Metamodel-Assisted Disciplining Algorithm for Detecting Spoofed GNSS Time Signals

O. Garitselov, PhD and D. Sohn
Spectracom, Rochester, NY

BIOGRAPHIES

O. Garitselov received his PhD in CSE from University of North Texas in 2012 and an M.S. degree in CS from Moscow Bauman State University in 2003. He is currently working as a Senior Engineer at Spectracom, Inc. in Rochester, NY.
D. Sohn earned a B.S. degree in CE at Penn State University in 2001. He is currently working as Engineering Manager at Spectracom.

ABSTRACT

The cost and technological barriers required to successfully spoof Global Navigation Satellite Positioning System (GNSS) signals have fallen greatly. At the same time, the use of GNSS signals for personal, civil and military applications has grown dramatically. The risk of a successful spoofing event to wreak havoc on commercial or civil interests is no longer a minor one, and in a world of 'big data' the consequences are potentially catastrophic in terms of financial or societal harm. The communication and data infrastructure we take for granted is at risk of attack, causing denial of use or even systemic failure. Due to technological restrictions and IP constraints it is hard to determine how GNSS receivers will react to spoofing of GNSS signals without having specialized equipment, and spending a considerable number of man hours. Unnoticed spoofing for an extended period of time can cause some GNSS receivers to drift by a substantial time difference. Therefore, it is essential to be able to identify if a GNSS receiver or signal has been tampered with, since many critical systems (e.g., stock markets, power grids, and communications) are dependent on its timing systems.
This research presents an algorithm that is used to detect and mitigate GNSS receiver time signal abnormalities that can arise by spoofing attempts altering GNSS signals before they reach the receiver. We have implemented a GNSS-receiver-independent solution that is added to the disciplining algorithm of a Global Positioning System Disciplined Oscillator (GPSDO) System in order to detect and filter abnormalities without introduction of any additional hardware to the timing system. A metamodeling approach is used to help accurately predict the behavior of the "un-spoofed" GNSS receiver and the internal oscillator of the system. To this end, parameters such as internal oscillator aging, temperature and previous error history are taken into account. The spoof prediction algorithm uses metamodels to enhance the disciplining process which is then able to reject bad data caused by spoofed GNSS timing data. Since the accuracy of the metamodel is a key to this approach, we are considering different variants of models: neural networks and other common multidimensional regression/interpolation models. Two different metamodels are used in the proposed algorithm: the short-term (coarse) model is created by using data from short time intervals to allow fast real-time calculations. A second, long-term (refined) model that is created from very long time intervals, takes into account overall behavior of an internal oscillator over an extended period of time. Using both models adds another protection layer in case that the offset of the spoofed timing signal is so minuscule that it is not detected by the coarse metamodel. The error offset introduced into a system by a spoofing attempt is greatly reduced by the algorithm even if it fails to detect a spoofing attempt significantly increasing the time it takes to induce a large scale error.

I. INTRODUCTION

Virtually all electronic communication infrastructures in one way or another depend on GPS, GLONASS or other navigation satellite system signals. Precise data on location and time-of-day are used in communication networks, for the transportation of goods, distribution of electricity, financial operations and rescue of people during disasters. Since the year 2000, the more accurate code division multiple access (CDMA) GPS satellite coarse/acquisition (C/A) code has become available for public and commercial use. The precise (P(Y)) code signal used by the Government is encrypted and unavailable to the general public.
The civilian version of GPS has a dangerous vulnerability with potentially serious consequences. Between recent inconsistencies in GPS availability and the GLONASS constellation data mathematical error, the generally accepted, freely available commercial GNSS data has proven to be unreliable.

To alleviate a temporary loss of the GNSS signal or data, multiple-constellation receivers have become more frequently used in sensitive systems. Other, increasingly popular solutions to bridge periods during which the GNSS signal is lost, are internal position data predictors, e.g., inertial navigation units (IMU), and time keeping hardware, e.g. high precision oscillators, such as Rubidium, Oven-Controlled Crystal Oscillator (OCXO), and Chip Scale Atomic Clocks (CSAC).

