Evaluation of Tracking Regimes for, and Security of PLI Systems

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Engineering

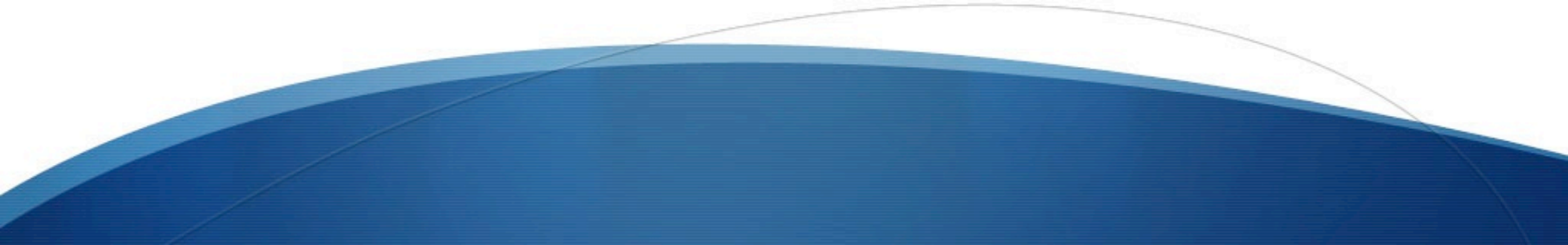
by

Shayan Taheri

Utah State University
Logan, Utah
Summer 2015

Overview

1. Introduction
2. Physical Layer Identification System
3. Tracking System Models
4. PLI System Security Evaluation
5. Conclusion
6. Acknowledgement
7. Questions



Introduction

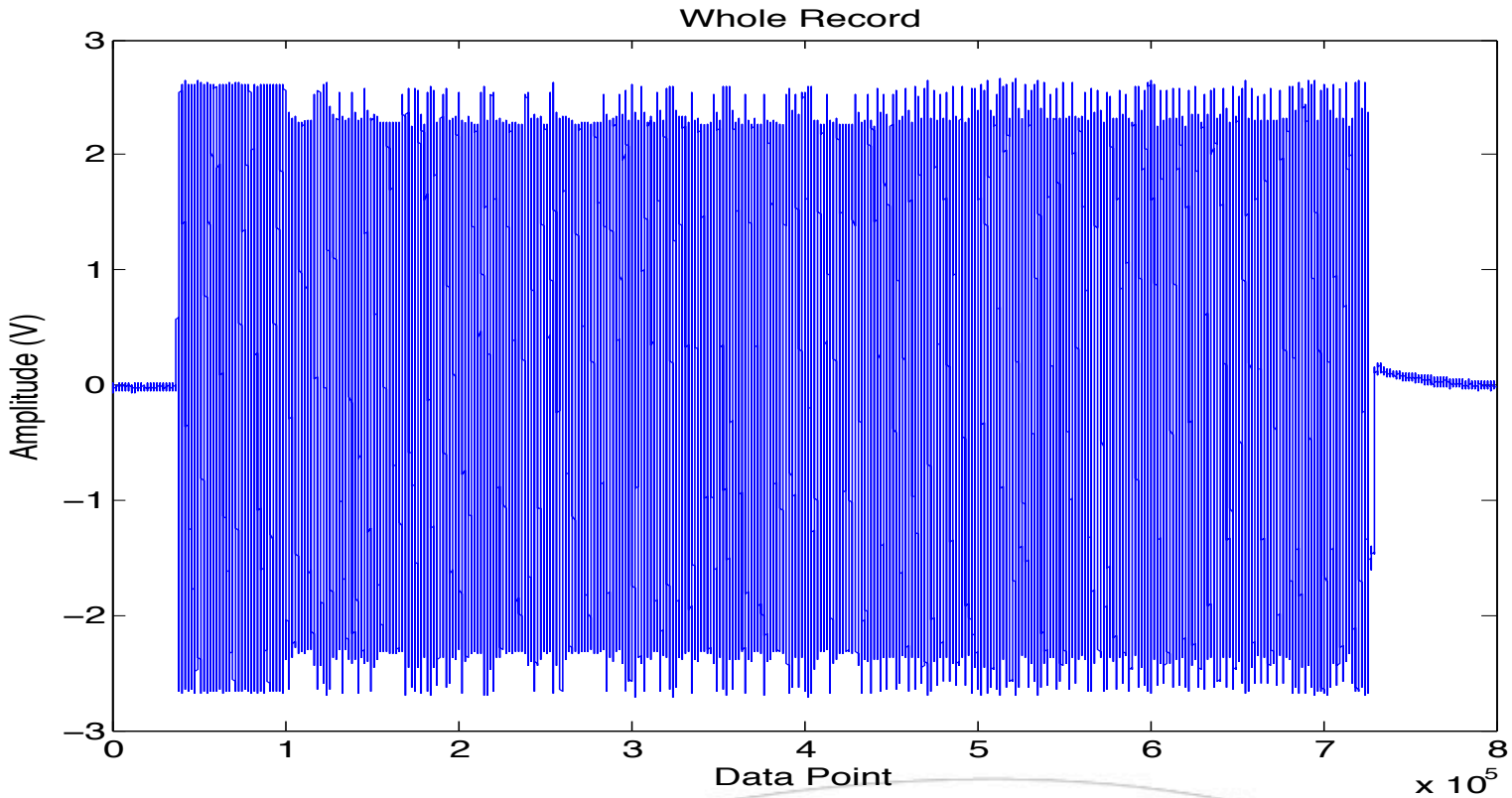
- Methods in the area of computer and network security:
  - Digital Domain
    - Flexible and controllable in design
    - **Problem**: Vulnerable in front of recent cyber attacks
    - Example: Leakage of identifiers (username and password)
    - **Solution**: Analog signaling behaviors
  - Physical Domain
    - Designing methods based on the analog representation of data
    - Unique variations for identification and monitoring purposes

Introduction …

- Physical Layer Identification (PLI)
  - Using the first layer of the OSI model for identification of the devices
- PLI Methodology
  - Identify and acquire a certain signal (i.e. Fingerprint)
  - Extract a set of meaningful features from the signal
  - Compare the test feature set with the reference feature set
  - Determination of the device identity

Introduction …

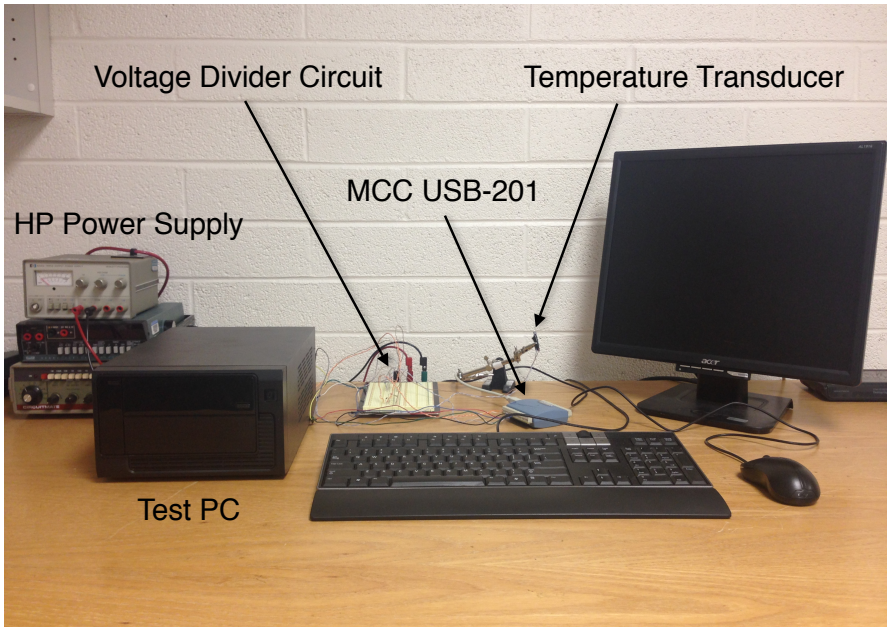
- Ethernet card record – Example for fingerprint:

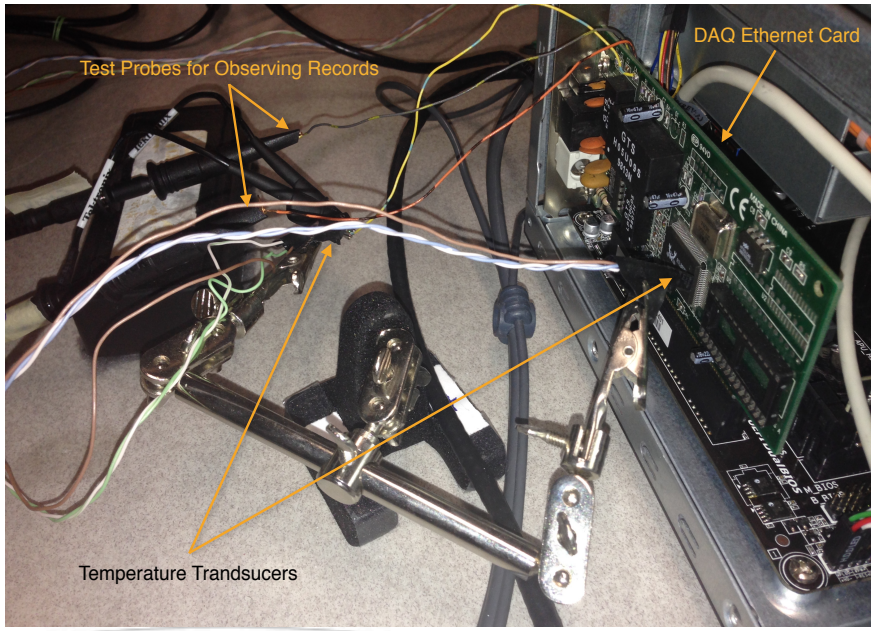


m5c1 card – Experiment 4 , Dataset 8

Introduction …

- PLI System
  - Any systematic approach for accomplishment of the PLI operation
  - Including: Equipment, Algorithm, …

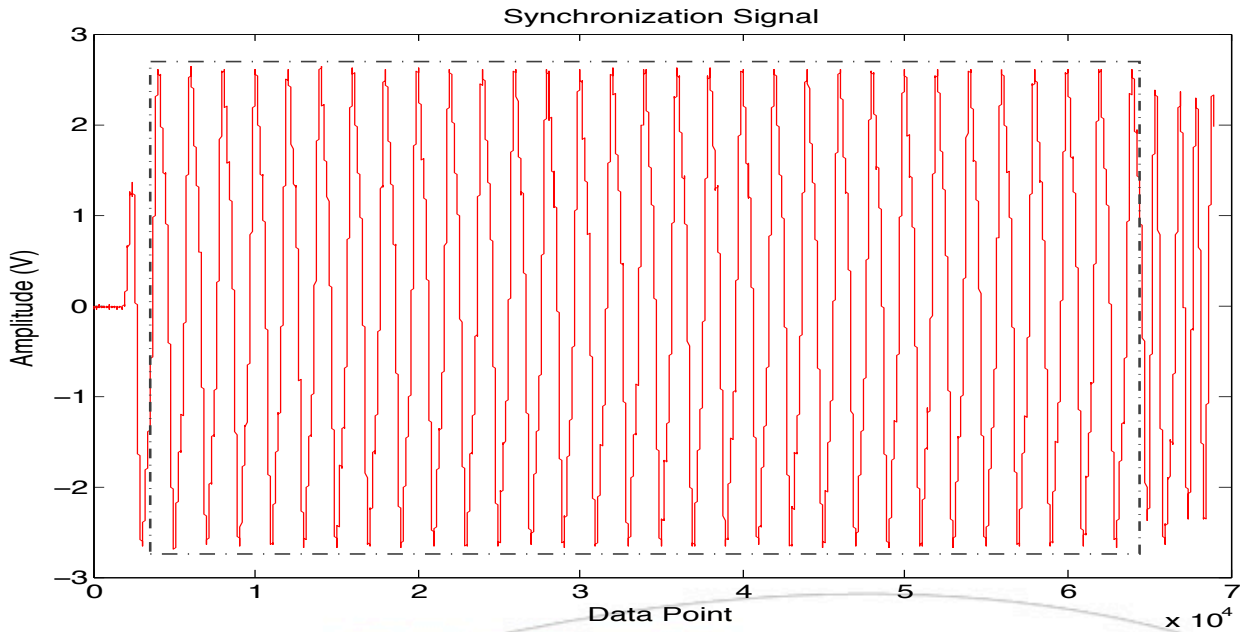




Experimental Setup for Test PC and DAQ PC

Introduction …

- Employing a PLI system, used for wired Ethernet cards:
  - Information profiles for identification of the devices
  - Using profiles for understanding the device behavior over time
  - Using the steady-state portion of the device's record for identification



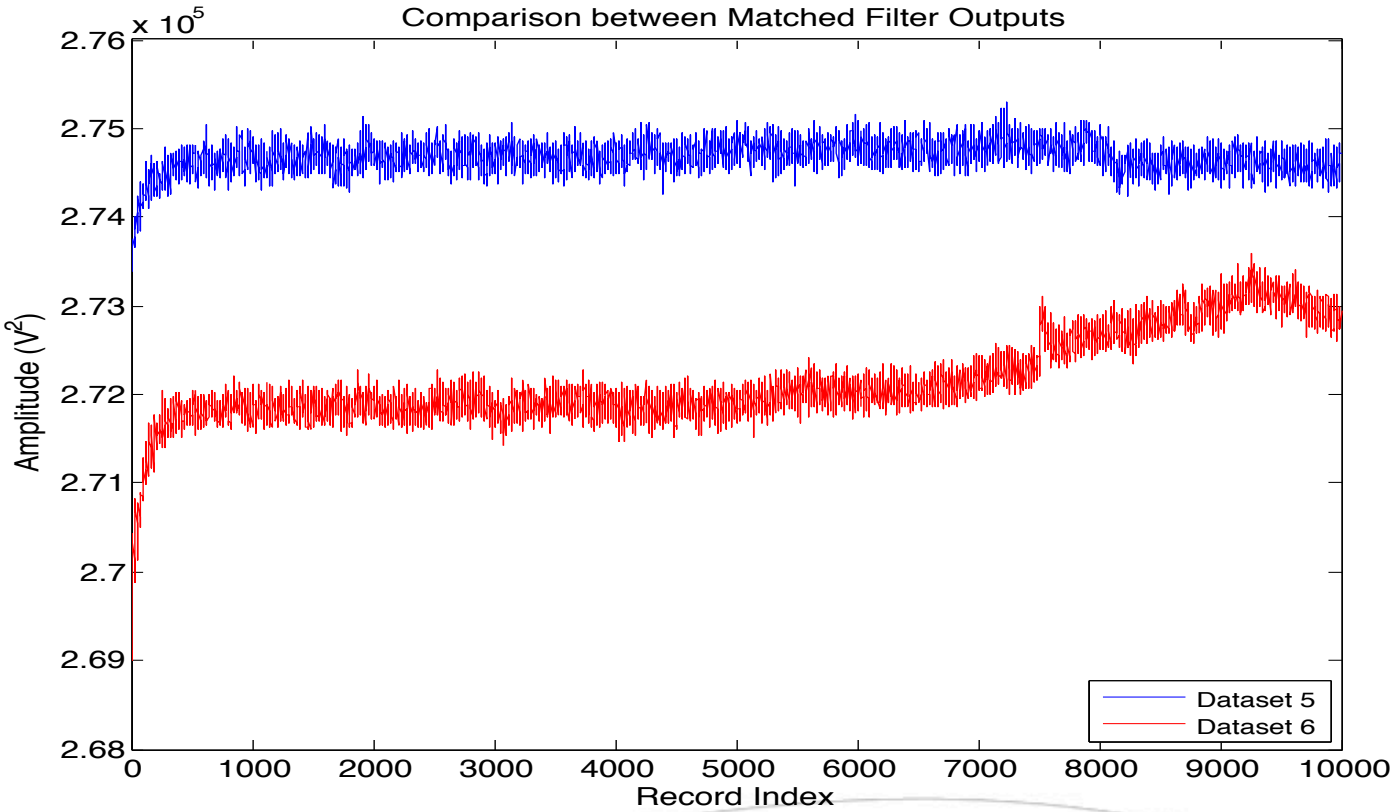
m5c1 card – Experiment 4 , Dataset 8

Introduction …

- Three main components in each information profile:
  - Matched Filter
  - Matched filter outputs for all of the collected records
  - All of the calculated threshold ranges for acceptance of the upcoming data
- The PLI system problem:
  - Lack of identification of the devices in different conditions
  - Significant changes in the features of the device's signal
  - **Solution**: Tracking these changes

Introduction …

- An example of the significant change in the device's signal



m5c3 card – Experiment 3 , Datasets 5 and 6

Introduction …

- Developing a Tracking System (a.k.a. tracking regime)
  - Capable of explaining the amount of variations of a device's signal
  - A Transductive Transfer Learning problem
  - **System performance**: Similarity between the predicted and actual data
- Security Evaluation of the PLI System
  - Exposing the system to different types of attack
  - **Attack Type**: Generating the forged version of a device's signal using an arbitrary waveform generator (AWG)

Physical Layer Identification System

1. Matched Filter
2. Adaptive Thresholding Strategy
3. APRS Metrics
4. Experimental Approach
5. Results

PLI System – Matched Filter

- Maximizing the signal-to-noise ratio of an input signal
- Transfer function of the filter:

$$H(\omega) = k.\frac{A^*(\omega)}{P(\omega)}.\exp(-j\omega t_0) \qquad h(t) = \begin{cases} \alpha(t_0 - t) & 0 \le t \le T \text{ (Period)} \\ 0 & \text{Otherwise} \end{cases}$$

- Matched Filter Output: A value that is a measurement of the closeness of the input signal to the reference signal

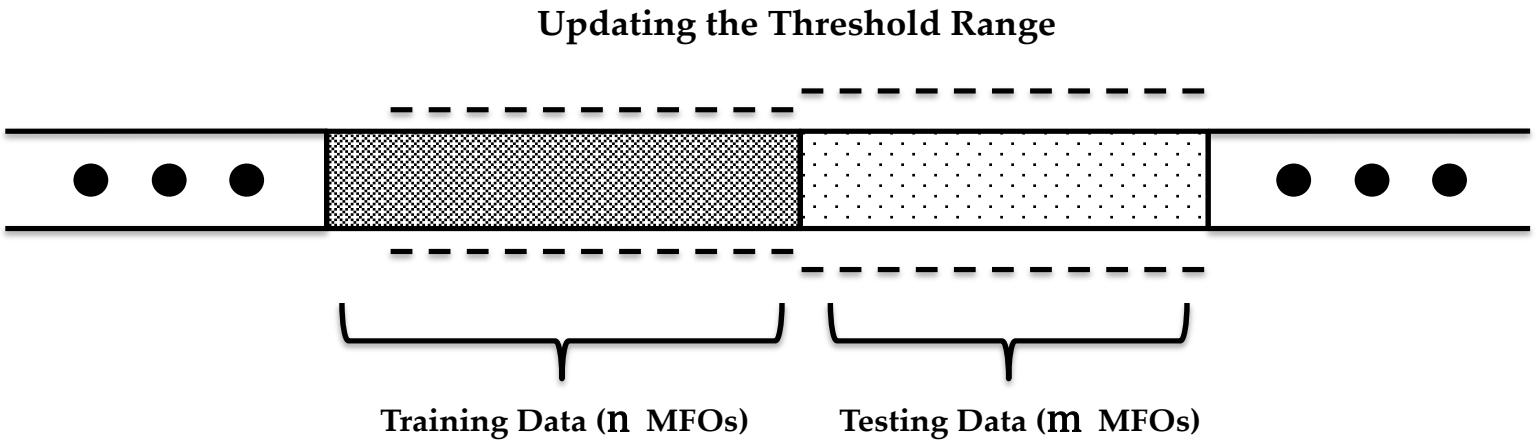
$$MFO : \lambda(t_0) = h(t_0) * \beta(t_0) = \int_{t_0-T}^{t_0} \alpha(\tau).\beta(\tau).d\tau$$

$\alpha(t)$: Reference Signal , P(ω): Power Spectral Density

$\beta(t)$: Input Signal , t_0 = Sampling Time

PLI System – Adaptive Thresholding Strategy

- Determining the threshold range for acceptance of the future signals
- Updating the threshold range periodically



$$threshold_{+/-}(MFO^{j},\ldots,MFO^{j+m-1}) =$$

$$\mu(MFO^{j-n},\ldots,MFO^{j-1}) \pm r_{(1-\gamma;m,n)} \times \sigma(MFO^{j-n},\ldots,MFO^{j-1}) \quad , \quad n < j$$

j = Record Index , r = Range Parameter

PLI System – APRS Metrics

- Metrics for evaluation of the system performance in device identification
- The result parameters for calculation of the system performance metrics:
  - **True Positive (TP)**: Correctly rejection of a record
  - **False Positive (FP)**: Wrongly rejection of a record
  - **True Negative (TN)**: Correctly acceptance of a record
  - **False Negative (FN)**: Wrongly acceptance of a record

PLI System – APRS Metrics …

- Accuracy (A), Precision (P), Recall (R), and Specificity (S):

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$S = \frac{TN}{TN + FP}$$

PLI System – Experimental Approach

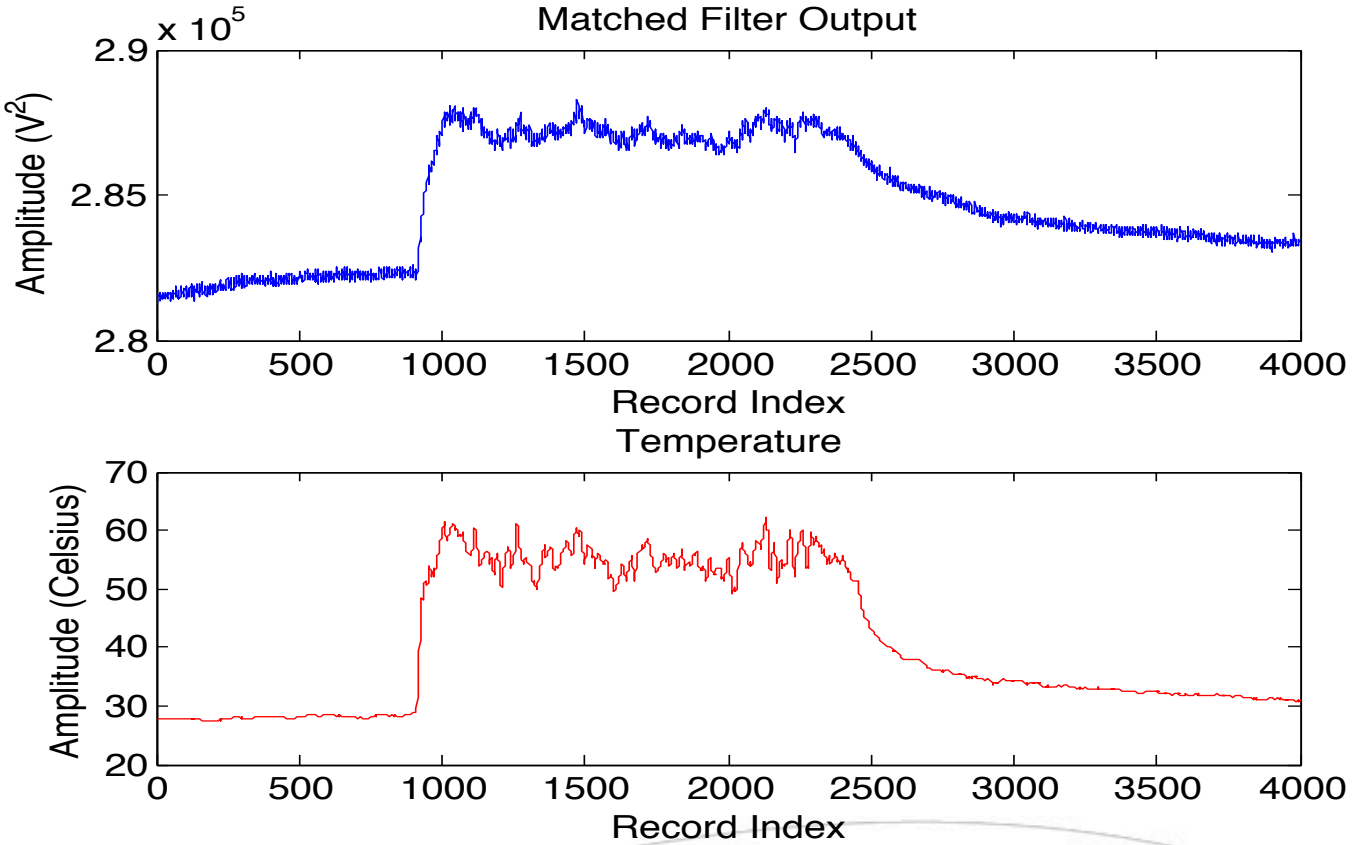
- The desired objectives for behavioral analysis of the devices:
  - **Experiment 1**: Analyzing the device behavior with respect to the **small** variations of its surrounding temperature
  - **Experiment 2**: Analyzing the device behavior with respect to the **large** variations of its surrounding temperature
  - **Experiment 3**: Analyzing the similarities and differences between the device's signal and the supplied voltages to the circuit's general bus and the mounted IC
  - **Experiment 4**: Analyzing the device behavior with respect to the **factual** variations of its surrounding temperature

