + 4\rho^2\sigma_r^2\sigma_i^2}$$

(Note that the signs in front of the square roots assume that $\sigma_r > \sigma_i$; if not, they should be reversed.) Under H_1

$$(z_r, z_i) \sim \mathcal{N}\left(0, 0, \sigma_1^2, \sigma_1^2, 0\right)$$

With this change of variables, the equivalent view in $\mathbf{z}$ is shown in Figure 7. In this figure we also show the best case and worst case locations for the distributions under H_0 (best at the top which would yield the smallest $\mathrm{P_{fa}}$, worst at the right). At this point we could attempt the integrals inside the red circle to yield the false alarm and detection probabilities. Unfortunately, if $\sigma_1^2 \neq \sigma_2^2$ no closed form solution exists. One could, of course, perform the integration numerically (probably easiest in polar coordinates).

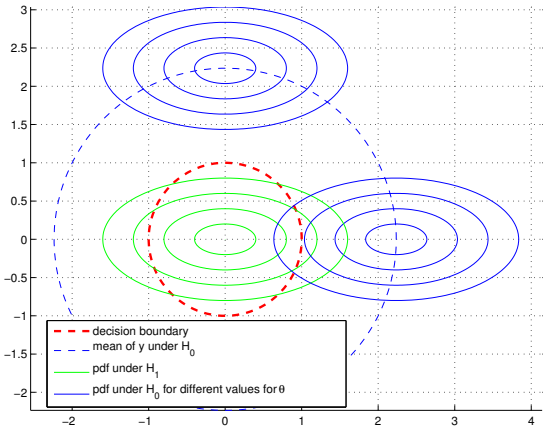

Figure 7: Distributions of z under H_0 and H_1 and the integration boundary.

An alternative, pursued here, is to bound the expressions. Examining the figures, we have

$$\mathrm{P_d} = \iint_{\Omega} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{z_r^2}{\sigma_1^2} + \frac{z_i^2}{\sigma_2^2}\right]} dz_r dz_i$$

and in the worst case

$$\mathrm{P_{fa}} = \iint_{\Omega} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(z_r-2r)^2}{\sigma_1^2} + \frac{z_i^2}{\sigma_2^2}\right]} dz_r dz_i$$

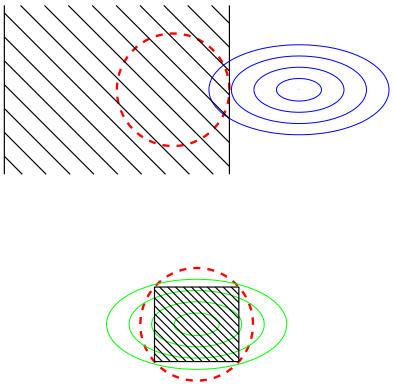

Figure 8: Bounding region for $\mathrm{P_{fa}}$ (top) and $\mathrm{P_d}$ (bottom). (Please ignore the vertical line on the left of the top figure; it's a MatLab artifact that I cannot get rid of!)

One simple upper bound for $\mathrm{P_{fa}}$ is to use a half-plane to encapsulate the spoofing decision region, Ω, as shown in the top subfigure of Figure 8. The result is

$$\mathrm{P_{fa}} \leq Q\left(\frac{2r - \lambda}{\sigma_1}\right)$$

Other bounds are obviously possible.

The detection probability can lower bounded with an inscribed box as shown in the lower subfigure of Figure 8. In general, we could use a box with corners $(\pm a, \pm b)$ where a and b satisfy $a^2 + b^2 = \lambda^2$ and maximize the result over a and b. Mathematically

$$\mathrm{P_d} \geq \left[1 - 2Q\left(\frac{a}{\sigma_1}\right)\right]\left[1 - 2Q\left(\frac{b}{\sigma_2}\right)\right]$$

At this point we can try several approaches. One would be to use the half-plane bound for $\mathrm{P_{fa}}$ to evaluate λ

$$\lambda = 2r - \sigma_1 Q^{-1}\left(\mathrm{P_{fa}}\right)$$

use $a^2 + b^2 = \lambda^2$ to solve for b in terms of λ

$$b = \sqrt{\lambda^2 - a^2}$$

and then optimize $\mathrm{P_d}$ (numerically?) over a and b. A second approach is to choose a and b to satisfy

$$\frac{a}{\sigma_1} = \frac{b}{\sigma_2}$$

so that the bound simplifies. The result is

$$\mathrm{P_d} \geq \left[1 - 2Q\left(\frac{\lambda}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\right]^2$$