Another major concern, however, is that the public GPS signal is not encrypted and available to everyone. Thus, it can be jammed or altered by an attacker. The method of substituting or replacing a GNSS signal is called spoofing. This is a deliberate attack on a particular object and its smooth pulling to one side of the planned trajectory, or slightly altering the frequency of transmission and time-of-day information. This is possible, because GNSS signals near the surface of the Earth are relatively weak; their level can be about -163 dB·watts, while the signal emitted by an attacker can be much stronger and hence can overpower the original signal. Standard GNSS receivers do not check the signal for plausible inconsistency [1]. In the worst case, a successful GNSS spoofing attack can cause distortions in navigation systems and landing aircraft, confusion concerning public trading up to and including a potential stock market crash, as well as collapsing energy systems.

This research work targets time and frequency synchronization systems that are used to keep accurate time and frequency, and keep other components of larger systems in sync with each other. Time and frequency synchronization information recovered from the GNSS signals are used to keep a high precision oscillator aligned in phase and frequency by means of a disciplining algorithm. The system can then provide a continuous precise time and frequency reference even if the GNSS signal is lost, referred to as holdover [2].

Frequency error induced in the oscillator will affect the system performance. Since the oscillator is locked in phase and frequency to the GNSS signal, frequency spoofing can cause the system to produce incorrect time and frequency. While the use of high priced multidirectional antenna/receiver systems, multi-constellation receivers, duplicate systems, etc. can slightly decrease the chances of a successful spoofing attempt, these solutions are rather expensive approaches to alleviate spoofing. If, however, a system is capable of detecting on an algorithm level that a reference signal has been altered, it can provide notification and stop tracking that reference signal, and reduce the error propagation into other system components.

In order to accomplish this, we need to identify what other conditions can alter the oscillator performance. One of the biggest conditions is temperature change: Since most oscillators are piezoelectric in nature, changes in temperature affect the way it behaves. Another factor is oscillator aging, which is natural degradation of the component over time, which has a smaller effect on the performance of oscillators, but due to its dynamic and nonlinear behavior is harder to statically predict. If we can measure the temperature and aging of the system, and the system has been tracking the reference signal, we can determine changes in the reference signal.

This paper is structured as follows: In Section II, we will show the spoofing setup for a GPSDO system. Section III will introduce a disciplining algorithm, and Section IV explains how to include metamodels for temperature and aging inside a disciplining algorithm. Section V will introduce a novel approach that includes a fiber optic delay loop inside a disciplining algorithm. Conclusions are discussed in Section VI.

II. SPOOFING THE SYSTEM

The GNSS synchronization system illustrated in Figure (1) is based on a GPSDO. The RF signal that is emitted by a GNSS satellite is transformed by a receiver into a 1PPS signal and streaming data. A 10 MHz signal from a controlled precision oscillator provides the frequency output. The oscillator frequency is steered by a disciplining algorithm that acts as phase detector and loop filter component of the PLL. The algorithm calculates a digital-to-analog converter (DAC) value from a 1PPS reference signal, based on the calculated oscillator frequency and phase offset.

It is possible to conduct a spoofing attack on this kind of system. The attacker needs to repeat the measurements for the location of the antenna, and then start broadcasting the same data they received from the GNSS signals, while increasing the signal strength. After the target receiver starts following the new (spoofed) signal, the latter can be modified. A GNSS signal generator can transmit correct constellation data for any location at any given time. The reference frequency, generated by the reference GPSDO, applied to the GNSS signal generator is adjusted gradually, while remaining within the receiver PLL

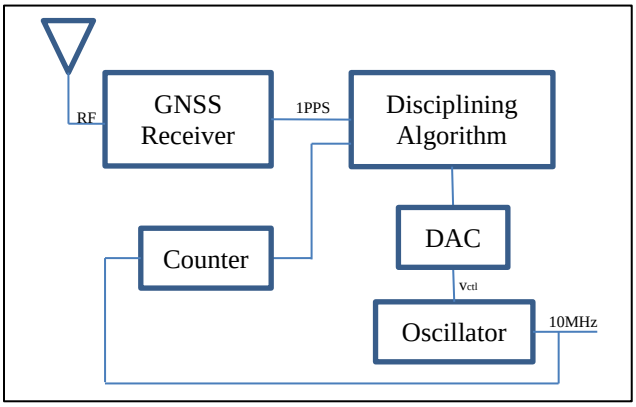

Figure 1: Block diagram of GNSS synchronization system

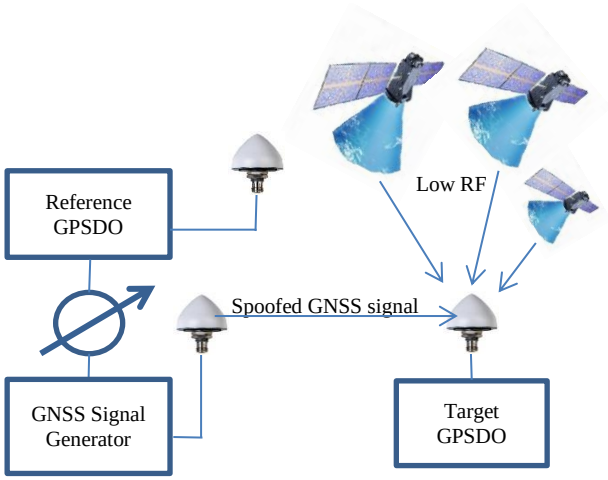

Figure 2: Spoofing setup

locking range, see Figure (2). This effect causes the RF data to transmit at a slower or faster interval, compared to the genuine signal.

Successfully spoofing a GNSS receiver with a high time offset will not necessarily be successful with a GPSDO system. The TCXO of the GNSS receiver has a much larger tuning range than the more precise oscillators that are used in a target GPSDO. Therefore, for targeted GPSDO systems, the maximum reference offset needs to be within the disciplined oscillator range. For a spoofing setup, we are using a reference GPSDO with the same kind of oscillator as the target GPSDO.

Also, the rate of the spoofing change cannot be too large due to filtering inside the disciplining algorithm. The frequency error between the oscillator and the spoofed system can accumulate and become large enough for the algorithm to stop tracking as expected. We have tested a slightly altered version of a disciplining algorithm that will be presented in Section (III), and concluded that step changes of larger than 0.03 Hz/sec were too large. Even though the GNSS receiver was still locked to the signal, the change was too large and the disciplining algorithm could not lock in phase and frequency with the 1PPS signal coming from the receiver.

Therefore, for GPSDO systems, the adjustments need to be done with a smaller offset to not lose the lock of the GPSDO system. The best result was accomplished by offsetting the 10 MHz signal of the system by 4.58 Hz, which corresponds to 4580 ns/sec or 39.57 ms/day.

III. DISCIPLINING ALGORITHM

The disciplining algorithm illustrated in Figure (3) attempts to lock a 10 MHz oscillator to a reference 1PPS signal by measuring the phase difference between the reference 1PPS signal and a generated 1PPS signal derived from the oscillator frequency.

The algorithm collects data for a certain period of time (measurement window). At the end of that window, it will provide a phase difference measurement between the

