Robust GNSS Spoof Detection using Direction of Arrival: Methods and Practice

Sherman Lo, *Stanford University*, Yu Hsuan Chen, *Stanford University*, Hridayangam Jain, *Stanford University*, Per Enge, *Stanford University*

BIOGRAPHY (IES)

Sherman Lo is a senior research engineer at the Stanford GPS Laboratory. He received his Ph.D. in Aeronautics and Astronautics from Stanford University in 2002. He has and continues to work on navigation robustness and safety, often supporting the FAA. He has conducted research on Loran, alternative navigation, SBAS, ARAIM, GNSS for railways and automobile. He also works on spoof and interference mitigation for navigation. He has published over 100 research papers and articles.

Yu-Hsuan Chen is a research engineer at the Stanford GPS Laboratory. He received his Ph.D. in Electrical Engineering from National Cheng Kung University, Taiwan in 2011.

Hridayangam Jain is an undergraduate majoring in Aeronautics and Astronautics at Stanford University

Per Enge is a professor of Aeronautics and Astronautics at Stanford University, where he is the Vance D. and Arlene C. Coffman Professor in the School of Engineering. He directs the Stanford GPS Laboratory, which develops satellite navigation systems. He has been involved in the development of the Federal Aviation Administration's GPS Wide Area Augmentation System (WAAS) and Local Area Augmentation System (LAAS).

ABSTRACT

GNSS spoofing is a growing concern due to the increasing use of GNSS in safety and economically important applications. The foremost task to manage the GNSS spoofing threat is detection. Indeed, many potential means and measurements have been proposed for spoof detection. However, there exist no panaceas. Rather, we should rely on many different detection means to provide robust detection. One measurement useful for detection is direction of arrival (DOA), which we can get from many sources such as array and dual polarization antenna (DPA). Even with accurate DOA measurements, good monitor tests are needed to make an accurate decision or determination of spoofing.

This paper develops and examines monitor tests suitable for DOA based spoof detection. Developed tests aim to have provable and conservative performance levels making them suitable for safety of life applications. Several tests are developed including signal pair comparisons based on statistical hypothesis tests, Bayesian estimates as well as comprehensive all satellites in view comparisons. On-air spoofing results from a DPA are used to validate performance. The developed techniques must detect simple cases such as when all examined signals are from a single spoofer ("all spoof" case) and challenging cases where there is a mix of genuine signals and spoofed signals ("mixed spoof" case). Another consideration is that measurements will typically be relative DOA as platform attitude may or may not be known. Solving the relative DOA problem also captures the simpler case where we have some knowledge of platform attitude (i.e. absolute DOA).

INTRODUCTION

The ubiquitous use of GNSS in many popular applications makes it an attractive target for attack such as deliberate interference and spoofing. The location based apps and games such as Uber and Pokémon Go has only increased interest and motivation in GNSS spoofing [1][2]. Additionally, the open source GNSS simulation code can be easily found meaning access to GNSS spoofing capability is readily available at lower and lower costs (< $1000). There are many potential means to detect spoofing however there are no silver bullets. Rather a combination of methods will likely need to be employed and hence several should be developed. One good method is to use the direction of arrival (DOA) to check the authenticity of the arriving GNSS signals. It is a useful compliment to many other commonly considered spoof detection techniques as its measurements are independent of those used by those techniques. Detection can be made because genuine satellite signals

arrive from many directions while a spoofer will generally employ one transmitter and hence all its signals will arrive from the same direction. If the signals are arriving from DOAs or have relative DOAs that are consistent with their actual locations in the sky, then we can generally assume they are from the genuine satellites. There are two challenges to this problem – first getting DOA and second developing a robust consistency check to decide if spoofing does exist. There are several means of getting DOA measurements suitable for spoof detection, as will be discussed later. This paper focuses on the latter issue. It develops and examines different tests that can be used as checks and testing them with data from on-air tests.

This paper first describes practical means of getting DOA for GNSS spoof detection. It then develops and overviews different approaches to DOA-based spoof detection test. It focuses the azimuthal direction as some means only provide good azimuth estimates. The performance of selected tests is analyzed. These tests are examined using experimental data from a live spoof exercise. Results for several key scenarios – nominal (no spoofing), all spoofed signals, and mixed genuine and spoofed signals are examined. The key goals of this paper are to 1) develop formal methods for using GNSS signal direction of arrival for spoof detection, 2) quantify the performance of the methods in terms of false alerts and missed detection and 3) demonstrated performance with on-air spoofing tests

BACKGROUND

Direction of arrival

There are many ways to make a direction of arrival measurement. Typically, these are relative DOA unless the antenna orientation relative to an earth fixed frame is known which can yield absolute DOA. The most commonly considered ways for a GNSS system to measure signal direction of arrival is with an array or controlled reception array antenna (CRPA). A CRPA can use carrier phase differences measured by individual antennas to provide information such as DOA for spoof detection [3][4][5][6]. Even two elements antennas can provide phase difference information to infer relative direction of arrivals of the various received satellite signal [3]. More elements can improve performance both in terms of resolution and mitigating the number of incoming interference sources. Four and seven elements CRPAs are commonly found.

There are several ways to use a CRPA to determine the direction of an incoming interference signal. One way is to use its null. If the incoming signal is above the noise floor, noise minimization algorithms can direct nulls in the direction [7]. Hence, the null direction indicates the direction of the signal. Another method is to use its beam and this can be done even without high incoming signal power. A CRPA can automatically direction find GNSS signals with a good choice of beam steering algorithm. Least mean squared (LMS) based algorithms form beams based on the best match of a reference signal, usually the desired GNSS signal. A LMS-based CRPA implementation will then form beams towards the strongest directions of that reference signal [8]. In the case of having a line of sight (LOS) and non-LOS (NLOS) or multipath signal, multiple beams will be formed provide there are enough elements [9][10]. Another method is to use carrier phase differences between the antenna elements [3].

A recent policy change has made CRPAs worth considering for civil spoof detection and mitigation applications. Due to its military applications, there are institutional limitations on civil use of array antenna technology from International Traffic in Arms Regulations (ITAR). Until recently, it was completely not allowed. Today, array antennas with three or less elements are no longer restricted. Fortunately, CRPA research for civil applications has been conducted for the last two decades providing useful resources for such a development [10][11][12].

Figure 1. Stanford PCB Dual Polarization Antenna

While DOA is generally thought of as requiring a multiple antenna set up, it can be achieved with a single antenna. The Stanford dual polarization antenna (DPA) was developed to mitigate spoofing by determining DOA [13]. A printed circuit board (PCB) version is shown in Figure 1. The advantages of the DPA is that it can have a small form factor, the same as a single patch, can be built using commercial off the shelf (COTS) components, and has been demonstrated to determine relative azimuth direction of arrival as well as rough elevation angle [14][15].

