Final Project

625.714 - Stochastic Differential Equations

Molina Nichols

Due: 14 May 2019

Contents

List of Figures

1 Introduction

Cyber physical systems are an increasingly important part of 21st century life. Whether through accidental failure or failure due to a malicious attack, errors induced in these systems can have huge ripple effects. Malicious attacks are of particular concern for public, private, and government infrastructure alike [1]. Perhaps the most well-known and the most damaging attack to date is the Stuxnet virus. Rumored to have been created by the United States and Israeli governments, Stuxnet was used to hack the cyber physical systems in Iranian nuclear facilities and gradually destroy physical components of the systems, causing great damage to the Iran nuclear program [2]. Older cyber materials or widely available software can be especially vulnerable, as hackers have access to learn about this software more easily and attacks on them may be widely known in online communities. Such a system on the Bowman Avenue Dam in Rye Brook, New York was attacked by Iranian hackers in 2013, though the attack was not disclosed by the United States government until 2015 [3]. Though the hackers did not cause damage and were nowhere near as sophisticated as the perpetrators of Stuxnet, the attack still served as a warning sign for ensuring the safety of cyber physical systems. Systems of most concern and value, especially when supporting public infrastructure, are those such as water supply and sanitation, electric grids, and pipelines. The threat Dragonfly, believed to be a state-sponsored attack from Eastern Europe, particularly targets such resources, as well as human exploitation in defense industries to attempt to hack physical systems [4].

Due to the high value and vulnerability of important cyber physical systems, various research efforts are continually underway across the public, private, and government sectors to develop solutions to detect and understand attacks. Some of the most common efforts include development of intrusion detection systems (IDS), enhanced authentication schemes [5], and understanding of network traffic patterns [6]. A proposed IDS in [7] will be examined and implemented in this project.

2 Proposed Intrusion Detection System

An intrusion detection system for a power grid is proposed in [7]. This approach could certainly be extended to other cyber physical systems, but we will consider the power grid formulation as an example to demonstrate cyber attacks on the workings of the grid.

First, we must understand the power grid. Sensors are installed throughout the cyber physical system and can measure various quantities about the system. These measurements are then used to determine if the system is operating as expected, or if anomalous behavior has been induced by an accident or by an attack. Once a decision has been made, the appropriate recovery systems or personnel can take action to recover normal system behavior.

The authors assert that an attack or accidental fault in the power grid will always be seen through a change in voltage, current, or phase on the system [7] [8]. By using a Kalman Filter [9], the anomalous behavior can be detected within these quantities. A Kalman Filter estimates these quantities given sensor measurements and is used extensively in applications for estimating and tracking dynamic processes. The Kalman Filter uses historical measurements as it steps through time in order to create its estimates. Thus, it is especially useful for this application because it is able to resist much of the noise seen in the regular measurements, particularly when there is an attack.

We now examine how to formulate these quantities over the power grid in Section 2.1, the construction of the intrusion detection system using a Kalman filter in Section 2.2, and the implementation adapted from [7] in Section 3.

2.1 Power Grid Formulation

Consider voltage to be the quantity to be measured and tracked on the power grid. Voltage sensors will be used to measure the state variables of the voltage - amplitude and phase. The measurements from these sensors can be represented as sinusoids. The single phase voltage signal can be represented as

$$V(t) = A_v cos(\omega t + \phi) \tag{1}$$

where A_v is amplitude, ωt is angular frequency, and ϕ is phase [7].

Using the assumptions stated in Section III.A of [7], we expand (1), assume that angular frequency is nearly constant over time, and obtain

$$V_1(t) = x_1 cos(\omega t) - x_2 sin(\omega t) \tag{2}$$

where $x_1 = A_v cos(\phi)$ and $x_2 = A_v sin(\phi)$. Note that I have removed A_v^* and A_v° to write simply A_v in these expressions. It is not evident in [7] how A_v^* and A_v° are derived from A_v, and I examined the effects of A_v on the voltage signal before making this assumption. A_v affects only the amplitude of the signal, and does not distort it in any other way. Thus if we are consistent throughout the formulation, simplifying A_v does not cause any computational inaccuracies, but would need to be resolved if this formulation were to be used on a true system.