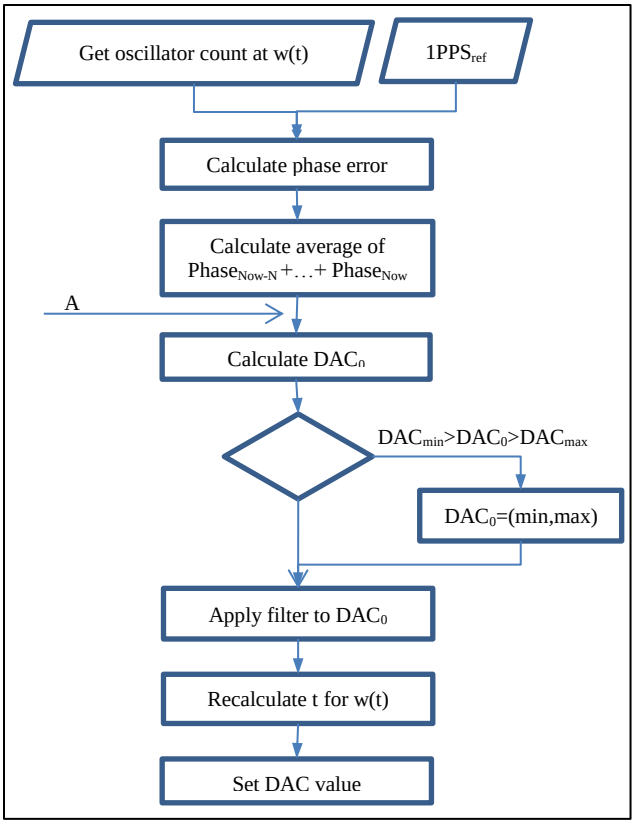

Figure 3: Flowchart for disciplining algorithm

reference 1PPS (filtered) and the generated 1PPS. From this measurement and the measurement from the previous window, a frequency error is calculated. Since we already know how much the frequency will change per step of the DAC and we know what the DAC value was for the latest window, we calculate what the DAC should be to have a frequency error of 0. Since this calculation may be affected by measurement error, it is averaged with the previous N measurements.

Now we can apply filtering to get the ideal DAC setting for the next window. Filtering is used to smooth the DAC setting changes, just like a loop filter in a PLL. This value is then bounded by the max change, the absolute max adjustment from the zero-freq-error DAC value, and ultimately by the DAC rails.

Depending on the history of the calculated error, we can adjust the window size to perform corrections at a higher or lower rate as needed.

A frequency error of the system is caused by the temperature and aging dependency, as well as the frequency error of the oscillator and the reference signal.

$$F_{err} = f_{osc\_err} + f_{osc\_temp} + f_{osc\_aging} + f_{ref\_err} \quad (1)$$

IV. TEMPERATURE AND AGING COMPENSATED DISCIPLINING ALGORITHM

Predicting oscillator behavior is commonly used in disciplining algorithms to control the oscillator more accurately. One of the largest contributors to error that

can be compensated for is temperature. Compensating for this error requires adding a temperature sensor to the system. In order to utilize the temperature reading, we need to create an accurate frequency behavioral model based on temperature.

The frequency reaction to temperature change can be extracted from the oscillator's data sheet characteristics. For our example GPSDO including an OCXO it is $\pm 5\times10^{-9}$ Hz from -20ºC to 90ºC. To obtain a more precise model, we placed the GPSDO in a temperature chamber, recording the frequency output as temperature changed over time. The recorded temperature is based on the temperature sensor that is part of the system.

The experimental setup is shown in Figure (3). The GNSS disciplined Rubidium oscillator produces a 10 MHz reference to the counter. The 10 MHz GPSDO signal is divided by 10e6 to bring the signal to the 1PPS range. The 1PPS signal from the receiver and the 1PPS from GPSDO are then compared.

The created metamodel is applicable only for that system's particular oscillator and temperature sensor combination. The generation of a more generic model will require a larger number of components and will include a statistical distribution of all frequency/temperature data, or use nominal characteristics as stated above. This will make the model less accurate for any particular system.

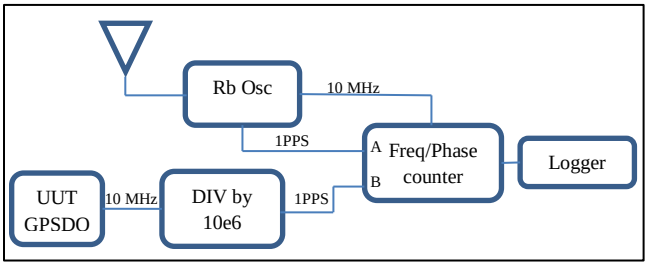

Figure 3: GPSDO behavioral test setup

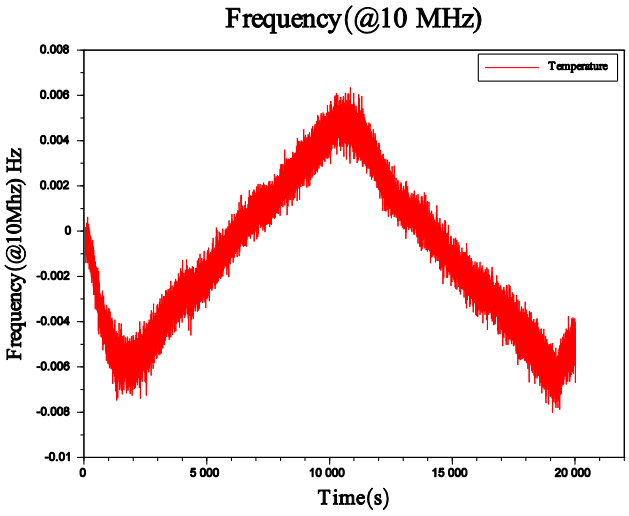

Figure 4: Frequency vs. time behavior

Figures (4) and (5) show the results for relative frequency and temperature variation during the experiment.

To generate an accurate metamodel, we need to apply the correct technique to fit the data to a function. To check the accuracy of the model we can use the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}SSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y(x_k) - \hat{y}(x_k))^2} \qquad (2)$$

where y are the actual simulation result values and $\hat{y}$ are the results of the created model for the same x. We are showing two metamodel methods: polynomial regression and feedforward neural networks (FFNN).

Starting with polynomial regression, where the function follows the form:

$$y = \sum_{i,j=0}^{k}(a_{ij} \cdot x_1^i \cdot x_2^j) \qquad (3)$$

where y is the respose frequency, a_{ij} are the coefficients and k is the order of the metamodel. The least square analysis is then used to find the coefficients that produce the smallest error in the sum of the squares for all points.

A FFNN is a no-cycle neural network. All the data moves from layer to layer.

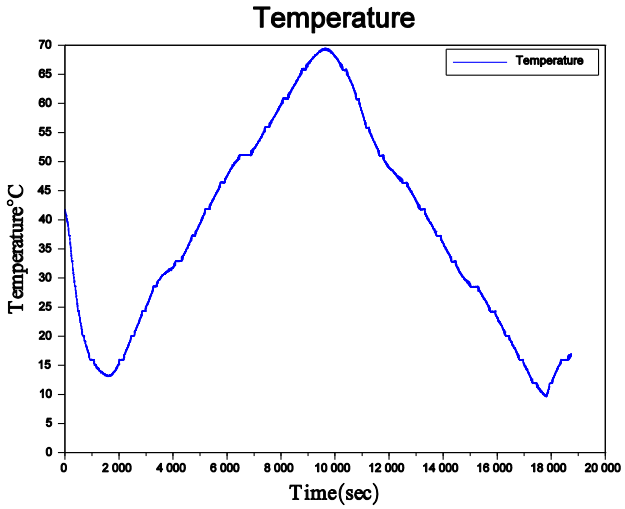

Figure 5: Temperature vs. time behavior

Each layer consists of a function and weight that is calculated through a training process of fitting the neural network to given data. These kinds of models require data to be normalized (scaled to the same degree, usually with ranges from 0 to 1). Otherwise, inputs that have much larger values can offset calculated weights. Furthermore, altering the amount of neurons in each internal layer, using different layer functions, as well as altering the accuracy and kind of NN training can alter the accuracy of FFNN dramatically.

For this example we have implemented a single hidden

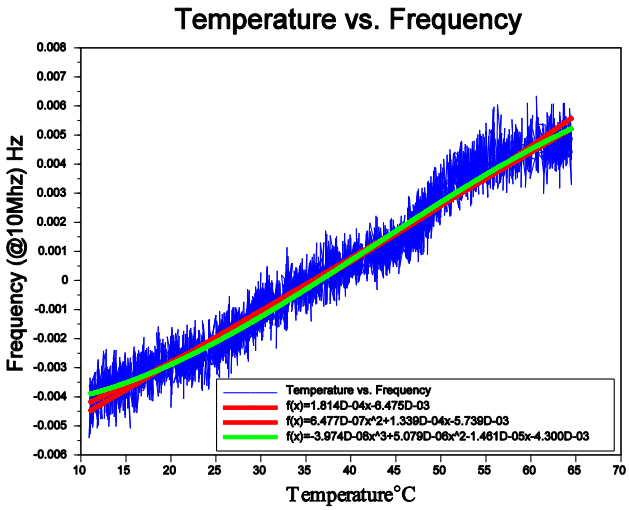

Figure 6: Frequency vs. temperature behavior. Fit data using first degree linear regression.

layer FFNN. The log-sigmoid function:

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{4}$$

is used in the hidden layer. The Latin hypercube sampling (LHS) is used to pick candidates for iteration of training. This method slightly alters the training dataset to stop overtraining [3].

Due to LHS, the random value of training data creates a slight change in the accuracy result of each FFNN, even if the same parameters are used for training and architecture of NN. To find the best NN fit, we generated over 100 NNs, altering the number of neurons in the hidden layer and the accuracy of training.