The Stanford DPA is capable of receiving both right and left hand circularly polarized (RHCP and LHCP, respectively) signals. The concept, shown in Figure 2, uses an insight about the relationship of these signals to find the direction of arrival. Regardless of initial polarization, when signals impact the antenna ground plane before entering the antenna, these signals become linearly polarized. A linearly polarized signal has equal RHCP and LHCP component. Hence the antenna can differentiate signals coming from low elevation, which is mostly linearly polarized as they impacts the ground plane, from RHCP GNSS signals. Just as important, the phase relationship between these two components of the linearly polarized signal depends on the azimuthal direction of arrival of the signal. So any spoofing signal that comes in from the level of the vehicle (automobile, aircraft) or below will impact on the ground plane first and an azimuth can be reasonably estimated.

The Stanford DPA has onboard electronics that processes these signals to estimate the azimuthal DOA. For the Stanford DPA, the azimuth φ is related to measured phase offset, ψ, between the RHCP and LHCP that results a null by Equation (1) [15]. φ_0 is the azimuth related to the orientation of an antenna feed. As seen the equation, the azimuth estimate has a 180 degree (°) ambiguity. This occurs because rotating DOA of the incoming signal rotates the LHCP and RHCP signals in opposite directions resulting in their phase difference rotates twice as fast as the DOA. An example measurement with genuine GPS satellites shows the variation of carrier to noise ratio (C/No) as the DPA rotates through different phase offsets is reproduced in Figure 3. Azimuth is then derived from the offset resulting in the null. Importantly, it shows that even with a RHCP signal (the genuine GPS signal), there will be significant linearly polarized component at elevation angles below about 60°. Hence, there are many scenarios that a DPA may provide useful azimuth measurements for spoof detection. For the analysis in this paper, we do not consider the 180° ambiguity so that the analysis is applicable to both CRPA and DPA azimuth. Handling the ambiguity will be covered in a future paper.

$$\psi = 2\left(\varphi - \varphi_0 + \varepsilon_\varphi\right) + 90° \quad (1)$$

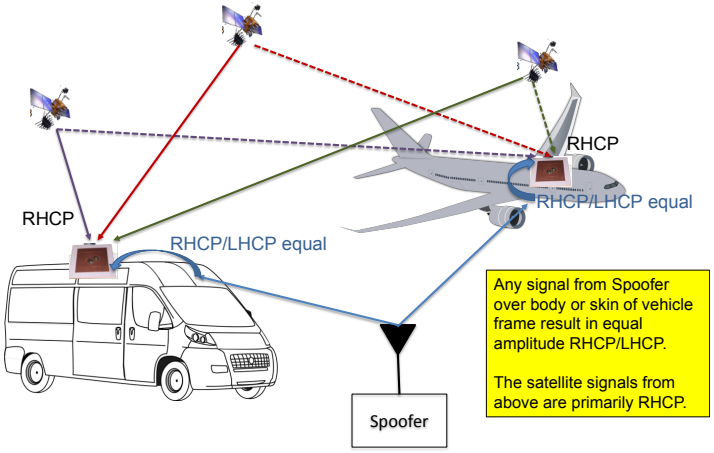


Figure 2. Dual Polarization Antenna Concept for Spoof Determination

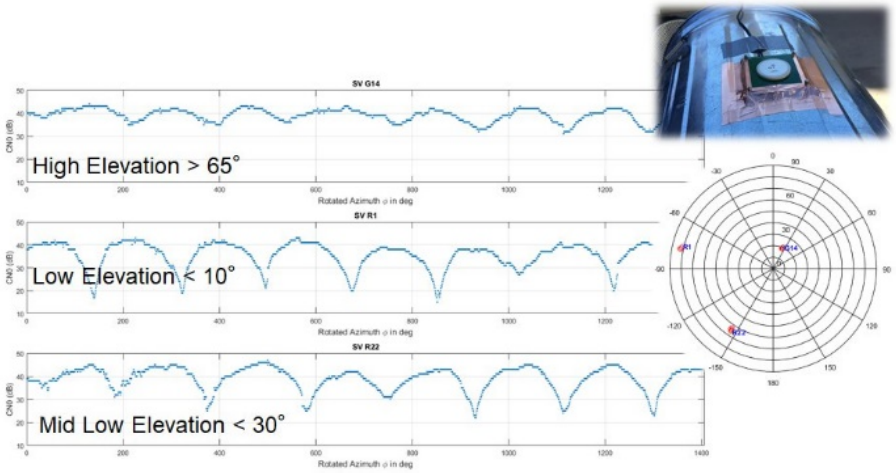


Figure 3. C/No vs. Rotation of LHCP component for different satellites: high (top), low (bottom), very low (middle). At low and very low elevation, rotation of LHCP component to the correct angle can significantly cancel out RHCP, ~ 180 second to complete all rotations [15]

There are some differences between the DPA-derived DOA and that from a CRPA with three or more elements. First, elevation angle cannot be accurately determined. Hence, only azimuth is used. Second, as seen previously, our DPA cannot estimate accurately azimuth from high elevation signals. Finally, the azimuth estimate has a 180° phase ambiguity. These are importance characteristics to remember as DPA-derived DOAs will be used to assess the developed tests.

Spoof detection methodology

Regardless of the source of the measured DOA (CRPA or DPA or otherwise), one will need detection tests to determine, to a targeted confidence level, if spoofing is present. For safety of life applications, it is important to have reliable and conservative tests for using DOA for spoof detection. Robustness is important as very low probability of false alerts (P_{fa}) is needed to limit availability loss and for users to trust the system. Additionally, safety applications require low probability of missed detection (P_{md}). So while we may be able to look at a plot of satellite azimuth (as seen in Figure 7) and see inconsistencies, a formal, conservative and assessable method is needed to accurately quantify detection probabilities as well as P_{md} and P_{fa}. The techniques must detect simple cases such as when all examined signals are from a single spoofer ("all spoof" case) and challenging cases where we have a mix of genuine signals and spoofed signals ("mixed spoof" case). Another consideration is that our measurements will typically be relative DOA as platform attitude may or may not be known or not known. Solving the relative DOA problem also captures the simpler case where we have some knowledge of platform attitude (i.e. absolute DOA). The tests assume only one dimensional DOA – specifically azimuth. Hence, for the analysis in the paper, DOA and azimuth will be used synonymously. However, the test can be used for other dimensions (elevation) or extended to multiple dimensions.

There are many possibilities for these tests. We start by examining individual comparisons, on a satellite pair basis, or and then move onto the collectively by using all satellites available. Reference [4] describes a general test for multiple satellites using angle differences measured by two antennas are inputs to a cost minimization test relative to expected differences for a nominal and spoofed assumption. In this paper, we pursue a similar concept where we start by separating our measurement into two non-distinct (i.e. they can have common satellites) subsets – assuming nominal and spoofed conditions. These subsets are created by eliminating outliers to the assumption. The outlier elimination depends on the test utilized. Then, for each subset, we develop tests to determine the relative likelihood of the subset being spoofed versus genuine. This section discusses the calculations required to determine the probabilities. Several tests based on different comparisons and different techniques to quantify probabilities of concern are developed.