Multi-Sample Tests

Imagine collecting position samples as the platform is moving. Indexing time samples by the variable n, let $x_k^{(n)}$ represent the n^{th} time sample at antenna k, $k = 1, 2, \ldots m$ and $n = 1, 2, \ldots N$. For the discussion below we assume that the samples are sufficiently far apart in time so that the measurement errors on each antenna are independent complex Gaussian variates.

We envision two different situation depending upon whether or not the platform rotates during the measurement period:

- No rotation — modifying the development of the test for independent spatial errors in [7] (the generalized likelihood ratio test for a single unknown angle of rotation) results in the test

$$T(x_1^{(1)}, \ldots, x_m^{(N)}) = -\left| \sum_{n=1}^{N} \sum_{k=1}^{m} d_k^* x_k \right|$$

  Combining the spatial matched filter outputs before the absolute value is a kind of *coherent* processing in which we exploit the fact that the random orientation is constant over the entire time.

- Unknown rotations — if we allow the rotation to be different for each data snapshot, then the generalized likelihood ratio test should optimize over each angle separately. Doing so yields the test

$$T(x_1^{(1)}, \ldots, x_m^{(N)}) = -\sum_{n=1}^{N} \left| \sum_{k=1}^{m} d_k^* x_k \right|$$

  Combining the spatial matched filter outputs after the absolute value is a *noncoherent* operation.

Recall that the summation over k is invariant to the actual location of the vessel; hence, the statistical distributions of each summation over k for different values of n are identical in the coherent case and identical except for a rotation in the non coherent case. Assuming that the $m > 2$ antennae are uniformly spread on a circle of radius d, then each sum over m is an independent complex Gaussian variate under both hypotheses. Specifically, for

$$y^{(n)} = \sum_{k=1}^{m} d_k^* x_k^{(n)}$$

then

$$y^{(n)} \sim \mathcal{CN}\left(e^{j\theta} m r^2, m d^2 \Gamma, 0\right)$$

and

$$y^{(n)} \sim \mathcal{CN}\left(0, m r^2 \Gamma, 0\right)$$

under H_0 and H_1, respectively. For the coherent test, we sum N such independent variable, so the change in the distributions is to increase the means and variances by a factor of N:

$$\sum_{n=1}^{N} y^{(n)} \sim \mathcal{CN}\left(e^{j\theta} N m r^2, N m r^2 \Gamma, 0\right)$$

and

$$\sum_{n=1}^{N} y^{(n)} \sim \mathcal{CN}\left(0, N m r^2 \Gamma, 0\right)$$

In this case, the performance results of the single snapshot analysis directly extend with the addition of the factor of N

$$\mathrm{P}_{\mathrm{d},coherent} = 1 - e^{-\frac{\lambda^2}{2 N m r^2 \Gamma}}$$

and

$$\mathrm{P}_{\mathrm{fa},coherent} = 1 - Q\left(\sqrt{\frac{2Nmr^2}{\Gamma}}, \lambda\sqrt{\frac{2}{Nmr^2\Gamma}}\right)$$

or

$$\mathrm{P}_{\mathrm{fa}} = 1 - Q\left(\sqrt{\frac{2Nmr^2}{\Gamma}}, \sqrt{-2\ln\left(1 - \mathrm{P}_{\mathrm{d}}\right)}\right)$$

From this last expression, it is clear that the metric has increased by a factor of N

$$\begin{aligned} \mathrm{metric}_{coherent}(m, r, N) &= \frac{2Nmr^2}{\sigma^2_{\mathrm{UERE}} \mathrm{HDOP}^2} \\ &= N \mathrm{metric}(m, r) \end{aligned}$$

and we observe that we can trade time (repeated sampling) for either number of receivers (m) or spacing (r). For example, employing 4 sets of time samples allows us to half the radius of the antennae array.

