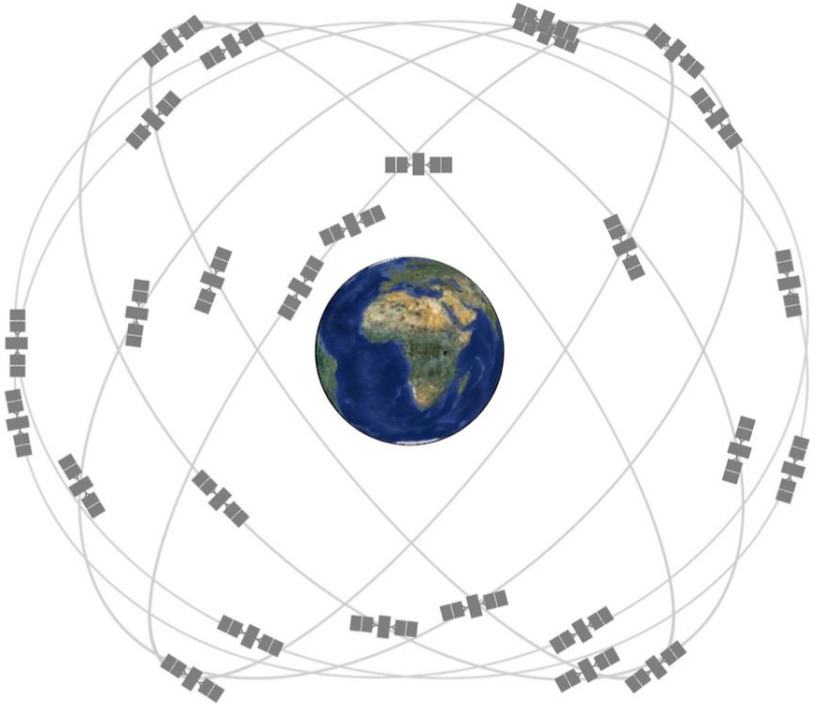
Anomaly detection with Autoencoder to detect Spoofing attack on GNSS

What is GNSS(Global Navigation Satellite Systems)

- US GLobal Positioning System
- Russian GLONASS
- European Galileo
- Chiese BeiDou

\+ Other augmentations



GPS Constellation

Principles

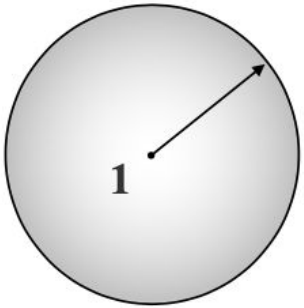
- Trilateration & Time of arrival

– If signal velocity is known, range from the transmitter to receiver is measured by v*(trec – ttrans).
– Position is determined by intersection of circular lines of positions centered on the transmitters.
– Must assume perfect clock for this work

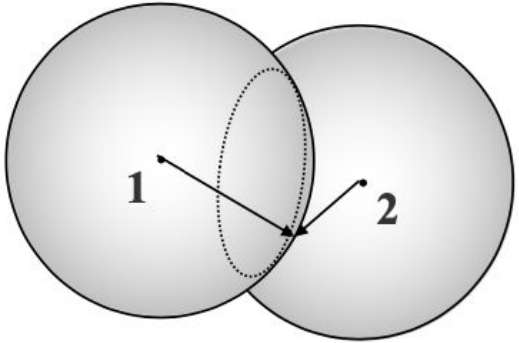
- Doppler positioning and frequency shift

– As the transmitter approaches the frequency is shifted higher due to Doppler effect (relativity).
– At the point of closest approach the shift is zero, and it goes lower as the transmitter moves away.
– If the trajectory of the transmitter is known, you can get 1D or 2D position information from this measurement and its rate of change.