Of course, other techniques may also be applied, including machine learning. Spoof detection is essentially a classification problem and hence is amenable to machine learning techniques. From the measurements of a given situation, we want to be able to classify whether that situation is experiencing spoofing or is under nominal conditions. This is not covered in this paper as it is beyond our scope and also because there are limitations to this approach for aviation safety. First, it is hard to

quantify the probability of missed spoof detection and false spoof alert. Second, even if the classification is good on the training data set, it is not clear how well it would perform on forms of spoofing that are new and not included in that set.

Individual Satellite & Pair tests

Hypothesis testing is a straight-forward way of testing if a measurement or test statistic is consistent with a given distribution associated with the null hypothesis (i.e. the condition being tested for). The basic concept is shown in Figure 4 where a distribution derived from the null hypothesis is given and a decision threshold is set. Values of measurement-derived test statistic that are above the threshold would result in rejecting the null hypothesis in favor of the alternative hypothesis (i.e. not the null hypothesis). In this case, the null hypothesis is that there is no spoofing with the alternative hypothesis being there is spoofing. This is used in the next paragraph. The threshold is set based on targeted probability of rejecting a correct hypothesis, typically termed type I error or false alert (fa). While the calculations and tests may be reasonably simple, it can form the basis of more sophisticated tests. With the case of spoof detection, assuming single spoofing source, we can go a step further as we can quantify both the null and alternative hypothesis distribution as seen a little later.

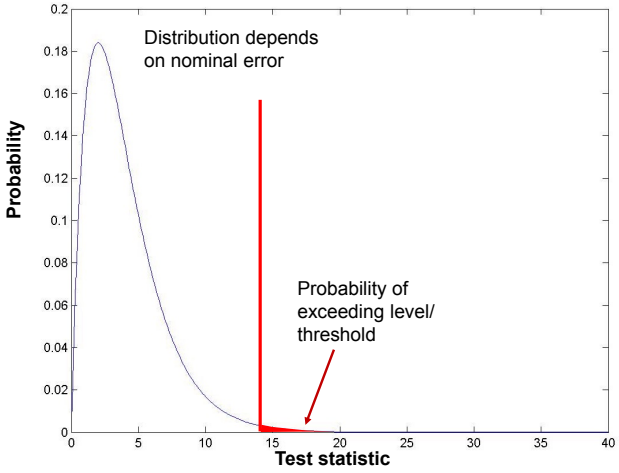


Figure 4. Statistical hypothesis testing based on null hypothesis derived distribution and detection threshold

For DOA, several different tests can be developed. If we have an absolute azimuth measurement, we can compare that measured azimuth (AZ) with the true azimuth (AZ_{true}) from satellite ephemeris and an approximate location. This difference shown in Equation (2), ε_{AZ}, can be used to decide if they are consistent. Ideally, ε_{AZ} should be distributed as zero mean with the variance of the azimuth measurement. In this analysis, we assume a Gaussian distribution for AZ measurement errors. The derivation can be modified if the distribution is different either by using a Gaussian overbound or modeling the actual distribution. Given this and assuming a probability distribution, we can solve the decision threshold to use based on the targeted probability of type I error, i.e. the probability of false alert (P_{fa}). This is shown in Equation (3). In the equation, normcdf is the cumulative distribution function (cdf) of a normal distribution with the three indices representing the decision threshold ($threshold_{decision}$) for the test statistic, mean and standard deviation of the distribution, respectively. A factor of two multiplies the normcdf as the equation assumes a two sided threshold such that the null hypothesis holds only if the absolute value of our measurement is below the threshold.

If we have relative azimuth measurements, then our measurement, m_{ij} or generically m, would be the difference in azimuth from a pair of different signals – as shown in Equation (4). The null hypothesis is that there is no spoofing. If the measurement errors are Gaussian and independent identically distributed (iid), the null hypothesis distribution (A) is given by Equation (5). Equation (5) states that the differential azimuth for an unspoofed measurement pair is distributed as normal with mean equal to the difference in the azimuth angle of the two signals and variance that is twice the variance of our azimuth measurement error, σ^2_{AZ}. We would decide that the null hypothesis is true if the measurement is below the decision threshold derived from Equation (6). If the measurement are from a single spoofer and Gaussian iid, the differential azimuth distribution is given by Equation (7) – a Gaussian with zero mean and variance that is twice the variance of our azimuth estimate. Spoof azimuth variance, $\sigma^2_{AZ,spoof}$, may be different from σ^2_{AZ} as all signals experience the same multipath error if coming from the same source. Hence multipath is common mode bias and should not factor into the variance. It will affect

the estimate of the incoming spoof direction but not our ability to detect signals from the same direction (spoofing). Knowing this distribution and our *threshold*$_{decision}$, we can determine P_{md} with missed detection being having a pair of spoofed signal but deciding that there is no spoofing. This is given by Equation (8). These equations can be modified should each azimuth measurement have different variances.

$$\varepsilon_{AZ,i} = AZ_i - AZ_{true\_i} \quad (2)$$

$$P_{fa} = 1 - 2 * normcdf(threshold_{decision}, 0, \sigma_{AZ}) = -erf\left(\frac{threshold_{decision}}{\sqrt{2}\sigma_{AZ}}\right) \quad (3)$$

$$\mathrm{m}_{ij} = \Delta AZ_{ij} = AZ_i - AZ_j \quad (4)$$

$$\Delta AZ_{ij} \sim N\left(\mu_{Ai,j} = AZ_i - AZ_j, 2\sigma_{AZ}^2\right) \quad (5)$$

$$P_{fa} = 1 - normcdf\left(threshold_{decision}, \mu_{Ai,j}, \sqrt{2}\sigma_{AZ}\right) = \frac{1 - erf\left(\frac{threshold_{decision} - \mu_{Ai,j}}{2\sigma_{AZ}}\right)}{2} \quad (6)$$

$$B \sim N\left(\mu_B = 0, 2\sigma_{AZ,spoof}^2\right) \quad (7)$$

$$P_{md} = 1 - normcdf\left(threshold_{decision}, \mu_B, \sqrt{2}\sigma_{AZ,spoof}\right) = \frac{1 - erf\left(\frac{threshold_{decision}}{2\sigma_{AZ,spoof}}\right)}{2} \quad (8)$$

Bayesian testing is useful when we have knowledge of the spoofed or faulted distribution. This can be the case with DOA-based detection provided the spoofed signals are coming from a single (or perhaps two) antenna. Assuming that spoofing is coming from one antenna, the nominal and spoofed distribution of the DOA difference between two satellites can be described. This is shown in Equation (5) and (7). These equations assume that the DOA and hence DOA difference measurements are the same for all satellites or satellite pairs. This may not be the case with genuine signals measured by the DPA as the variance is elevation dependent. This can be accounted for in the above equations.

