2014

Cyber-Physical Security Strategies

Sarah Davis
Louisiana State University and Agricultural and Mechanical College, sdavi93@tigers.lsu.edu

CYBER-PHYSICAL SECURITY STRATEGIES

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Electrical and Computer Engineering

by
Sarah Davis
B.S.E, Tulane University, 2008
May 2014

Acknowledgements

I'd like to take this opportunity to thank my graduate advisor, Dr. Guoxiang Gu, as well as my examination committee members, Dr. Morteza Naraghi-Pour and Dr. Martin Feldman. I'd also like to thank the AFRL/Clarkson Aerospace for partial funding.

Table of Contents

Abstract

Cyber-physical security describes the protection of systems with close relationships between computational functions and physical ones and addresses the issue of vulnerability to attack through both cyber and physical avenues. This describes systems in a wide variety of functions, many crucial to the function of modern society, making their security of paramount importance. The development of secure system design and attack detection strategies for each potential avenue of attack is needed to combat malicious attacks. This thesis will provide an overview of the approaches to securing different aspect of cyber-physical systems. The cyber element can be designed to better prevent unauthorized entry and to be more robust to attack while its use is evaluated for signs of ongoing intrusion. Nodes in sensor networks can be evaluated by their claims to determine the likelihood of their honesty. Control systems can be designed to be robust in cases of the failure of one component and to detect signal insertion or replay attack. Through the application of these strategies, the safety and continued function of cyber-physical systems can be improved.

Chapter 1: Introduction

1.1 Motivation

Technological advances have greatly eased modern life, but as infrastructure becomes increasingly dependent on technology, its vulnerability to malicious attack also grows. More and more technological systems interface thoroughly with the physical world, leaving them vulnerable not only to traditional cyber attacks but also to attacks through physical avenues. Cyber-physical security as a field applies to systems with close relationships between computational functions and physical ones, such as the example system shown in Figure 1. Examples of cyber-physical systems include the smart grid, process control systems, air traffic control systems, and medical monitoring. This applies to systems in a wide variety of functions, many of great importance to the function of modern society. Attacks can occur through the network, through replacement or control of sensors, or through providing false data to sensors by manipulating the conditions at the sensor site.

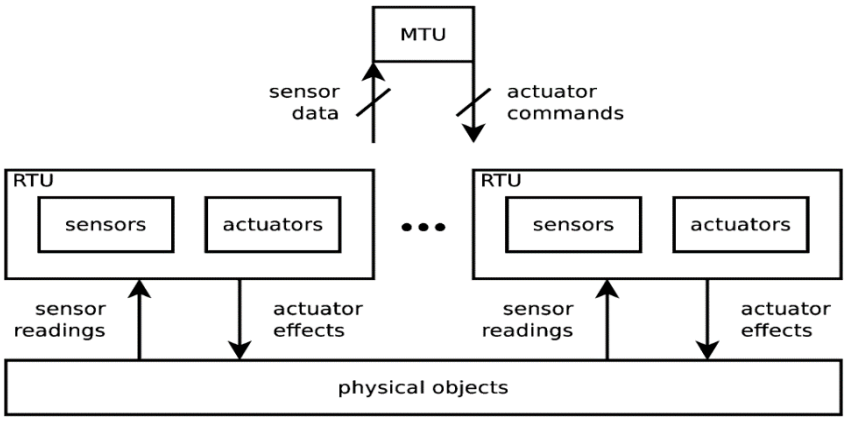

Figure 1: Example diagram of a cyber-physical system. [Mitchell, 2013]

Attacks in control systems have been reported in electric power control systems, including those for transmission, generation and distribution in fossil, gas turbine, and nuclear plants, and in business such as water, oil and gas, chemicals, paper and agribusiness [Turk, 2005]. An attack

on the communication system used by a railway company shut down all rail traffic in the Washington, D.C. area, including morning commuter traffic, for twelve hours [Turk, 2005]. A former consultant at a waste water plant in Queensland, Australia, used the system to release up to 1 million liters of sewage into nearby waterways [Turk, 2005]. One of the most famous attacks against a control system is the Stuxnet worm, a virus that affects specific Supervisory Control and Data Acquisition (SCADA) systems, and was allegedly used against nuclear facilities in Iran. Stuxnet was designed to reprogram industrial centrifuges, leading them to failure while remaining almost undetectable due to the replay data disguising the attack [Chabukswar, 2011]. Stuxnet continued to spread through the internet and through thumb drives, and its presence in the United States was reported by Chevron in 2012 [Kushner, 2013]. Cyber–physical security encompasses the detection of attacks against such systems and the design of these systems to continue to function in the event of an attack. Replacement or control of sensors can be used in attacks such as denial-of-service (DoS) attacks, which overload the system with requests to deny access to users.

1.2 State of the Research Field

Research into cyber-physical security has become better prioritized as examples of well-publicized cyber-physical attacks force industries and governments to recognize their vulnerabilities. Many organizations were reluctant to report security incidents out of embarrassment, and others denied that the risk even existed, believing that the distinctness of their systems provided safety [Byrnes, 2004]. This was a reasonable conclusion prior to 2001, when reported cyber incidents were largely the result of accidents or disgruntled employees within a company, with only thirty-one percent resulting from outside attack [Byrnes, 2004]. A huge shift in reported cyber attacks occurred in a short period, so that by 2003 seventy percent of

cyber attacks were external. These attacks made vulnerabilities previously unnoticed, especially in industrial control systems, unavoidable. Attacks such as the Slammer Worm, in which a common frame relay used for internet traffic as well as the power grid was overwhelmed, blocking traffic to the substations, highlighted the vulnerability of modern infrastructure. As the Slammer Worm showed, attacks occurring through the internet can impact even systems that don't use the internet for their function [Byrnes, 2004].

The cyber security field has many strategies for the defense from network attacks from outsiders, but remains especially vulnerable to malicious attack by insiders and loss of function due to denial-of-service attacks, to which there is no sufficient protection. Sensor network security has a robust set of strategies for the determination of malicious nodes in ideal conditions, but remains vulnerable to the ejection of honest nodes from the network resulting from communication issues common in real world situations. Significant progress has been made in the determination of the ideal controller in the presence of DoS attack but remains more vulnerable to an attacker with an ideal attack strategy. Detection and design strategies for signal insertion to control systems have been well developed, but control systems remain in large part quite vulnerable to replay attack in that current detection strategies have a low success rate and significant loss of function.

1.3 Outline of the Thesis

Cyber-physical security can be approached in a number of ways, and we will attempt to divide these approaches by the avenue of attack which they seek to defend from. Chapter 2 will outline the study of traditional cyber attacks, which remain of concern for cyber-physical systems; even those not connected to a network can become vulnerable through carelessness on the part of users. Design of the system through the correct use of keys and network topology can

help protect the system, as can the monitoring of system function for anomalous use indicative of an incursion. Chapter 3 elaborates on the security of sensor networks; cyber physical systems generally require sensors to gather physical data to perform their functions, leaving them vulnerable to attacks which alter the environment, fool the sensor, or insert a false sensor into the system. Through evaluation of the data provided by sensors across the network over time, the reliability of individual sensors in the network can be evaluated. Chapter 4 discusses the protection of control systems, which can be harmed through DoS attack, manipulated through the insertion of malicious signals, and deceived through data readings used to disguise an attack. Filters can be used to detect signal insertion, but defense against more complicated attacks remain challenging.

Chapter 2: Approaches to Cyber Security

2.1 Cyber Attacks

Cyber physical systems have many points of vulnerability due to the many components that these systems employ. For instance, a smart grid system is particularly vulnerable to cyber incursions due to the physically distributed nature of the system, leaving many entry points on the consumer side such as meters or appliances physically vulnerable [Fadlullah, 2011]. With entry points vulnerable to cyber attack from outsiders across the system and to the possibility of misuse by system users, the pricing data, integrity of commands and software, and system availability of the smart grid must be protected [Mo, 2012]. These attacks from entry points can come in the form of infected devices, intrusion via a network such as the internet, preinstalled vulnerabilities left during the supply chain, or a malicious insider. Network attacks have a number of possible sources, including backdoors in the IT infrastructure, direct access to remote terminal units, exploiting trusted peer utility links, or hijacking the Virtual Private Network (VPN) connection employed by a legitimate user. Once an attacker is inside the system, it is vulnerable in a range of ways, from more trivial attempts to reduce electricity prices to terrorist attacks. Once the cyber-physical system has been infiltrated, the adversary's access to the control system has the potential for enormous physical consequences. In the case of a smart grid, an attacker could cause widespread blackouts, putting lives in danger when elements of critical infrastructure, such as hospitals or traffic lights, do not receive power. An example model of one such smart grid network is shown in Figure 2. Attackers to such a system endanger the confidentiality of system information, the integrity and the availability of the system information and the system function to legitimate users [Mo, 2012].