The analysis for the non-coherent case is slightly harder. Since we take the absolute value first, before summing over n, we have a set of N independent Rician random variable (Rayleigh under H_0). Unfortunately, there appear to be no closed form results for computing P_{fa} and P_{d}. However, we conjecture that the metric increases by a square root of N

$$\begin{aligned} \mathrm{metric}_{non-coherent}(m, r, N) &\approx \frac{2\sqrt{N}mr^2}{\sigma^2_{\mathrm{UERE}} \mathrm{HDOP}^2} \\ &\approx \sqrt{N} \mathrm{metric}(m, r) \end{aligned}$$

Extending to 3 Dimensions

For some platforms of interest (e.g. aircraft), limiting the problem statement to a horizontal placement of the antennae seems appropriate. For others, however, such as ships, the option to place an antenna on the top of a superstructure might be available and might provide some

improvement to performance. To consider this, we add an altitude measurement to our situation, maintaining the option to rotate the array in the horizontal plane by angle θ. Letting g_k and z_k be the actual and measured altitude for antenna k, respectively, the first question is how to extend the test statistic. Assuming that we define our origin so that we have a vertical centroid condition

$$\sum_{k=1}^{m} g_k = 0$$

then extending the white measurement error analysis from [7] yields the test

$$T(x_1, z_1, \ldots, x_k, z_k) = \left| \sum_{k=1}^{m} d_k^* x_k \right| + \delta \sum_{k=1}^{m} g_k z_k \underset{\mathrm{H}_1}{\overset{\mathrm{H}_0}{\gtrless}} \lambda$$

where we have added the scalar δ

$$\delta = \frac{\sigma^2}{\sigma_z^2}$$

(the ratio of the horizontal position error variance, σ^2, to the vertical error variance, σ_z^2) to take into account that vertical accuracy is usually worse than horizontal accuracy. Clearly an analysis of performance involves the combination of a non-central chi distribution (from the horizontal term) and a Gaussian distribution (from the vertical term). Because of this, we do not expect to be able to find closed form results, but expect that bounds might be possible.

To continue the discussion, consider an example of a platform with 2 antennae mounted at different heights

$$d_1 = -r, g_1 = -\frac{h}{2}$$

and

$$d_2 = +r, g_2 = +\frac{h}{2}$$

where $2r$ and h are the length and height of the platform, respectively. With this choice, the test statistic is

$$T = |-rx_1 + rx_2| + \delta \left(-\frac{h}{2} z_1 + \frac{h}{2} z_2 \right)$$

Without loss of generality, we can modify the coefficients and use the test

$$\begin{aligned} T &= |x_2 - x_1| + \frac{\delta h}{2r} (z_2 - z_1) \\ &= |y_h| + y_v \end{aligned}$$

To analyze this test we need statistical models for the measurements. For simplicity, we characterize y_h and y_v separately. Invoking the white assumption:

- Under H_0 we have

$$x_1 \sim \mathcal{CN}\left(-re^{j\theta}, 2\sigma^2, 0\right) \qquad z_1 \sim \mathcal{N}\left(-\frac{h}{2}, \sigma_z^2\right)$$

$$x_2 \sim \mathcal{CN}\left(re^{j\theta}, 2\sigma^2, 0\right) \qquad z_2 \sim \mathcal{N}\left(\frac{h}{2}, \sigma_z^2\right)$$

- Under H_1 we have

$$x_1 \sim \mathcal{CN}\left(0, 2\sigma^2, 0\right) \qquad z_1 \sim \mathcal{N}\left(0, \sigma_z^2\right)$$

$$x_2 \sim \mathcal{CN}\left(0, 2\sigma^2, 0\right) \qquad z_2 \sim \mathcal{N}\left(0, \sigma_z^2\right)$$

all mutually independent. The distributions of the terms in the test statistic are

- Under H_0

$$y_h \sim \mathcal{CN}\left(2re^{j\theta}, 4\sigma^2, 0\right) \qquad y_v \sim \mathcal{N}\left(h, 2\left(\frac{\delta h}{2r}\right)^2 \sigma_z^2\right)$$

  With the non-zero mean of y_h, the distribution of $|y_h|$ is non-central chi.

- Under H_1

$$y_h \sim \mathcal{CN}\left(0, 4\sigma^2, 0\right) \qquad y_v \sim \mathcal{N}\left(0, 2\left(\frac{\delta h}{2r}\right)^2 \sigma_z^2\right)$$

  In this case, the distribution of $|y_h|$ is central chi.