Now given these two model distributions, we want to know first if we can develop a reasonable test and second if we can create a conservative test. In this test, we are trying to determine the likelihood of having a pair of spoofed signals (B) or not (A*). We use A* to indicate this to be a broader definition of A and includes cases where there is potentially one spoofed signal out of the pair. This derivation can also apply to condition A instead of A* provided the likelihood of having cases of one spoofed signal is effectively zero. We can use Bayes rule to express the probability of A* given a measurement m, *P(A*|m)*, which we do not know, as a function of the likelihood of A*, m, and probability of m given condition A. This is shown in Equation (9). A similar expression can be made for condition B. Now we need to relate the likelihood of condition A* and B. First, these two conditions are mutually exclusive and collectively exhaustive. This is the reason why we use A* rather than A. Second we can relate the probability of A* and B by a ratio N. These two are seen in Equation (10). From these equations, we can derive the following relationship for *P(A*|m)* and *P(B|m)* as a function of the probabilities that we know – probability of *m* given condition A* or B. This is shown in Equations (11) and (12). For this paper, we assume N = 1 which results in the prior probabilities of a genuine and spoof event being equal. While not true, this setting should provide for a good balance so that detection can occur during a spoofing event while keeping false spoof alerts low during non-spoof periods. If we want a spoof detector that is effective during a spoofing event, a significant likelihood of spoofing needs to be assumed. While the absolute likelihood of spoofing is low (i.e. P(B) very small), if that assumption is used then almost any measurement will result in a high likelihood of no spoofing given the prior probability of A*.

$$P(A^*|m) = \frac{P(m|A^*) * P(A^*)}{P(m)} \quad (9)$$

$$P(A^*|m) + P(B|m) = 1 \quad P(A^*) = P(B) * N \quad (10)$$

$$P(A^*|m) = \frac{N * P(m|A^*)}{P(m|B) + N * P(m|A^*)} \quad (11)$$

$$P(B|m) = \frac{P(m|B)}{P(m|B) + N * P(m|A^*)} \quad (12)$$

Under certain assumptions, it turns out that by inflating the estimated error variances, the desirable property of overbounding the probability of false alerts and missed detection is produced. This is an important property for many reasons. First, this yields a conservative estimate of P_{fa} and P_{md} for small P_{fa} and P_{md} as is desired. Second, even if the measurement error is not Gaussian, we can use a Gaussian overbound for the variance to create a conservative estimate.

Examine what happens when an overbounding measurement variance or standard deviation is assumed. Start by deriving the ratio between the likelihood of condition A* and B given a measurement, *m*. Assuming N = 1, Equation (11) and (12) simplifies to Equations (13) and (14). The ratio of the conditional probabilities in (14) lead to Equation (15) where R denotes the ratio of the mean to standard deviation. The equation shows the ratio of the calculated probability of having unspoofed signal to the probability of having spoofed signals for a Gaussian distribution. Given this result, Table 1 shows what happens to the ratio should the standard deviation used be increased. The table shows that there is less certitude in the dominant hypothesis, the hypothesis with the greater probability, as the standard deviation is increased. This generally is the conservative property that we desire as it would imply that the true P_{fa} and P_{md} be smaller than that calculated for small values of P_{fa} and P_{md} (< 0.5 for N =1).

$$P(m) = P(m|A^*) + P(m|B) \qquad (13)$$

$$P(A^*|m) = \frac{P(m|A^*)}{P(m)}, P(B|m) = \frac{P(m|B)}{P(m)} \qquad (14)$$

$$\frac{P(A^*|m)}{P(B|m)} = \frac{P(m|A^*)}{P(m|B)} = exp\left(-\frac{\mu_A^2 - 2m\mu_A}{2\sigma^2}\right) = exp\left(-\frac{R^2}{2} + \frac{mR^2}{\mu_A}\right) \qquad (15)$$

Table 1. Effect of increasing model error standard deviation on likelihood of each hypothesis as a function of the measurement

m (fraction of μ_A)	Ratio	Effect of increasing σ (decrease R)	Dominant hypothesis
0	$\exp(-R^2/2)$	Increases ratio (more likelihood of A)	B
¼	$\exp(-R^2/4)$	Increases ratio (more likelihood of A)	B
1/2	0	Same	Neither
3/4	$\exp(R^2/4)$	Decreases (less likelihood of condition A)	A
1	$\exp(R^2/2)$	Decreases (less likelihood of condition A)	A

Composite tests

The above tests examined the signals individually or in pairs. It is more useful to use all signals to together to provide a more confident sense of whether spoofing exists. Several forms of composite tests are formulated and examined. As part of these tests, it is important to consider outliers. Outliers will exist and can come from several sources. Multipath can cause DOA measurements that differ significantly from expected in nominal conditions. A few genuine signals during spoofing conditions also represent outliers. Such outliers can significantly affect our test statistics and perhaps provide an incorrect level of certitude on our detection or lack thereof. Hence, outlier elimination for composite testing is important. Also, an outlier to the test depends on which of two hypothesis is tested – either the signals are nominal (genuine) or they are spoofed. As a result, in our analysis, two subsets of satellites are used with one based on assuming that the signals are genuine and the other assuming that they are spoofed.

One way to determine the overall spoofing situation is to combine the individual pairwise tests to form a pairwise composite test. Care must be taken in the combination and two examples are provided as illustration. To illustrate, a simple but naive combination is shown in Equation (16) with *S* being the subset of satellite signals being tested for spoofing. This is not a good method as it has the property that more measurements can only reduce P_{spoof} and hence our confidence. For this paper, a simple approach is used with the average of the individual probability from each pair used to calculate the overall probability of the subset used being spoofed or nominal. Equation (17) and (18) are the calculated probabilities for the spoofed case subset, S, assuming a base satellite j. M is the number of satellites in the subset. A satellite i is eliminated from the set *S* if their probability of spoof, $P(B|m_{ij})$, is too low over several base satellites *j* thus removing potential outliers. By eliminating outliers, we assume that A = A*. Also, Equations (17) and (18) are already the results of applying Equation (11). So they represent both possible cases and the sum of the two probabilities should equal 1. Similarly, the approach is also used the

nominal assumption with *G* being the subset of satellite signals being tested as genuine and K being the number of satellites in the set. This is shown in Equation (19) and (20). For the paper, we term results of this test as pairwise composite.

$$P_{spoof} = \prod_{i,j\in S, i\neq j} P\left(B\middle|m_{ij}\right) \quad (16)$$