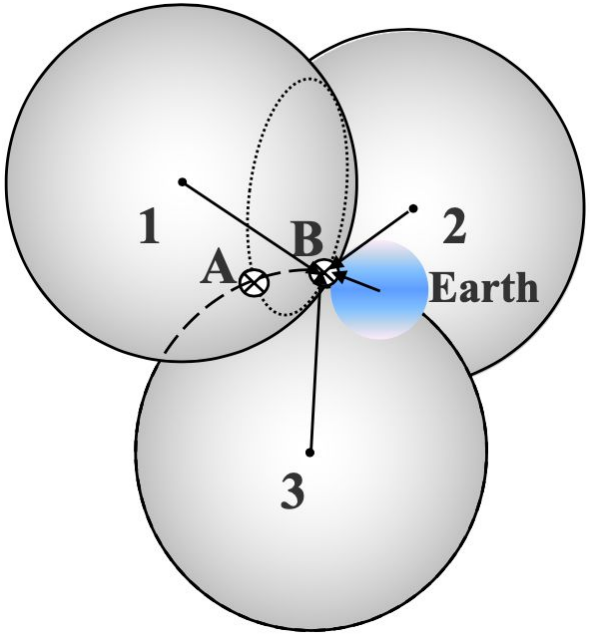
Trilateration for GNSS



1 satellite

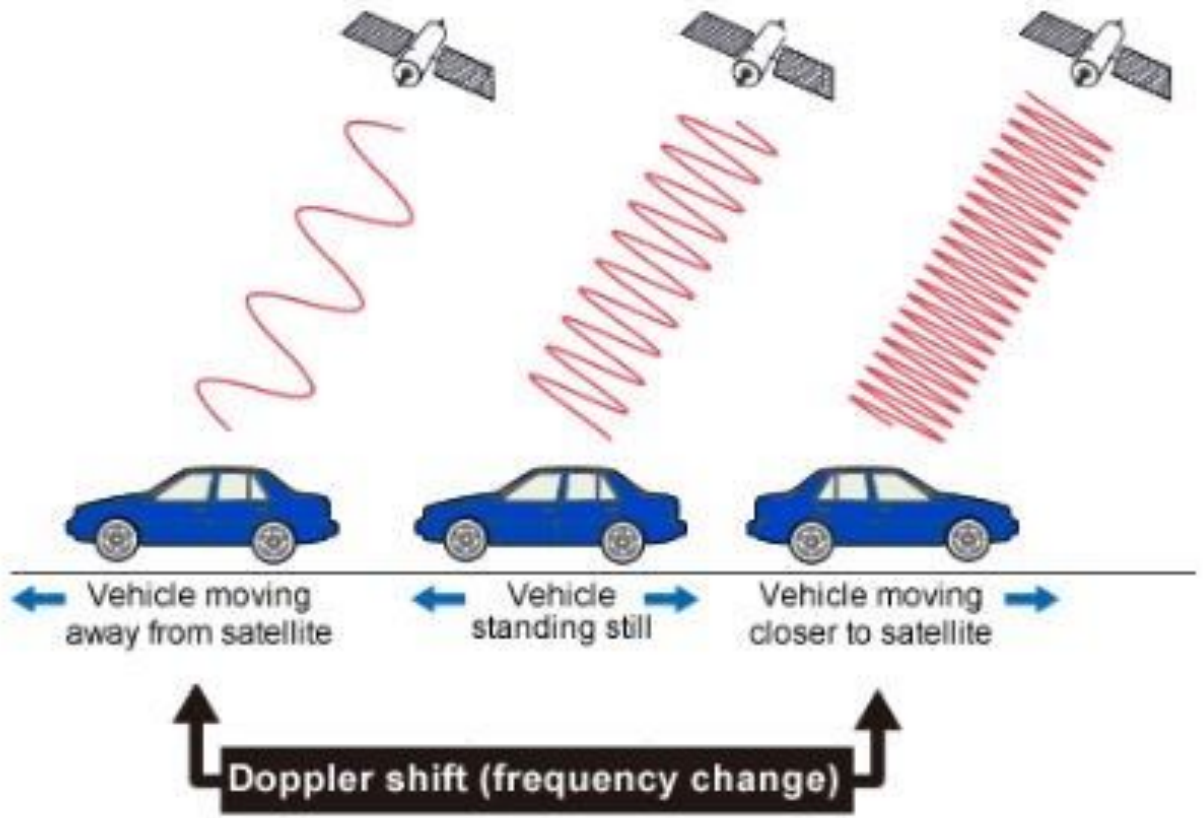


2 satellites



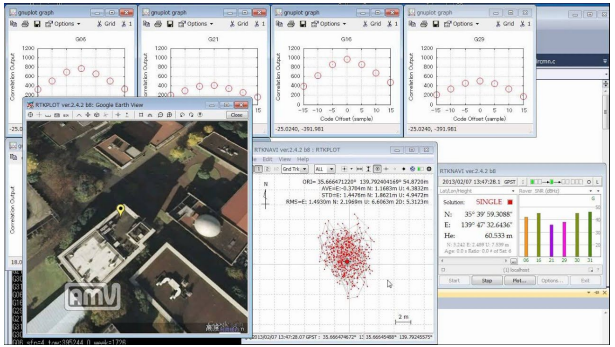
3 satellites

Doppler Shift for GNSS



Why GNSS-SDR

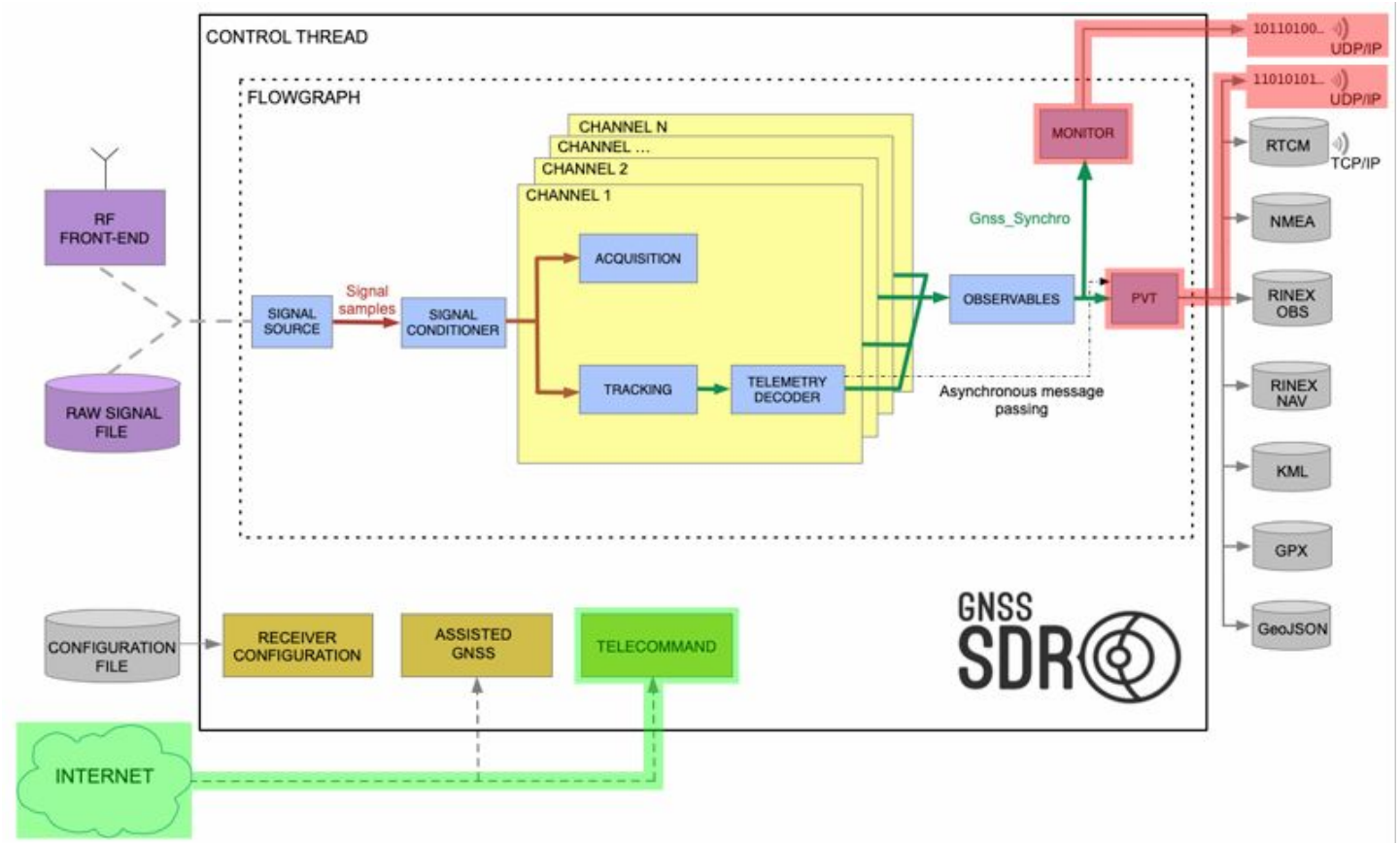
GNSS SDR



GNSS-SDR is

- Hosted by non-profit research foundation(cttc)
- It enables Software defined radio to receive and process GNSS signals

GNSS-SDR Structures (= General GNSS receiver)



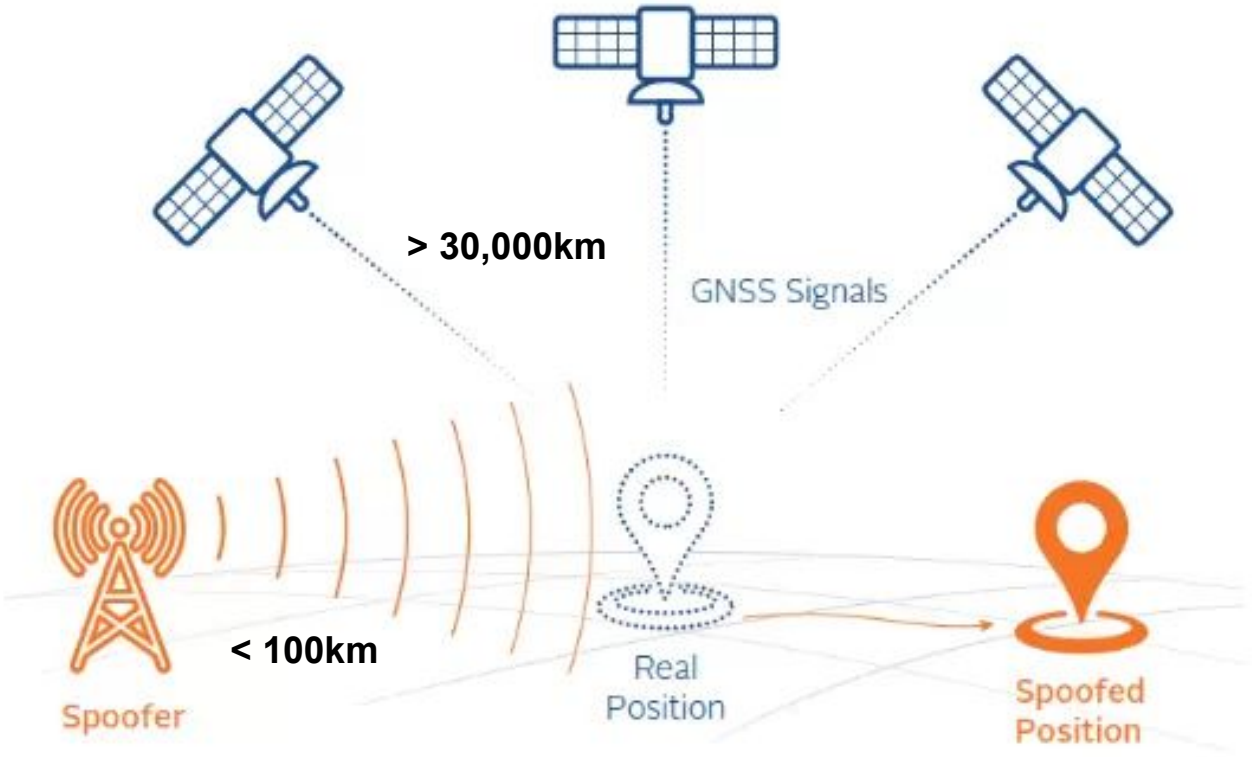
Threats in GNSS

Spoofing :
Deceive receivers who are tracking a channel with strong signal

Jamming :
Propagate the signal which has same frequency with current receiving signal which potentially overwhelm the original signal

Meaconing :
Record the normal signal and replay it

Spoofing in GNSS



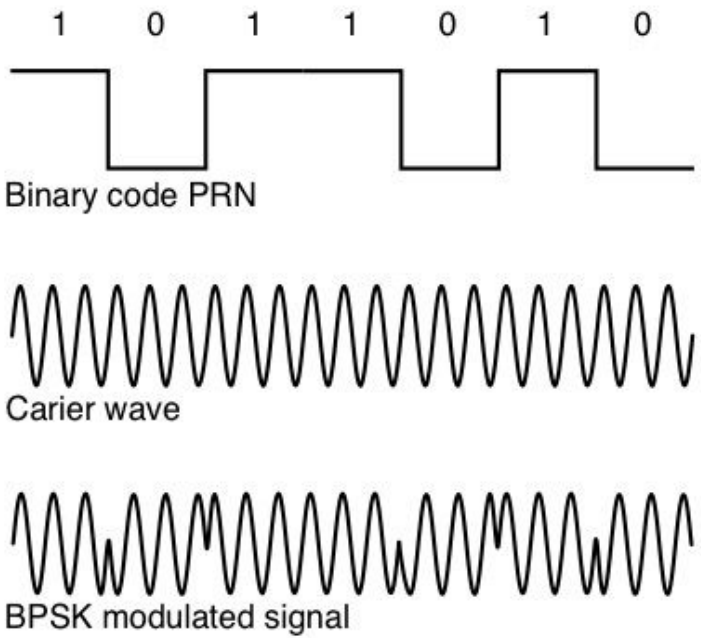
Why GNSS is vulnerable to Spoofing attacks

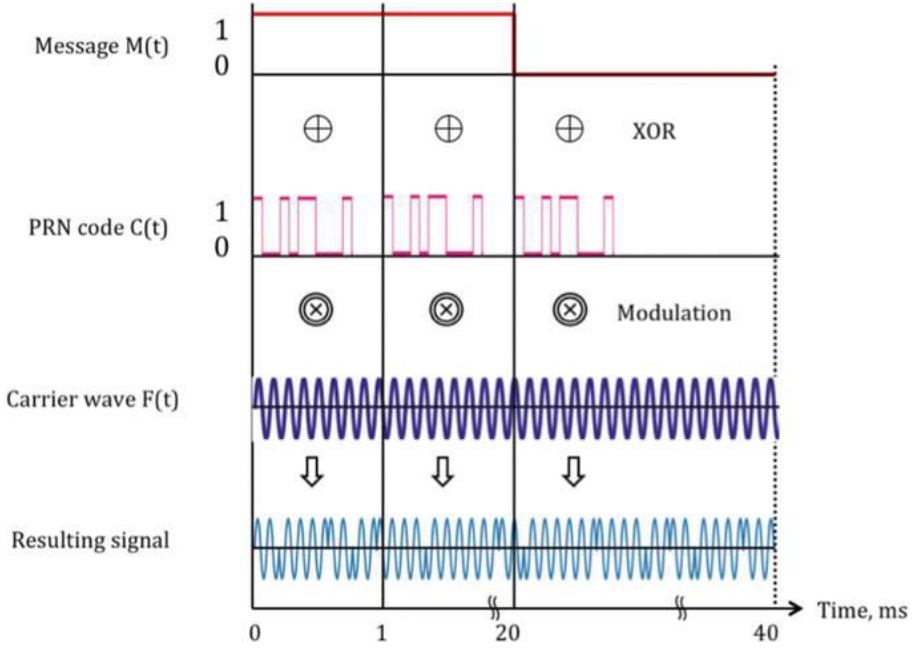
Distance :
Signal from satellite (30,000km) cannot beat spoofing signal(100km)

Publicity :
GNSS is generally open to public(PRN code)

Authentication :
Once start tracking, the GNSS receiver do not check the signal source

Signal modulation in GNSS





Why Spoofing attack is detectable

So many variables in GNSS signal :

- Position info : Satellite’s trajectory, position info
- Time info : The time when the signal is sent
- Signal feature info
  ex) C/N0 : Signal to noise ratio
  Pseudorange : Calculated distance from Doppler shift

➡️**What if utilize multiple variables to build anomaly detection model?**

Variables from GNSS-SDR

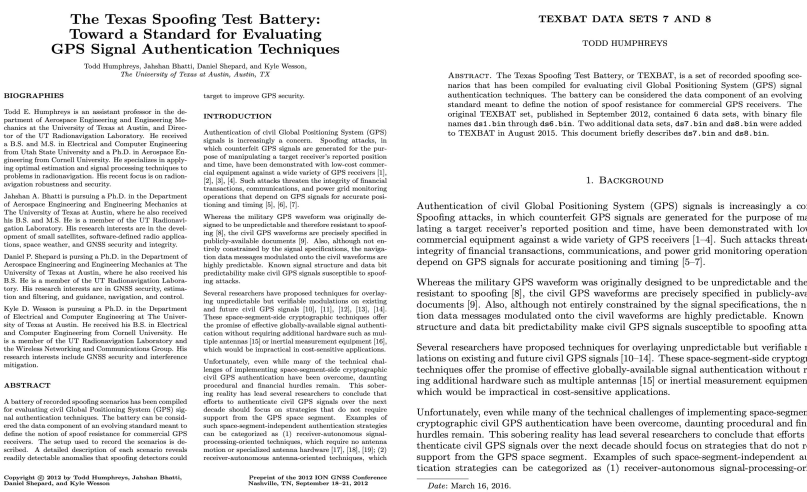
Name	Type	Description
Acq_delay_samples	double	Coarse code delay estimation, in [samples].
Acq_doppler_hz	double	Coarse Doppler estimation in each channel, in [Hz].
Acq_samplestamp_samples	uint64_t	[samples] at signal SampleStamp.
Acq_doppler_step	uint32_t	Frequency bin of the search grid, in [Hz].
Flag_valid_acquisition	bool	Acquisition status in each channel.

Name	Type	Description
Flag_valid_word	bool	Indicates the validity of the decoded word for pseudorange computation.
TOW_at_current_symbol_ms	uint32_t	Time of week of the current symbol, in [ms].
Flag_PLL_180_deg_phase_locked	bool	Indicates if the PLL got locked at 180 degrees, so the symbol sign is reversed.

Name	Type	Description
Pseudorange_m	double	Pseudorange computation in each channel, in [m].
RX_time	double	Receiving time in each channel after the start of the week, in [s].
Flag_valid_pseudorange	bool	Pseudorange computation status in each channel.
interp_TOW_ms	double	Interpolated time of week, in [ms].

Name	Type	Description
fs	int64_t	Sampling frequency, in [Hz].
Prompt_I	double	In-phase (real) component of the prompt correlator output.
Prompt_Q	double	Quadrature (imaginary) component of the prompt correlator output.
CN0_dB_hz	double	Carrier-to-Noise density ratio, in [dB-Hz].
Carrier_Doppler_hz	double	Doppler estimation in each channel, in [Hz].
Carrier_phase_rads	double	Carrier phase estimation in each channel, in [rad].
Code_phase_samples	double	Code phase in [samples].
Tracking_sample_counter	uint64_t	Sample counter as an index (1,2,3,..etc) indicating number of samples processed.
Flag_valid_symbol_output	bool	Indicates the validity of the tracking for each channel.
correlation_length_ms	int32_t	Time duration of correlation-integration, in [ms].

Dataset : TEXBAT(Texas Spoofing Battery)



The Texas Spoofing Test Battery: Toward a Standard for Evaluating GPS Signal Authentication Techniques

Todd Humphreys, Jahshan Bhatti, Daniel Shepard, and Kyle Wesson,
The University of Texas at Austin, Austin, TX

BIOGRAPHIES

Todd E. Humphreys is an assistant professor in the department of Aerospace Engineering and Engineering Mechanics at the University of Texas at Austin, and Director of the UT Radionavigation Laboratory. He received a B.S. and M.S. in Electrical and Computer Engineering from Utah State University and a Ph.D. in Aerospace Engineering from Cornell University. He specializes in applying optimal estimation and signal processing techniques to problems in radionavigation. His recent focus is on radionavigation robustness and security.

Jahshan A. Bhatti is pursuing a Ph.D. in the Department of Aerospace Engineering and Engineering Mechanics at The University of Texas at Austin, where he also received his B.S. and M.S. He is a member of the UT Radionavigation Laboratory. His research interests are in the development of small satellites, software-defined radio applications, space weather, and GNSS security and integrity.

Daniel P. Shepard is pursing a Ph.D. in the Department of Aerospace Engineering and Engineering Mechanics at The University of Texas at Austin, where he also received his B.S. He is a member of the UT Radionavigation Laboratory. His research interests are in GNSS security, estimation and filtering, and guidance, navigation, and control.

Kyle D. Wesson is pursuing a Ph.D. in the Department of Electrical and Computer Engineering at The University of Texas at Austin. He received his B.S. in Electrical and Computer Engineering from Cornell University. He is a member of the UT Radionavigation Laboratory and the Wireless Networking and Communications Group. His research interests include GNSS security and interference mitigation.

ABSTRACT

A battery of recorded spoofing scenarios has been compiled for evaluating civil Global Positioning System (GPS) signal authentication techniques. The battery can be considered the data component of an evolving standard meant to define the notion of spoof resistance for commercial GPS receivers. The setup used to record the scenarios is described. A detailed description of each scenario reveals readily detectable anomalies that spoofing detectors could target to improve GPS security.

INTRODUCTION

Authentication of civil Global Positioning System (GPS) signals is increasingly a concern. Spoofing attacks, in which counterfeit GPS signals are generated for the purpose of manipulating a target receiver's reported position and time, have been demonstrated with low-cost commercial equipment against a wide variety of GPS receivers [1], [2], [3], [4]. Such attacks threaten the integrity of financial transactions, communications, and power grid monitoring operations that depend on GPS signals for accurate positioning and timing [5], [6], [7].

Whereas the military GPS waveform was originally designed to be unpredictable and therefore resistant to spoofing [8], the civil GPS waveforms are precisely specified in publicly-available documents [9]. Also, although not entirely constrained by the signal specifications, the navigation data messages modulated onto the civil waveforms are highly predictable. Known signal structure and data bit predictability make civil GPS signals susceptible to spoofing attacks.

Several researchers have proposed techniques for overlaying unpredictable but verifiable modulations on existing and future civil GPS signals [10], [11], [12], [13], [14]. These space-segment-side cryptographic techniques offer the promise of effective globally-available signal authentication without requiring additional hardware such as multiple antennas [15] or inertial measurement equipment [16], which would be impractical in cost-sensitive applications.

Unfortunately, even while many of the technical challenges of implementing space-segment-side cryptographic civil GPS authentication have been overcome, daunting procedural and financial hurdles remain. This sobering reality has lead several researchers to conclude that efforts to authenticate civil GPS signals over the next decade should focus on strategies that do not require support from the GPS space segment. Examples of such space-segment-independent authentication strategies can be categorized as (1) receiver-autonomous signal-processing-oriented techniques, which require no antenna motion or specialized antenna hardware [17], [18], [19]; (2) receiver-autonomous antenna-oriented techniques, which

Preprint of the 2012 ION GNSS Conference
Nashville, TN, September 18–21, 2012

TEXBAT DATA SETS 7 AND 8

TODD HUMPHREYS

Abstract. The Texas Spoofing Test Battery, or TEXBAT, is a set of recorded spoofing scenarios that has been compiled for evaluating civil Global Positioning System (GPS) signal authentication techniques. The battery can be considered the data component of an evolving standard meant to define the notion of spoof resistance for commercial GPS receivers. The original TEXBAT set, published in September 2012, contained 6 data sets, with binary file names `da1.bin` through `da6.bin`. Two additional data sets, `da7.bin` and `da8.bin` were added to TEXBAT in August 2015. This document briefly describes `ds7.bin` and `ds8.bin`.

1. Background

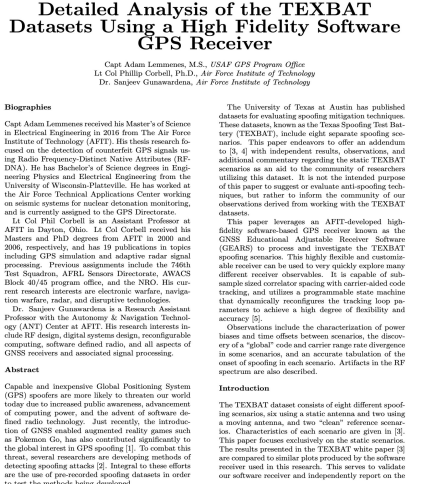
Authentication of civil Global Positioning System (GPS) signals is increasingly a concern. Spoofing attacks, in which counterfeit GPS signals are generated for the purpose of manipulating a target receiver's reported position and time, have been demonstrated with low-cost commercial equipment against a wide variety of GPS receivers [1–4]. Such attacks threaten the integrity of financial transactions, communications, and power grid monitoring operations that depend on GPS signals for accurate positioning and timing [5–7].

Whereas the military GPS waveform was originally designed to be unpredictable and therefore resistant to spoofing [8], the civil GPS waveforms are precisely specified in publicly-available documents [9]. Also, although not entirely constrained by the signal specifications, the navigation data messages modulated onto the civil waveforms are highly predictable. Known signal structure and data bit predictability make civil GPS signals susceptible to spoofing attacks.

Several researchers have proposed techniques for overlaying unpredictable but verifiable modulations on existing and future civil GPS signals [10–14]. These space-segment-side cryptographic techniques offer the promise of effective globally-available signal authentication without requiring additional hardware such as multiple antennas [15] or inertial measurement equipment [16], which would be impractical in cost-sensitive applications.

Unfortunately, even while many of the technical challenges of implementing space-segment-side cryptographic civil GPS authentication have been overcome, daunting procedural and financial hurdles remain. This sobering reality has lead several researchers to conclude that efforts to authenticate civil GPS signals over the next decade should focus on strategies that do not require support from the GPS space segment. Examples of such space-segment-independent authentication strategies can be categorized as (1) receiver-autonomous signal-processing-oriented

Date: March 16, 2016.



Detailed Analysis of the TEXBAT Datasets Using a High Fidelity Software GPS Receiver

Capt Adam Lemmenes, M.S., *USAF GPS Program Office*
Lt Col Phillip Corbell, Ph.D., *Air Force Institute of Technology*
Dr. Sanjeev Gunawardena, *Air Force Institute of Technology*

Biographies

Capt Adam Lemmenes received his Master's of Science in Electrical Engineering in 2016 from The Air Force Institute of Technology (AFIT). His thesis research focused on the detection of counterfeit GPS signals using Radio Frequency-Distinct Native Attributes (RF-DNA). He has Bachelor's of Science degrees in Engineering Physics and Electrical Engineering from the University of Wisconsin-Platteville. He has worked at the Air Force Technical Applications Center working on seismic systems for nuclear detonation monitoring, and is currently assigned to the GPS Directorate.

Lt Col Phil Corbell is an Assistant Professor at AFIT in Dayton, Ohio. Lt Col Corbell received his Masters and PhD degrees from AFIT in 2000 and 2006, respectively, and has 19 publications in topics including GPS simulation and adaptive radar signal processing. Previous assignments include the 746th Test Squadron, AFRL Sensors Directorate, AWACS Block 40/45 program office, and the NRO. His current research interests are electronic warfare, navigation warfare, radar, and disruptive technologies.

Dr. Sanjeev Gunawardena is a Research Assistant Professor with the Autonomy & Navigation Technology (ANT) Center at AFIT. His research interests include RF design, digital systems design, reconfigurable computing, software defined radio, and all aspects of GNSS receivers and associated signal processing.

Abstract

Capable and inexpensive Global Positioning System (GPS) spoofers are more likely to threaten our world today due to increased public awareness, advancement of computing power, and the advent of software defined radio technology. Just recently, the introduction of GNSS enabled augmented reality games such as Pokemon Go, has also contributed significantly to the global interest in GPS spoofing [1]. To combat this threat, several researchers are developing methods of detecting spoofing attacks [2]. Integral to these efforts are the use of pre-recorded spoofing datasets in order

The University of Texas at Austin has published datasets for evaluating spoofing mitigation techniques. These datasets, known as the Texas Spoofing Test Battery (TEXBAT), include eight separate spoofing scenarios. This paper endeavors to offer an addendum to [3, 4] with independent results, observations, and additional commentary regarding the static TEXBAT scenarios as an aid to the community of researchers utilizing this dataset. It is not the intended purpose of this paper to suggest or evaluate anti-spoofing techniques, but rather to inform the community of our observations derived from working with the TEXBAT datasets.

This paper leverages an AFIT-developed high-fidelity software-based GPS receiver known as the GNSS Educational Adjustable Receiver Software (GEARS) to process and investigate the TEXBAT spoofing scenarios. This highly flexible and customizable receiver can be used to very quickly explore many different receiver observables. It is capable of sub-sample sized correlator spacing with carrier-aided code tracking, and utilizes a programmable state machine that dynamically reconfigures the tracking loop parameters to achieve a high degree of flexibility and accuracy [5].

Observations include the characterization of power biases and time offsets between scenarios, the discovery of a "global" code and carrier range rate divergence in some scenarios, and an accurate tabulation of the onset of spoofing in each scenario. Artifacts in the RF spectrum are also described.

Introduction

The TEXBAT dataset consists of eight different spoofing scenarios, six using a static antenna and two using a moving antenna, and two "clean" reference scenarios. Characteristics of each scenario are given in [3]. This paper focuses exclusively on the static scenarios. The results presented in the TEXBAT white paper [3] are compared to similar plots produced by the software receiver used in this research. This serves to validate our software receiver and independently report on the

Scenario Designation	Spoofing Type	Platform Mobility	Power Adv. (dB)	Frequency Lock	Noise Padding	Size (GB)
1: Static Switch	N/A	Static	N/A	Unlocked	Enabled	43
2: Static Overpowered Time Push	Time	Static	10	Unlocked	Disabled	42.5
3: Static Matched-Power Time Push	Time	Static	1.3	Locked	Disabled	42.6
4: Static Matched-Power Pos. Push	Position	Static	0.4	Locked	Disabled	42.6
5: Dynamic Overpowered Time Push	Time	Dynamic	9.9	Unlocked	Disabled	38.9
6: Dynamic Matched-Power Pos. Push	Position	Dynamic	0.8	Locked	Disabled	38.9

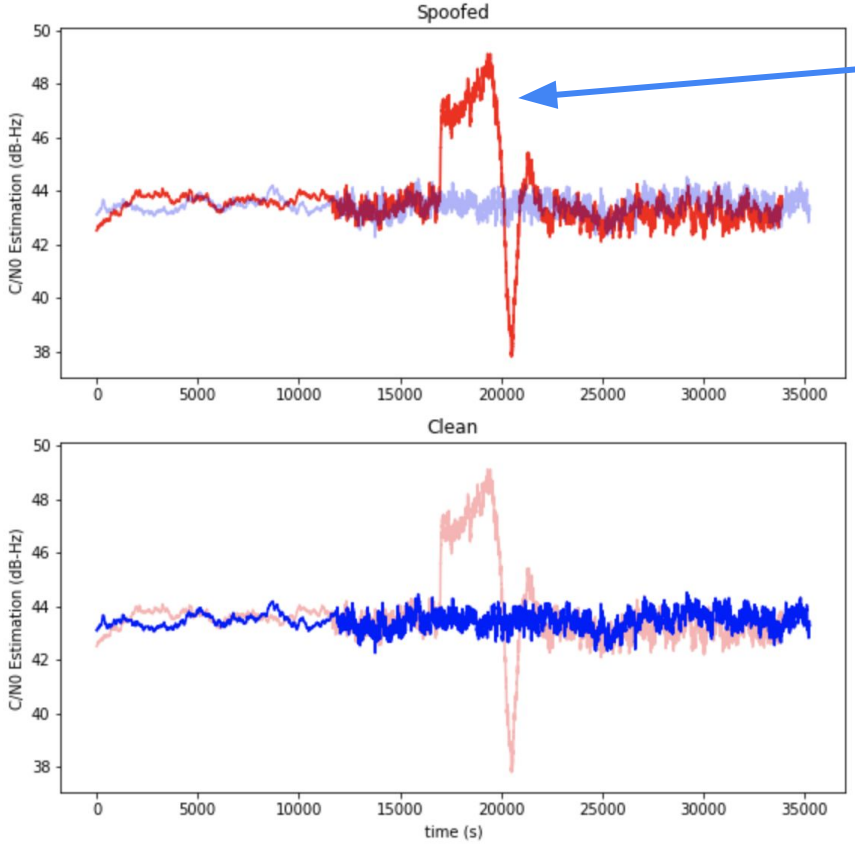
Exploratory Data Analysis in TEXBAT

The dataset is composed of binary files :
required to process it through gnss-sdr software

Based on the setup(config file), it generates log of variables from different modules.

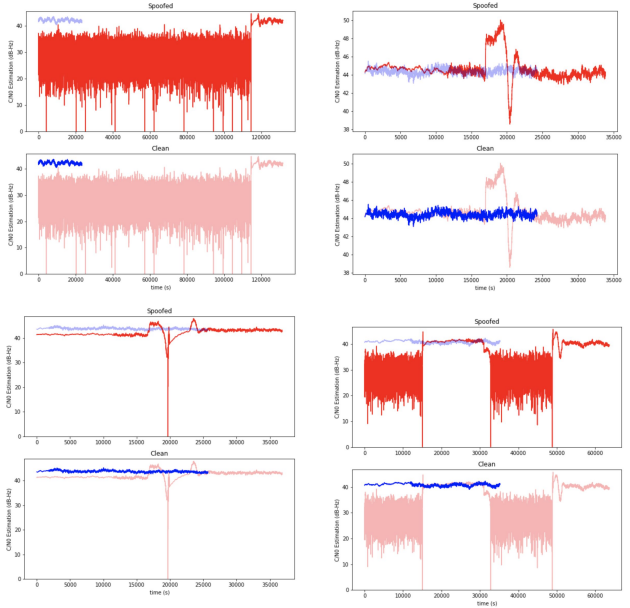
From the extracted geojson file, I visualized variables with comparison between normal vs. spoofed data.

Exploratory Data Analysis in TEXBAT - C/N0

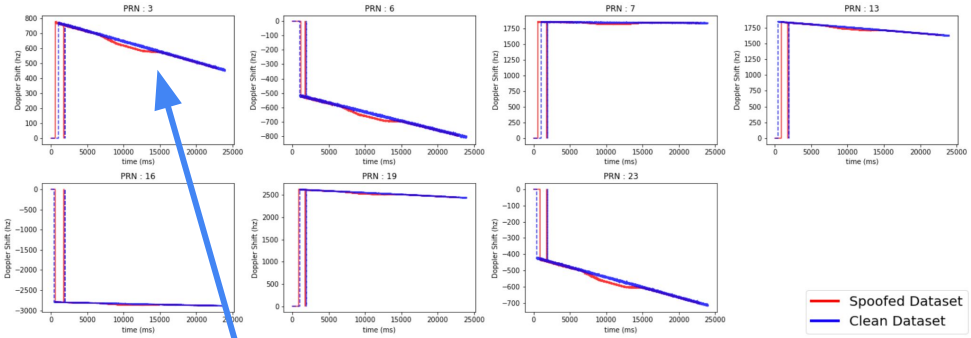


C/N0 is totally out of control in all datasets

- C/N0 is stable in the stationary status
- Shows gradual increase/decrease in motion

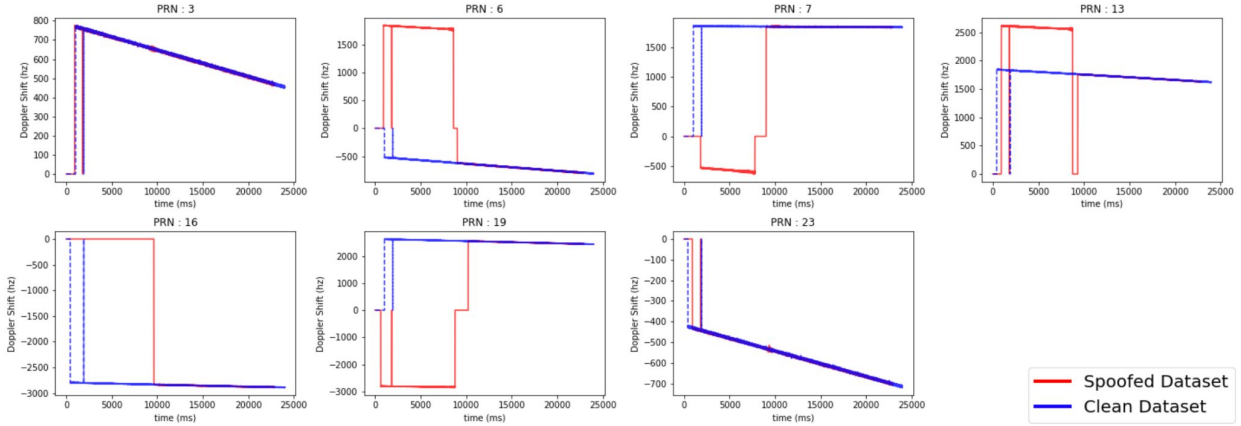


Exploratory Data Analysis in TEXBAT - Doppler Shift

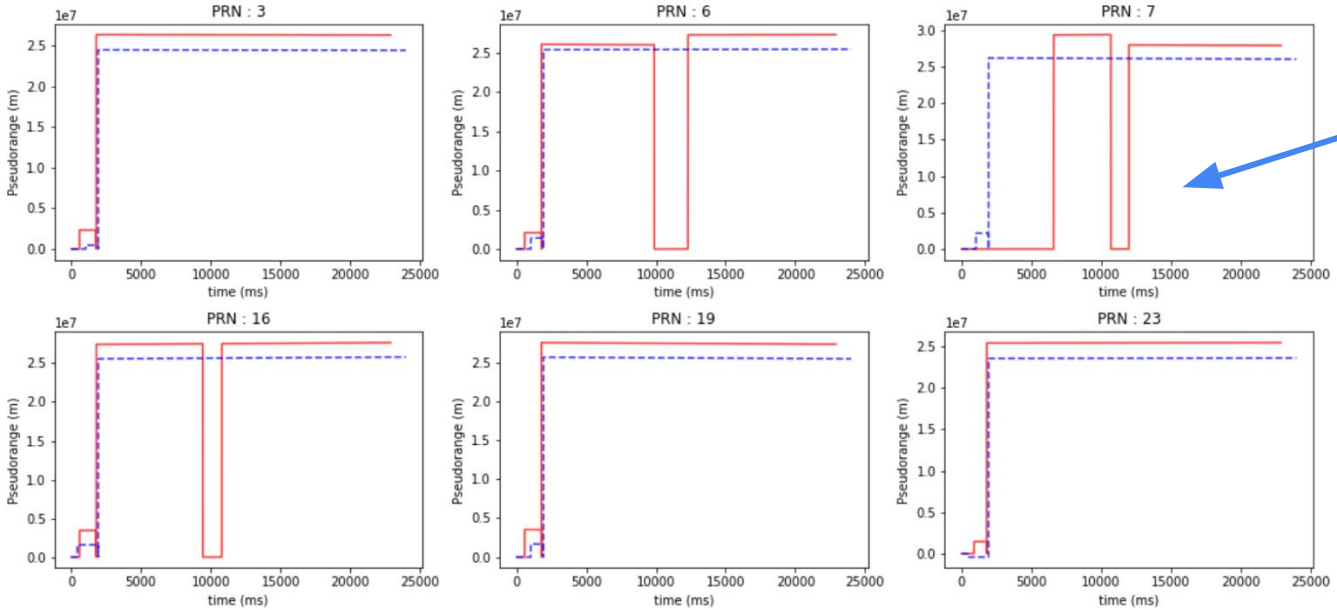


Pattern 1 :
Unexpected fluctuation

Pattern 2 :
Abrupt extreme jump/drop



Exploratory Data Analysis in TEXBAT - Pseudorange



Abrupt extreme jump/drop has been observed

Exploratory Data Analysis in TEXBAT - PVT (To be spoofed)

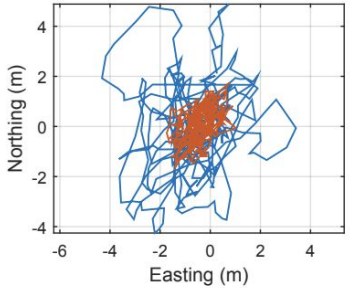


Figure 7. Scenario 2 (blue) horizontal position track overlayed on the clean scenario's track (orange) as calculated by the software receiver. The origin is located at 30°17′15.068″N, 97°44′08.642″W.

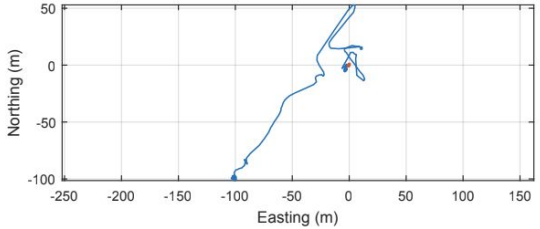


Figure 9. Scenario 4 (blue) horizontal position track overlayed on the clean scenario's track (orange) as calculated by the software receiver. The origin is located at 30°17′15.068″N, 97°44′08.642″W.

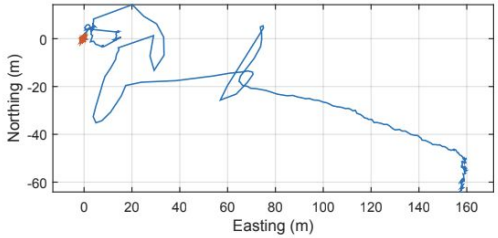


Figure 8. Scenario 3 (blue) horizontal position track overlayed on the clean scenario's track (orange) as calculated by the software receiver. The origin is located at 30°17′15.068″N, 97°44′08.642″W.

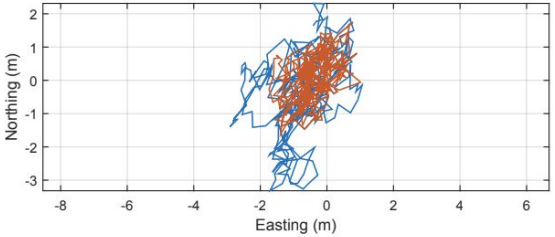


Figure 10. Scenario 7 (blue) horizontal position track overlayed on the clean scenario's track (orange) as calculated by the software receiver. The origin is located at 30°17′15.068″N, 97°44′08.642″W.

Motivation

Spoofing signal cannot be exactly the same with the original :
It’s almost impossible to make every variables as same as the original signal is while distorting the target variable(mostly PVT).

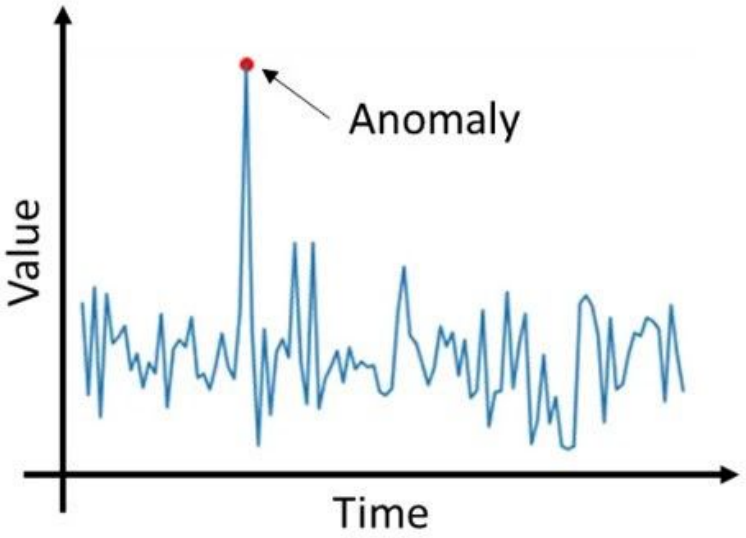
Monitoring individual variable may not work :
Aggregate multiple variables from different modules

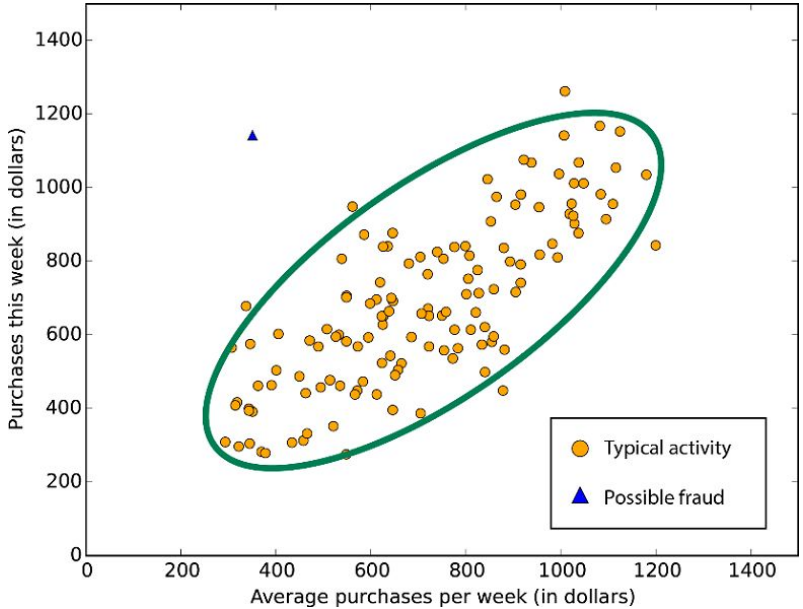
Anomaly detection typically works well on this :
Frequently used in industrial / netsec area

Variables used for the task

- prompt_i : Value of the Prompt correlator in the In-phase component.
- prompt_q : Value of the Prompt correlator in the Quadrature component.
- cn0_db_hz : C/N0 estimation, in dB-Hz.
- carrier_doppler_hz : Doppler shift, in Hz.

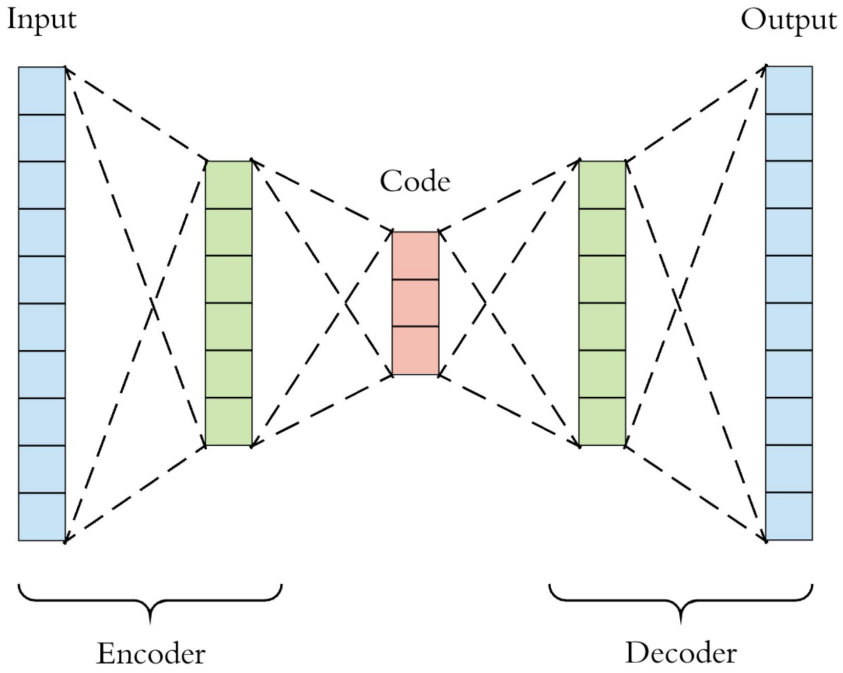
Anomaly Detection





Anomaly detection :
Detecting the abnormal pattern which doesn’t go along with its neighboring data.

Deep Learning model for the task : Autoencoder



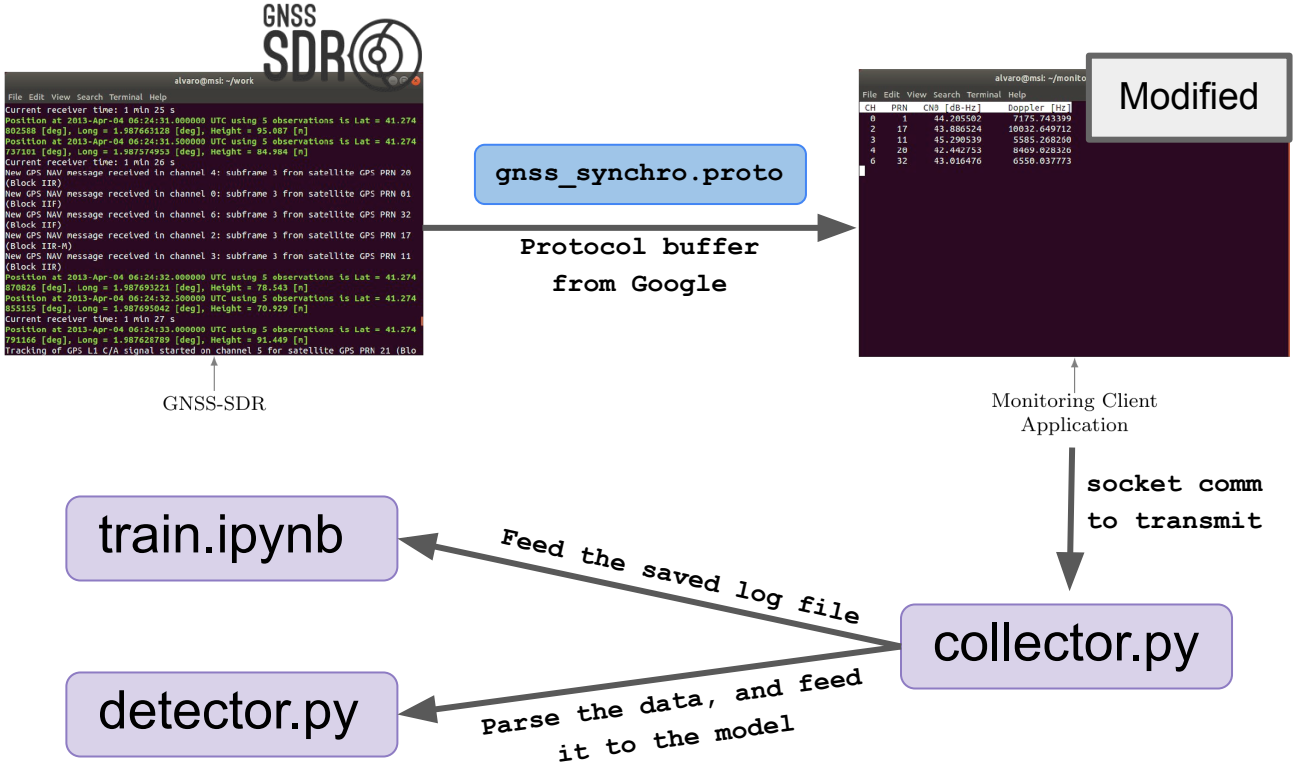
Autoencoder for Anomaly detection :
Training the model to predict as same as input

Then, the model is trained to have less error for the normal input.

On the other hand, the model has relatively huge error on abnormal pattern.

By setting the threshold of error,
it's possible to set its sensitivity.

System overview



Functions

monitoring_client.cc

+ socket communication with python code
+ select the variables to send
+ print out the streamed data

collector.py

+ receive data from the socket comm
+ parse the raw data
+ save the parsed data into csv file

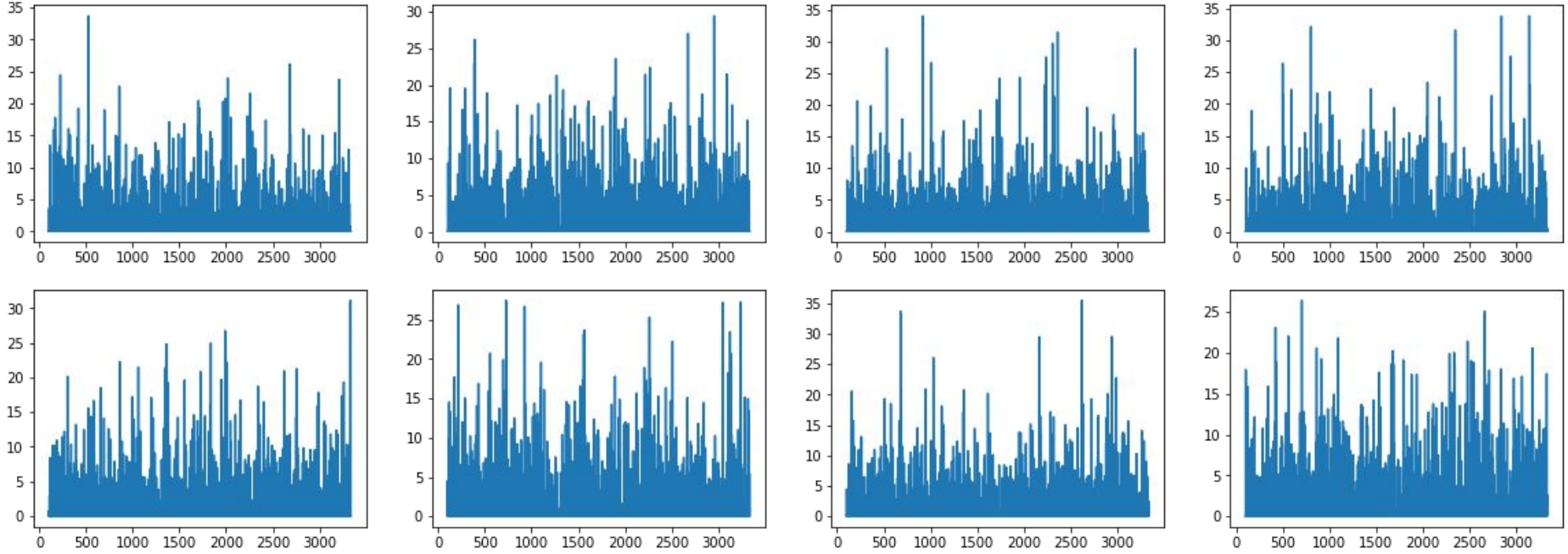
train.ipynb

+ read the data from csv file
+ preprocess data - diff()
+ normalize data - standardscaler()
+ build and train a model - tensorflow
+ save the model

detector.py

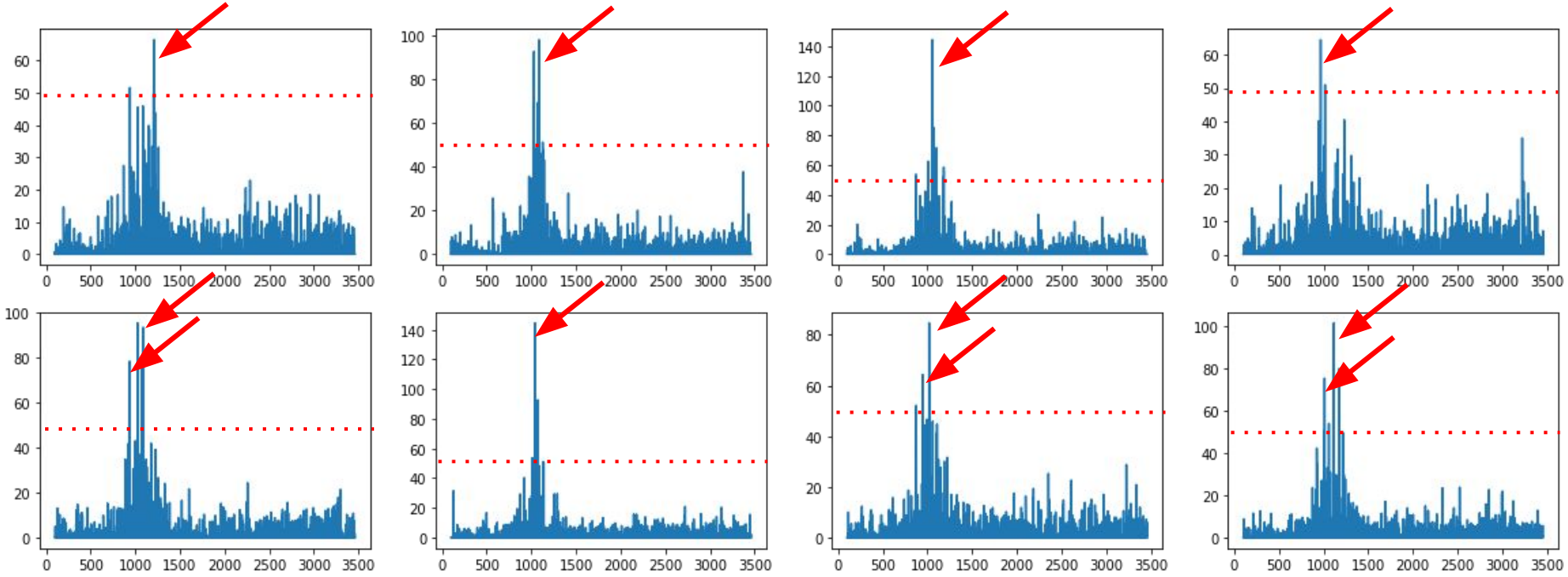
+ receive data from socket comm
+ parse the raw data
+ load and restore the model
+ feed the data to the model
+ set alert based on error

Result - errors from clean data



➡️ Set threshold = 50 based on the observed error graph above

Result - errors from spoofing data



➡️ Overall, the model detected 34 out of 40 spoofing cases (85%)

Limitation

Hard to test with real data :
Testing with spoofing signal is not easy

Limited variables :
Monitor block doesn’t have all variables from GNSS-SDR

Processing speed :
Not working robustly in real-time