Metamodel	Mean Error	Std Error	RMSE
Polynomial k=1	4.42×10^{-4}	3.32×10^{-4}	5.53×10^{-4}
Polynomial k=2	4.27×10^{-4}	3.26×10^{-4}	5.37×10^{-4}
Polynomial k=3	4.18×10^{-4}	3.16×10^{-4}	5.24×10^{-4}
FFNN	4.15×10^{-4}	2.97×10^{-4}	5.19×10^{-4}

Table 1: Metamodel results (Hz)

Table (1) shows the results for different metamodels that were used to fit the temperature data. Even though the FFNN showed the best overall results, we selected a polynomial regression model with k=3 for our algorithm since it is much easier to implement in an embedded system and does not require large processing power to compute.

The aging of an SC cut OCXO follows this logarithmic function:

$$F(t) = A[\ln(Bt + 1)] + fo \tag{5}$$

Due to this logarithmic aging, modeling accurate oscillator aging behavior is generally not possible for oscillators with less than 30 days of operation [4]. However, OCXOs are usually shipped from the factory when the daily aging rate of the oscillator has already stabilized after a sufficient period of operation. It is possible to keep track of the total time the oscillator has been operational, but the starting point of time including factory operation is unknown. We will use nominal specifications for oscillator aging: $\pm5.0\times10^{-10}$ Hz/day which is $\pm5.787\times10^{-15}$ Hz/second. Since the OCXO aging slope has a tendency to change sign dynamically [5], we are unable to model it with the available information and it is left included in the error of the oscillator.

We add temperature checking at entry point A of Figure (3) to the disciplining algorithm. The average temperature that is calculated over a window period is compared to the previous temperature window average:

$$\Delta f = f(temp_{avg}) - f(temp_{avgold}) \tag{6}$$

Then, the calculated frequency error for the period of $w(t)$ is checked to be within the allotted constant error offset f_{offset}:

$$f_{offset} = \Delta f \pm (f(temp_{tserr}) + w(t) \cdot f_{osc} \tag{7}$$

where $temp_{tserr}$ is the nominal temperature sensor error and f_{osc} represents the oscillator error including aging characteristics. If it is outside the bounds, the algorithms will go into an error state. Otherwise it will replace the old temperature average with the new, and continue to DAC calculation.

Even though the w(t) period can be chosen to be large (10s to 100s of seconds) it is comparatively small to the total run time of the system. The statistical data that is observed for the running system can be very useful to identify abnormal system behavior. For example: A GPSDO system was running for 30 days continuously without any issues. The model that was generated has enough points to calculate where the next w(t)'s should be in terms of performance (f_{err}). When a small error is introduced to the system by spoofing, the calculations for the short term metamodel can fall within the allowed frequency error range. After some iterations in calculations of each w(t) period, the long term metamodel gains more weight behind its predictions. The system can then identify that a small drift error is applied to GPSDO. Therefore, the long term metamodel can identify the abnormal behavior of the unit in comparison to the overall performance of the system. We can gather the frequency error of the model from each run of the disciplining algorithm and calculate the weight of each run since the window size is variable by the disciplining algorithm.

The metamodel is regenerated every w(t) period. The weights of the current point and the metamodel play a large role, since the new frequency error affects previously created long term model directly proportional to the ratio of weights.

$$f'_{errap}(t) = \left(\frac{\omega_{cp} \cdot f_{err} - \omega_{ap} \cdot f_{errap}(t - w(t))}{w(t)}\right) f_{errap}$$

where ω_{cp} is the weight of the current point, ω_{ap} is the weight of all collected points, and $f_{errap}(t)$ is the value for the model that was calculated from that data of all collected points. The predicted frequency error is calculated from the metamodel that was generated from

the previous run. Therefore, we can calculate the model error:
ModelError = $f'_{err}(t) - f_{err}$

V. FIBER OPTIC ASSISTED DISCIPLINING ALGORITHM

In order to make the reference frequency error detection more stable, we introduce a constant delay loop into the system that can be used to compare GPSDO performance as well as reference. Proposed is a novel approach [6] that measures constant propagation delay, which is achieved by adding a fiber optic propagation device consisting primarily of a fixed length fiber optic cable. The device's propagation delay is measured and recorded during the system calibration stage. Figure (9) shows the statistical data for measurement of the propagation delay of an example fiber optic propagation device consisting of a fixed length fiber optic cable of 1 km over a period of 6 days. The calculated standard deviation for the measurement is 127.9 ps with a mean of 10.447 µsec. calculated short term stability with an Allen deviation $\sigma_y(\tau)$ =348.74 fs. The data was not collected in a temperature controlled environment; therefore some deviation is due to variations caused by room temperature changes.
Using a fiber optic based propagation device has many bonuses: There is propagation delay variation due to temperature for fiber optic cable due to the material expanding and contracting, changing the length of fiber optic wire, but it is considered small and can be reduced further with specialized fiber optic cable and temperature control. The propagation delay is not subject to parasitic electrical/RF effects, unlike coax cables. There are also no frequency aging effects in fiber optic signaling. Figure (8) shows the diagram of the GPSDO system temperature

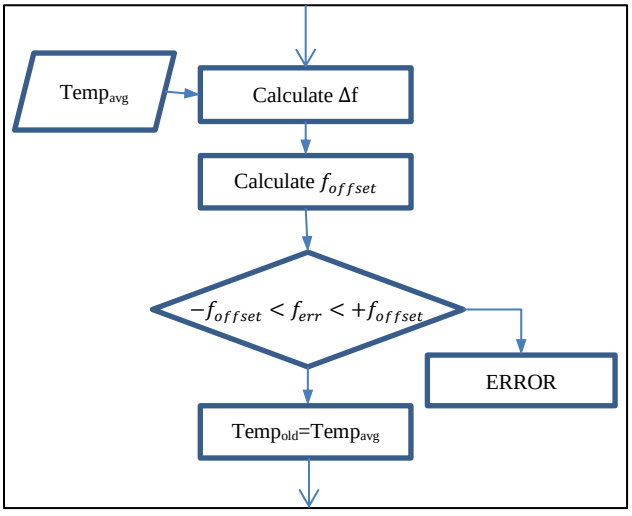

Figure 7: Flowchart for a metamodel assisted temperature and aging compensated algorithm, addition to Figure (3)

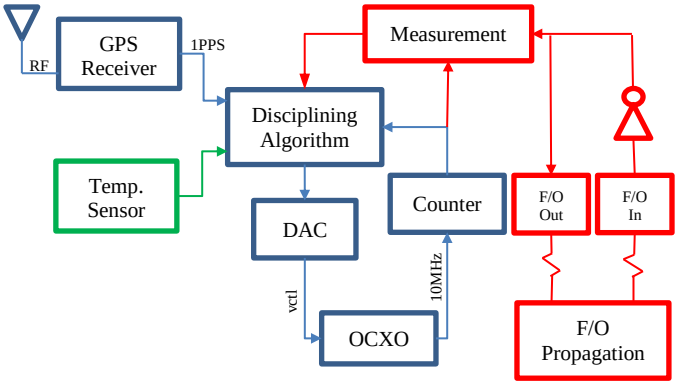

Figure 8: Block diagram of a GPS disciplined oscillator system with temperature sensor (green) and fiber optic propagation device (red).

sensor and fiber optic propagation device. The disciplining algorithm collects measurement data every window period from the fiber optic measurement component. The component measures propagation delay using counters by using a steered oscillator as a reference. The disciplining algorithm calculates the difference between the known count, which was identified during calibration stage, and from the current count.
The calculation for f_{fo_temp} is done using a nominal thermal coefficient based on the fiber optic cable type. The calculated frequency error consists of the temperature dependency and the oscillator error:

$$F_{err} = f_{osc\_err} + f_{osc\_temp} + f_{osc\_aging} + f_{fo\_temp} \quad (8)$$

A small error window for the calculations to allow free play for frequency and phase adjustment is given. This error window needs to be increased during the OCXO warm up phase and then tighten when the oscillator is locked.
If the count is off by more than the allotted error, this indicates an error condition with either reference signal or the oscillator. Using the error calculated with the fiber