For clarity in the derivation, we rewrite (2) with x_1 and x_2 substituted into the expression

$$V_1(t) = A_v cos(\phi) cos(\omega t) - A_v sin(\phi) sin(\omega t) \tag{3}$$

We see that an equivalent formulation for this expression is

$$V_1(t) = \begin{bmatrix} cos(\omega t) & -sin(\omega t) \end{bmatrix} \begin{bmatrix} A_v cos(\phi) \\ A_v sin(\phi) \end{bmatrix} \tag{4}$$

Substituting x_1 and x_2 back into the expression, we obtain

$$V_1(t) = \begin{bmatrix} cos(\omega t) & -sin(\omega t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{5}$$

To model true system behavior, we must assume that there is some small amount of measurement error in the voltage sensors. Thus we obtain an expression for voltage signal from the sensors, including measurement noise $\gamma(t)$.

$$y(t) = \begin{bmatrix} cos(\omega t) & -sin(\omega t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \gamma(t) \tag{6}$$

2.2 Kalman Filter Formulation

To implement a Kalman Filter, we must first write the equation for the state variables x_1 and x_2. From [7] we have

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + w(t) \tag{7}$$

where $w(t)$ is the process noise in the system. We can rewrite this equation using $x(t)$ to represent x_1 and x_2 together as follows

$$x(t+1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x(t) + w(t) \tag{8}$$

Letting $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, we have

$$x(t+1) = Ax(t) + w(t) \tag{9}$$

as our state equation.

In a similar fashion, we must write the equation for the measurements provided by the voltage sensor; these measurements will be fed into the Kalman filter. Adapting (6) [7] we obtain

$$y(t) = C(t)x(t) + v(t) \tag{10}$$

where $y(t)$ are the measurements from the voltage sensors, $C(t) = \begin{bmatrix} cos(\omega t) & -sin(\omega t) \end{bmatrix}$, and $v(t)$ is the measurement noise on the sensor. We will define $v(t)$ to be Gaussian with mean $\mu = 0$ and standard deviation σ [7].

We may now write equations for each of the steps of the Kalman filter. First is the time update, consisting of updating the state x and covariance P

$$\hat{x}(t+1|t) = A\hat{x}(t) \tag{11}$$

$$P(t|t-1) = AP(t-1)A^T + Q \tag{12}$$

where A is the $2x2$ identity matrix as stated previously, Q is the process noise covariance matrix, and the T superscript indicates the matrix transpose.

Second is after receiving a voltage measurement from the sensors on the system, we update the state and covariance using the new measurement information incorporated into the Kalman gain matrix, K.

$$K(t) = P(t|t-1)C(t)^T(C(t)P(t|t-1)C(t)^T + R)^{-1} \tag{13}$$

$$P(t) = P(t|t-1) - K(t)C(t)P(t|t-1) \tag{14}$$

$$\hat{x}(t) = \hat{x}(t|t-1) + K(t)(y(t) - C(t)\hat{x}(t|t-1)) \tag{15}$$

where C is as given in (10), R is the measurement noise covariance matrix, and the -1 superscript indicates the matrix inverse.

We have now full fully formulated the voltage equations, voltage measurements, state equations, and Kalman filter process.

3 Implementation and Results

We must first set up the power grid, using the parameters stated in [7]. We have a sampling frequency of the sensor at 2 kilohertz, frequency of the voltage signal at 60 hertz, and amplitude of 1 volt. Phase was not stated, so I chose 1 for simplicity. After several runs it was determined that 2 kilohertz was a quite high sample rate for effective runtime on a laptop, so the sampling frequency was lowered to 1 for the results following.

Timesteps in the simulation are determined by the frequency at which the sensor provides a measurement, so we must convert the sampling frequency of 1 kilohertz to samples per second. Then, we may generate timesteps within a given internal. At each of these time steps, we generate a measurement according to (10), to which we add measurement noise according to a Gaussian distribution with a mean of 0 and variance of 1. We now have simulated the measurements which are received from the sensor. These measurements can be seen in Figure 1, and in more detail in Figure 2 where the sinusoidal shape can be seen.