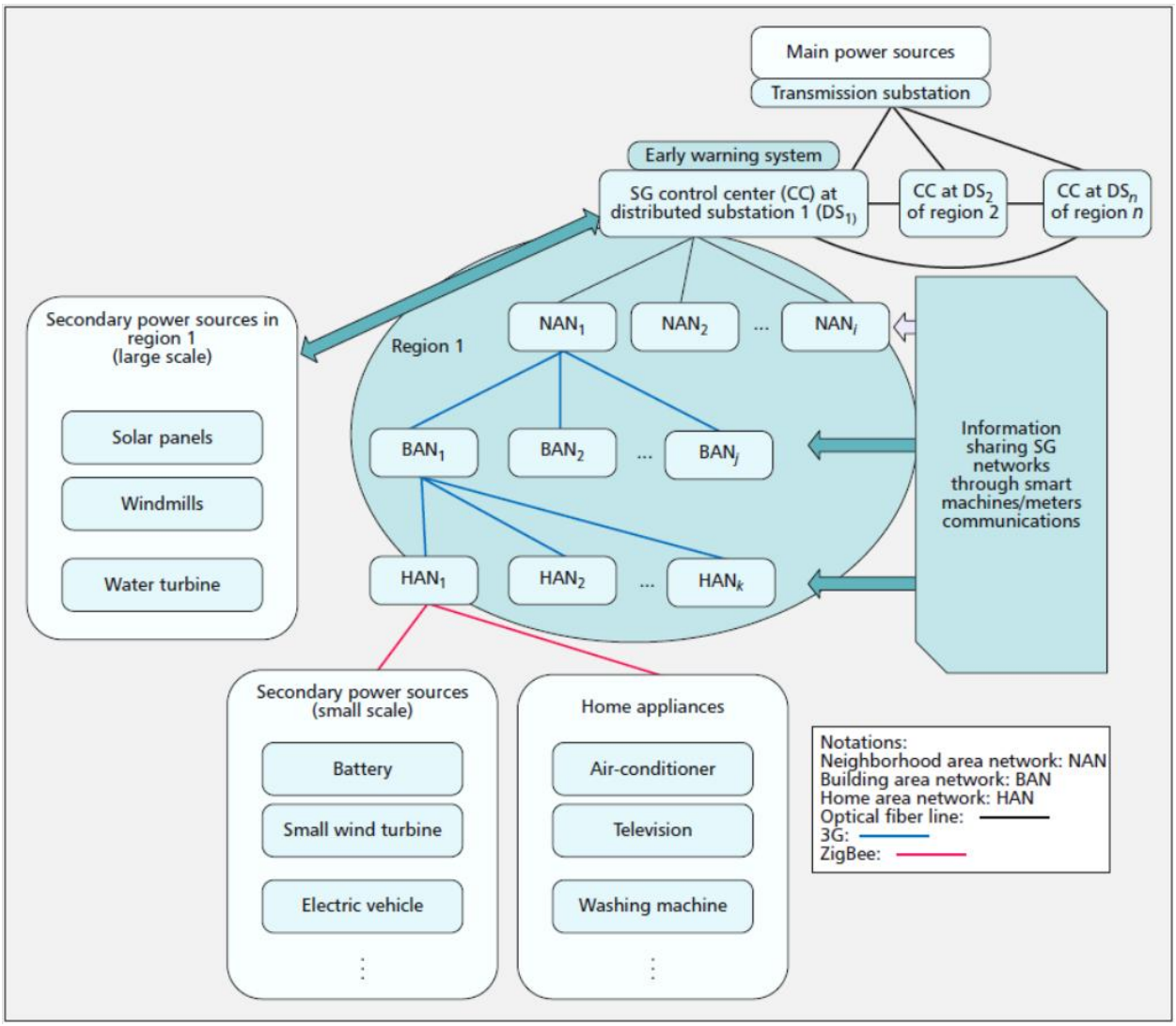

Figure 2: Model of a smart grid's hierarchical network. [Fadlullah, 2011]

Another common type of cyber attack is a Denial of Service (DoS) attack, which brings down a server or network by overloading the system with fake requests so that no resources are available for legitimate requests. Distributed DoS (DDoS) attacks in a smart grid employ compromised meters or appliances and can be used to prevent pricing data from being updated accurately, leading to inaccurate information about the demand on the system. These attacks are especially common in the types of wireless implementation common to widespread grids, in which the scale often makes lower cost equipment and protocol necessary [Zhang, 2011]. The

efficiency of the smart grid is reliant on the accurate information about power use and princing, which is used to determine the amount of power to be generated and the running of various home appliances to be sure that all aspects of the system run in the most energy and cost effective manner possible. Inaccurate information from DDoS attacks could become very costly to the power company or the consumer [Mo, 2012].

2.2 Design of a Secure Cyber System

The traditional approach to the protection of a cyber system is to create a figurative wall around it, in which the system can't be entered except by legitimate users [Mukherjee, 1994]. Firewalls are designed to secure the entrances to the system by examining packets as they pass through and determining whether or not they should be allowed to proceed. This evaluation takes place using a series of rules to which the packet is compared [Gouda, 2004]. The interface in which the packet arrives at the firewall, the packet's original source, its final destination, and the transport protocol it uses are all used to determine whether or not the packet can be allowed through. Appropriate key management is another important aspect of securing the cyber system under this approach. Secure key encryption can be a valuable barrier to outside intrusion. The difficulty with encryption keys in cyber-physical systems is managing them over a large and varied infrastructure. For instance, most smart meters have a single key for each meter, which lasts over the life of the device. While this simplifies the management of a large meter system, the ability to revoke keys and to update them periodically would be a simple step to ensure far greater security to the entire grid [Hadley, 2010].

It cannot be assumed that such methods will prevent all attackers from gaining entrance, leaving a need for the design of the communication architecture which achieves maximum security. The network topology, or the design of the connections between nodes, can be designed

to be more resilient to attack. The routing protocol used should be designed to avoid vulnerabilities in which a single compromised router could bring down the system [Mo, 2011]. Lack of consideration when constructing a communication network leaves the system unnecessarily vulnerable; it has been shown that the internet is even more vulnerable to attack than a random network topology would be [Lee, 2006]. In order to design a network topology that is most robust against cyber attacks, the way in which an attack is spread through the network must be properly understood. This spread can be modeled in the same way as many other phenomena, such as rumor routing, or the spread of a virus. Such occurrences have been studied at length, providing a variety of existing models to choose from. The work of [Roy, 2012] assumes the adversary of the cyber physical system network seeks to measure or modify points in the network in order to estimate and or actuate the network and attempt to design the network topology to be more resistant to such measures. The security of the system is measured in terms of the discoverability of attacks and recoverability of the system in the face of attacks. When considering spread dynamics for an n^{th} order system with the state space equations

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

the security of which is considered to be equivalent to the concept of observability in control theory, or rank $\{\mathcal{O}_n\} = n$, where $\mathcal{O}_n$ is the observability matrix defined by

$$\mathcal{O}_n = \begin{pmatrix} C \\ CA \\ . \\ . \\ . \\ CA^{n-1} \end{pmatrix}$$

[Roy, 2012] theorizes firstly that the less partial knowledge of the system the attacker requires the lower the rank of the observability matrix must be. Secondly, they theorize that the

estimation goal at time state k is secure "if and only if there is a nonzero vector in the range of A^k that is in the null space of $\mathcal{O}$" [Roy, 2012].

The network's graph topology can be evaluated in terms of the ease of estimation of its state and the security of different locations within the network. The spread model can employ either the probability of an infected node coming into contact with an infected node or by tracing the infection itself.

Network analysis can also be applied to the design of the power grid to minimize the damage done by a physical attack. Such design considers power flow to the most crucial system components and the largest possible disruption to the grid to reduce the damage done [Salmeron, 2004]. Examples include [Salmeron, 2004] and [Pinar, 2010].

2.3 Approaches to Cyber Attack Detection

Traditional cyber defense works to ensure the availability, confidentiality, and integrity of the cyber system [Mo, 2012]. It seeks to ensure the system is available for use when needed, that the data in the system cannot be viewed by outsiders without permission, and that the system data cannot be altered by those without permission. Analysis of the use of the cyber-physical system can be used to determine likelihood of an intrusion, by invasion or by malicious insider. There are two categories for intrusion detection, misuse detection and anomaly detection. Misuse detection is used to search for specific patterns, events or data associated with a known system weakness and anomaly detection looks at changes in patterns of utilization that can indicate an attack [Balasubramaniyan, 1998]. Anomaly detection presupposes that an attack on the system will involve patterns that do not occur during normal system operation, and searches for such unusual patterns [Denning, 1987]. This has the advantage of being able to detect attacks even if the system vulnerability is not previously known. The normal behavior patterns in activities such

as logins, commands executed, and files and devices accessed can be observed for later comparison in which a statistically significant anomalous behavior would indicate a possible attack [Mukherjee, 1994]. A number of approaches to such detection have been developed, among them agent based intrusion detection and bioinformatics inspired intrusion detection.

In a large scale network, intrusion detection must be adapted from the single detector model to something better suited analyzing the number of packets used on such a scale. Agent based intrusion detection was developed to meet this need [Chatzigiannakis, 2004]. Software agents are designed to operate independently of other programs or user input, in different areas of the network. In addition to detecting individual attacks across the network, when implemented properly, agent based intrusion detection can determine when a cooperative intrusion is occurring across the network [Benattou, 2004]. [Balasubramaniyan, 1998] proposed the intrusion detection architecture known as Autonomous Agents for Intrusion Detection (AAFID), using a hierarchal structure of agents to perform detection tasks. Using AAFID, any number of agents can be distributed over any number of hosts in a network. All agents in a single host report to a single transceiver, and transceivers report their results to one or more monitors, each of which oversee several transceivers and perform high level detection through correlating data across the network. This is not specific to any one detection technique, and can be adapted for different approaches [Balasubramaniyan, 1998]. Benattou and Tamine combine the concept of a network of agents with that of mobile agents. Agents are categorized by their varying analysis functions and can be dispatched by the Specialized Local Agent, or SLA, which is responsible for coordinating agent activity and determining where different analysis agents are needed. Mobile Agents collect data, which together with the Correlate Agents determine if attacks are widespread. Interpreter Agents select specific local events to be considered by the Analyzer

Agents to detect complex local attacks. This keeps all detection functions from being run at all times and minimizes the processing power required for the intrusion detection system [Benattou, 2004].