The next question is to find the distribution of T under the two hypotheses. Unfortunately, even in the independent case, the convolution of the Gaussian and non-central chi densities does not appear to lead to a closed form expression. Instead, we resorted to simulation.

1. We selected system parameters of $r = 10$, $h = 4$, $\sigma = 4$, and $\sigma_z = 8$ ($\delta = 0.25$)
2. We generated $200,000$ sets of the 6 Gaussian random variables needed under H_0
3. We computed T for each trial; note that without loss of generality we can set $\theta = 0$
4. We repeated steps 2 and 3 for H_1 (the independence of the trials allows us to better describe the results)
5. We sorted the values of T, counted them to estimate $\mathrm{P_{fa}}$ and $\mathrm{P_d}$, and plotted the resulting ROC

For comparison, we also simulated a second platform with $h = 0$ (and ignored the vertical component, i.e. $T = |y_h|$). Figure 9 shows portions of the two ROCs for comparison. Clearly the vertical separation helps. At this time we are unable to characterize how much gain a specific separation yields.

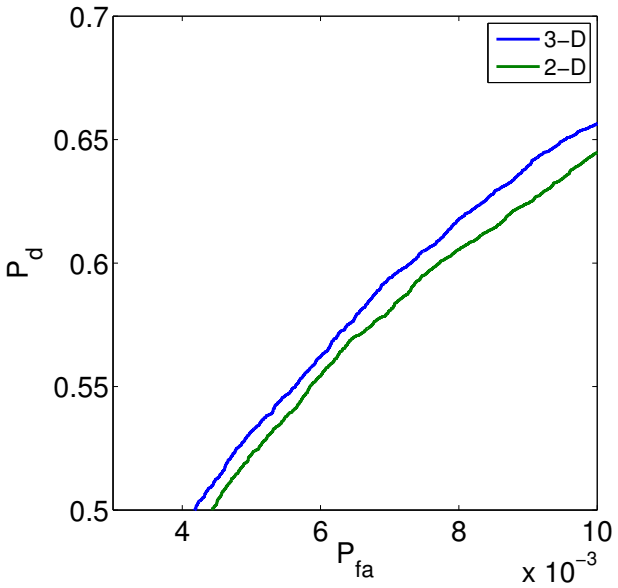

Figure 9: Sample ROCs with and without vertical separation.

References

[1] J. S. Warner and R. G. Johnston, "GPS spoofing countermeasures," *Homeland Security Jour.*, Dec. 2003.

[2] M. L. Psiaki, B. W. OHanlon, J. A. Bhatti,. D. P. Shepard, and T. E. Humphreys, "Civilian GPS spoofing detection based on dual-receiver correlation of military signals," *Proc. ION GNSS*, Portland, OR, Sept. 2011.

[3] K. D. Wesson, D. P Shepard, J. A. Bhatti, and T. E. Humphreys, "An evaluation of the vestigial signal defense for civil GPS anti-spoofing," *Proc. ION GNSS*, Portland, OR, Sept. 2011.

[4] B. M. Ledvina, W. J. Bencze, B. Galusha, and I. Miller, "An in-line anti-spoofing device for legacy civil GPS receivers," *Proc. ION ITM*, San Diego, CA, Jan. 2010.

[5] S. Daneshmand, A. Jafarnia-Jahromi, A. Broumandon, and G. Lachapelle, "A low-complexity GPS anti-spoofing method using a multi-antenna array," *Proc. ION GNSS 2012*, Nashville, TN, Sept. 2012.

[6] P. F. Swaszek, R. J. Hartnett, M. V. Kempe and G. W. Johnson, "Analysis of a simple, multiple receiver GPS spoof detector," *Proc. ION ITM*, San Diego, CA, Jan. 2013.

[7] P. F. Swaszek and R. J. Hartnett, "Spoof detection using multiple COTS receivers in safety critical applications," *Proc. ION GNSS+*, Sept. 2013, Nashville, TN.

[8] H. L. Van Trees, **Detection, Estimation, and Modulation Theory, Part I**, New York: Wiley, 1968.

[9] **Global Positioning System Standard Positioning Service Performance Standard**, 4^{th} Ed., Sept. 2008.