We now proceed to the Kalman Filter time loop. We initialize the state of the Kalman filter as stated on page 375 of [7]. Values for the matrices Q and R were not stated in [7], so Q was chosen to represent a fairly accurate sensor with $Q = \begin{bmatrix} .01 & .01 \\ .01 & .01 \end{bmatrix}$ and R chosen for simplicity as $R = 1$. Stepping through the timesteps we previously generated, we run the Kalman filter according to (11), (12), (13), (14), (15), saving off the state x and covariance matrix P from the Kalman Filter at each timestep. At the end, we use (10) to calculate the voltage value at each timestep based

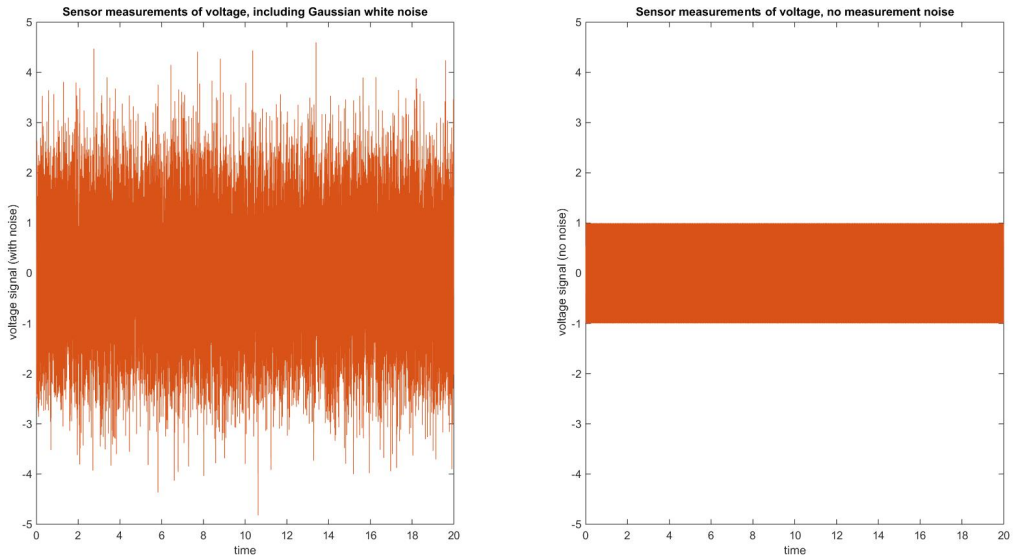

Figure 1: Voltage measurements.

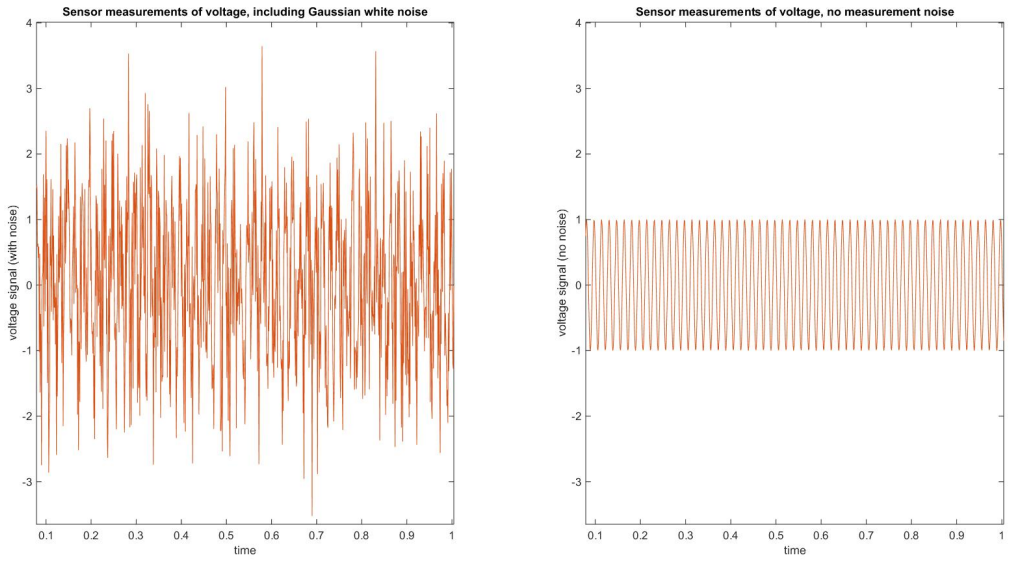

Figure 2: Voltage measurements showing sinusoidal detail.