$$P_{spoof,S,j} = \frac{\sum_{i,j\in S, i\neq j} P(B|m_{ij})}{M-1} \quad (17)$$

$$P_{nom,S,j} = \frac{\sum_{i,j\in S, i\neq j} P(A|m_{ij})}{M-1} \quad (18)$$

$$P_{nom,G,j} = \frac{\sum_{i,j\in G, i\neq j} P(A|m_{ij})}{K-1} \quad (19)$$

$$P_{spoof,G.j} = \frac{\sum_{i,j\in G, i\neq j} P(B|m_{ij})}{K-1} \quad (20)$$

Instead of combining individual tests, all signals in view can be compared at once using a least squared fit. First outliers are eliminated. A comparison with the expected case is made. The comparison of the azimuth to the ephemeris-derived azimuth and a single azimuth is used for the nominal and single spoofer cases, respectively. In each case, an angular offset from true zero azimuth is solved as the measurement are relative azimuth. Solving the problem minimizing deviation leads for finding the angular offset through an averaging of the error from expected. Then outliers are excluded based on extreme errors relative to the standard deviation of the measurements. Standard statistical testing is then employed for the test by taking the difference or error in the measured DOA from expectations as a whole. A straight-forward way combination is to use the sum squared error (SSE) or weighted sum squared error (WSSE) where the error is weighted by the inverse of its estimated standard deviation. Provided that the error is distributed as a standard normal, then the resulting distribution is chi squared with *n* degrees of freedom where *n* is the number of satellite signals used. As it is assumed that error is normally distributed, we assert that the WSSE, with a correct standard deviation, is chi squared. The results for probability of spoof and nominal using the spoof subset *S* is shown in Equation (21) and (22), respectively. Bayes rules from Equation (11) can then be applied to yield the overall probability of that subset being spoofed. This overall probability of the signals in *S* being spoofed, provided there are only two possibilities (spoofed or genuine), and is shown in Equation (23). Again, for this paper, we use N = 1. Equation (23) is insightful as it uses the likelihoods of the set *S* being from an all spoofed or nominal set. Thus, it provides a measure of the ability of set *S* to validate that its signals are spoofed. For example, while $P_{spoof,S}$ may have high probability (e.g. 0.999), if $P_{nom,S}$ also has high probability (e.g. 0.75), result would be that $P_{spoof\ overall,S}$ would not be very large (0.57 using the prior example probabilities and N = 1). This may happen if the true locations of the satellites line up well with a given direction. Such a true geometry is not good for distinguishing between spoofed and genuine signal hence a good test should not provide high confidence. Equation (24) shows the probability of finding a WSSE value greater than our calculated one which can be used for a standard hypothesis test described in the beginning of the section. The nominal subset results are similar and shown in Equations (25)-(28). For the paper, we term results of this test as least squared composite.

$$P_{spoof,S} = \chi^2_M\ pdf\left(\sum_{i,\in S}\left(\frac{AZ_i - AZ_{spoof}}{\sigma_{AZ,spoof}}\right)^2\right) \quad (21)$$

$$P_{nom,S} = \chi^2_M\ pdf\left(\sum_{i,\in S}\left(\frac{AZ_i - AZ_{true}}{\sigma_{AZ}}\right)^2\right) \quad (22)$$

$$P_{Spoof\ overall,S} = \frac{P_{spoof,S}}{P_{spoof,S} + N * P_{nom,S}} \quad (23)$$

$$P_{spoof,S,exceed} = 1 - \chi^2_M\ cdf\left(\sum_{i,\in S}\left(\frac{AZ_i - AZ_{spoof}}{\sigma_{AZ,spoof}}\right)^2\right) \quad (24)$$

$$P_{spoof,G} = \chi^2_M \, pdf\left(\sum_{i,\in G}\left(\frac{AZ_i - AZ_{spoof}}{\sigma_{AZ,spoof}}\right)^2\right) \quad (25)$$

$$P_{nom,G} = \chi^2_K \, pdf\left(\sum_{i,\in G}\left(\frac{AZ_i - AZ_{true}}{\sigma_{AZ}}\right)^2\right) \quad (26)$$

$$P_{nom\ overall,G} = \frac{P_{spoof,G}}{P_{spoof,G} + N * P_{nom,G}} \quad (27)$$

$$P_{nom,G,exceed} = 1 - \chi^2_K \, cdf\left(\sum_{i,\in G}\left(\frac{AZ_i - AZ_{true}}{\sigma_{AZ}}\right)^2\right) \quad (28)$$

The calculations from the combined processing provide many results which can be used to decide on detection. It yields 1) number of satellites used (and outliers), 2) overall probability of spoofing calculated using Bayes rule and 3) a χ^2 statistic on spoofing (for the batch all in view test). While incorporated implicitly in the least SSE composite, the goodness of the geometry is for spoof detection is another metric to utilize. Later, in Figure 12, we show a general processing for detection using results from the least squared composite. No attempt was made to optimize the use of these results for detection. One way it determines spoofing seeing if the number of satellites used in the spoof case is adequate and the overall probability of spoofing is adequately high. The former requirement is to prevent detection from cases where we sometimes get a subset of several genuine satellites in the same general direction. For the processing, the requirement is at least five satellites or more than half the total, whichever is greater. The latter is determined by the targeted probability of missed detection and typically means overall probability of spoofing needs to be greater 99.9% for detection. A similar calculation is done for the nominal subset and the results are compared. Even if the DOA spoof detection does not yield an definite, it can still help overall spoof detection. For example, if a subset has been determined to be genuine, then that subset can then be used as a trusted subset by other tests such as receiver autonomous integrity monitoring (RAIM) to examine if satellites not in the subset are spoofed.

EXPERIMENTAL SET UP

Assessment and validation of these tests are performed using on-air test data and the Stanford DPA. Nominal on-air data was collected at Stanford University Durand Building which houses the Department of Aeronautics & Astronautics. Spoof data from many scenarios was collected at the US government sponsored spoofing exercise in 2017. These exercises allowed us to first determine DPA direction finding capabilities [15] in many different scenarios. For the tests at Stanford with the genuine GNSS signal, the DPA was set up on a steel cylinder with a one foot by one foot ground plane [14]. This set up was placed on the roof of the four story tall Durand building. For the government exercise, DPA is set up on the roof of a sports utility vehicle (SUV) with the same 1 feet squared (ft^2) ground plane. It went through multiple scenarios of on-air spoofing at three different distances from the spoofing source. Figure 5 shows the antenna on top of the SUV and its size relative to a 1 foot ruler.

Figure 5. Dual Polarization Antennas (DPA) on SUV roof: Front using u-blox, Rear using USRP (Left) & DPA in lab (Right)

This antenna has been tested using genuine GNSS broadcasts and in on-air spoofing tests. We test the performance of our developed algorithms on data from these field experiments. The DOA estimate error statistics from actual measurements will be used to inform our algorithms. Analysis and simulations are performed for comparison and as a baseline.

Several steps are needed to determine DOA using the DPA. The PCB Stanford DPA generates a signal that is the combination of the LHCP and a phase shifted RHCP signal to find lower elevation signals and direction find. The DPA is built with a scan mode that uniformly step a variable phase shifter through the full range of phase shifts. A u-Blox receiver processes the combined signal to determine its carrier to noise ratio (C/No). DOA is determined by finding the phase shift that results in the minimum C/No [7].