PLI System – Experimental Approach …

- Data collection for five Ethernet cards
- Having two test runs for each experiment
- Acquiring the related auxiliary data along with the devices records
- Collecting 10,000 numbers of records for each device in each test run

PLI System – Results

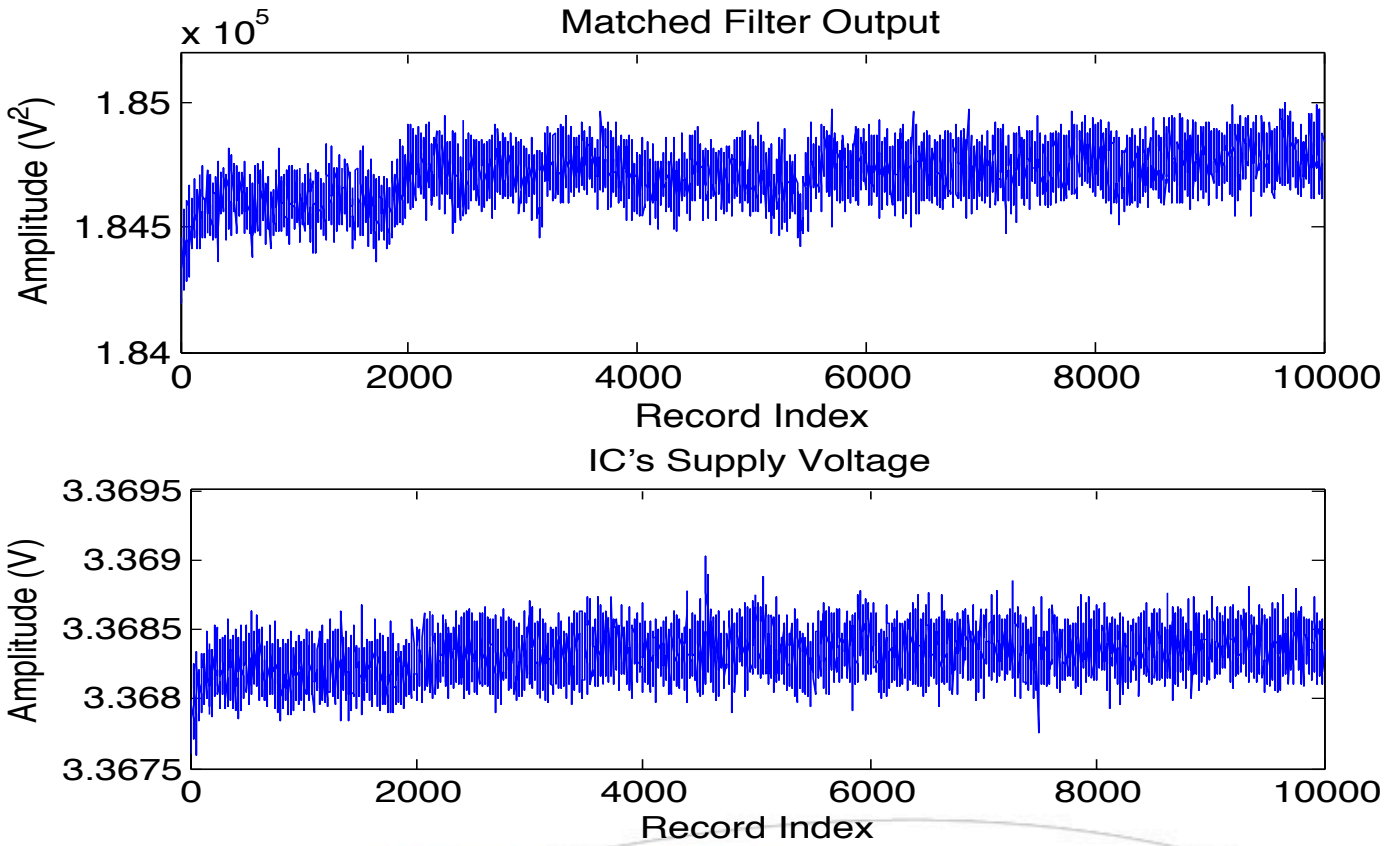
- Similarity between the matched filter output and the surrounding temperature



m5c7 card – Experiment 2 , Dataset 4

PLI System – Results …

- Similarity between the matched filter output and the IC's supply voltage



m5c1 card – Experiment 4 , Dataset 8

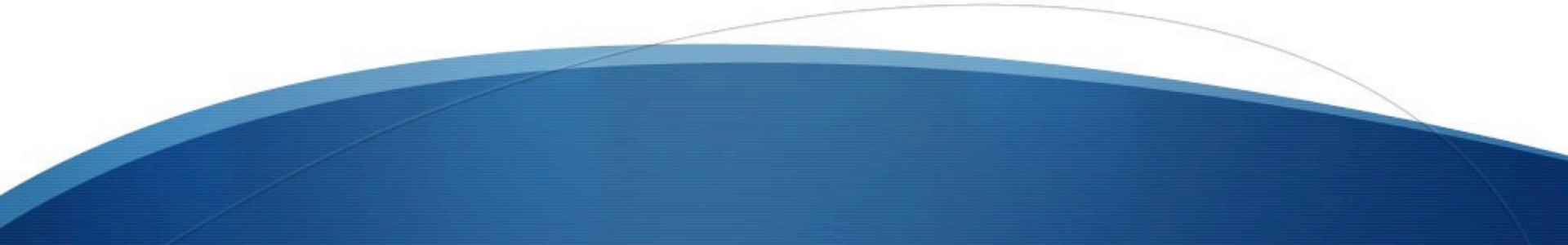
PLI System – Results …

- In Datasets 3-4: Effect of large temperature variations
- Getting a large number of false positives and low specificity values

Experiment	Dataset	A	P	R	S
1	1	0.941	1.000	0.926	1.000
	2	0.927	1.000	0.909	0.999
2	3	0.838	0.853	0.964	0.331
	4	0.832	0.870	0.932	0.431
3	5	0.938	1.000	0.923	0.999
	6	0.933	1.000	0.916	0.999
4	7	0.924	1.000	0.906	0.998
	8	0.924	0.999	0.906	0.998