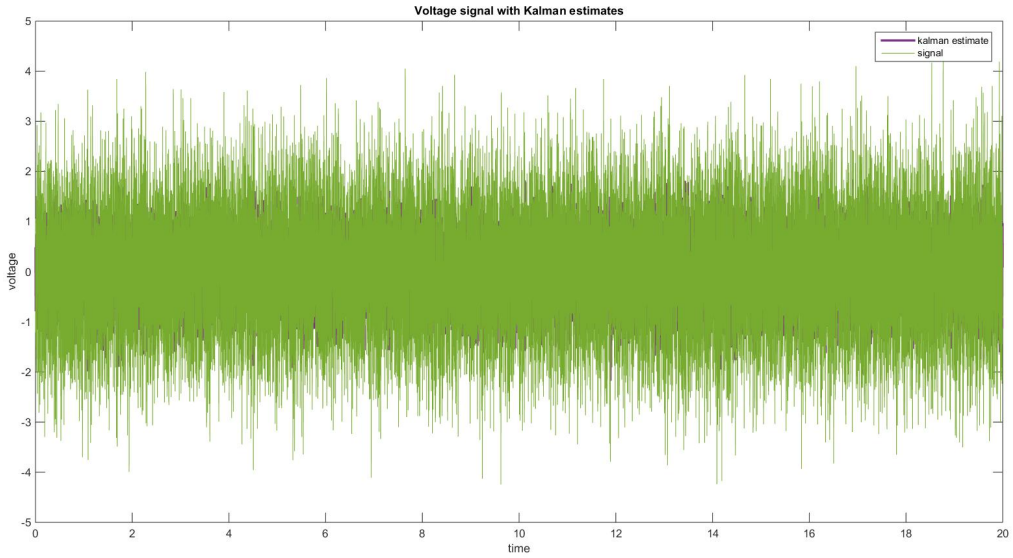

Figure 3: State estimates using Kalman Filter.

on the Kalman estimates. Figure 3 shows the noisy sensor measurements along with the Kalman Filter estimate for each timestep, with Figure 4 showing more detail of how the Kalman Filter resists noise.

Now consider an instance of a random noise attack. We add randomly generated numbers to the true measurements coming from the voltage sensor. We see that the Kalman filter does not perfectly maintain the previously non-noisy sinusoidal voltage, but it does remain closer to those previous values than the extremely noisy measurements now coming from the attacked sensor. If we were to calculate the simple difference between the Kalman values and the sensor values, we could set a threshold after which the disparity would be so large that we would suspect an attack or fault in the system. A more sophisticated approach is described in [7], using a χ^2 test to compare the values. The random noise attack can be clearly seen in Figure 5. It is evident in Figure 6 and even more so in Figure 7 how much the Kalman Filter resists the exacerbated noise of the attack, and how effective examining the difference between the values reported by the system and the values estimated by the Kalman Filter can be.