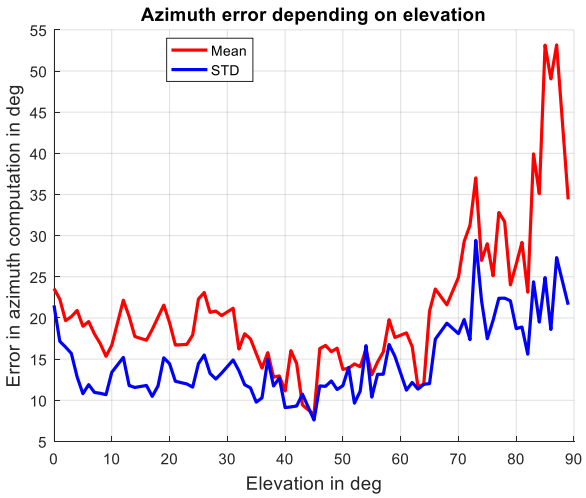


Figure 6. Mean and standard deviation (STD) of the error on estimated azimuth as a function of satellite elevation angle during static testing

To use the developed methodologies, an estimate of the error statistics of the DOA measurements is also needed. For nominal measurements, this is gathered by assessing GNSS satellites in nominal conditions. Baseline data from Stanford is used and provides an independent set from the US government sponsored spoofing exercise scenarios. Figure 6 shows a plot of the performance statistics of DPA angle estimates as a function of satellite elevation using a scan period of 51.2 seconds. The measurements is used to provide the statistics for nominal azimuth measurement error. An overall overbounding standard deviation of 35 degrees (15 degree standard deviation, 20 degree mean). This is decreased by a factor of square root of 5 as the collected data had five times longer dwell period (256 sec) resulting in σ_{AZ} = 15.65 degrees being used for the paper.

The sponsored spoofing exercise provided scenarios to quantify the DPA angular errors and assess our spoof detection techniques. At this exercise, the DPA was set up to complete a scan every 256 seconds. This long dwell time improved azimuth estimates. After processing the results, relative azimuth is calculated. The left side of Figure 7 presents this with the ephemeris-derived elevation angle shown to help separate out each satellite. For comparison, the ephemeris-derived azimuth and elevation is shown on the right hand side of Figure 7. All comparison plots that are shown later will follow this format. Measurement from many scenarios showed the relative azimuth error for spoofing having a smaller standard deviation with $\sigma_{AZ,spoof} < 5$ degrees. This is shown in the histogram in Figure 8. This likely is due three sources: 1) common mode error, 2) slightly higher spoof power than genuine satellites and 3) generally lower elevation angle of spoofer.

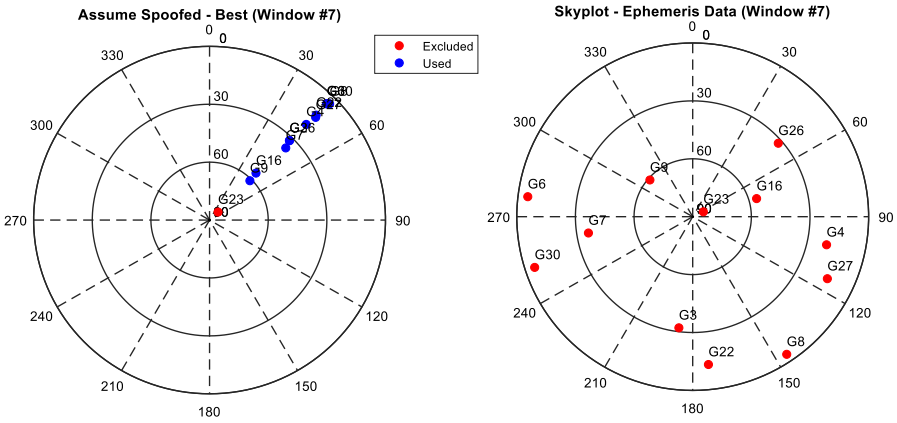


Figure 7. Comparison of azimuth and elevation from ephemeris (right) and azimuth from antenna DOA estimate and elevation from ephemeris (left) for GPS satellites during on air spoof.

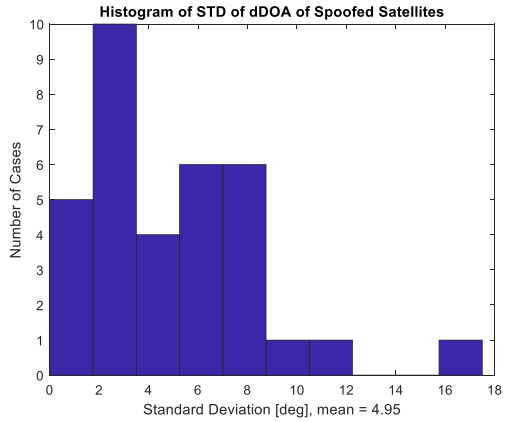


Figure 8. Standard deviation of azimuth measurements from spoofing signal over 1 day.

RESULTS FROM VARIOUS SCENARIOS

Three different scenarios are analyzed for illustration. The processing of the azimuth data is shown in Figure 9. First, the 180° ambiguity from the DPA is eliminated based on knowledge of the actual scenario. As mentioned earlier, handling the ambiguity will be covered in a future paper. Next, high (> 65°) and zero elevation satellites are removed. The former is removed as the DPA does not provide accurate azimuth for those elevations. The latter is removed as zero elevation is usually an indication that we did not have the correct ephemeris data. Next, the spoof and nominal cases are assumed and outliers excluded based on that assumption. Finally, the probabilities discussed in the prior section are calculated.

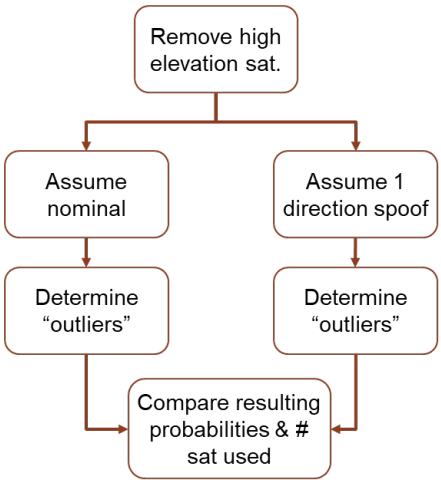


Figure 9. General Process for Assessing Azimuth and Calculating Spoof Detection Metrics

Scenario 1 has all spoofed GPS signals and is shown in Figure 7. For the left skyplot, it shows the satellites retained by the calculations assuming spoofed condition in blue and the not used ones in red. All 11 satellites below 65° are retained while only one satellite is not used as it has elevation angle above 65°. As all signals are spoofed and coming from the same source, the error statistics on each should be the same. Both the pairwise and least SSE composite values are calculated and shown in Table 2. For clarity, the "Equations" column references the equations used for the calculation to its immediate left. Excluding the title row, the top and bottom three rows are results assuming nominal and spoof conditions, respectively. Nominal condition results in few satellites used, hence many "outliers". Thus it is easy to exclude this case. Spoof conditions results in all available satellites (i.e. not high elevation) used. The pairwise and least SSE (Bayes) spoof probabilities are high with the latter having greater confidence (essentially 100%). This is both because the geometry is good for distinguishing between single direction spoof and the actual azimuths and because the χ^2 test statistic from our measurements is low relative to expected variations with nearly all true values being larger than it (> 99.999999%, see row 5). So the developed tests clearly indicated spoof though the pairwise composite is not as confident with its probability is only 98.75%. This is not surprising as the construction of this test does not necessarily make the test significantly more confident with additional measurements.