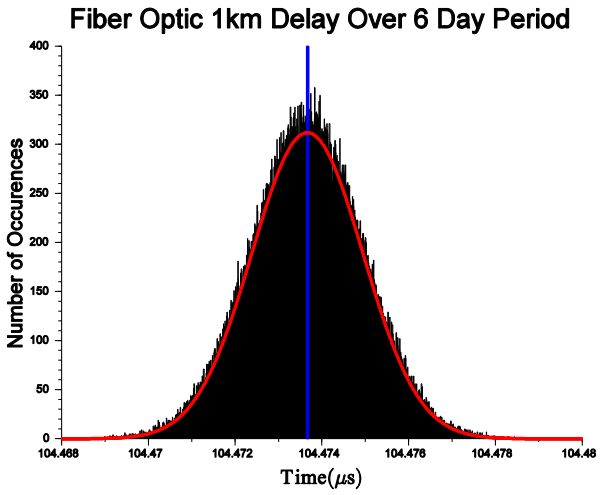

Figure 9: Fiber optic 1km delay statistics over a 6 day period.

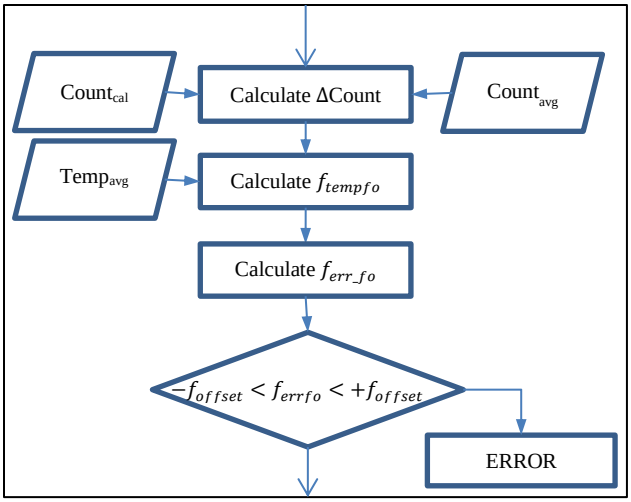

Figure 7: Flowchart for a GPDO system with a fiber optic delay loop, addition to Figure (3)

optic propagation device (eq 8) and the error calculated within the disciplining algorithm (eq 1), we can determine the reference error, f_{ref_err}, and use that to determine the combined oscillator and oscillator aging error, $(f_{osc\_err} + f_{osc\_aging})$.

VI. CONCLUSION

We showed how a GPSDO system can be spoofed if there is no spoofing detection in the system. It was demonstrated that large time offsets, can remain undetected by a GPSDO, hence falsely maintaining receiver and disciplining algorithm integrity. We showed two different ways how spoofing can be identified in the system without being fully dependent on a GNSS receiver or other outside components. A small alteration to the disciplining algorithm, by including a temperature model, allowed the system to roughly identify reference inconsistencies. We have introduced the long term metamodel, which is dynamically generated every iteration cycle of the disciplining algorithm. Another progression to the algorithm was to add a fiber optic propagation device. A constant known propagation delay through the device allowed the system to identify if a frequency error is present in the oscillator or reference, causing a deviation from the nominal 10MHz frequency.

REFERENCES

[1] Annex, A. "Global Positioning System Standard Positioning Service Signal Specification." June 1995, 2nd edition, Web. http://www.navcen.uscg.gov

[2] Lombardi, Michael A. "The use of GPS disciplined oscillators as primary frequency standards for calibration and metrology laboratories." *Measure: The Journal of Measurement Science* 3.3, pages 56-65, 2008.

[3] Oleg Garitselov, Saraju P. Mohanty, and Elias Kougianos. "Fast-accurate non-polynomial metamodeling for nano-CMOS PLL design optimization." *VLSI Design (VLSID), 2012 25th International Conference on*. IEEE, pages 316-321, 2012.

[4] O. Lebfried, B. Neubig. "Correlation of predicted and real aging behavior of crystal oscillators using different fitting algorithms", *Proc. 11th European Forum Time Frequency*, pages 268-272, 1997.

[5] Manish Vaish. "Very Long Time Scale Aging Performance Results of Quartz Crystal Oscillators." Web. <http://www.mti-milliren.com>

[6] D. Sohn. "Independent Fiber-Optic Reference Apparatuses and Methods Thereof." Spectracom, US Patent Application 14/532,620 Unpublished (filing date November 4, 2014)