Similar works have been completed by [Helmer, 1998], [Blanc, 2006], [Chatzigiannakis, 2004], and [Zhao-wen, 2007].

One approach to intrusion detection uses a biological model of the immune system's response to viral invasion as inspiration for a cyber strategy. There are a variety of immune based strategies that can be employed, including negative selection algorithms, immune network algorithms, danger theory, and clonal selection algorithms. In a biological immune system, antigens, or foreign proteins indicative of a virus or harmful bacteria, are detected by antibodies chemically binding to the specific part of a protein they are designed to recognize, after which the antigen is destroyed. These techniques seek to mimic the way in which a biological immune system trains itself to detect such invaders. One of the easiest to implement examples of this is the process developed by [Forrest, 1997], in which the specific processes run by a computer were considered the antigens. Given a collection of digital data to be monitored for changes, a set of detectors that did not match the data was generated. The detectors were then continuously compared to the data, and match detectors were used to indicate a change had occurred with a known location. Like with a biological immune system, matches between the detector and the data that were close but not precise were also considered matches, and the closeness of the match was determined using Hamming distances. The primary weakness in this system was the generation of detectors, which could be improved over a random set using dynamic programming methods but was still quite imprecise [Forrest, 1997].

The work of [Zhang, 2011] on artificial immune system (AIS) algorithms is another example of such method, employing the clonal selection algorithm, a machine learning based technique that trains the system over time to recognize attacks. The AIS is a machine learning based system where positive examples alone can be used to train the system to recognize attacks. The clonal algorithm is based on antigen recognition in the immune system, in which only cells that recognize the antigens are allowed to reproduce themselves. Immune cells contain a wide variety of receptors that can bind to a bacterial invader, and so when an attack occurs cells with a matching receptor will bind to the invader. Cells that bind to a receptor are stimulated to reproduce, resulting in an immune response prepared to deal with that immune attack. The CLONALG algorithm is an unsupervised algorithm that uses cycles of maintenance, selection, cloning and mutation to mimic an immune response and train the system to respond to attacks. An initial group of 'antibodies' is generated randomly and divided into a group of memory cells and a reservoir pool, as well as a set of antigenic patterns. A single antigen is chosen and compared to every member of the set of antigens, and the antibodies with the greatest similarities to the antigen are selected and reproduced in numbers varying linearly with the degree of similarity between the antibody and antigen. This population is then mutated and again compared to the antigen. The antibody with the highest affinity is saved and the previous set of antibodies is replaced by the new group with higher affinity. By repeating this process, a memory pool is generated that can recognize the desired 'antigen'. The Artificial Immune Recognition System (AIRS) builds on the clonal algorithm by the addition of affinity maturation and affinity recognition balls (ARB). Like CLONALG, AIRS buildings 'antibodies' over time, but it differs in that it is a supervised system. After initialization, AIRS operates in cycles of antigen training, competition, memory cell selection, and classification of the dataset. In initialization, two sets,

the ARBs and the memory pool are made after the dataset has been normalized and the variable for the affinity threshold calculated. Each antigen is then compared to the memory cell pool in the antigen training step and the best memory cell is selected. Like in the CLONALG, the cell is then mutated and placed in the ARB pool. The ARB cells with the lowest similarity are discarded and the ones with the highest go into the memory cell pool. This is then repeated until all antigens are tested.

The weakness of AIS techniques is the need for a large set of attack samples to be used during the training process in order to ensure that the system will be able to recognize an attack. Without a sufficient number, AIS is no more efficient than traditional machine learning techniques such as support vector machine (SVM). SVM is a machine learning system that uses large margin separation, in which the distance between a data point and a line drawn to maximize the distance between the data points on each side of the line separate data into two separate areas.

These are only a few examples of the use of bioinformatics methods for intrusion detection. Other examples include the work of researchers such as [Coull, 2003] and [Janakiraman, 2006].

Chapter 3: Detection of Compromised Sensors

3.1 Attacks on Sensor Networks

Networks of sensors are useful in a variety of applications, from straightforward information gathering to large projects such as a smart grid. They are prized for their low cost and quick deployment, but these same factors make the network vulnerable to incursions such as node replication attack. Such an attack is considered in terms of the classic mathematical problem of the Byzantine Generals, first considered by Lamport et al. in 1982 [Vempaty, 2013]. In this problem, a group of generals of the Byzantine Empire work to plan an attack on a city, but one or more of the generals may be a traitor. Using the information each general reports and the attack plan suggested by each one, the group must nonetheless come to an accurate consensus as to the correct information or best attack plan. The same problem applies to a network of sensor nodes in which some may be either controlled or added by an adversary. Compromised sensors function much like a traitorous general, impeding the decisions made by the group, as shown in Figure 3. Node replicas can be used to inject false data into the system, suppress legitimate data, revoke legitimate nodes and disconnect the network by using the correct protocols [Vempaty, 2013]. For instance, if sensors in a smart grid are made to send false data about the state estimation of the grid, it will alter decisions made by the grid based on the state estimation data. This could change power availability to the grid, pricing based on power usage, and charges to the users [Mo, 2011]. Thus a system must be developed to detect such attacks while also using low cost hardware.

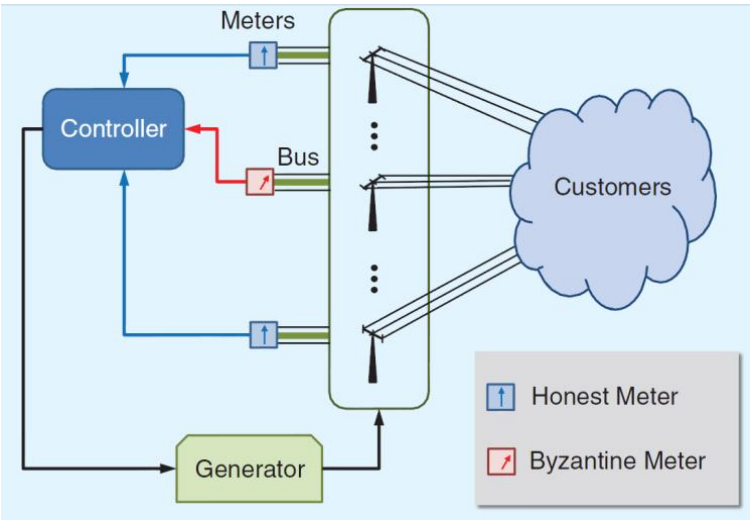

Figure 3: Compromised smart grid with dishonest reports in the network. [Vempaty, 2013]

The two general approaches to detecting node replicas are centralized detection and localized voting systems. Centralized detection uses a single central point of contact, or fusion center which receives and processes all claims from sensors and is responsible for determining whether or not they are compromised, as depicted in Figure 4.

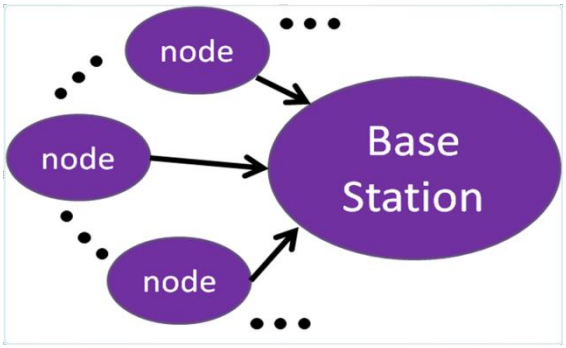

Figure 4: A centralized detection scheme.

This has the advantage over localized voting of being able to detect distributed node replications more easily. False nodes can also work in unison to better change the consensus decision of the network, or they can be working independently. If working independently, the ratio of Byzantine nodes to honest nodes must be at least one half, but if the Byzantine nodes

work in concert this ratio can be lower and still deceive the network. The more nodes working together the easier the deception will be to achieve. Similarly, if the attacking nodes are aware of the true decision they are working against they can better prevent the system from reaching that conclusion.

3.2 Centralized Detection

There are a variety of approaches to centralized detection, including reputation based scheme and adaptive learning schemes. A reputation based scheme such as the one developed by [Rawat, 2011] uses the fusion center to determine a value for each node's reliability based on discrepancies between the individual node's values and input from other nodes. For a system with sensors transmitting every time instant t, the reputation metric is defined as

$$K_i = \sum_{t=1}^{T} I_{(u_i[t] \neq u_o[t])}$$

after time interval *T*, where $u_i[t]$ is the i^{th} sensor's decision at time instant *t*, $u_0[t]$ is the decision made at the fusion center at that time and is the indicator function over the set *S*. Thus the greater the difference between the values the less the node is trusted, and if the pattern holds over time the node will be trusted less and less. This could be an issue if there are honest outliers in the nodes readings or if there are communication issues between the node and the fusion center. The fusion center could become suspicious of a node due to communication issues and has the potential to eventually drop honest nodes. Reputation based schemes can be problematic if they do not take into account the possibility of communication errors that can cause readings which appear suspicious. If the nodes under suspicion are removed from the decision making process, over time the network can remove many honest nodes due to naturally occurring errors, leaving few or no nodes left in the system. The combination of reputation based schemes with other

security detection strategies to improve the security of the sensor network is discussed further in Section 3.4.