Table 2. Performance of Pairwise and Least Sum Squared Error Composite Calculations on Spoof Case shown

	Pairwise Composite	**Equations**	**Least SSE Composite**	**Equations**
Number of Sats Retained for Nominal Subset	0 of 11		2 of 11	
Probability nominal subset χ^2 metric being exceeded	N/A		51.524%	(28)
Probability of Nominal (given nominal subset)	N/A	(19)	34.058%	(25), (26), (27)
Number of Sats Retained for Spoof Subset	11 of 11		11 of 11	
Probability spoofed subset χ^2 metric being exceeded	N/A		99.9999999658414%	(24)
Probability of Spoof (given spoof subset)	98.751%	(17)	100%	(21), (22), (23)

Scenario 2 examines a case with mixed genuine and spoofed signals as shown in Figure 10. Again, the left side shows the skyplot based on DPA measured the ephemeris derived elevation angle shown to help separate out each satellite. The right side shows the skyplot with ephemeris-derived azimuth and elevation. In this case, there is spoofing but a few genuine satellites are tracked – specifically G17, G11 and G18 (note that G11 is about 180 degrees off due to the DPA ambiguity). The results of the detection calculation is shown in Table 3. In this scenario, a few satellites are used in the assumed nominal condition – too few to make a proper assertion. However, the nominal satellites retained result in high confidence of nominal (see row 3 "Best Nominal Prob (Ave) – Bayes"). Looking assume spoof condition, many, but not all satellites, are retained. In this case, eight and seven out of eleven for the pairwise and least squatted composite cases respectively. With this subset, both calculations indicated high (> 99%) probability of spoofing.

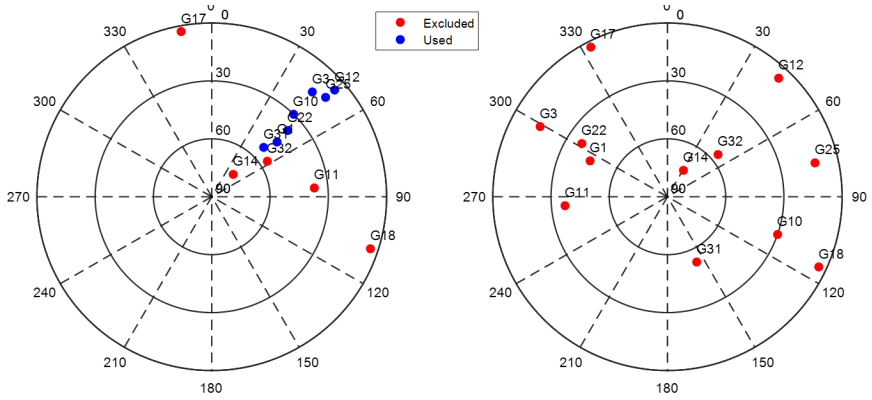


Figure 10. Comparison of azimuth and elevation from ephemeris (right) and azimuth from antenna DOA estimate and elevation from ephemeris (left) for GPS satellites (with PRN number) during on air spoof (mixed with some genuine signals)

Table 3. Performance of Pairwise and Least Sum Squared Error (SSE) Composite Calculations on Mixed Spoof Case shown

	Pairwise Composite	**Least SSE Composite**
Number of Sats Retained for Nominal Subset	3 of 11	4 of 11
Probability nominal subset χ^2 metric being exceeded	N/A	89.038%
Probability of Nominal (given nominal subset)	100%	100%
Number of Sats Retained for Spoof Subset	8 of 11	7 of 11
Probability spoofed subset χ^2 metric being exceeded	N/A	98.592%
Probability of Spoof (given spoof subset)	99.439%	100%

Scenario 3 examines a nominal case with no spoofed signals. Figure 11 shows the skyplot of the ephemeris-derived azimuth and elevation on the right and the DPA-derived azimuth and ephemeris-derived elevation on the left. The detection calculation for this case is shown in Table 4. In this case, eight of 11 satellites are retained for testing as nominal by both techniques with both yielding a high probability of those being nominal based on DOA. The excluded satellites had larger than expected errors in their azimuth measures (> 30°).

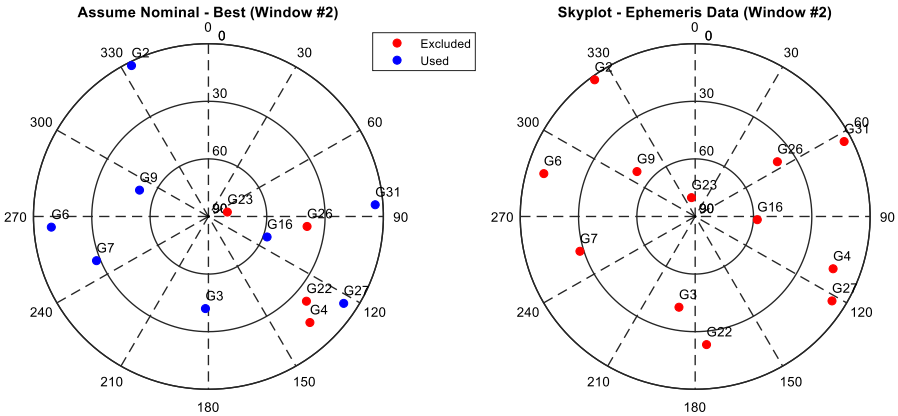


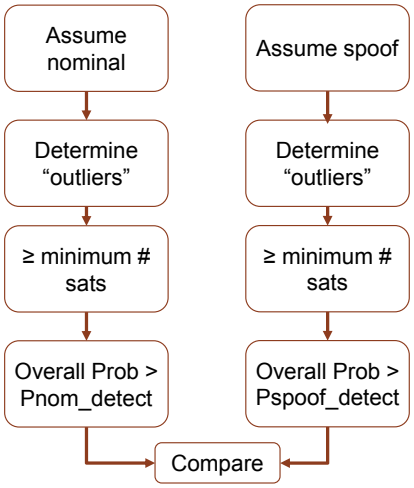
Figure 11. Comparison of azimuth and elevation from ephemeris (left) and azimuth from antenna DOA estimate and elevation from ephemeris (right) for GPS satellites (with PRN number) with no spoofing (nominal).

Table 4. Performance of Pairwise and Least Sum Squared Error Composite Calculations on Nominal Case shown

	Pairwise Composite	**Least SSE Composite**
Number of Sats Retained for Nominal Subset	8 of 11	8 of 11
Probability nominal subset χ^2 metric being exceeded	N/A	56.333%
Probability of Nominal (given nominal subset)	99.9999998097329%	100%
Number of Sats Retained for Spoof Subset	2 of 11	2 of 11

Probability spoofed subset χ^2 metric being exceeded	N/A	37.423%
Probability of Spoof (given spoof subset)	99.687%	96.101%

These are only select cases and reasonably good ones at that. So far, only a portion of the spoof test scenarios have been processed using the developed techniques. Table 5 shows results from three different tests based on the spoof detection process shown in Figure 12. For the thresholds P_{nom_detect} and P_{spoof_detect}, 99 and 99.9% were used, respectively. The different thresholds did not affect the results in the table. The minimum number of satellites was the maximum of five and number of useable satellites (< 65°) divided by two. The DPA analyzed nominal, mixed, and spoofed data that occurred at different time windows of during each test.