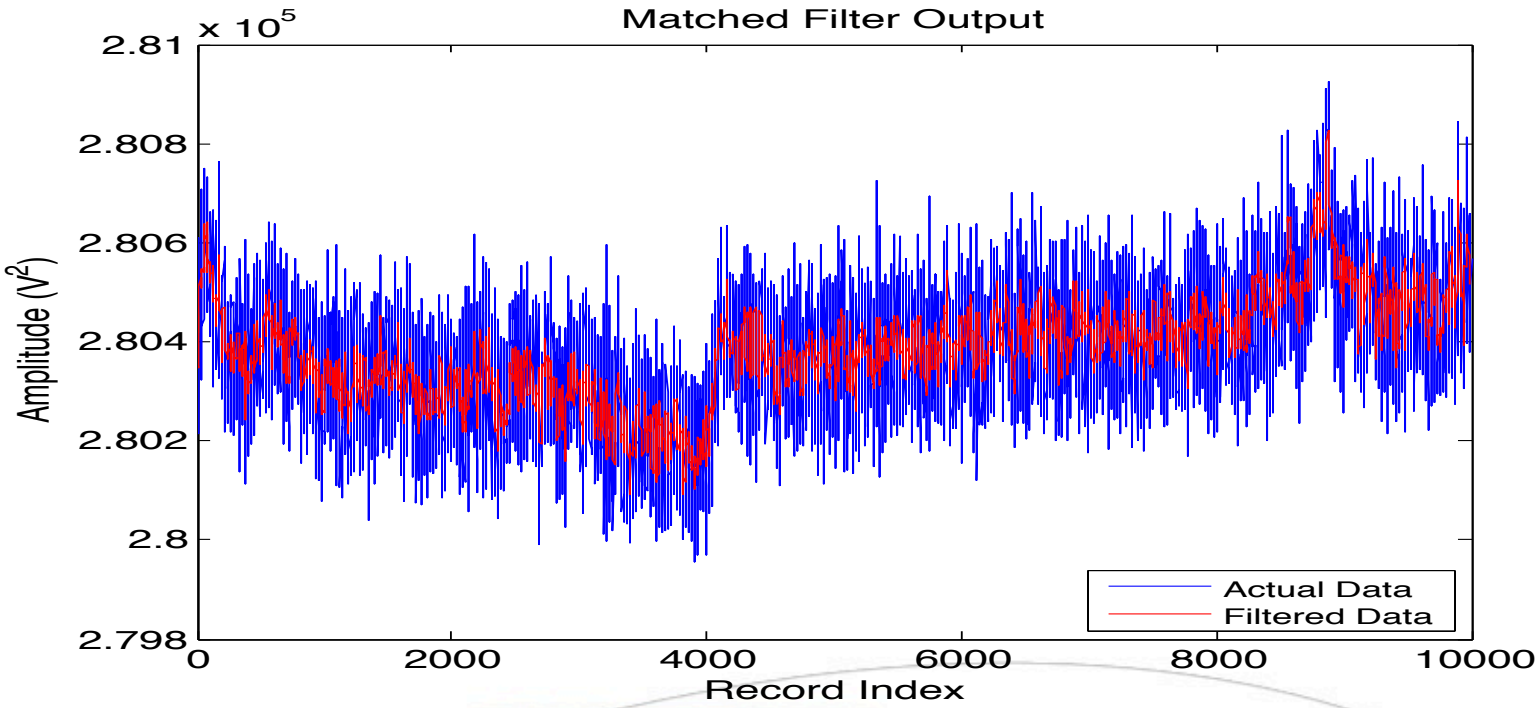
Tracking System Models

1. Data Preparation Concepts
2. Data Modeling Techniques
3. Results



Tracking System Models – Data Preparation Concepts

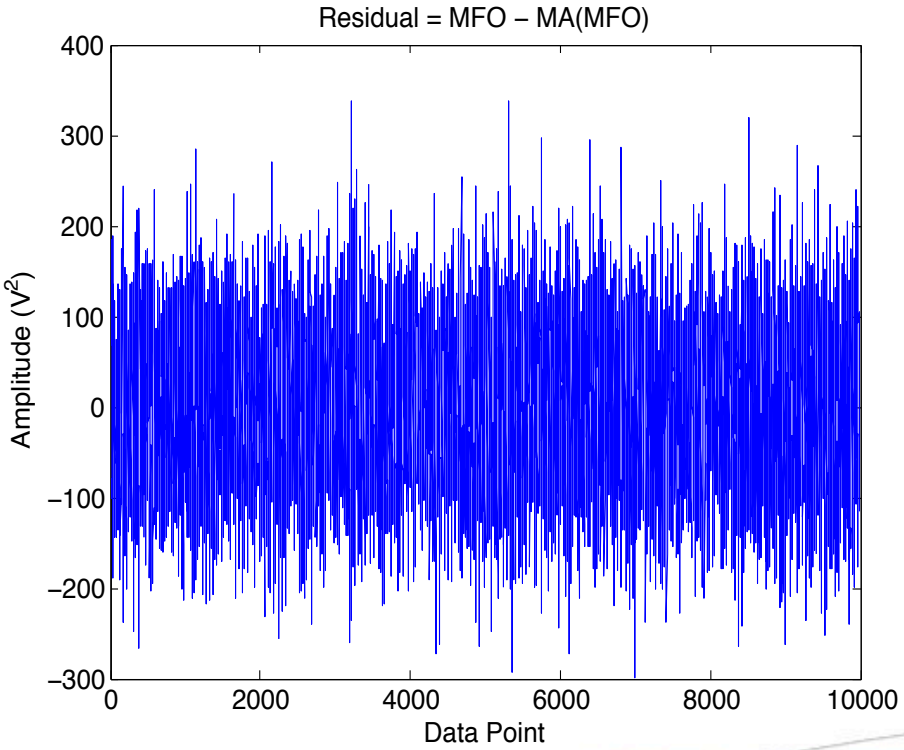
- Moving Average
  - A filter for smoothing out short-term fluctuations of a signal
  - **Filtering process**: Averaging the signal data points in equal length subsets

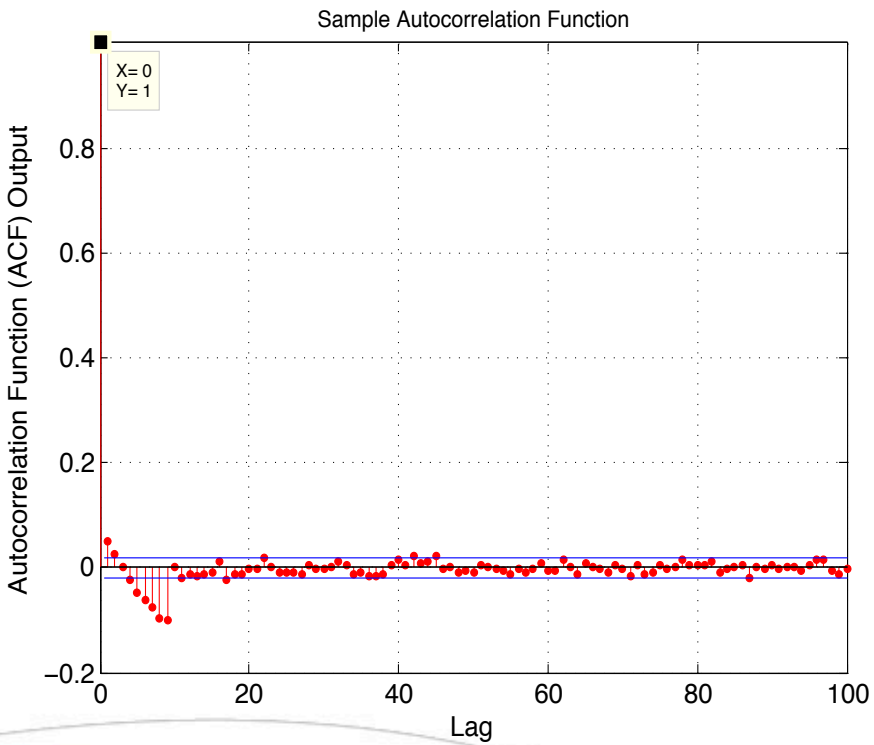


m5c2 card – Experiment 1 , Dataset 1

Tracking System Models – Data Preparation Concepts

- Moving Average
  - Eliminated noise from the signal





m5c2 card – Experiment 1 , Dataset 1

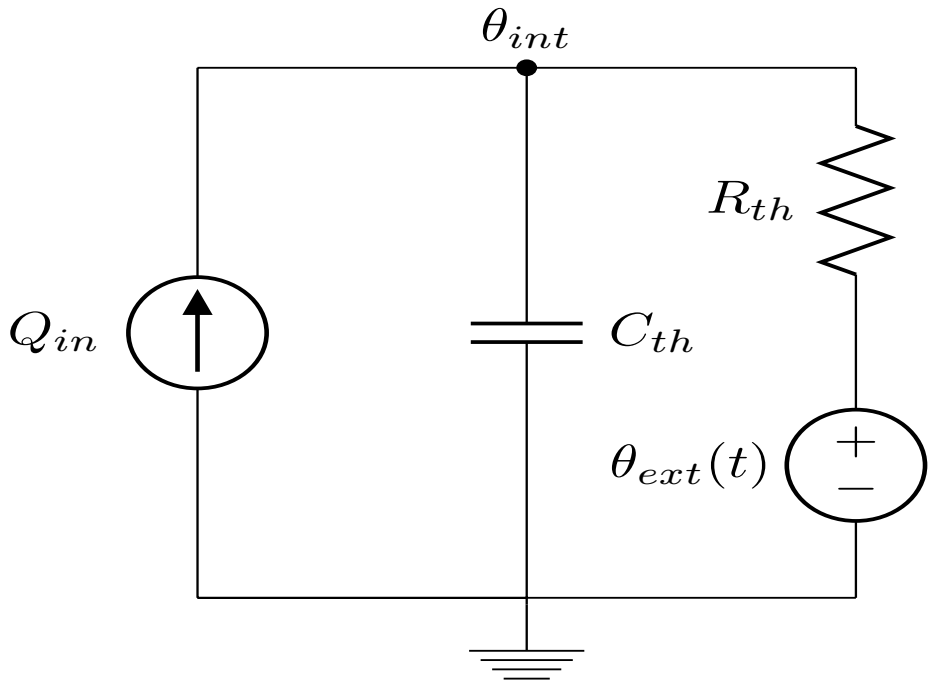
Tracking System Models – Data Preparation Concepts …

- Charging/Discharging Trend
  - MFO signal shows charging/discharging trend in the device behavior
  - **Simple Modeling**: Using the formulations of series RC circuit
  - **Advanced Modeling**: Using Thermal System Modeling
- Thermal System Modeling
  - Thermal system operation: Storing and/or transferring the heat flux
  - Three mechanisms for heat flow: **Conduction**, Convection, Radiation
  - **Energy Balance Theory**: Entered Heat = Emitted Heat + Stored Heat
  - Similar to electrical circuit modeling

Tracking System Models – Data Preparation Concepts …

Q_{in} = Heat Source , R_{th} = Thermal Resistance , C_{th} = Thermal Capacitance

θ_{int} = Internal Temperature , $\theta_{ext}(t)$ = External Temperature



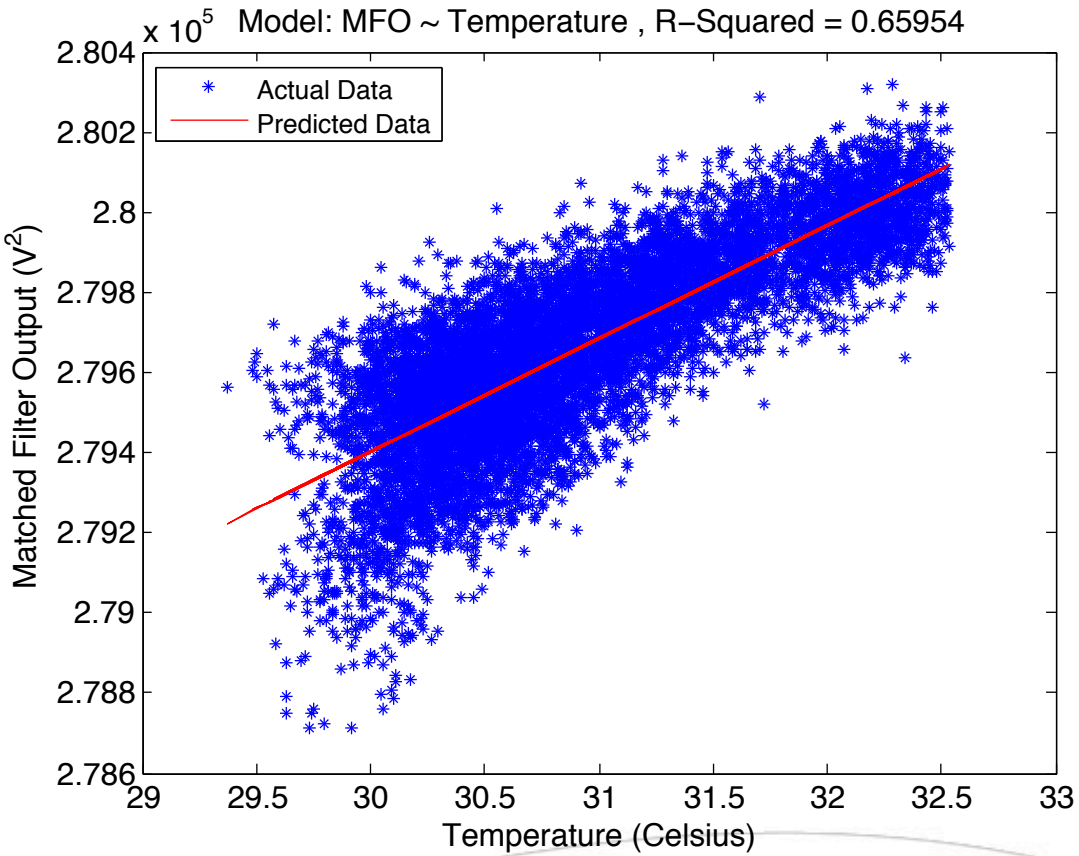
Employed Thermal System Model

Tracking System Models – Data Modeling Techniques

- Linear Regression
  - To find a linear relationship between a response variable and one or more explanatory variable(s)
  - Predicting future response data using the constructed model
  - **Residuals**: Actual Data – Predicted Data
  - **R-Squared**: To show the model accuracy in predicting data
- Smoothing Spline
  - Using as a low-pass filter for the system's output signal

Tracking System Models – Data Modeling Techniques …

- An example for scatter plot of linear regression:



m5c2 card – Experiment 1 , Dataset 1

Tracking System Models – Data Modeling Techniques …

- Robust Regression
  - Common data modeling technique
    - ✓ Having certain underlying assumptions
    - ✓ Example: Normal distribution for error data in linear regression
  - Dissatisfaction of these assumptions → Insensibility to the outliers
  - Robust Regression → Reduce the influence of outliers
  - A common method for this purpose: **Least Absolute Residuals**
    - ✓ Finding a line or curve that minimizes the absolute difference of the residuals

Tracking System Models – Results

- 50 models were tried to find the most efficient data predictive model

- Device's Surrounding Temperature

  - The best option for predicting the device behavior
  - An intruder into the device's signal
  - Trend of MFO signal ~ Trend of temperature signal
  - Larger variations of the temperature signal → Higher R-Squared value

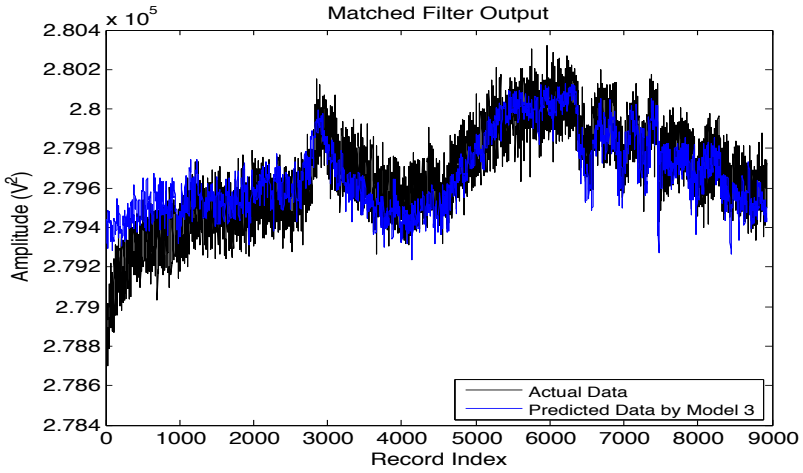
- Strengthening the data prediction process, using:

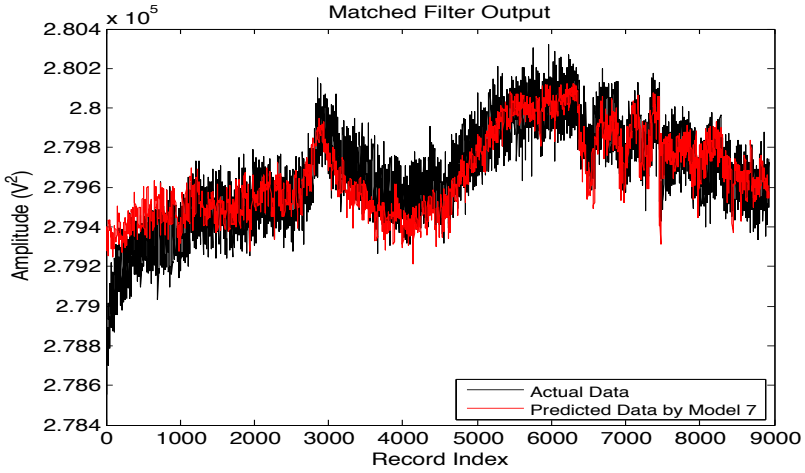
  - IC's supply voltage signal
  - Charging/discharging trend data

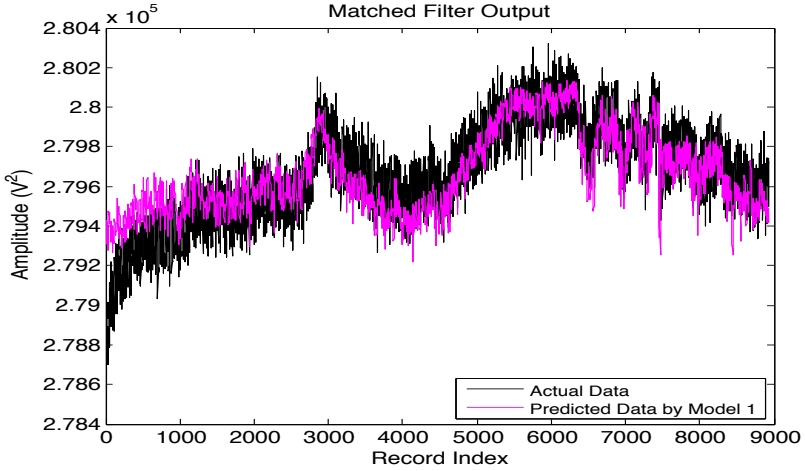
Tracking System Models – Results …

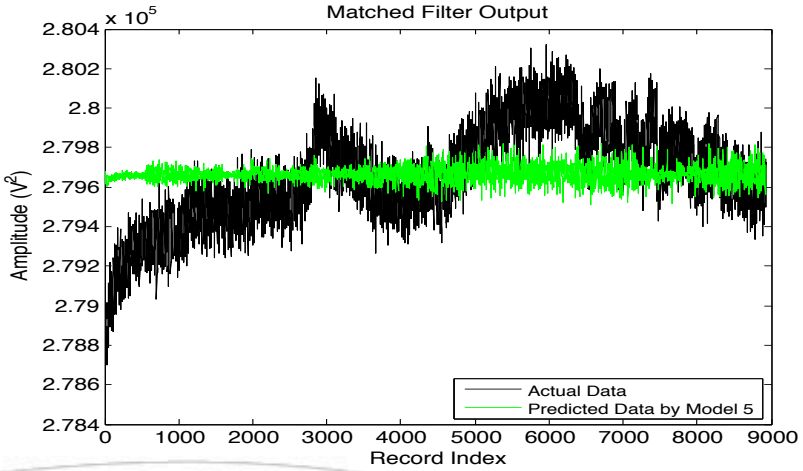
Number	Model	Experiment	R-Squared Values (%)			
			Minimum	Mean	Median	Maximum
1	MFO ~ Temperature	1	2.2955	11.0092	7.8574	24.7855
		2	88.9905	92.0434	92.518	94.4505
		3	0.0406	16.9359	6.7058	42.0475
		4	6.6724	21.3197	13.3645	38.9905
2	Moving Average(MFO) ~ Moving Average(Temperature)	1	11.251	41.7109	39.622	69.4915
3	Spline(MFO) ~ Temperature Robust Mode = Least Absolute Residuals	1	81.4355	92.5322	94.6145	97.8655
		3	44.2801	79.7687	84.657	98.6255
4	MFO ~ Device's Internal Temperature	4	6.7102	21.3772	16.6935	39.139
5	MFO ~ IC's Supply Voltage	3	0.0699	1.7399	0.5696	5.0276
		4	1.3272	6.6198	5.0743	13.7611
6	MFO ~ Bus's Supply Voltage	3	0.2070	3.0754	2.2827	8.4348
		4	0.7487	4.5355	3.9416	9.1063
7	MFO ~ Temperature + IC's Supply Voltage + Charging Trend Data	3	29.071	47.8629	44.3565	68.227
8	MFO ~ Temperature + IC's Supply Voltage + Discharging Trend Data	3	39.8315	53.1384	54.346	65.412

Tracking System Models – Results …









m5c2 card – Experiment 3 , Dataset 6

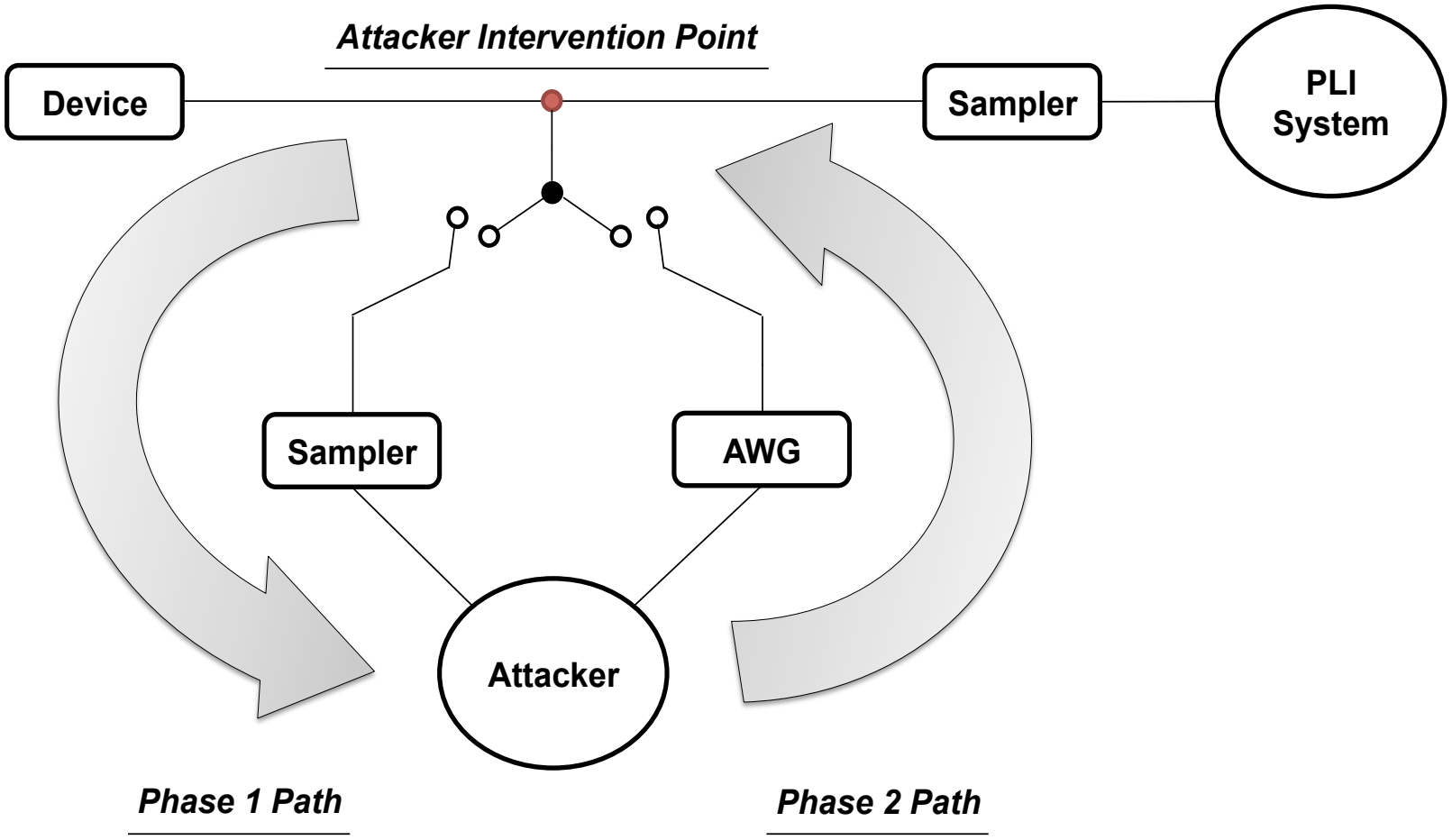
PLI Systems Security Evaluation

1. Architecture of Attack
2. AWG Configuration
3. Results

PLI System Security Evaluation – Architecture of Attack

- Attacking the PLI system by:
  - **Signal Replay**: Generating the forged version of a device's signal using an arbitrary waveform generator (AWG)
- The attacking process consists of two phases:
  - **Phase 1**: Delivering the sampled version of the device's signal to the attacker and PLI system
  - **Phase 2**:
    - ✓ Generation of a forged version of the acquired signal by the attacker
    - ✓ Sending the forged signal to the PLI system.

PLI System Security Evaluation – Architecture of Attack …



PLI System Security Evaluation – AWG Configuration

- Three main components in an AWG structure:
  - **Source Memory**: Location of the sampled version of the device's signal
  - **Digital-to-Analog Converter**: For generation of a stepped analog waveform with certain characteristics
  - **Low-Pass Filter**: For smoothing the generated signal
- Evaluating the AWG's output quality based on the DAC performance
- The most important parameters for the DAC performance analysis:
  - Settling Time
  - Resolution
  - Total Harmonic Distortion
  - Signal-to-Noise Ratio

PLI System Security Evaluation – AWG Configuration …

- Settling Time
  - The required time for the steady-state of a system's output (approximately)
  - Having inverse relationship with the sampling rate
  - Considering a very low value for this parameter

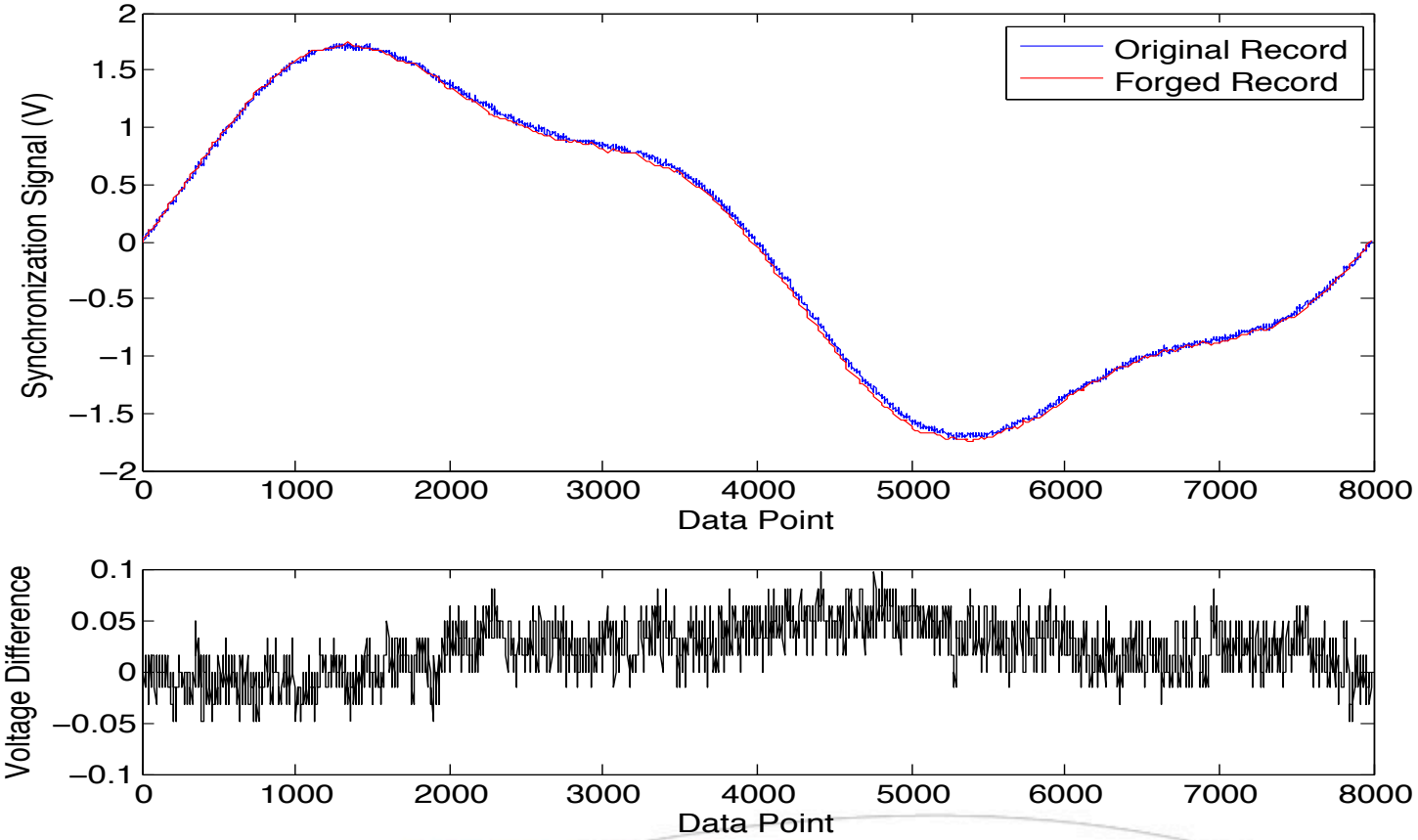
- Resolution
  - The number of discrete values for the range of values of an analog signal
  - **Discretizer**: For matching between the voltage levels of the signals

PLI System Security Evaluation – Experimental Approach

- Forged Record Generation:
  - Having 10,000 collected records for each of the 26 tested Ethernet cards
  - Selecting 25 records (i.e. 1001-1025) to be forged
- Using two different sampling rates:
  - 2.0 GS/s (the highest possible value)
  - 1.0 GS/s
- The discretizer outcome – close to the actual record voltages:
  - Maximum Voltage = 1.982 (V)
  - Minimum Voltage = -1.968 (V)
- Aligning the original and forged records before applying the matched filter

PLI System Security Evaluation – Results

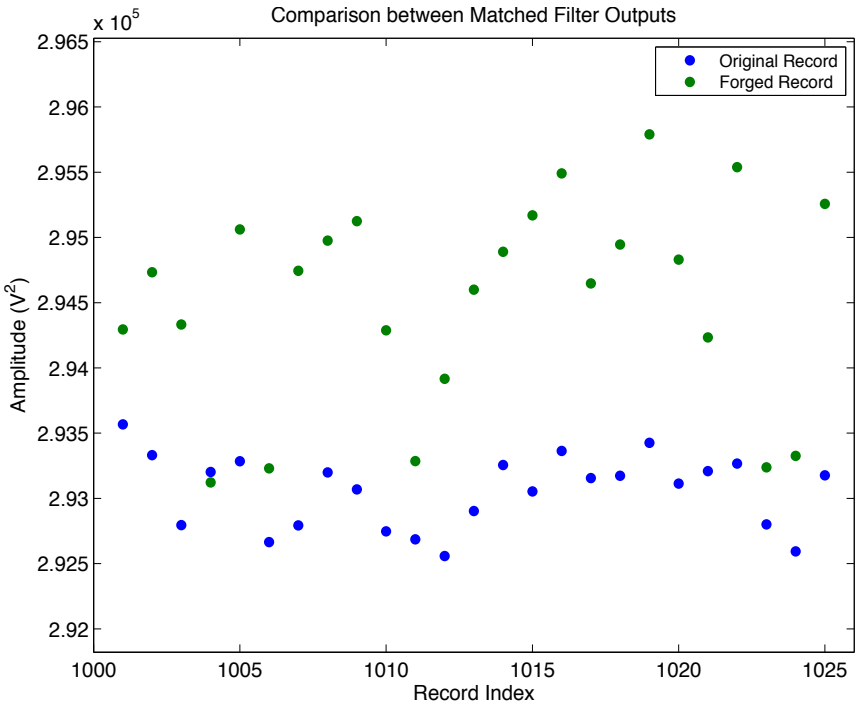
- Small difference between the original and forged records → Mean = 0.021 (V)

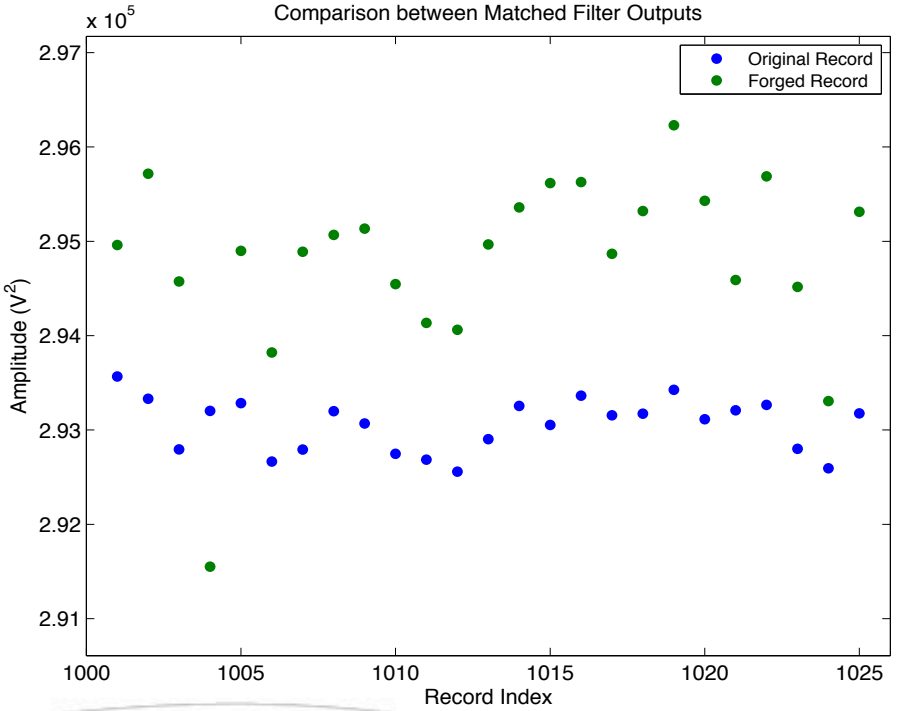


1023th record of the m5c8 card - Sampling Rate: 1.0 GS/s rate

PLI System Security Evaluation – Results …

- Forged Records → Larger expected value and dispersion of data → Defended !
- Sampling Rate Effect → Not deterministic in getting larger or smaller dispersion





m5c8 card
The generated forged records at 1.0 GS/s (left) and 2.0 GS/s (right) rates

Conclusion

- PLI System and Tracking System Models
  - Device's surrounding temperature
    - Its trend is similar to the matched filter output
    - Having large variations → Disruption in identification of the devices
    - The best option for prediction of the MFO signal
    - **Role**: Intruder into the circuit activities

Conclusion …

- PLI System Security Evaluation
  - MFO signals of forged records
    - Having larger expected value and dispersion of data
  - The PLI system isn't consistently defeatable

Acknowledgment

- My advisor: Dr. Ryan Gerdes
- The committee members: Dr. Tam Chantem and Dr. Reyhan Baktur
- My family and friends

Questions