[Vempaty, 2013] worked on an adaptive learning scheme, in which the system identifies attacking sensors by comparing their behavior with what would be expected of an honest sensor. Knowledge of the dishonest sensors and the information they send is used to adapt global decision making. To determine the behavior expected of an honest sensor, the sensor's previous behavior over a period of time must be used to determine the probability of the sensor sending a one as its value. This demonstrates the most likely behavior of the sensor, which is then used to determine the likelihood of its honesty later. This has the advantage of working even when Byzantine sensors are in the majority but can only work if the system has knowledge of the honest sensor's behavior. This is a flaw that could prevent it from being useable in many circumstances; if any behavior deviates from what is expected ahead of time then honest readings could be unfairly disregarded by the detection system.

The work of [Soltanmohammadi, 2013] describes a centralized detection system that seeks to identify the nature of a node's false data, in this case termed misbehavior, so that false data due to attack may be distinguished from false data due to hardware or software degradation. The source of the node's misbehavior used to determine the type of decision it will make, and by analyzing the node's decision over time the fusion center can determine whether or not it is the result of an attack. [Soltanmohammadi, 2014] describes a system for the classification of misbehavior by cognitive radios based on the expectation maximization algorithm. Assuming that the majority or the cognitive radios are honest, a set of the CR's decisions in relation to the hypothesis are used to determine whether or not it is dishonest or not working properly.

3.3 Distributed Decision Making

The primary flaw of a centralized detection design is that it also makes the system vulnerable; the whole system fails if the fusion center is compromised. Alternative approaches involve the collaborative determination of an incursion by the sensor network. More recent advances in algorithms to detect node replication allow for the collaborative detection of distributed node replication. One such approach to determining the false node employs the use of witness nodes to verify the location claims of each sensor node. There are different methods of selecting witness nodes, including deterministic multicast, randomized multicast, and line selected multicast. In deterministic multicast, a node broadcasts its location information, considered a claim until verified, to a deterministically selected set of nodes called witnesses, which are chosen using the node's identification number in the network. Therefore, if an adversary attempts to replicate the node and claim its identification number, the imposter will broadcast to the same witness nodes, which will be able to find the false node based on the conflicting claims. In randomized multicast, nodes notify a neighbor of their location claims, and the neighboring nodes randomly send these location claims to other nodes in the network. In a network of n nodes, as long as at east $\sqrt{n}$ nodes are witnesses to each, the birthday paradox predicts that there will be a high probability of a collision. This approach is considered more resilient than the deterministic multicast. An approach with lower communication cost than randomized multicast is line selected multicast, in which nodes send their location claims in lines through their neighbors, as shown in Figure 5. This is based on the fact that nodes in sensor networks work as both sensors and routers, and to send locations claims to other nodes the information must pass from node to node. Each node sends its location claim to an immediate

neighbor, which sends the claims to its neighbor, and so on, creating a line of witness nodes which can detect false claims when the line is crossed.

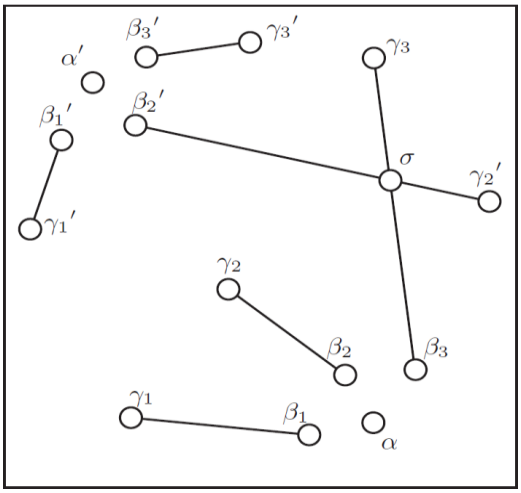

Figure 5: An illustration of line selected multicast, in which the node α has been replicated as α‘ and neighbors β1, 2 and 3 report location claims to γ, resulting in an intersection at σ. [Parno, 2005]

When considering approaches to distributed detection in terms of their efficiency versus the cost in memory to each node and communication cost in the system, line selected multicast is the most efficient system.

Another approach to witness nodes which reduces the necessary memory is the timed distribution of location claims. In the strategy known as the high noon approach, nodes devote their computing power to the detection of Byzantine nodes for a periodic length of time *t* and then spend the rest of the time on non-detection tasks [Parno, 2005]. During the detection period nodes determined to be Byzantines have their privileges revoked and then the data is forgotten at the end of the period. During the next detection period the process begins again. In the time slot approach, a length of time is divided into units of nodes by their identification number, and during each period one group of nodes broadcasts their location claims, leaving the network with a far smaller pool of nodes to check during each time period. All of these techniques are valuable

to reduce the necessary equipment costs without impairing the effectiveness of the byzantine detection system.

3.4 Resilience to False Sensor Claims

Security can also be designed for a sensor network in terms of making it more robust against node replication attacks, such as noise enhanced signal processing or weighted sequential ratio probability testing. These are signal processing methods used to increase the likelihood of accurate sensor data being used regardless of whether or not byzantine nodes have been identified. While these methods do take into account the likelihood of a sensor being a byzantine, this is used only for the network to come to a final decision rather than to remove a sensor from the network, as might occur in a different security strategy. Noise-enhanced signal processing, such as the work done by [Gagrani, 2011] uses stochastic resonance to make the system more resistant to Byzantine attack by the introduction of noise. Stochastic resonance is a physical phenomenon in nonlinear systems in which the signal output can be enhanced by the addition of noise to the input. This is most helpful to the system when the Byzantine nodes do not also apply this strategy; using the ratio of honest nodes needed to come to the correct decision as a metric, stochastic resonance when both the honest and byzantine nodes employ it makes no difference to the system's performance. Stochastic resonance can be used in coordination with other metrics to determine honesty, in which a node deemed honest is told to employ stochastic resonance from then forward and the nodes under suspicion are not. The stochastic resonance noise can also be added at the fusion center, as is the case in the model shown in Figure 6. The optimal function of such a system occurs is dependent on the system's knowledge of the channel state information and the local sensor detection performance indices. The approach which requires the least information is the equal gain combiner, or ECG.

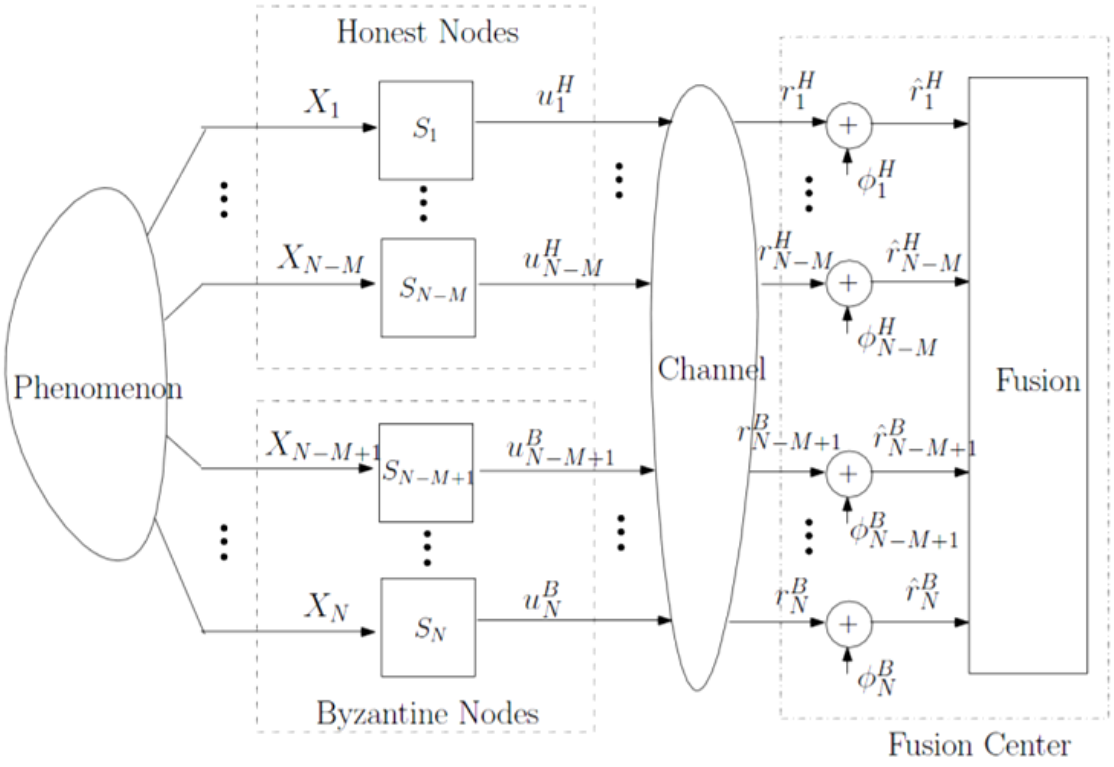

Figure 6: Inference network model for stochastic resonance added at the fusion center. [Gagrani, 2011]

Though it does not have a performance as good as other stochastic resonance approaches, addition at the fusion center can still be a valid approach depending on the circumstances under which the system operates. As a whole, noise enhanced signal processing is primarily of interest in addition to other approaches to sensor network security rather than as a singular method of protection, but when used in conjunction with a robust method of byzantine detection it can be valuable.