Calculation Result Comparison

	Spoof Detect Pass	Spoof Detect Not Pass
Nominal Detect Pass	Compare prob, # sats	Nominal
Nominal Not Pass	Spoof	No decision (leave for other tests)

Figure 12. Calculation process for spoof detection decision using least squared composite results

Table 5. Probability of Spoofing calculated based on various tests

Test	1	2	3
Measurement set (Windows)	9	11	7
% Spoof Detected Correct	100% (5 of 5)	100% (5 of 5)	66.7% (2 of 3)
% Spoof Detected Incorrect	0%	0%	0%
% Nom Detected Correct	75% (3 of 4)	83.3% (5 of 6)	75% (3 of 4)
% Nom Detected Incorrect	0%	0%	0%
% All - No Decision	11.1% (1 of 9)	9% (1 of 11)	28.6% (2 of 7)

The results are encouraging. An important result is that there were no false alerts. There are several no decisions, this occurred where we had a strong mix (roughly equal numbers) of genuine and spoofed signals. In Test 1, the no decision window was with mostly nominal signals with likely some spoofing signals. In Test 2, the no decision window occurred about when spoofing started which resulted in significant errors in azimuth measured as the DPA, as implemented, requires a 256 second window to estimate azimuth. In Test 3, there were two strong mixed cases where there were many satellites spoofed and genuine satellites. In one of the mixed case, no decision resulted from only having 5 of 11 satellites selected for the spoofed subset. So while the calculations showed high confidence that these satellites were spoofed, there was not enough satellites to make that decision. The other no decision was a window that contained the start of spoofing. It had only six satellites with large azimuth error such that it was unclear as to which were spoofed and which were not. These are cases where we generally would want the detection to indicate no decision rather than false alert.

SUMMARY

This paper develops and demonstrates several spoof detection tests based on direction of arrival measurements. It provides a procedure for robustly utilizing two developed statistically based detection techniques: 1) pairwise comparisons and 2) least

sum squared error composites using on air spoofing data gathered by the DPA. The developed techniques provide three important contributions for spoof detection using DOA: 1) many key metrics for a spoof detection monitor, 2) outliers detection, and 3) statistical comparison between two hypotheses (nominal and spoof) based on relative probabilities. The metrics generated for spoof detection includes as determining the satellites that are consistent with the hypotheses and the confidence of the hypotheses given those satellites. Both techniques utilize outlier detection based on design as they examine two possible hypotheses, satellites are genuine (nominal) and satellites are spoofed, and determines outliers based on that. Hence not only are they robust to outliers, they also formulate two subsets representing possible spoofed and possible genuine satellites. The outlier removal and separation into the two hypothesis allows us to use a Bayes rule derive probability of the likelihood of spoof relative genuine for each condition. Not only does this quantify the likelihood of spoofing, it is also powerful as the probability accounts for how good the geometry is for spoof detection.

This work on represents a start at developing these detection test. The test need to be verified on the additional data sets not tested. Better utilization of calculation outputs also should be examined to improve performance. Additionally, the challenge of using DPA measurements, with its 180° ambiguity, should and will be addressed in the future.

ACKNOWLEDGMENTS

The authors thank the Federal Aviation Administration (FAA) and the Stanford Center for Position Navigation and Time (SCPNT) for sponsoring this research. The authors also thank the US government providing us with an opportunity to test under live GPS spoofing.

REFERENCES

[1] Yemisi Adegoke, “Uber drivers in Lagos are using a fake GPS app to inflate rider fares,” Quartz Africa, November 13 2017, https://qz.com/1127853/uber-drivers-in-lagos-nigeria-use-fake-lockito-app-to-boost-fares/

[2] “GPS Spoofing A Growing Problem for Uber,” Solid Driver, June 9, 2017, http://soliddriver.com/GPS-Spoofing-A-Growing-Problem-for-Uber

[3] P. Y. Montgomery, T. E. Humphreys, B. M. Ledvina, "Receiver-Autonomous Spoofing Detection: Experimental Results of a Multi-Antenna Receiver Defense against a Portable Civil GPS Spoofer," Proceedings of the 2009 International Technical Meeting of The Institute of Navigation, Anaheim, CA, January 2009, pp. 124-130.

[4] M. Psiaki, Brady W. O'Hanlon, Steven P. Powell, Jahshan A. Bhatti, Kyle D. Wesson, Todd E. Humphreys, Andrew Schofield, “GNSS spoofing detection using two-antenna differential carrier phase," Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2014), Tampa Bay, Florida

[5] Yu Hsuan Chen, Sherman Lo, Per Enge, " Using Multi-Antenna Iridium Measurements for Rapid Spoof Detection," Proceedings of the Institute of Navigation Joint Navigation Conference, Orlando, FL, June 2015 (Presentation Only)

[6] M. Appel, A. Konovaltsev, M. Meurer, “Joint Antenna Array Attitude Tracking and Spoofing Detection Based on Phase Difference Measurements,” Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016). ION GNSS+ 2016, Portland, Oregon.

[7] Yu-Hsuan Chen, Sherman Lo, Dennis Akos, David De Lorenzo, Per Enge, “Validation of a Controlled Reception Pattern Antenna (CRPA) Receiver Built from Inexpensive General-purpose Elements During Several Live Jamming Test Campaigns,” Proceedings of the Institute of Navigation ITM Conference, San Diego, CA, January 2013

[8] A. Konovaltsev, F. Antreich, A. Hornbostel, "Performance Assessment of Antenna Array Algorithms for multipath and Interference Mitigation", Proceedings of the ESA Workshop on GNSS Signals 2007, ESTEC, April 2007

[9] M.H. Keshvadi , A. Broumandan, G. Lachapelle, “Analysis of GNSS Beamforming and Angle of Arrival Estimation in Multipath Environments,” Proceedings of the Institute of Navigation ITM Conference, San Diego, CA, January 2011

[10] David De Lorenzo, “Navigation Accuracy and Interference Rejection for GPS Adaptive Antenna Arrays,” Ph.D. Dissertation, Stanford University, August 2007

[11] Alison Brown, Dale Reynolds, Huan-Wan Tseng, John Norgard, "Miniaturized GPS Antenna Array Technology," Proceedings of the 55th Annual Meeting of The Institute of Navigation (1999), Cambridge, MA, June 1999, pp. 243-251.

[12] Andrew J O’Brien, “Adaptive Antenna Arrays for Precision GNSS Receivers,” Ph.D. Dissertation, Ohio State University, 2009

[13] Emily McMilin, "Single Antenna Null Steering for GPS & GNSS Aerial Applications," Ph.D. Dissertation, Stanford University, March 2016

[14] Y.-H. Chen, F. Rothmaier, D. Akos, S. Lo, P. Enge, "Towards a Practical Single Element Null Steering Antenna," Proceedings of the Institute of Navigation International Technical Meeting, Monterrey, CA, January 2017

[15] Yu Hsuan Chen, Fabian Rothmaier, Dennis Akos, Sherman Lo, Per Enge, “Demonstrating Single Element Null Steering Antenna Direction Finding for Interference Detection,” Proceedings of the Institute of Navigation International Technical Meeting, Reston, VA, January 2018