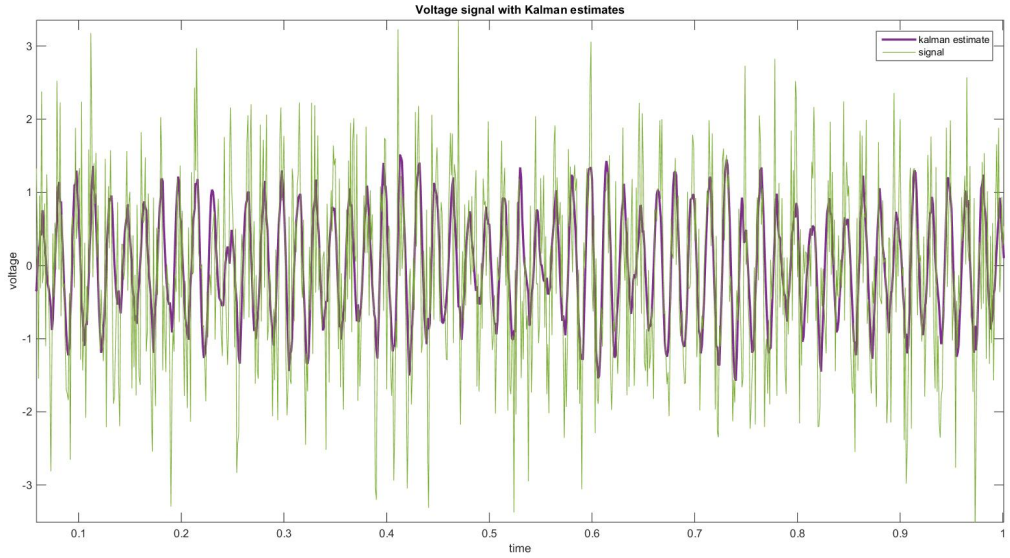

Figure 4: State estimates using Kalman Filter showing detail.

H]

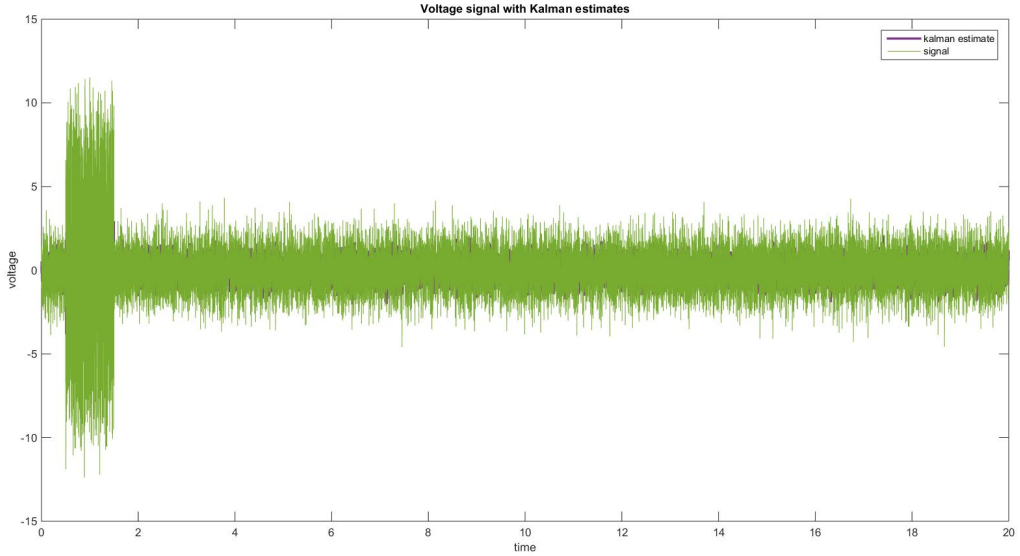

Figure 5: Random noise attack on voltage sensors.

H]

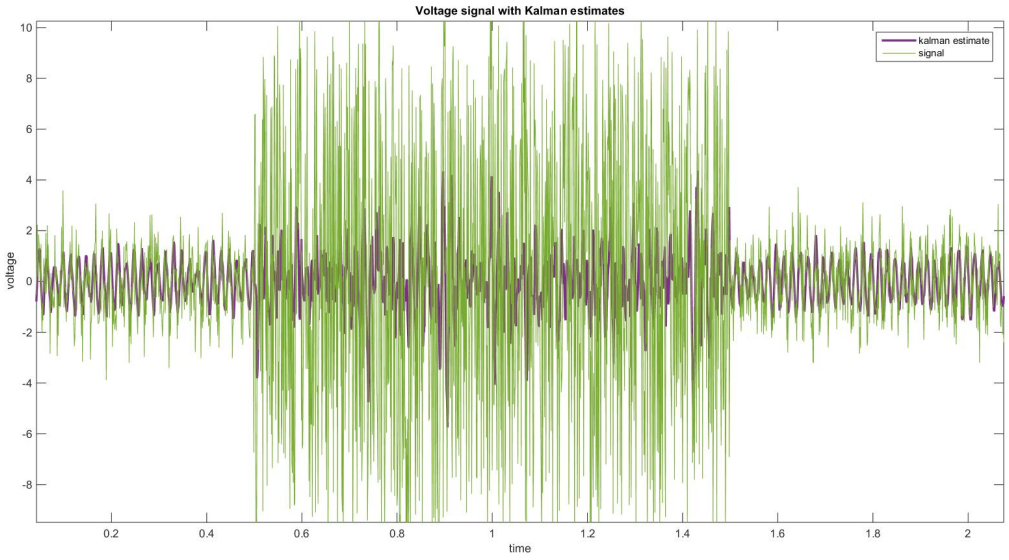

Figure 6: Random noise attack with Kalman Filter detail.