Weighted sequential probability ratio test (WSPRT) is another method to make the sensor network more robust to attack. One example of WSPRT is the work of [Chen, 2008], which considers the detector design and the data fusion process using WSPRT in cognitive radio network sensor systems. WSPRT has two steps; the first is a reputation based action and the second is the hypothesis test. The reputation step is similar to other reputation schemes in that it

judges the accuracy and honesty of a sensor based on how closely its data reflects the final local decision. The second step is based on the Sequential Probability Ratio Test (SPRT), a hypothesis test used for sequential analysis which allows for the sampling of a variable number of operations [Chen, 2008]. In a non-adversarial environment, SPRT has both bounded false alarm probability and bounded miss detection probability. WSPRT is essentially SPRT with the reputation of the sensors taken into account. The likelihood ratio of SPRT is defined as

$$S_n = \prod_{i=0}^{n} \frac{P[u_i|H_1]}{P[u_i|H_0]}$$

where H_1 and H_0 are the hypotheses to be chosen between. The likelihood ratio of WSPRT is defined as

$$W_n = \prod_{i=0}^{n} \left(\frac{P[u_i|H_1]}{P[u_i|H_0]}\right)^{w_i}$$

where w_i is the weight of N_i and is a function of r_i, the reputation value of sensor N_i,$w_i = f(r_i)$. When designing f several factors must be taken into account, including outputting $w_i \in [0,1]$, accepting arbitrary r_i values and giving proper weight to even those sensors which have a slightly negative reputation metric, since a slightly low reputation could be due to factors such as temporary interference. [Chen, 2008] use the function

$$w_i = f(r_i) = \begin{cases} 0 & r_i \leq -g \\ \dfrac{r_i + g}{\max(r_i) + g} & r_i > g \end{cases}$$

where g is a number greater than zero which increases with each decision and is selected to ensure that sensors are weighted properly based on their reputation.

Because it incorporates a reputation scheme, WSPRT is only as effective as the reputation scheme it is based on. With correct evaluation of the reputation metric and selection of the weighting function this can be a valuable method of analyzing sensor data.

Chapter 4: Security in Control System

4.1 Attacks on Control Systems

Like other cyber-physical systems, control systems are vulnerable to malicious attacks in the form of deception or denial-of-service (DoS) attacks. DoS attacks can be used in cyber-physical control systems with communication systems which can be jammed, leading sensor and control packets to be dropped [Amin, 2009]. As shown in [Long, 2005], DoS attacks cause a significant increase in overshoot, settling time, rise time, and error. Designing the control system with safety constraints that reduce its vulnerability is one straightforward way of protecting from DoS attack [Amin, 2009]. Deception attacks occur through compromising the integrity of the control system's sensor and control data packets, resulting in the receipt of false data which the system believes to be true. This data can be used to cause the system to damage itself or to manipulate its function for the benefit of the attacker. Through the use of filtering this signal insertion can be detected and dealt with. Detection becomes more difficult in the case of a replay attack, in which false sensor readings are relayed to the controller, disguising the sensor data which could be used to detect the insertion. Such an attack is depicted in Figure 7. The system becomes open loop without input from the system's sensors, so the performance of the control system can no longer be guaranteed. This also leaves the system vulnerable to other attacks that cannot be detected while the sensor data is obfuscated. The most straightforward method of disguise is to record the sensors' previous data and play it back to the system, but this can also be done using sensor data generated to resemble normal operations without duplicating it exactly [Mo, 2009].

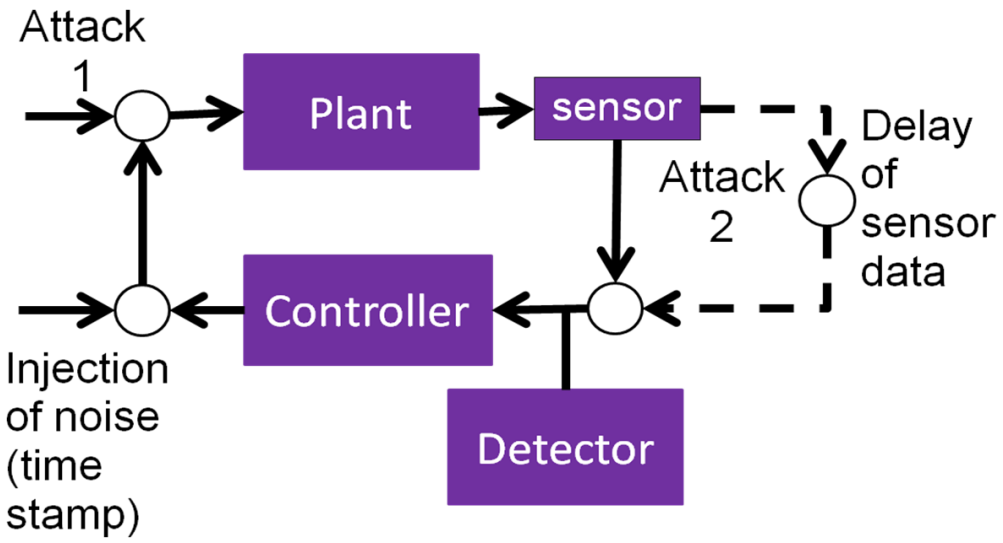

Figure 7: System diagram of a replay attack.

Replay attacks had only been theorized before the use of the Stuxnet worm against nuclear facilities in Iran [Chabukswar, 2011]. Stuxnet highlighted the vulnerability of cyber-physical infrastructure in general and in control systems in particular and brought new urgency to the research into protection against replay attacks.

4.2 Protecting Against Denial-of-Service Attacks

The first line of defense against DoS attack in control systems is a design that minimizes the danger due to packet drop. [Long, 2005] suggest a straightforward method of DoS detection which assumes any packet load over a certain threshold outside of normal operating procedure to be a DoS attack. Once this threshold is cleared, the router begins to increase the probability of packet drop. This was shown through simulation to greatly mitigate the effects of DoS attack [Long, 2005].

[Amin, 2009] consider optimal controller design in the presence of DoS attack for the case of a linear time invariant (LTI) stochastic system

$$x_{k+1} = Ax_k + Bu_k^a + w_k$$

$$u_k^a = v_k u_k$$

$$x_k^a = \gamma_k x_k$$

over time horizon *k=0,…, N-1* and with measurement and control packets (γ_k, v_k) subject to DoS attacks. The authors sought to produce a causal feedback control law that for the given system would minimize the finite horizon objective function

$$J_N(\bar{x}, P_0, {u_0}^{N-1}) = E[x_N^T Q^{xx} x_N + \sum_{k=0}^{N-1} \begin{pmatrix} x_k \\ u_k \end{pmatrix}^T \begin{pmatrix} I_n & 0 \\ 0 & v_k I_m \end{pmatrix} Q \begin{pmatrix} x_k \\ u_k \end{pmatrix} | {u_0}^{N-1}, \bar{x}, P_0]$$

where for $Q^{xx} \succ 0$ and $Q \succcurlyeq 0$

$$Q = \begin{pmatrix} Q^{xx} & 0 \\ 0 & Q^{uu} \end{pmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

and power constraints in an expected sense on both the state and input

$$E\left[\begin{pmatrix} x_k \\ u_k \end{pmatrix}^T \begin{pmatrix} I_n & 0 \\ 0 & v_k I_m \end{pmatrix} H_i \begin{pmatrix} x_k \\ u_k \end{pmatrix}\right] \leq \beta_i, for\ i = 1, \dots, L, and\ k = 0, \dots, N-1$$

and safety specification probabilistic constraints on state and input

$$P\left[\begin{pmatrix} x_k \\ u_k \end{pmatrix}^T \begin{pmatrix} I_n & 0 \\ 0 & v_k I_m \end{pmatrix} H_i \begin{pmatrix} x_k \\ u_k \end{pmatrix} \leq \alpha_i\right] \geq (1-\varepsilon), for\ i = 1, \dots, T, and\ k = 0, \dots, N-1$$

The system is analyzed under the Bernoulli packet drop model, in which the system is subjected to an attacker randomly jamming a measurement or control packet according to independent Bernoulli trials and with a probability of success $\bar{\gamma}$ and $\bar{v}$ respectively. The attack has the admissible attack actions

$$\mathcal{A}_{Ber(\bar{\gamma},\bar{v})} = \{({\gamma_0}^{N-1}, {v_0}^{N-1}) | P(\gamma_k = 1) = \bar{\gamma}, P(v_k = 1) = \bar{v}, \mathrm{k} = 0, \dots, \mathrm{N}-1\}$$

The use of Kalman filter for state estimate $\hat{x}_k = E[x_k|\mathcal{I}_k]$ and state estimation error $e_{k|k} = (x_k - \hat{x}_{k|k})$ leads to the update step

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + v_k B u_k$$

$$e_{k+1|k} = Ae_{k|k} + w_k$$

and correction step

$$\hat{x}_{k+1|k+1} = \gamma_{k+1}x_{k+1} + (1 - \gamma_{k+1})\hat{x}_{k+1|k}$$

$$e_{k+1|k+1} = (1 - \gamma_{k+1})e_{k+1|k}$$

$$E_\gamma[\Sigma_{k+1|k}] = AE_\gamma[\Sigma_{k|k}]A^T + W$$

$$E_\gamma[\Sigma_{k+1|k+1}] = (1 - \bar{\gamma})E_\gamma[\Sigma_{k+1|k}]$$

leading to Kalman filter equations

$$\hat{x}_{k+1} = A\hat{x}_k + v_k Bu_k^a + \gamma_k Ae_k$$

$$e_{k+1} = (1 - \gamma_k)Ae_k + w_k$$

$$E_\gamma[\Sigma_{k+1}] = (1 - \bar{\gamma})AE_\gamma[\Sigma_k]A^T + W$$

The optimal control model for the Bernoulli attack case is found to be $u_k^* = -L_k\hat{x}_{k|k}$ with $\hat{x}_{k|k}$ taken from the Kalman filter equations and $L_k = B^T S_{k+1} + Q^{xx} - R_k$ with $S_k = A^T S_{k+1}A + Q^{xx} - \bar{v}R_k$ and$R_k = L_k^T(B^T S_{k+1}B + Q^{uu})L_k$. It is also noted that an attacker may apply an optimal attack instead, having no incentive to comply with a Bernoulli model of attack, complicating the ability to design a controller to handle such an attack.

4.3 Detection of Signal Insertion

Many different approaches have been taken to the filtering of malicious signal insertion in a control system. Mo and Sinopoli evaluated the 'attackability' of a control system with a sensor network evaluating its state and a χ^2 failure detector, [Mo, 2010]. By evaluating the way in which the system fails under different types of attacks they were able to determine that the failure detector would sound immediately in a case of an attack that injects large values. They were also able to determine that through evaluation of unstable eigenvectors sensors needed by

an attacker in order to cause the network to fail could be identified, providing a way to increase system reliability by increasing the number of sensor nodes along those vectors.

Pasqualetti et al developed a centralized filter based on a modified Luenberger observer designed to detect attacks in a control system [Pasqualetti 2012, Part II]. Given a time invariant system

$$E\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

where *E* is a possibly singular matrix and $Bu(t)$ and $Du(t)$ describe unknown signals attributed to disturbances affecting the plant. The centralized attack detection filter

$$E\dot{w}(t) = (A + GC)w(t) - Gy(t)$$

$$r(t) = Cw(t) - y(t)$$

where *w(0) = x(0)* and output injection gain *G* is such that the pair *(E, A+GC)* is regular and Hurwitz. If $u_k(t) = 0$ at all times then it can be concluded that $r(t) = 0$ at all times. In the absence of attacks, the error *e(t) = w(t)-x(t)* is exponentially stable.

Pasqualetti et al have also developed a distributed attack detection system in control systems, employing knowledge of the decentralized stabilization of the filter's error dynamics and the waveform relaxation method to produce a distributed attack detection filter. The decentralized system is the interconnection of *N* subsystems with the state and output $x_i(t)$ and $y_i(t)$ and neighbors N_i^n for the *i*-th subsystem. It can be represented by

$$E_i\dot{x}_i(t) = A_i x_i(t) + \sum_{j \in N_i^{in}} A_{ij}\, x_j(t)$$

$$y_i(t) = C_i x_i(t), i \in \{1, \dots, N\}$$

Each subnetwork with control center G_t^i has the local residual generator

$$E_i\dot{w}_i(t) = (A_i + G_iC_i)w_i(t) + \sum_{j\in N_i^{in}} A_{ij}\, x_j(t) - G_i y_i(t)$$

$$r_i(t) = y_i(t) - C_i w_i(t), i \in \{1, \dots, N\}$$

where $w_i(t)$ is the *i*-th estimate of $x_i(t)$. This gives the overall filter dynamics of

$$E\dot{w}(t) = (A_D + GC)w(t) + A_C w(t) - Gy(t)$$

$$r_i(t) = y(t) - Cw(t)$$

where $r(t) = 0$ when $u_k(t) = 0$

Waveform relaxation is used to obtain the waveform relaxation iteration

$$E\dot{w}^k(t) = A_D w^k(t) + A_C w^{k-1}(t) - Gy(t)$$

where *k* is the iteration index. This leads to the distributed attack detection filter

$$E\dot{w}^k(t) = (A_D + GC)w^k(t) + A_C w^{k-1}(t) - Gy(t)$$

$$r(t) = y(t) - Cw^k(t)$$

Similar to the case of the centralized filter, $\lim_{k\to\infty} r^k(t) = 0$ at all times if and only if $u_k(t) = 0$ at all times, and without the presence of an attack, the asymptotic filter error $\lim_{k\to\infty} w^k(t) - x(t)$ is exponentially stable [Pasqualetti, 2012, Part II].

4.4 Detection of Replay Attacks

While many studies assume, perhaps falsely, that failure in a control system is due to random events or benign ones, the issue of failure due to an attack must be considered differently than in the typical failure detection algorithm. Mo and Sinopoli developed one technique to detect replay data in a linear time invariant (LTI) Gaussian system with an infinite horizon Linear Quadratic Gaussian (LQG) Controller, assuming the system has a χ^2 failure detector [Mo, 2009]. One strategy to detect replay attacks is to inject time-stamped noise into the system. The LTI system's dynamics at time *k* are described as

$$x_{k+1} = Ax_k + Bu_k + w_k$$

where w_k is the process noise with Gaussian distribution and x_0 is the initial state. The network is assumed to be monitored by a sensor network with observation equation

$$y_k = Cx_k + v_k$$

where y_k is the vector of measurements from the sensors and v_k is the measurement noise with Gaussian distribution. The system uses a Kalman filter, the optimal estimator for such a system, providing the minimum variance unbiased estimate of state, denoted by $\hat{x}_{k|k}$ indicating the estimate of x_k based on measurements up to time k. Because Kalman gain converges over time and control systems run for long period, it is assumed that to be in a steady state, and that the Kalman filter will be a fixed gain estimator. An LQG controller is designed for systems in the presence of the Gaussian noise, and is the result of a Linear Quadratic Estimator (LQE), or Kalman filter, with a Linear Quadratic Regulator (LQR). The controller works by minimizing the function

$$J = \min \lim_{T\to\infty} E\frac{1}{T}\left[\sum_{k=0}^{T-1} {x_k}^T W x_k + {u_k}^T U u_k\right]$$

based on the state estimation $\hat{x}_{k|k}$. This minimization leads to the fixed gain controller

$$u_k = u_k^* = -(B^T S B + U)^{-1} B^T S \hat{x}_{k|k}$$

where u_k^* is the optimal control input. The system uses a failure detector which operates by detecting when the normalized estimation error rises above a certain probability threshold, indicating that an error has occurred. For a replay attack, it is assumed that an attacker can inject a control signal at any time, knows all of the sensor readings, and can modify them. This leads to the conclusion that given these capabilities the attacker will first record a period of sensor data and then replay it to the system while injecting a sequence of desired control input. The injected

signal could also be generated through observation of legitimate values and machine learning to resemble accurate data without actually being a replay of previous recordings. In either case the attack would be disguised by the false data. The controller is redesigned to reflect this, and becomes

$$u_k = u_k^* + \Delta u_k$$

where u_k^* is the optimal LQG control signal and Δu_k is taken from an i.i.d. Gaussian distribution and zero mean and is independent of u_k^* [Mo 2009]. The covariance of Δu_k is indicated by $\mathcal{Q}$, and the system with the addition of Δu_k is depicted in Figure 8.

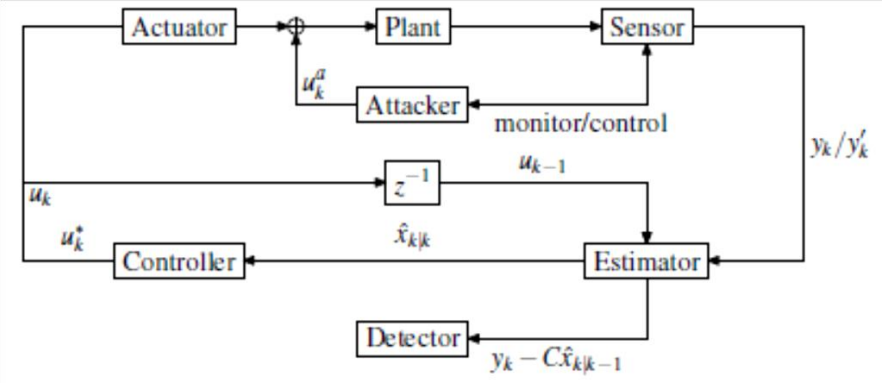

Figure 8: System with insertion of authentication signal. [Mo, 2011]

It functions as an authentication signal, though it does mean that there is a loss of performance as a result of its addition to the optimal LQG signal.

The expectation of the normalized error covariance while not under attack is

$$E\left[(y_k - C\hat{x}_{k|k-1})^T \mathcal{P}^{-1} (y_k - C\hat{x}_{k|k-1})\right] = m$$

when under attack expectation is asymptotically given by

$$\lim_{k\to\infty} E\left[(y'_k - C\hat{x}_{k|k-1})^T \mathcal{P}^{-1} (y'_k - C\hat{x}_{k|k-1})\right] = m + 2trace(C^T \mathcal{P}^{-1} C\mathcal{U})$$

where $\mathcal{U}$ is the solution to the Lyapunov equation

$$\mathcal{U} - BQB^T = \mathcal{A}\mathcal{U}\mathcal{A}^T$$

The asymptotic expectation can be used to set a threshold for failure detection, in which values that rise above the set threshold indicate that the system is under attack. The value set for the threshold varies by the importance placed on detection rate versus false alarm rate. The greater the covariance is set to the greater the detection rate and the loss of performance will be. In order to achieve a detection rate of over 35%, the covariance must be 0.6, which means a sacrifice in performance of 91% with respect to the optimal performance [Mo, 2009]. Chabukswar, Mo and Sinopoli developed ways to improve on this sacrifice in performance by optimizing the covariance according to the desired performance constraints, but there is still great loss of performance [Chabukswar, 2011].

To improve on this loss of performance, [Miao, 2010] developed a game theoretic approach to replay attacks designed to minimize losses due to detection. Like Mo and Sinopoli, Miao et al consider the case of a LTI system with a LQG Controller, and a χ^2 failure detector. In this case, however, the authors design a system in which two controllers are switched between depending on the system dynamics. One is the optimal controller which operates without the addition of Δu_k and thus no loss of performance, and one is the controller designed by Mo and Sinopoli for replay attack detection with the addition of Δu_k. The controller is chosen by considering the system and the attacker as opponents within a game theory framework with a game divided into *K* stages corresponding to the time steps *n* considered in the previous work. Within the game, the attacker is considered the maximizer, or row player, and the system is considered the minimizer, or column player. Both may view the current state of the game but neither has complete knowledge of the other player's previous decisions. If the detector alarm is triggered by the attacker than the system is considered to have won, but the system is strictly penalized for false alarms. The game space with three game states is denoted $S = \{s_1, s_2, s_3\}$ and

the action spaces are $A_{tk}$for the attacker and A_{sk} for the system. The game states are shown in Figure 9. In each stage *k* the attacker's action space A_{tk}includes *m* options and the system action space A_{sk}has two, the two controller options. The safe space, in which the system has detected an attack, is denoted s_1. The no detection space in which no alarm has been triggered is s_2 and the state for a false alarm being triggered is s_3.

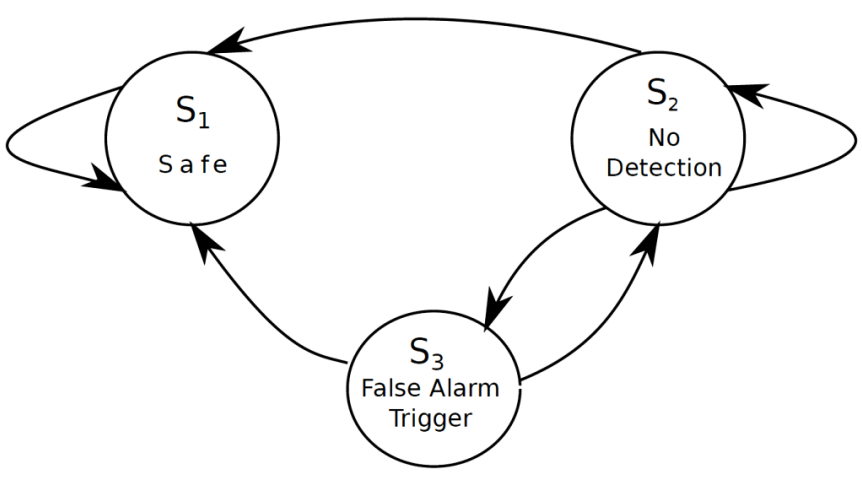

Figure 9: The stochastic game model developed by [Miao, 2010], with s1 in an absorbing state.

The state transition probability matrix is $\mathbb{P}_k$, r_k is the immediate payoff matrix, and the set strategies in each state *k* of the attacker and system for each system state are F_k and G_k. The state transition probability $\mathbb{P}_k$ is determined by $\tilde{P}_k^{ij}(s'|h_k, s)$, the probability provided by the detector of an alarm being set off and changing the system state given the system history h_k. At stage *k* with strategies f_k and g_k and probability that system is at state s_l given as $p(s_k^l)$, the probability of the system being at state $s_l$in stage *k+1* is

$$p(s_{k+1}^l) = \sum_{l=1}^{3} p(s_k^l)[f_k(s_l)]\,\tilde{P}_k(s_l|h_k, s_l)g_k(s_l)$$

These can be used to determine an optimal game strategy based on the state of the game and the expected payoff and a suboptimal game strategy which does not require knowledge of

the previous history of actions. Use of the suboptimal strategy can be shown to improve the control cost by approximately fifty percent compared to using only the controller with constant detection while the detection rate only decreases by approximately five percent when switching between controllers. The addition of game theoretic strategy increases the performance of the control system, but a trade-off between effectiveness of detection and function remains. The improvement in control cost is greater than the loss in detection rate, but given the initial detection rate still leaves room for improvement.

An alternate strategy for the detection of replay attacks using the addition of time stamped noise while reducing the loss of performance was proposed by [Tran, 2013] for use in smart grids. When applying the additional noise to the smart grid there will be discrepancies between the power usage and the usage that is measured by smart meters and possibly lead to a waste of power. To reduce this loss, this technique calls for the random signal to be added periodically for a relatively small span of time, allowing equipment to operate normally during the longer periods in which the signal has not been added.

Chabukswar, Mo and Sinopoli also worked on another approach to intrusion detection in the case of power grids, in which sudden problem with load could easily be due to changes in demand rather than an attack or a faulty sensor, making the Gaussian noise previously proposed by Mo and Sinopoli impractical. A wide sense stationary (WSS) Gaussian authentication signal is added to the set point of the generator while the controller runs a model of the generator and the effect of the new signal of the grid frequency. The actual grid frequency is then observed and changes are compared to the predicted change from the added signal. If the correlation is 0, the absence of the authentication signal is easily detected, signaling an attack on the system [Chabukswar, 2011].

These replay attack detection techniques, while an improvement over approaches that do not take the sensor data obfuscation of replay attacks into account, still do not adequately address the danger such attacks present, particularly given the loss of efficiency that results. Further directions of investigations should address these issues.

Chapter 5: Conclusion

5.1 Summary of Thesis

There are many avenues of attack through which a cyber-physical system may be vulnerable, and protection for each of those routes must be provided for. We have approached these avenues divided into the system that must be defended against.

The study of traditional cyber attacks was outlined in Chapter 2, which covered the ways of ensuring the integrity, availability, and confidentiality of the system. Cyber systems are vulnerable DoS attacks and network attacks that can be used to access unauthorized information, alter information within the system, or directly damage the system function. The system can be designed to be robust in the face of an attack through the correct use of keys and network topology can help protect the system. Incursions to the system can be detected by monitoring of quantified and learned values of system function for anomalous use indicative of an unauthorized use.

Chapter 3 discussed the security of sensor networks. Sensor networks are vulnerable to the addition of false nodes, the control of an honest node by an attacker to make it behave dishonestly, and the manipulation of the environment around a sensor to cause it to provide incorrect data. Such incursions can be detected in a sensor network through the examination of the values given by the sensors. Using these values as provided by sensors across the network over time, the reliability of individual sensors in the network can be evaluated. This can be evaluated by a single centralized decision maker or across the network by nodes acting together. Sensor values can also be weighted to provide maximum resilience to false node claims.

Chapter 4 covered the protection of control systems from damage or manipulation due to the DoS attacks, insertion of malicious signals, and deception through data readings used to

disguise an attack. Designing the system to minimize the damage caused by packet drop helps to detect against DoS attack. Appropriate filters can be used to detect signal insertion, but defense against replay attacks remain challenging, requiring the insertion of time-stamped noise which destabilizes the system. Switching between such a system and a more stable control system based on a game theoretic model allows the system performance and attack detection to be optimized. These attack detection and secure system design approaches may be combined to combat attack, ensuring, dependent on what attacks a system might face, a robust defense for cyber-physical systems can be developed.

5.2 Cyber-Physical Security Outlook

Despite the great strides made in the field of cyber-physical security, many weaknesses in the defensive approaches must be dealt with more thoroughly. Future research in cyber-physical security may address issues such as the reduction of cost and computing power necessary to detect attacks, improvement of system performance while attack detection is in operation, or the adaptation of existing security approaches to new or more specific kinds of cyber-physical systems. Some improvements in security are simply a matter of integrating known methods of protection to the systems currently in use.

Cyber security has a great deal of completed work in identification and defense from network attacks, but remains vulnerable to malicious attack by insiders and to denial-of-service attacks. Weaknesses in network topology have yet to be fully addressed. The internet has been shown to be even more vulnerable to path based attacks such as DoS [Lee, 2006], demonstrating that the great potential of a system whose network topology has not been optimized for security to be exploited by an attacker. While the need for this work has been shown [Lee, 2006], more

research must be completed in determining the optimum network topology for different types of attacks, including both DoS attacks and physical attacks such as physical sabotage to equipment.

A large portion of the research into sensor network security works well in optimal conditions, but has a tendency to false alarms due communication errors that occur in real world conditions. This leads to the expulsion of normally functioning sensors. Further work should be done on techniques that take such normal errors into account. An additional weakness in the field is that while detection of smaller numbers of false nodes in a sensor network is well documented, the protection of the network when a larger proportion of nodes are false becomes more difficult as the number of false nodes increases. A future direction of research in sensor networks is the detection of attack in the case of larger scale attacks.

Design of a controller in the presence of a non-optimal DoS attack has been formulated, but the formulation of such a controller in the face of an ideal attack strategy is more challenging. There is still a great deal of research to be done on the design of control systems to be resilient to DoS attacks in which the attacker has knowledge of the system and is attacking in the optimal way to cause damage to the system. Detection and design strategies for signal insertion to control systems have been well developed, but control systems are still vulnerable to replay attack. The best of current strategies have a low success rate and significant loss of function. In the area of detecting replay attacks in control systems, the improvement of the system performance when attempting to detect replay attacks is crucial to its functionality in real use. The improvement of system functionality when incorporating an added signal for time verification is crucial to the practical use of such techniques. Improvement of the rate of detection of such a system would also be valuable.

References

[1] Amin, S., Cárdenas, A. A., & Sastry, S. S. (2009). Safe and secure networked control systems under denial-of-service attacks. In *Hybrid Systems: Computation and Control* (pp. 31-45). Springer Berlin Heidelberg.

[2] Balasubramaniyan, J. S., Garcia-Fernandez, J. O., Isacoff, D., Spafford, E., & Zamboni, D. (1998, December). An architecture for intrusion detection using autonomous agents. In *Computer Security Applications Conference, 1998. Proceedings. 14th Annual* (pp. 13-24). IEEE.

[3] Benattou, M., & Tamine, K. (2005). Intelligent agents for distributed intrusion detection system. *World Academy of Science, Engineering and Technology*, *6*, 190-200.

[4] Byres, E., & Lowe, J. (2004, October). The myths and facts behind cyber security risks for industrial control systems. In *Proceedings of the VDE Kongress* (Vol. 116).

[5] Blanc, M., Briffaut, J., Clemente, P., El Rab, M. G., & Toinard, C. (2006, May). A collaborative approach for access control, intrusion detection and security testing. In *Collaborative Technologies and Systems, 2006. CTS 2006. International Symposium on* (pp. 270-277). IEEE.

[6] Chatzigiannakis, V., Androulidakis, G., Grammatikou, M., & Maglaris, B. S. (2004). An Architectural Framework for Distributed Intrusion Detection Using Smart Agents. In *Security and Management* (pp. 193-199).

[7] Chabukswar, R., Mo, Y., & Sinopoli, B. (2011, 9).*Detecting integrity attacks on scada systems*. Proceedings of the 18th world congress international federation of automatic control, Milano, Italy.

[8] Chen, R., Park, J., & Bian, K. (2008). Robust distributed spectrum sensing in cognitive radio networks.*Proceedings of the 27th Conference Computer Communication, IEEE INFOCOM*, 1876-1884.

[9] Coull, S., Branch, J., Szymanski, B., & Breimer, E. (2003, December). Intrusion detection: A bioinformatics approach. In *Computer Security Applications Conference, 2003. Proceedings. 19th Annual* (pp. 24-33). IEEE.

[10] Denning, D. E. (1987). An intrusion-detection model. *IEEE Transaction on Software Engineering*, *SE-13*(2), 222-232.

[11] Fadlullah, Z. M., Fouda, M. M., Kato, N., Shen, X., & Nozaki, Y. (2011). An early warning system against malicious activities for smart grid communications.*IEEE Network*, 50-55.

[12]Forrest, S., Hofmeyr, S. A., & Somayaji, A. (1997). Computer immunology.*Communications of the ACM*, *40*(10), 88-96.

[13] Gagrani, M., Sharma, P., Iyengar, S., Nadendla, V., Vempaty, A., Chen, H., & Varshney, P. (2011). On noise-enhanced distributed inference in the presence of byzantines. *Proceedings of the 49th Annual Allerton Conference Communications Control Computing*, 1222-1229.

[14] Gouda, M. G., & Liu, X. Y. (2004). Firewall design: Consistency, completeness, and compactness. In *Distributed Computing Systems, 2004. Proceedings. 24th International Conference on* (pp. 320-327). IEEE.

[15] Hadley, M., Lu, N., & Deborah, A. (2010). Smart-grid security issues. *IEEE Security and Privacy*, *8*(1), 81-85.

[16] Helmer, G. G., Wong, J. S., Honavar, V., & Miller, L. (1998, September). Intelligent agents for intrusion detection. In *Information Technology Conference, 1998. IEEE* (pp. 121-124). IEEE.

[17] Janakiraman, S., Vasudevan, V., & Radhakrishnan, S. (2006, September). Agent based Intrusion Detection System: A Computational Biology Approach. In *India Conference, 2006 Annual IEEE* (pp. 1-4). IEEE.

[18] Kushner, D. (2013). The real story of Stuxnet. *Spectrum, IEEE*, *50*(3), 48-53.

[19] Lee, H. E. E., Jong, K. I. M., & Lee, W. Y. (2006). Resiliency of network topologies under path-based attacks. *IEICE transactions on communications*, *89*(10), 2878-2884.

[20] Liu, Y., Ning, P., & Reiter, M. K. (2011). False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, *14*(1), 13.Mo, Y, Kim, THJ, Brancik, K, Dickinson, D, Lee, H, Perrig, A, & Sinopoli, B. (2011). Cyber-physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 1-15.

[21] Long, M., Wu, C. H. J., & Hung, J. Y. (2005). Denial of service attacks on network-based control systems: impact and mitigation. *Industrial Informatics, IEEE Transactions on*, *1*(2), 85-96.

[22] Miao, F., Pajic, M., & Pappas, G. J. (2010). Stochastic Game Approach for Replay Attack Detection.

[23] Mitchell, R., & Chen, I. R. (2013). Effect of intrusion detection and response on reliability of cyber physical systems.

[24] Mo, Y., & Sinopoli, B. (2009, September). Secure control against replay attacks. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on* (pp. 911-918). IEEE.

[25] Mo, Y., & Sinopoli, B. (2010, April). False data injection attacks in control systems. In *Preprints of the 1st Workshop on Secure Control Systems* (pp. 1-6).

[26] Mo, Y., Kim, T. H., Brancik, K., Dickinson, D., Lee, H., Perrig, A., & Sinopoli, B. (2012). Cyber–physical security of a smart grid infrastructure. *Proceedings of the IEEE*, *100*(1), 195-209.

[27] Mukherjee, B., Heberlein, L. T., & Levitt, K. N. (1994). Network intrusion detection. *Network, IEEE*, *8*(3), 26-41.

[28] Parno, B., Perrig, A., & Gligor, V. (2005). Distributed detection of node replication attacks in sensor networks. *IEEE Symposium on Security and Privacy*, 49-63.

[29] Pasqualetti, F., Dorfler, F., & Bullo, F. (2011, December). Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on* (pp. 2195-2201). IEEE.

[30] Pasqualetti, F., Dörfler, F., & Bullo, F. (2012). Attack Detection and Identification in Cyber-Physical Systems--Part I: Models and Fundamental Limitations. *arXiv preprint arXiv:1202.6144*.

[31] Pasqualetti, F., Dörfler, F., & Bullo, F. (2012). Attack Detection and Identification in Cyber-Physical Systems--Part II: Centralized and Distributed Monitor Design. *arXiv preprint arXiv:1202.6049*.

[32] Pinar, A., Meza, J., Donde, V., & Lesieutre, B. (2010). Optimization strategies for the vulnerability analysis of the electric power grid. *SIAM Journal on Optimization*, *20*(4), 1786-1810.

[33] Rashvand, H., Salah, K., & Calero, J. (2010). Distributed security for multi-agent systems-review and applications. *IET Information Security*, *4*(4), 188-201

[34] Rawat, A. S., Anand, P., Chen, H., & Varshney, P. K. (2011). Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks. *IEEE Transactions on Signal Processing*, *59*(2), 774-786.

[35] Roy, S., Xue, M., & Das, S. K. (2012). Security and Discoverability of Spread Dynamics in Cyber-Physical Networks. *Parallel and Distributed Systems, IEEE Transactions on*, *23*(9), 1694-1707.

[36] Salmeron, J., Wood, K., & Baldick, R. (2004). *Analysis of electric grid security under terrorist threat*. NAVAL POSTGRADUATE SCHOOL MONTEREY CA DEPT OF OPERATIONS RESEARCH.

[37] Soltanmohammadi, E., Orooji, M., & Naraghi-Pour, M. (2013). Decentralized hypothesis testing in wireless sensor networks in the presence of misbehaving nodes. *Information Forensics and Security, IEEE Transactions on*, *8*(1), 205-215.

[38] Soltanmohammadi, E., & Naraghi-Pour, M. (2014). Fast Detection of Malicious Behavior in Cooperative Spectrum Sensing.

[39] Sundaram, S., & Hadjicostis, C. N. (2011). Distributed function calculation via linear iterative strategies in the presence of malicious agents. *Automatic Control, IEEE Transactions on*, *56*(7), 1495-1508.

[40] Tran, T. T., Shin, O. S., & Lee, J. H. (2013, January). Detection of replay attacks in smart grid systems. In *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on* (pp. 298-302). IEEE.

[41] Turk, R. J. (2005). *Cyber incidents involving control systems*. Idaho National Engineering and Environmental Laboratory.

[42] Vempaty, A., Tong, L., & Varshney, P. K. (2013). Distributed inference with byzantine data. *IEEE Signal Processing Magazine*, *30*(5), 65-75.

[43] Zhang, Y., Wang, L., Sun, W., Green, R., & Alam, M. (2011). Distributed intrusion detection system in multi-layer network architecture of smart grids.*IEEE Transactions on Smart Grid*, *2*(4), 796-808.

[44] Zhao-wen, L., Xing-tian, R., & Yan, M. (2007, August). Agent-based Distributed Cooperative Intrusion Detection System. In *Communications and Networking in China, 2007. CHINACOM'07. Second International Conference on* (pp. 17-22). IEEE.

Vita

Sarah Davis, a native of Slidell, Louisiana, received her bachelor's degree at Tulane University in 2008. She went on to become a graduate student at Louisiana State University in the Department of Electrical and Computer Engineering. She will receive her master's degree in May 2014 and plans to find work as an engineer upon graduation.