H]

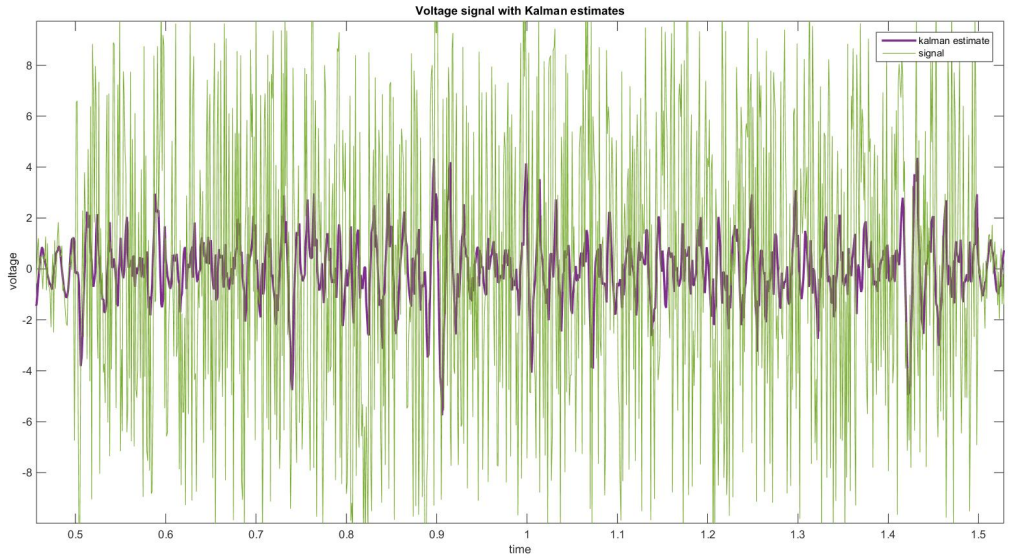

Figure 7: Kalman Filter resistant to extra measurement noise from random noise attack.

References

[1] Loukas, George (2015). *Cyber-Physical Attacks: A Growing Invisible Threat*. Retrieved from https://dl.acm.org/citation.cfm?id=2818550.

[2] McAfee (Retrieved 2019, April 28). *What is Stuxnet?*. Retrieved from https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/what-is-stuxnet.html

[3] CNN (2015, December 22). *Former official: Iranians hacked into New York dam*. Retrieved from https://www.cnn.com/2015/12/21/politics/iranian-hackers-new-york-dam/index.html

[4] Symantec (2014, June 30). *Emerging Threat: Dragonfly/Energetic Bear - APT Group*. Retrieved from https://www.symantec.com/connect/blogs/emerging-threat-dragonfly-energetic-bear-apt-group

[5] Fouda, M., Fadlullah, Z., Kato, N., Lu, R., Shen, X. (2011, August 15). *A Lightweight Message Authentication Scheme for Smart Grid Communications*. Retrieved from https://ieeexplore.ieee.org/document/5983424

[6] Sikdar, B., Chow, J. (2011, October 6). *Defending Synchrophasor Data Networks Against Traffic Analysis Attacks*. Retrieved from https://ieeexplore.ieee.org/document/6035753

[7] Manandhar, K., Cao, X., Hu, F., Liu, Y. (2014, September 12). *Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using Kalman Filter*. Retrieved from https://ieeexplore.ieee.org/document/6897944

[8] Qi, H., Wang, X., Tolbert, L., Li, F., Peng, F., Ning, P., Amin, M. (2011, August 30). *A Resilient Real-Time System Design for a Secure and Reconfigurable Power Grid*. Retrieved from https://ieeexplore.ieee.org/abstract/document/6003812

[9] Welch, G., Bishop, G. (2001). *An Introduction to the Kalman Filter*. Retrieved from http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf