M-TRA: A Multi-Tiered Resilient Architecture for Cyber-Physical Systems

by

Janice Cañedo

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 2017

Keywords: Internet of Things, Architecture, Security, Cyber-Physical Systems

Approved by

Anthony Skjellum, Professor of Computer Science and Software Engineering

Abstract

The Internet of Things (IoT) is a growing network of devices connected to the Internet. IoT devices range from wearables including smart watches and heart rate monitors, home automation including thermostats, locks, appliances, and Cyber-Physical Systems (CPS) including smart grids and smart healthcare. As IoT/CPS has grown, security, scalability, and resiliency are all critical points of interest for practical,trustworthy systems. At present, many systems contain weak security practices that result in hacks, failures, threats to safety, economic losses, and/or compromised data.

To address these issues, we propose M-TRA: A Multi-tiered Resilient Architecture for IoT/CPS. M-TRA would allow for a secure, scalable, and resilient system. To provide resiliency, tiers of gateways will form a coalition to determine how edge device traffic flows throughout the system in an effort to distribute the load over the system. For security, trust will be established between devices before a new device enters the system. Once trust is establish, encryption will be used to maintain safe communication between the devices. As the system runs, active monitoring will be deployed to detect anomalies in the system including device failures and hacks using GP-GPU enhanced machine learning.

The goal of this research is to create a secure, scalable, and resilient IoT/CPS architecture. We will analyze trade-offs in performance as the system scales in size range from 10 devices to 1000s of devices.

Contents

List of Figures

List of Tables

Chapter 1

Overview

1.1 Introduction

It is estimated that by 2020 there will be over 20 billion devices connected over the Internet [1, 2, 3]. This collective of connected devices is called the Internet of Things (IoT). IoT systems can range from home monitoring, appliances and smart watches to smart hospitals and smart buildings. Some IoT systems can also be categorized as Cyber-Physical Systems (CPS), such as the smart grid and smart healthcare [4, 5, 6]. CPS are mission-critical systems that interact with their surrounding environment [7, 8, 9]. IoT forms a foundation for Internet connected CPS [6]. As more devices become connected, there will be a need to understand how the IoT and CPS Systems can rapidly scale, be secure, and function with resiliency [10, 11].

Within an IoT and CPS systems devices can be divided into two primary groups: edge devices and gateway devices. Edge devices are low resource, simple devices that usually perform a single purpose, for instance reading ambient temperature. Gateway devices are more powerful than edge devices. They are used to aggregate data from edge, perform analysis on the data, make informed decisions, and connect data back to the Internet when required. IoT and CPS systems share many commonalities. There are two distinct differences between these systems. First, IoT systems do not always make changes to the surronding environment, while CPS do interact with an alter their physical environment [5, 6]. Second, CPS systems do not always connect back to the Interent. There can be closed systems CPS. Figure 1.1 provides a representation between IoT and CPS systems. We will refer to the connection of these systems as IoT/CPS. The goal of this dissertation is to address scalability,

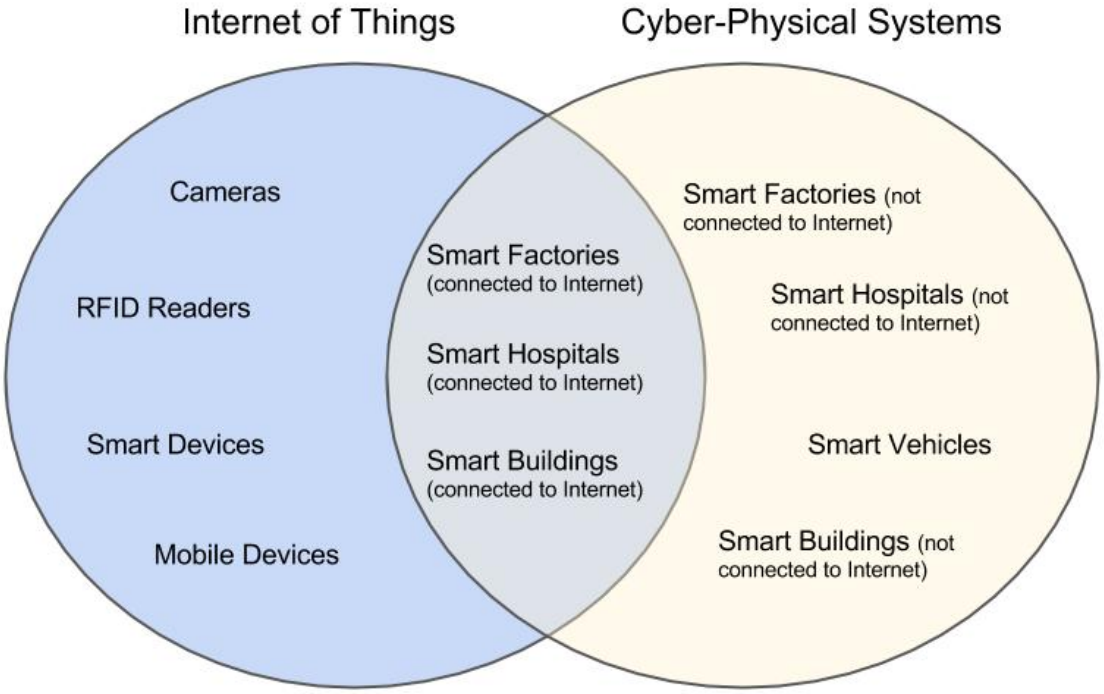

Figure 1.1: Internet of Things and Cyber-Physical Systems

security, and resiliency separately and jointly in IoT/CPS architectures in an effort to create a robust, adaptable architecture.

The remainder of this chapter is organized as follows: Section 1.2 defines the key attributes of an IoT/CPS Architecture. Section 1.3 overviews key technologies used in this dissertation. Section 1.4 examines current challenges and opportunities. Section 1.6 is the statement of contribution provided in this dissertation. Section 1.7 outlines our innovative claims. Finally, Section 1.8 overviews the broader impact associated with this dissertation.

1.2 Key Architecture Attributes

There are three key attributes within a cohesive IoT/CPS Architectures for a robust system including security, scalability, and resiliency. Below we define and identify key points of each area.

1.2.1 Secure Architecture

Within an IoT system, there are multiple levels of security including network security, secure device handshakes, and protection against intrusion into the devices. According to Höller, the Security Model for IoT consists of communication security that focuses mostly on the confidentiality and integrity protection of interacting entities, and functional components such as Identity Management, Authentication, Authorization, and Trust & Reputation [12].

The SANS Institute defines network security as the process of taking physical and software preventative measures to protect the underlying networking infrastructure from unauthorized access, misuse, malfunction, modification, destruction, or improper disclosure, thereby creating a secure platform for computers, users and programs to perform their permitted critical functions within a secure environment." [13]

First, verifying the validity of a device before it is allowed to enter the system is fundamental for a trusted system. Secure device handshakes can provide this security to a system. For an IoT system, the handshake needs to be light-weight enough to run on end devices, yet robust enough to not be easily replicated or faked. Second, monitoring of each device is important to enable verification that a device has not become corrupt.

Along with security within the IoT system, if an intrusion or defect is detected, it is necessary to have the ability to determine who or what caused the failure. Digital Forensics can be implemented to provide a means to find this information.

1.2.2 Scalable Architecture

When identifying scaling within an individual computer, there are two types of scaling: strong scaling and weak scaling. Strong scaling is CPU bound scaling. In strong scaling the problem size stays fixed while the number of processing elements are increased [14]. Weak scaling is memory bound scaling. In this case the problem size (workload) assigned to each processing element stays constant and additional elements are used to solve a larger total problem (one that wouldn't fit in RAM on a single node, for example) [14]. Scalability in parallel

computing systems is the measure of its capacity to increase speedup in proportion to the increase in the number of processing elements [15].

Determining the speedup and efficiency within the systems as scaling occurs can be done by using Amdahl's law and Gustafson's law. Amdahl's law analyzes the speedup of a program given the addition of parallelization [16]. Gustafson's law analyzes the increase in workload that can occur during a given time with the addition of parallelization. These two laws provide a foundation for determining the efficiency within the program is maximizing the use of the system.

In IoT architecture, scalability includes scalability of operations within a gateway and scalability of increasing the number of devices included within the system. By analyzing weak and strong scaling within the gateway, we can investigate the ability to scale up the number of devices by using parallel processing with the CPU and GP-GPU.

1.2.3 Resilient Architecture

According to Delic, Resiliency in an IoT system is defined in three parts [17]. First, an IoT system needs the ability to resist external perturbances and internal failures. Second, an IoT system should be capable of recovering and reentering a stable state after a failure occurs. Third, an IoT system should adapt its structure and behavior to constant change.

1.3 Overview of Technology

The following subsections provide an overview of each key area we explore with our work.

1.3.1 The Internet of Things

In 1999, Kevin Ashton coined the term *Internet of Things* for a presentation for Proctor and Gamble when linking RFID technology in Proctor and Gamble's supply chain to the Internet [18, 19]. Ashton used the term IoT to describe the network connecting devices together to the Internet [20]. IoT has subsequently achieved a broad reach and garnered a broader

still definition. Consumer IoT includes the connection of a device to the Internet including smart phones, watches, thermostats, refrigerators, alarm systems, cameras, cars, lights, etc [21, 22, 23]. Industrial IoT (IIoT) is the use of IoT devices in industries including factories, companies, plants to increase revenues by boosting production [24]. IoT has grown drastically since it's notional inception in 1999. For instance, between 2014 and 2016, it is estimated that there was a 68% increase in the number of devices connected online [1]. It is expected that between 2016 to 2020 there will be a further increase of 225% with an estimated 20 billion devices connected online [1].

System Examples

IoT environments are clearly becoming increasingly popular, and they are moving to large-scale adoption. Two fast growing IoT areas are Smart Lighting and Smart Hospitals. The following section explore these two "verticals" and the need for resiliency, security and stability for such applications.

Smart Lights Smart Lighting has becoming increasingly more accessible and popular in businesses and homes. Smart Lighting is rapidly growing and it is estimated it will be worth $8.14 billion by 2020 [25]. According to Digital Lumens, Smart Lighting is defined as LED fixtures with intelligence and sensors whose data can be harvested for a range of purposes, primarily to control when and where light is on [26].

Smart lighting is an increasingly popular area of IoT. As more fixtures installed in the workplace and home are smart, the need for increased security and scalability become correspondingly more important. There is a need to ensure that as device interconnect, they connect both securely and that each device is a part of a trusted system. Given these needs, a scalable architecture, device handshakes, lifecycle management, low latency between commands, and system monitoring are important needs for such deployments.

Common Large-scale IoT Themes The following are common themes of large-scale IoT deployments, adding directly to in many cases the complexity of supporting scale, security, and resiliency. These include:

- geographic distribution (*e.g.*, buildings, campuses, cities)
- mobility and transience of some of the devices
- redundancy and fungibility of certain ubiquitous devices (like temperature sensors)
- uniqueness or low-redundancy of high value devices
- inaccessibility of many devices
- low-resource, low power devices are in the vast majority
- heterogeneity of device types, sources, purposes, and ages.

1.3.2 Cyber-Physical Systems

CPS are closely integrate, context aware, mission critical systems that interact with the physical world [9, 7, 8]. These distributed systems are used in Healthcare and Medicine, Electric Power Grid, Unmanned Vehicles, and Next Generation Air Transportation Systems [7, 8]. Edge devices within a CPS actuate and affect the physical environment. One example of CPS are Healthcare and Smart Medicine.

Healthcare and Smart Medicine In view of the rate in which technology has advanced, healthcare providers are beginning to adopt addition technology in their environment to help expedite service, improve patient care, and provide additional assistance to patients as needed. In 2014, Healthcare Design undertook an overview of how technology could be used to create a Smart Hospital [27]. Healthcare Design determined that currently planning for technology in new buildings or renovations is necessary for hospitals [27]. A Smart Hospital is a hospital that includes the following characteristics [27, 7]:

- Check-in Kiosks
- Connective Furniture
- Hybrid Exam/Consult Modules
- Smart operating rooms
- Telemedicine
- Electronic Medical Records Scanning Room
- Interactive Features to promote movement in patients

Within this particular environment, scalability, resiliency, and security are imperative to ensure the best care for patients and protection of patient information.

1.3.3 Fog Computing

A fog architecture is an architecture organization in which local clusters of devices are used in conjunction with the cloud to expand storage and networking services between the cloud and edge devices [28]. This brings computing power down from the cloud service to the gateway and edge device layer. This allows for scalability to occur more easily due to increased architecture at the gateway layer.

1.3.4 High Performance Computing

High Performance Computing (HPC) is the use of parallel processing to running advanced programs, particularly in clusters of devices. In HPC, parallel computation of data provides a methodology for performing computations on large data sets across multiple threads in multiple CPUs and/or General Purpose Graphical Processing Units (GP-GPUs). Parallel computing has applications from data mining to engineering applications including combustion engines and high-speed circuits [15].

1.3.5 Machine Learning

Machine Learning is an area of Artificial Intelligence (AI) in which computer programs are enabled to learn from experience, examples, and analogies [29]. As learning occurs, the capabilities within a program become more intelligent and the program becomes capable of making informed decisions. Within machine learning, two of the most popular approaches are artificial neural networks (ANN) and genetic algorithms.

ANNs mimic the neurons and synapses within the brain to transfer data for communication, learning, and decision making [29]. Current ANNs are much simpler than the brain, but still follow the same principles. Within the brain, neurons are a dense set of interconnected nerves that are connected by synapses. The synapses send information between the neurons. The neurons learn from the information and are used for decision making. Within the brain there are approximately 10 billion neurons and 60 trillion synapses [29]. Within a computer, neural networks mimic this structure for learning where neurons are mimiced by nodes and synapses are mimiced by weighted connections. Each ANN contains an input layer of nodes connected to additional hidden layers using weighted connections then to an output layer. The hidden layers of neurons within an ANN help provide additional tailoring of the learning. Each level adds additional weights to the connections that provide the basics of long-term learning within the system [29].

The first neural network was built in the 1950s by two Harvard students, Marvin Minsky and Dean Edmonds [30]. Minsky and Edmonds' ANN simulated a rat learning to navigate a maze. Since 1950, neural networks have expanded from single perceptron learning, to multilayer perceptron learning algorithms. Several major advancements in ANN occurred in the 1980's. One major advancement was backpropogation was rediscovered. The backpropogation learning algorithm allowed for multilayer neural networks to learn and has become one of the most popular learning algorithms for multilayer neural networks [29]. A second major advancement in ANN were recurrent neural networks (RNN) . A RNN adds feedback loops to the network that connect the output to the input [29]. These feedback loops

allows the RNN to form short term memory and learn sequences [31]. In 1997, Hochreiter and Schmidhuber proposed Long Short-Term Memory(LSTM) , which provide the addition of a memory cell with gradient based learning within the feedback loop to allow for long term memory to occur within a RNN [32]. This approach allows for a long term memory as well as short term memory to occur as the system learns sequences. This can be used in time-series predictions, robotics, handwriting prediction, and many other applications.

Now, artificial intelligence and neural networks are present in many technologies including medicine, image compression, and stock market analysis [33]. ANNs are used within IoT systems to monitor the state of IoT devices and to make informed decisions [34]. We propose the addition of neural networks to help secure an IoT system.

1.3.6 Active Monitoring and Attack Prevention

Active monitoring is the accomplished when a tool or hook is placed within a system and when the execution reaches the hook it will cause and interrupt and control will be handed to the security program [35]. Active monitoring can detect anomalies as a system runs. This allows for detection of intrusions or anomalies to provide the ability to prevent attacks within a system.

1.3.7 Digital Forensics

When an anomaly or intrusion occurs, there is a need to determine what or who was the cause. This can be done with digital forensics. For an IoT system, there are two areas of digital forensics to investigate: device forensics and network forensics. Device forensics includes forensics of device memory and device storage. Network forensics focuses on network intrusion detection, attack analysis, and collecting evidence using network packets and other network analysis techniques [36]. Combining the device forensics and network forensics within an IoT system allows for comprehensive forensics of IoT systems.

1.4 Challenges and Opportunities

In IoT/CPS system, there are several challenges that must be addressed. As the system scales there becomes a challenge due to additional latency in packet transfer. It is necessary to maintain low latency in an effort to have a responsive system to ensure adequate QoS. Additional challenges exist in managing the low resources in edge devices efficiently. Low resources include limited memory, low power, limited storage, and slower processing. Identifying a means of establishing trust between these low resource edge devices and the gateway provides an opportunity for additional effective methods of establishing trust.

1.5 Current Limitations and Problems

When analyzing current IoT systems, several limitations exist including weakness in security, connecting heterogeneous devices, and lack of system-wide resiliency. Below we further investigate each of these limitations.

1.5.1 Weakness in Security

IoT Security is universally weak. Many implementations of IoT systems treat security as an afterthought [37]. This causes difficulty in adding security measures later because underlying code and architecture are not designed correctly. Many security experts also believe security measures are actively resisted and circumvented [37].

In a survey conducted by HP in 2015, they found that 80% of devices, along with their application components, failed to require passwords of a sufficient complexity and length [38]. One specific problem is there is a lack of light-weight, secure hand-shake protocols between IoT devices and gateways. This means that either devices must be or are made larger than necessary to use heavier protocols or the protocols that are being used are simpler to hack.

A second problem is IoT systems rarely contain alerts for detection of intrusion or malicious activity. This means that if malicious activity or intrusions occur and do not

cause a device failure, it will take longer to detect because there are no automated detection techniques implemented across the system.

A third weakness in IoT security is a limited quantity of methods for performing forensics investigation to determine who or what caused an intrusion. From previous research we have determined that there is a need to adapt current tools for IoT forensics in a effort to perform forensics investigation with limited data from the physical devices while waiting for data from the cloud [39].

1.5.2 Heterogeneous Systems

Many IoT systems require heterogeneous devices can contain different types of sensors, sensor data, networking, and data transfer rates. Chipset and device specifications are also different. Since there is no standard in architecture or networking, it can become extremely difficult to connect devices together. There are also limitations in current routing algorithms to balance how the data transfers between each of these different protocols and devices.

1.5.3 Lack of Resiliency

Many systems contain numerous single failure points due to single gateway design. This means edge devices talk to a single gateway and the single gateway talks to the cloud. This provides a single failure point within a system. A lack of sensor redundancy can also cause failures due to sensors breaking or malfunctioning.

1.6 Statement of Contributions

The goal of this research is to create a secure, scalable,and resilient IoT/CPS architecture by using analytics within the gateway to distribute edge device connection among gateways for resiliency, using machine learning to monitor a secure system state, provide feedback regarding anomalies to the user, and collect relevant data to determine who or what caused the anomaly. To secure the system further, device handshakes between devices will be investigated in an

effort to determine a secure and scalable method for devices to connect into collectives. Finally, architectures include mesh, point-to-point, single-tier gateway, and multi-tier interconnection (network) architectures will be investigate to determine their benefits and weaknesses(singly and in composition) as a system scales from a few devices to thousands of devices and beyond.

1.7 Innovative Claims

In examining the scalability, resiliency and security of IoT systems, there are key innovations we are investigating.

1.7.1 Need for HPC in IoT

Given that IoT systems are a group of devices clustered together, HPC principles can be adapted including parallel computation of edge device data within the gateways. We investigated the use of parallel computation of edge device data and found that in most cases parallel computation outperformed sequential computation by up to a factor of 10 times, which has head room over dedicating all the ARM cores to the problem by at least a factor of 2.5.

We also propose that machine learning can be implemented within the gateways and used to learn the healthy state of the system. By continuous monitoring of the system, as the system grows, the model can be recalculated for the new state of the system. This can allow for gateways to provide device checking and security to the IoT system. We believe this will provide a resilient system by reducing single failure points, providing redundancy, and implementing a self-healing network.

1.7.2 Lightweight Handshake for ARM devices

We propose the use of certificates to provide a lightweight handshake between end devices and gateways, as well as adapting algorithms for secure handshakes between gateways and

end devices. This will allow smaller ARM architecture devices to authenticate securely to gateways. Secure authentication is necessary for a trusted IoT system.

1.7.3 Multi-Tiered Resilient Architecture

Many current IoT/CPS implementations lack the redundancy and resiliency necessary for large scale implementations. We are convinced that a multi-tiered architecture will provide resiliency and scalability to a system by allowing reduced single failure points and providing an infrastructure that can monitor itself. A smart gateway or coalition of gateways can be used to route traffic in an efficient manner from edge devices to lower latency, ensure quality of service (QoS), and provide ability to reroute traffic as necessary to maintain an efficient, resilient system.

1.8 Broader Impact

This dissertation and research aim to provide a scalable architecture for large-scale IoT/CPS applications including smart grids, smart hospitals and smart factories. The use of embedded HPC in a scalable IoT framework will allow for more powerful programming and computing opportunities including machine learning to provide active monitoring on a systems, smart routing to enable more efficient flow of network traffic, and light-weight handshakes for low resource devices. As IoT/CPS application spread from home automation and security, to smart healthcare, and to the smart grid, IoT/CPS system are increasingly integrated into everyday life and therefore it becomes more necessary and import to provide and secure, resilient, and scalable system.

1.9 Summary

IoT/CPS is quickly evolving and becoming integrated more in our everyday world. These systems require the ability to scale from small home automation systems to large-scale smart grids. With CPS, it is necessary that these devices and systems are secure and resilient.

There is a need to investigate means to provide security within these systems including lightweight device handshakes and active monitoring for anomalies and intrusions. Along with security, resiliency within a IoT/CPS system is of concern. Resiliency requires redundancy of communication and ability to recover from an anomaly. Finally, the ability to maintain a resilient and secure system must remain be scalable as a system grows from 10 devices to 1000s of devices. The goal of this dissertation will be to address these issues and outline a scalable, resilient and secure architecture for IoT/CPS Systems.

The remainder of this dissertation proposal is organized as follows: Chapter 2 is a review of literature covering CPS, security, scalability, and resiliency in IoT/CPS systems. Chapter 3 is an overview of research questions we intend to answer during this research. Finally, Chapter 4 explains the approach and metrics we will use during the research process.

Chapter 2

Review of Literature

The following sections overview the current literature in each area of focus related to this dissertation.

2.1 Cyber-Physical Systems

Cyber-Physical Systems are systems of embedded devices that are interact with the physical world. The physical world provides the information that is necessary for the system to operate [9]. For instance, with unmanned vehicles integrated intelligent roads are used for navigation [7]. Below we investigate the current state of the art in CPS and CPS challenges including security concerns.

2.1.1 State of the Art

CPS are currently evolving and expanding. Many applications are currently in the research and development stage. Smart healthcare and medicine is one leading area of CPS. Smart healthcare includes electronic patient record initiatives, home care, intelligent operating room, image-guided surgery and therapy, and fluid flow control [8, 7, 9]. GE currently has an initiative for smart healthcare including smart cardiac and smart dose [40].

Next Generation Air Transportation Systems are CPS systems that impact the future of flight and aviation [8]. These systems include the use of integrated flight deck systems, functionality to achieve greater safety and become more efficient.

Another major application of CPS is the Smart Grid. IBM is currently investigating Smart Grid technologies [41]. The goal of the smart grid is to provide an infrastructure capable of handling distributed generation, renewable energy sources, electric vehicles, and

demand-side management electricity [42]. There are seven key requirements that were identified by the Department of Energy for the smart grid to meet demands. These seven are self healing, motivates and includes the consumer, resists attacks, provides power quality, accommodates generation and storage options, enables markets, and optimizes assets for efficient operation [42, 43]. Smart Grid is a CPS with far reaching impact and purpose, however, there are many challenges to a secure and reliable smart grid.

2.1.2 CPS Challenges

Many challenges still exist within CPS. First, they interact with the physical world and can make changes to the physical environment. These changes must maintain a safety requirements of the system [44]. The physical world introduces the physics of timing constraints and these must be obeyed [44]. Second, CPS are distributed systems [44, 7]. These systems can also be mobile. The mobility adds additional challenges dues to Internet access and connectivity, availability of energy, and the context of location [45]. Third, CPS are vulnerable to many security threats. These threats include cyber criminals, activists, disgruntled employees, or others who intend to crack the system [46]. Other security threats include device authentication, adapting to various environments, preserving privacy including sensor data and location information [45, 46, 8]. Fourth, systems must be resilient and reliable. With the mission critical nature of CPS, it is essential for these systems to be resilient, reliable and secure.

2.2 Security

We investigate three key areas for use within IoT security: machine learning, network/device security and forensics.

2.2.1 Using Machine Learning to Secure Systems

In *Computer Security and Machine Learning: Worst Enemies or Best Friends?*, Rieck investigated the problems, challenges and advantages of using Machine Learning to help secure a system[47]. According to Riech, there are five factors that impact the efficacy of using machine learning for securing a system: effectively, efficiency, transparency, controllability, and robustness. One area of research in which machine learning is effective in practice is intrusion detection within a network [47]. Another promising area discussed is automated analysis of threats. Machine learning can be used to provide an instrument for accelerating threat analysis.

The effectiveness of machine learning within networks and accelerating threat analysis suggests machine learning could be an effective manner of securing an IoT system. We intend to further this study by implementing neural networks to secure a system.

2.2.2 Network and Device Security

In *Research on the Basic Characteristics, the Key Technologies, the Network Architecture and Security Problems of the Internet of Things*, Xingmei, Jing, and He introduce the key concepts, network architecture, and security problems within the Internet of Things [48]. IoT requires intelligent processing and reliable transmissions within the network. To provide this, the network architecture contains three layers: the application layer, the transport layer, and the sensing layer. The application layer contains the link between the user and the Internet through intelligent application, (for example, intelligent architecture or intelligent home furnishings). The application layer uses machine learning, data mining, data processing, and other analytics to process information from the system and provide an output. The transport layer consists of various networks including Wi-Fi, Bluetooth, ZigBee, and 802.15.4. The transport layer contains the gateways that process the information and send it across the network. The sensing layer contains end devices that are composed of a variety of sensors and actuators. The data from these sensing layer is sent through the transportation layer to the

application layer for analysis. This approach provides a framework for reliable communication, however, as noted by the authors, there are many security threats present in the transport layer. Our approach is to add machine learning within the transport layer to help determine if there are interruptions in the data transfer and to monitor the end devices from the sensing layer.

One key feature of IoT systems is they contain a heterogeneous combination of networking techniques and devices. Grabovica, et al., provided a summary of networking techniques for connecting IoT devices in *Providing Security Measures of Enabling Technologies in Internet of Things (IoT): A Survey* [49]. There are four techniques summarized: ZigBee, Bluetooth, RFID, and WiFi. ZigBee, which is a personal area network created by the ZigBee alliance, is placed on top of the physical and MAC layer of the 802.15.4 stack. This provides an application layer and network layer. The security within the ZigBee protocol is provides the MAC layer of 802.15.4 protocol. Additionally, there is cryptography which is based on a 128-bit key and AES encryption. Bluetooth is an open standard for short range radio frequency that was created in 1994 [50]. Bluetooth provides the developer with four security modes to choose from and three encryption modes. For a secure Bluetooth connection, authentication, encryption, and authorization are required. There are threats against Bluetooth technology. The authors mentioned Bluejacking, Car Whispering, and Bluesnarfing are examples of threats that occur by exploiting a Bluetooth connection. Radio Frequency Identification is used for automatic identification of people or objects. There are two types of RFID tags, active and passive. Active require power a source and passive do not. There are three frequency ranges that RFID tags operate on, four security modes, and three levels of encryption. Potential threats against RFID tags include clandstien tracking or scanning, skimming, and cryptography weaknesses. The final technology discussed by Graboviac, et al., was WiFi. WiFi enables access to the Internet through radio signals. WiFi is enable with the ability to select 6 different security modes and four levels of encryption. The main threats against WiFi include weaknesses in WEP encryption, search wireless signal attacks,

and eavesdropping. As noted, there is no singular secure method of communication between devices. Given the difference in each communication method, there is a need to monitor communication in an effort to provide a more secure system.

One method of securing communication includes encryption in communication. IoT devices are lightweight and therefore need lightweight encryption methods [51]. In *Pass-IoT: A Platform for Studying Security, Privacy, and Trust in IoT*, Arseni, et al., presents a testbed for implementing IoT systems and enhancements to existing encryption algorithms are reviewed. The encryption techniques reviewed by the authors in Present, Tea, and Trivium. Present is substitution based block cipher that is implement for lightweight and super lightweight cryptography. Tea was developed by Roger Needham and David Wheeler. It is a block cipher with 64-bit data block sizes and 128-bit key. Previously there were exploits within the key schedule that were addressed. Trivium was designed by exploring how stream ciphers can be simplified. The goal was to simplify stream ciphers without sacrificing security, speed, or flexibility. Given this, Trivium is a lightweight synchronous stream cipher that generates up to 264 bit stream, 80 bits of secret key, and 80 bits of initial value. These techniques allow us to further analyze methods for securing device handshakes and communicating between the devices. We also propose further investigation of lightweight encryption techniques and certificates to securely connect edge devices to gateways.

Once devices connect securely, it's important the devices themselves are secure. In *Hardware Security Assurance in Emerging IoT Applications*, Dofe, et al., reviewed to hardware attacks: hardware trojans and side-channel analysis attacks [52]. Hardware Trojans are malicious modifications on the original chip that are used to corrupt the chips normal operation. Side-channel analysis attacks occured when side channels are analyzed of a cryptographic device to guess the secret key. Power is one side-channel that is often analyzed. To address these issues, the authors propose the use of dynamic permutation method. The dynamic permutation method changes the original order of the information received by the sensor. Due to the original order being changed, attackers can't precisely find a predefined

condition to perform a Trojan hardware attack. This approach provides a method for reducing hardware trojans and side channel attacks. We propose the addition of machine learning to such a model could help detect not only these forms of attacks, but other attacks within an IoT system as well.

In *Neural Network Approach to Forecast the State of the Internet of Things Elements*, Katenko, et al., investigated the use of neural networks to forecast the state of an IoT element [34]. They're approach combined a multilayered perceptron network along with a probabilistic neural network. They discovered that by using the multilayer perceptron network to look at similar values throughout the past, they could then use a probabilistic neural network to determine the state of the element. They found they were able to reduce the labor costs of the IoT administration and emergency resolution through this technique.

While their technique did reduce labor costs and allowed for monitoring and forecasting an element in an IoT network, the need to forecast the entire state of an IoT system was still needed. We propose the use of machine learning techniques as mentioned above in both of the gateways to monitor subsystem components, and in the application layer of the whole system to monitor the state of the entire system.

2.2.3 Forensics

Forensics for IoT systems can be challenging. Recent studies have suggested new methods for DF in IoT systems to handle the short comings in current techniques. One major concern in IoT forensics is jurisdiction [39]. Currently, most IoT data is sent to cloud storage which can be in a different region or state from the actual system. If a crime does occur, the data must be subpoenaed from the cloud service provider. Waiting on the completion of paperwork and appropriate subpoenas can have a negative impact on the investigation because of the loss of time. To help minimize the wait time and provide improve forensics, several models have been suggested. In *FAIoT: Towards Building a Forensics Aware Eco System for the Internet of Things*, Zawoad and Hasan propose a Forensics-aware IoT (FAIoT) model for support

forensics in IoT systems [53]. Within their model, constant monitoring occurs for each device and evidence is stored in a secure shared repository. This allows for the evidence to be access using API's. A second approach to IoT forensics was suggest by Oriwoh and Sant in *The Forensics Edge Management System:A Concept and Design* [54]. Within the Forensics Edge Management System (FEMS) , smart devices within a home autonomously providing basic security and forensic services by monitoring itself and reporting back to the user.

Both approaches that there is a need for continuous monitoring within an IoT system. With our proposed security approach of using machine learning to monitor the health of the system, we also propose the addition of identifying key data elements within the IoT system that would be stored locally to help aid in forensics investigations.

2.3 Scalability

When studying scalability, we investigate the current use of scalability within high performance computing and within IoT systems. The following sections provide a brief overview of the literature in each of these areas.

2.3.1 Scalability in High Performance Computing

In the study of scalability, Amdahl's law is a key area of focus. Amdahl's law focuses on the speedup of a program given the proportion of the program that can be parallelized[16, 55]. A second law of focus in scalability is Gustafson's law which workload can scale up and execution time remain the same based on the addition of parallelization of processing[16, 56, 55]. Amdahl's law and Gustafson's law provide the foundation of determining the effectiveness of scaling within a system.

In *Reliability-Aware Scalability Models for High Performance Computing*, Zheng and Lan investigate the impact of failures on Amdahl's law and Gustafsons law [55]. By extending Amdahls law and Gustafsons law, the study was able to derive a s study reliability-aware scalability models. Trace simulations were used to determine the accuracy of the models.

The trace logs, which also contained failures, used with the enhanced Amdahl's law and Gustafson's law was able to better represent application performance and scalability when failures are present. They were also able to demonstrate models that provided proactive failure prevention by using process mitigation and provided a fast recovery.

It is necessary for IoT systems to have the ability to provide faster recovery and provide proactive failure prevention. By adding additional information within our study, we have the ability to use the extended models provided by Zheng and Lan to analyze our ability to prevent and recover as our system continues to scale.

2.3.2 Scalability in IoT Systems

In *Internet of Things Scalability: Analyzing the Bottlenecks and Proposing Alternatives*, Gomes, et al., proposed (i) a microbenchmark for to evaluate IoT systems; (ii) a preliminary architecture redesign for future IoT middleware [23]. For the proposed microbenchmark, the measure of scalability was performed by combining data regarding different number of threads, number of requests per thread, access analysis of components, and different load situations. Fosstrack middleware was implemented to analyze performance. The results demonstrated that CPU usage, network traffic, and response time increased as the number of threads and requests increased. The bottlenecks in the experiment included the number of process. The ALE module implemented by the researched could only handle 200 threads simultaneously.

Given their results, changes were proposed to be included in IoT middleware. First, the use of NoSQL database should be included because it can provide a solution for for handling large values of data, fault tolerance, scalability, and high availability. Second, they suggested the use of High Performance Computing (HPC) including parallel processing within the module to handle more simultaneous requests. Finally, provide templates for virtual machines for capturing data and querying interfaces. Templates can be used to to instantiate new prepared virtual machines.

In *A Scalable Distributed Architecture Towards Unifying IoT Applications*, Sarkar, et al., proposed a scalable distributed architecture to address issues of scalability, interoperability, and heterogeneity within an IoT application [10]. The approach divides IoT architecture into a virtual object layer, composite virtual object layer, and a service later. The first layer, virtual object layer, is responsible for virtualizing the physical objects. By virtualizing the objects, the inherit difficulties provided by heterogeneity. The second layer, composite virtual object layer, groups like objects and to perform a unified task. The final layer, the service layer, is respondsible for creating and managing services.

We believe Gomes, et al., proposal for HPC use in the middleware will provide the increased scalability and Sarkar, et al., provides an overview of a layered architecture to address issues of scalability and heterogeneity. We intend to combine HPC within the gateways to perform more complex data calculations, parallel machine learning, and adding the ability to increase the number of edge devices each gateway can handle. We believe blending these techniques will provide a scalable IoT architecture with the ability to scale from a few devices to thousands of devices.

2.4 Resiliency in IoT Systems

Resilience is the ability for an IoT system to resist and recover from a malfunction. In *On Resilience of IoT Systems: The Internet of Things*, Delic defines resilience, resilience in IoT systems, and methods for research in resilience in the future [17]. Delic identifies resiliency in a system as the capability to resist external perturbances and internal failures, recover and enter a stable state, and adapt structures and behaviors to constant change. The article proposes that resiliency in IoT systems should be studied in at least three large domains including infrastructure, nomadic users, and digital economy [17].

In *Enabling Resilience in the Internet of Things*, Benson proposes a middleware framework for resilient and stable communications between IoT devices and services [57]. Previous research conducted by Benson investigated creation of testbeds including using accelerometers

to detect earthquakes on the University of California, Irvine Campus. While creating testbeds there were several lessons learned. The use of MQTT lacked range and made connecting to specific sensor readings difficult. Off the shelf equipment caused inaccurate readings and multiple device failures. The proposed approach is to investigate resilient communication and establish location aware overlay network, develop a resilient data exchange middleware and resilient application deployment.

This dissertation will have a focus on resiliency and stability within IoT infrastructure. There will be focus on identifying methods of resisting and recovering from internal and external failures and adapting to constant changes within the structure and behaviors in the system. Benson's approach is to create a middleware framework to address each issue [57]. We propose the use of machine learning to be included within the gateway layer. Machine learning will allow for adapting to changing system structure and devices. This approach would include the use of incorporating an ARM multicore plus GP-GPU capability to enhance performance for data-parallel computations.

2.5 Summary

Extensive research is occurring in IoT/CPS systems. As these systems scale up, scaling occurs within the gateway to increase the amount of computations that occuring. Scaling also occurs within the number of edge devices that the system can maintain and process information from. There is the need to achieve low latency during data transfer and maintain privacy of data.

There is a need throughout these systems for security. Providing security within a device and network includes the use of lightweight certificates, encryption techniques, and machine learning. Lightweight certificates are used to ensure trust between the edge devices and gateways. Encryption techniques provide the ability to ensure data privacy is ensured as data is transferred from edge devices and gateways. Along with providing privacy of the data, neural networks can be used to probabilistically predict the next state of an IoT application.

Using Neural Networks for security can be expanded to include the ability to detect anomalies and provide active monitoring.

Along with scalability and security, resiliency is required to resist and recover from a malfunctions. If a malfunction does occur, it is necessary in IoT/CPS systems to be able to return to a stable state. There are proposals for this to occur within middleware and application deployment. We propose that the additional of multi-tiered architecture will allow for a reduction of single failure points and allow to recover more quickly if an error does occur.

Chapter 3

Research Questions

This dissertation focuses on developing a secure, scalable, stable, resilient IoT Architecture by investigating single-tier or multi-tier gateway architectures, determining secure device handshakes between devices, and using machine learning to detect anomalies and determine who or what caused the anomaly. The following research questions drive the activities proposed:

1. Can a multi-tier gateway provide the framework for a more resilient and scalable architecture?

   We will investigate using a multi-tier gateway versus a single tier gateway to analyze the difference in resiliency and scalability. A single-tier gateway architecture means each edge device is connected to a single gateway level, then the gateway connects to the cloud. A multi-tiered gateway architecture means edge devices would be connected to one or more gateways for redundancy and each gateway will connect to a higher tier gateway, based on the number of tiers within the design. This means if gateway failures occur, a second tier gateway exists to which the device can connect. In theory, this would allow for scaling and resiliency.

2. What are the optimal number of gateways and tiers of gateways based on the number of end devices?

   As we investigate the use of a multi-tier system, we will also research the benefits or disadvantages of the number of tiers within the IoT architecture. For instance, a four-tier architecture might provide a more resilient system, however, the cost of the system might make it undesirable for industry applications. We will analyze the

trade-offs in cost, performance, and resilience to define a means of optimizing number of tiers and gateways based on the number of edge devices.

3. Can certificates be used to provide a lightweight handshake between edge devices and gateways?

   The method in which edge devices and gateways authenticate should be lightweight, stable, and secure. Certificates can provide a method for handshakes between the devices. We intend to investigate the use of certificates as an option for a lightweight handshake, and determine its relevance in an IoT system.

4. Can performance be maintained throughout the system as the number of devices increase?

   System performance should be maintained or minimally impacted as more devices are added. Network communication and packet transfer rate is one of the important aspects to be investigated as more packets flood the gateways. Within the gateway we intend to investigate resource usage including processing speed and network traffic as more edge devices are added to the system. For each edge device, we will examine the number of packets sent versus received by gateway as more devices are added.

5. Can a gateway or group of gateways be used to make decisions regarding where edges devices should connect to create a more resilient system?

   We will investigate the capabilities of a gateway or group of gateways to determine how edge devices traffic should be routed by making informed decisions based on device location, data frequency, and data redundancy. As the system scales, we intend to provide analysis as on how a coalition gateways can be used to make decisions to provide more reliable data communication and route traffic in an efficient manner throughout the system.

6. Can the use of parallel processing in each gateway allow for data aggregation and device checking to be performed more quickly for a more secure system?

   In previous work, we demonstrated that with GP-GPU computing within the gateway we were able to compute a simple regression model up to ten times more quickly than within the 4-core arm processor. Given our results, we intend to investigate further the use of GP-GPU programming within the gateway to perform more complicated computations in an effort to provide a secure system.

7. Can the use of machine learning be used to monitor the health and performance of a system in an effort to detect anomalies or failures in the system?

   Machine learning is an increasingly important field within IoT. Currently, most machine learning within an IoT system is to perform an action, not monitor the health of the system. We propose using machine learning within the gateways to monitor the health of the system by learning the system. This would allow the gateways to detect anomalies within the system and alert the user.

8. How do the machine learning and data aggregation scale as a system scales?

   As we continue to add more devices to the system, we will analyze how scaling effects both the machine learning and data aggregation within the gateway. We monitor the system performance in an effort to determine advantages and disadvantages to adding devices.

9. What data should be collected at anomaly or failure detection for forensics to occur?

   Upon detecting an anomaly, reporting appropriate information back to the user to investigate what has occurred is essential. This will allow for the user to determine what type of error has occurred, whether someone caused it, and how it can be fixed. As we build anomaly detection into the system, we also intend to determine what appropriate

information we should provide to the user to help enable quick response to fixing the error.

As we analyze and answer these questions, we maintain the goal of developing an architecture for IoT/CPS systems that is secure, scalable, and resilient.

Chapter 4

Approach and Metrics

When identifying the approach and metrics, we divided our research into three primary categories: architecture, device handshake, and system monitoring and security. Architecture encompasses the physical hardware architecture, how devices are connected together, how edge device traffic is routed and structure. Device handshake analyzes ways of creating trust between edge devices and gateways. System monitoring and security focuses on security between devices and gateways, as well as, active monitoring devices and gateways in an effort to maintain a healthy system.

4.1 Approach

The approach overview describes how we will investigate each area of focus we defined.

4.1.1 Architecture

Internet of Things architecture can be divided into three primary layers: application layer, transport layer, and sensing layer. The sensing layer is the bottom layer. In the sensing layer, edge devices including sensors and actuators, will report their environment to the transport layer or perform specified task provided by the transport layer.

The middle layer is the transport layer. The transport layer includes gateways and networking protocols that are responsible for transferring data between the sensing layer and the application layer. Current architectures include a single tier of gateway devices within the transport layer, which is displayed in figure 4.1. The application layer, the top layer, includes cloud services where data can be stored, user applications, and data mining services.

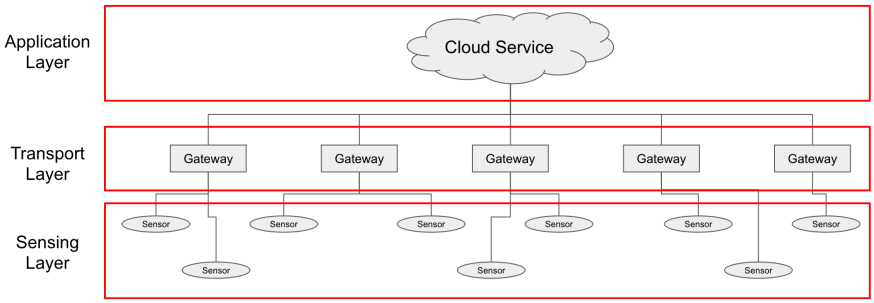

Figure 4.1: Single-tiered Architecture

We propose the expansion of a fog architecture to CPS by adding decision-making within the transport layer and multiple tiers of gateways. The decision layer would contain machine learning and other predictive techniques to determine the health of the subsystem within the sensing and transport layers. Multiple tiers of gateways would provide redundancy within the transport layer to help reduce single failure points. Figure 4.2 displays the multi-tier gateway approach. As the system scales up, a coalition of gateways will be used to determine the most effective manner of routing edge device traffic. We will investigate systems of size 10, 100, 1000, and 10000 to understand how systems change as the number of devices increase. We will investigate the optimal number of gateways as the system scales, as well as, the optimal number of tiers for these gateways.

We will begin by creating single-tiered and multi-tiered CPS and monitor normal health by collecting sensor data, transmission speed, number of transmissions, and number of active sensors in the system. After the baseline data has been created, single failure points, intrusion, and other malfunctions/attacks will be introduced to the system to test effectiveness of both architectures. For each test, we will calculate the number of active sensors, the time to return to a functioning state, and the resources used to return to a functional state.

4.1.2 Routing Protocol

For routing, we will begin by investigating mobile ad-hoc network (MANET) routing protocols including Optimized Link State Routing Protocol (OLSR) , Destination Sequenced Distance

Vector Routing Protocl (DSDV) , and Ad-hoc On Demand Distance Vector Protocol (AODV) . OLSR and DSDV are proactive routing protocols that maintain a routing table that is periodically updated through control message [58, 59]. AODV is a reactive routing protocol that uses Bellman-Ford Distance Vector Routing Algorithm to determine the shortest connection between the devices [58]. These protocols provide a method for routing traffic, however, they do not take into account the current load on the gateway or secondary connections for the edge devices. We will investigate adapting these algorithms for our routing protocol.

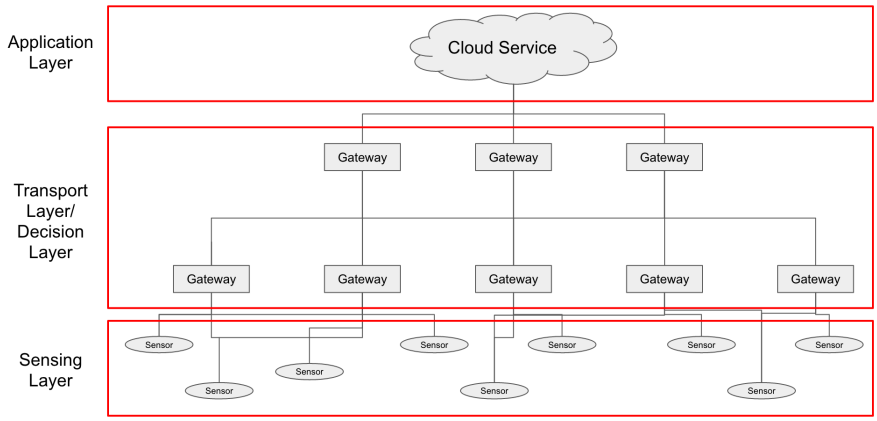

Figure 4.2: Multi-tiered Architecture

4.1.3 Device Handshake

Devices within an IoT architecture need a secure, quick, and efficient way to connect with each other. To provide this level of authentication, we will test multiple methods of device handshakes in an effort to determine the effectiveness of each handshake. We will investigate the use of certificate authentication between end devices and gateways and multiple encryption techniques and keys to authenticate the devices. To begin, we will analyze AWS IoT X.509 Certificates and investigate creating our own certificates [60, 61].

For effectiveness, we will measure resources necessary to store and authenticate the handshake on each device and time it takes to authenticate. We will also investigate the security strengths and weaknesses of each method.

4.1.4 System Monitoring and Security

System monitoring and security will be beneficial to monitoring the health and stability within the system. We will begin by collecting system process information including *task_struct*, number of running processes, data transfer information including packet size, number of packets per sender, time between packets, etc. We will perform Exploratory Data Analysis to determine relevant data points in each area. From these results, we can create Neural Networks at system level and gateway level that can be used to monitor the health of the system. To test our monitoring, we will cause faults within the system and monitor a) time to detect fault, b) time to learn new system for future monitoring. Once we complete these tests, we will investigate the necessary information to determine who or what caused the fault.

4.1.5 Machine Learning

For machine learning, we will begin by creating ANNs from device information, sensor information, time data, and other relevant data points.

We will train these networks experimenting with the number of layers. Each hidden layer provides a different measure of transformation to the model. We will investigate various numbers of hidden layers in an effort to create a trained system capable of detecting anomalies.

For computing our neural networks, will begin by using R, a statistical computing software. The *neuralnet* package for R allows us to create a neural network to use for predictions [62]. We will process our preliminary networks sequentially. Once we have completed preliminary analysis, we will expand to parallel computing of the networks using the CPU and then GPU. For deep learning in R with the CPU and GPU, we will investigate the use of MXNET. MXNET is a a flexible library for deep learning [63]. The library can be used within R, C++, or python.

4.1.6 Summary

We will divided our work into three sections: architecture, device handshake, and system monitoring and security. As we investigate the architecture we will analyze how multi-tiered architecture scales with systems of 10, 100, 1000, and 10000 devices. We will examine the most efficent method of providing trust between the edge devices and gateway devices. Finally, we will implement active monitoring using machine learning to detect anomalies or intrusions within the IoT/CPS system.

4.2 Metrics

To determine the level of success within this study, metrics need to be defined. In the following sections, we define metrics for how to analyze system architecture changes, device handshakes, and system monitoring and security.

4.2.1 Architecture

To validate architecture performance, we will monitor packet transmission rates, number of devices online (percentage of fully functioning system) and time it requires to return to a functioning state after a failure occurs. Data will be gathered from the system while in normal use to set a baseline of performance. When a baseline performance and a normal threshold have been established, then faults will be added to the system and the same data will be gathered. The data from each test will then be analyzed for similarities and differences. The hypothesis is that a multi-tiered gateway architecture will provide a more resilient and scalable system in which recovering from single failure points is done more quickly.

4.2.2 Routing Protocol

To analyze performance of the routing protocols, we will analyze number of packets transmitted, number of packets dropped, delay within the wireless network, and memory usage. We

will investigate the changes in bandwidth, latency and throughput to see trade-offs with each routing protocol.

4.2.3 Device Handshake

For device to device authentication, we will measure the amount of resources necessary for authentication including storage space, memory, and time to authenticate. We will also analyze the security of the handshake.

4.2.4 System Monitoring and Security

For device security, we will measure the amount of time it takes to detect an intrusion and the amount of time that is required to recover. Baseline measurements will include the number of packets and the packet transmission rate, number of correctly functioning devices, the amount of resource in use in each device, etc. The goal is to demonstrate a significant improvement over current detection and recovery time, and minimize system interruption due to a malfunction.

Upon detection of an each fault, we use collected data to determine who or what caused the fault. We will explore the time, resources, and quantity of data needed to determine who or what cased the failure.

4.2.5 Machine Learning

To measure the success of our neural networks, we will monitor the amount of time to compute each network. We will compare the time, as well as success in prediction, to determine the whether creating the network sequentially, in parallel over the CPU, or parallel of the GPU.

We also will measure the number of false positives, false negatives, incorrect predictions and correct predictions to determine the success of the model. A false positive is defined a determining an anomaly is not an anomaly. A false negative is defined as a correct process

being flagged as an anomaly. Our intention is to minimize the number of false positives and false negatives within our system.

4.2.6 Summary

As we test our systems, we will measure system performance including latency, resources, and successful packet transfers. To examine our active monitoring, we will measure the number of positive detections, as will as, analyze the false positive and false negatives within the system. The goal is to determine the trade-offs between performance, security, and resiliency as the system scales up in size.

Bibliography

[1] Gartner Research. Gartner says 6.4 billion connected "things" will be in use in 2016, up 30 percent from 2015, 2015. http://www.gartner.com/newsroom/id/3165317.

[2] Juniper Research. internet of things connected devices to almost triple to over 38 billion units by 2020, 2015. http://www.juniperresearch.com/press/press-releases/iot-connected-devices-to-triple-to-38-bn-by-2020.

[3] Statista. Internet of things (iot): number of connected devices worldwide from 2012 to 2020 (in billions), 2016. http://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/.

[4] Ideas 2020. Questions and answers. http://www.ideen2020.de/en/2993/.

[5] Elseiver. Special issue on cyber-physical systems (cps), internet of things (iot) and big data. http://www.journals.elsevier.com/future-generation-computer-systems/call-for-papers/special-issue-on-cyber-physical-systems-cps-internet-of-thin.

[6] Kate Carruthers. Internet of things and beyond: Cyber-physical systems. http://iot.ieee.org/newsletter/may-2016/internet-of-things-and-beyond-cyber-physical-systems.

[7] Jianhua Shi, Jiafu Wan, Hehua Yan, and Hui Suo. A survey of cyber-physical systems. In *Wireless Communications and Signal Processing (WCSP), 2011 International Conference on*, pages 1–6. IEEE, 2011.

[8] Radhakisan Baheti and Helen Gill. Cyber-physical systems. *The impact of control technology*, 12:161–166, 2011.

[9] Ayan Banerjee, Krishna K Venkatasubramanian, Tridib Mukherjee, and Sandeep Kumar S Gupta. Ensuring safety, security, and sustainability of mission-critical cyber–physical systems. *Proceedings of the IEEE*, 100(1):283–299, 2012.

[10] C. Sarkar, S. N. A. U. Nambi, R. V. Prasad, and A. Rahim. A scalable distributed architecture towards unifying iot applications. In *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, pages 508–513, March 2014.

[11] S. F. Abedin, M. G. R. Alam, N. H. Tran, and C. S. Hong. A fog based system model for cooperative iot node pairing using matching theory. In *Network Operations and Management Symposium (APNOMS), 2015 17th Asia-Pacific*, pages 309–314, Aug 2015.

[12] Höller Jan. *From machine-to-machine to the Internet of things: introduction to a new age of intelligence*. Elsevier Academic Press, 2014. pg. 197.

[13] SANS Institute. Network security resources, 2015. https://www.sans.org/network-security/.

[14] sharcnet. Measuring parallel scaling performance, 2016. https://www.sharcnet.ca/help/index.php/Measuring_Parallel_Scaling_Performance.

[15] Geprge Karypis Vipin Kumar Anatha Grama, Anshul Gupta. *Introduction to Parallel Computing, Second Edition*. Addison-Wesley, 2003. pg. 211.

[16] James Reinder Michael McCool, Arch Robison. Amdahl's law vs. gustafson-barsis' law, 2013. http://www.drdobbs.com/parallel/amdahls-law-vs-gustafson-barsis-law/240162980.

[17] Kemal A. Delic. On resilience of iot systems: The internet of things (ubiquity symposium). *Ubiquity*, 2016(February):1:1–1:7, February 2016.

[18] Kevin Ashton. That 'internet of things' thing, 2009. http://www.rfidjournal.com/articles/view?4986.

[19] Kevin Maney. Meet kevin ashton, father of the internet of things, 2015. http://www.newsweek.com/2015/03/06/meet-kevin-ashton-father-internet-things-308763.html.

[20] Arik Gabbai. Kevin ashton describes "the internet of things", 2015. http://www.smithsonianmag.com/innovation/kevin-ashton-describes-the-internet-of-things-180953749/?no-ist.

[21] Jacob Morgan. A simple explanation of 'the internet of things', 2014. http://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#1d620e4a6828.

[22] IoT Analytics. Why the internet of things is called internet of things: Definition, history, disambiguation. https://iot-analytics.com/internet-of-things-definition/.

[23] M. Gomes, R. da Rosa Righi, and C. A. da Costa. Internet of things scalability: Analyzing the bottlenecks and proposing alternatives. In *2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pages 269–276, Oct 2014.

[24] Walid Negm Paul Daugherty, Prith Banerjee and Allan E. Alter. Driving unconventional growth through the industrial internet of things, 2015. https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Driving-Unconventional-Growth-through-IIoT.pdf.

[25] Markets and Markets. Smart lighting market worth 8.14 billion usd by 2020, 2015. http://www.marketsandmarkets.com/PressReleases/smart-lighting.asp.

[26] Tom Pincince. Part 2. what is smart about smart lighting?, 2015. https://www.digitallumens.com/resources/blog-post/part-2-smart-smart-lighting/.

[27] Evan Weremeychik. How to design a smart hospital, 2014. http://www.healthcaredesignmagazine.com/article/how-design-smart-hospital.

[28] Cássio V. S. Prazeres and Martin Serrano. Soft-iot: Self-organizing FOG of things. In Leonard Barolli, Makoto Takizawa, Tomoya Enokido, Antonio J. Jara, and Yann Bocchi, editors, *30th International Conference on Advanced Information Networking and Applications Workshops, AINA 2016 Workshops, Crans-Montana, Switzerland, March 23-25, 2016*, pages 803–808. IEEE Computer Society, 2016.

[29] Michael Negnevitsky. *Artificial Intelligence: A Guide to Intelligent Systems*. Pearson, 2011.

[30] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (3rd Edition)*. Pearson, 2009.

[31] Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernockỳ, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Interspeech*, volume 2, page 3, 2010.

[32] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[33] Standford. Applications for neural networks, 2016. https://cs.stanford.edu/people/eroberts/courses/soco/projects/2000-01/neural-networks/Applications/index.html.

[34] I. Kotenko, I. Saenko, F. Skorik, and S. Bushuev. Neural network approach to forecast the state of the internet of things elements. In *Soft Computing and Measurements (SCM), 2015 XVIII International Conference on*, pages 133–135, May 2015.

[35] Bryan D Payne Martim Carbone Monirul and Sharif Wenke Lee. Lares: An architecture for secure active monitoring using virtualization.

[36] L. Jiang, G. Tian, and S. Zhu. Design and implementation of network forensic system based on intrusion detection analysis. In *Control Engineering and Communication Technology (ICCECT), 2012 International Conference on*, pages 689–692, Dec 2012.

[37] Mahendra Ramsinghani. How the 'insecurity of things' creates the next wave of security opportunities. https://techcrunch.com/2016/06/26/how-the-insecurity-of-things-creates-the-next-wave-of-security-opportunities/.

[38] Hewlett Packard Enterprise. Internet of things research study, 2015 report, 2015. http://www8.hp.com/h20195/V2/GetPDF.aspx/4AA5-4759ENW.pdf.

[39] Edewede Oriwoh, David Jazani, Gregory Epiphaniou, and Paul Sant. Internet of things forensics: Challenges and approaches. In Elisa Bertino, Dimitrios Georgakopoulos, Mudhakar Srivatsa, Surya Nepal, and Alessandro Vinciarelli, editors, *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, Austin, TX, USA, October 20-23, 2013*, pages 608–615. ICST / IEEE, 2013.

[40] GE Healthcare. Smart technologies, 2016. `http://www3.gehealthcare.com/en/products/categories/computed_tomography/smart_technologies`.

[41] IBM Research Zurich. Smartgrid, 2016. `http://www.research.ibm.com/labs/zurich/smartgrid/`.

[42] S. Sridhar, A. Hahn, and M. Govindarasu. Cyber 2013;physical system security for the electric power grid. *Proceedings of the IEEE*, 100(1):210–224, Jan 2012.

[43] U.S. Department of Energy. A systems view of the modern grid, 2007. National Energy Technology Laboratory.

[44] G. Karsai, D. Balasubramanian, A. Dubey, and W. R. Otte. Distributed and managed: Research challenges and opportunities of the next generation cyber-physical systems. In *2014 IEEE 17th International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing*, pages 1–8, June 2014.

[45] K. Nahrstedt, H. Li, P. Nguyen, S. Chang, and L. Vu. Internet of mobile things: Mobility-driven challenges, designs and implementations. In *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 25–36, April 2016.

[46] Alvaro Cardenas, Saurabh Amin, Bruno Sinopoli, Annarita Giani, Adrian Perrig, and Shankar Sastry. Challenges for securing cyber physical systems. In *Workshop on future directions in cyber-physical systems security*, page 5, 2009.

[47] K. Rieck. Computer security and machine learning: Worst enemies or best friends? In *SysSec Workshop (SysSec), 2011 First*, pages 107–110, July 2011.

[48] X. Xingmei, Z. Jing, and W. He. Research on the basic characteristics, the key technologies, the network architecture and security problems of the internet of things. In *Computer Science and Network Technology (ICCSNT), 2013 3rd International Conference on*, pages 825–828, Oct 2013.

[49] M. Grabovica, S. Popi, D. Pezer, and V. Kneevi. Provided security measures of enabling technologies in internet of things (iot): A survey. In *2016 Zooming Innovation in Consumer Electronics International Conference (ZINC)*, pages 28–31, June 2016.

[50] Bluetooth. Story behind bluetooth technology, 2016. `https://www.bluetooth.com/what-is-bluetooth-technology/bluetooth`.

[51] C. Arseni, M. Mioi, and A. Vulpe. Pass-iot: A platform for studying security, privacy and trust in iot. In *2016 International Conference on Communications (COMM)*, pages 261–266, June 2016.

[52] J. Dofe, J. Frey, and Q. Yu. Hardware security assurance in emerging iot applications. In *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2050–2053, May 2016.

[53] Shams Zawoad and Ragib Hasan. Faiot: Towards building a forensics aware eco system for the internet of things. In *2015 IEEE International Conference on Services Computing, SCC 2015, New York City, NY, USA, June 27 - July 2, 2015*, pages 279–284. IEEE Computer Society, 2015.

[54] Edewede Oriwoh and Paul Sant. The forensics edge management system: A concept and design. In *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing, UIC/ATC 2013, Vietri sul Mare, Sorrento Peninsula, Italy, December 18-21, 2013*, pages 544–550. IEEE Computer Society, 2013.

[55] Z. Zheng and Z. Lan. Reliability-aware scalability models for high performance computing. In *2009 IEEE International Conference on Cluster Computing and Workshops*, pages 1–9, Aug 2009.

[56] Kwon. Multi-processor laws, 2013. `http://www.d.umn.edu/~tkwon/course/5315/HW/MultiprocessorLaws.pdf`.

[57] K. Benson. Enabling resilience in the internet of things. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*, pages 230–232, March 2015.

[58] N. Kumari, S. K. Gupta, R. Choudhary, and S. L. Agrwal. New performance analysis of aodv, dsdv and olsr routing protocol for manet. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 33–35, March 2016.

[59] S. Kukreja and P. Singh. Performance metrics of aodv and olsr in wireless mesh network. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 3182–3185, March 2016.

[60] Amazon Web Services. Authentication in aws iot. `http://docs.aws.amazon.com/iot/latest/developerguide/identity-in-iot.html`.

[61] Eustace Asanghanwa. How to create root and other certificates for iot devices, 2016. `http://www.atmel.com/images/How-to-Create-Root-and-Other-Certificates-for-IoT_Article.pdf`.

[62] Michy Alice. Fitting a neural network in r; neuralnet package, 2015. `https://www.r-bloggers.com/fitting-a-neural-network-in-r-neuralnet-package/`.

[63] MXNET. Mxnet, 2016. `http://mxnet.readthedocs.io/en/latest/`.

Appendices

Appendices

Appendix A
Publication Plans

- Adding Scalability to Internet of Things Gateways using Parallel Computation of Edge Device Data - 3rd Quarter 2016 - Accepted by: High Performance Extreme Computing Conference - Peer Reviewed IEEE Conference

  Abstract: The Internet of Things (IoT) is a rapidly growing area with an estimated 25 billion connected devices anticipated by the year 2020. As more devices join the IoT landscape, the ability to scale from small to large deployments is becoming paramount. In this paper, we investigate the ability to scale an IoT system above the leaf-level by using parallel computing within the gateway devices. The initial task identified for gateway parallel computing is to aggregate and analyze data from end devices. This approach provides a scalable architecture for IoT Systems. Devices such as the Jetson TX1 and TK1 incorporate an ARM multicore plus GP-GPU capability to enhance performance for data-parallel computations. We demonstrate the value of this type of hybrid architecture on an example IoT system test bed with non-trivial speedup. This points clearly to the need for gateway devices to move to highly parallel architectures, rather than simply serving as small servers with multiple network adapters.

- Using Machine Learning to Secure IoT Systems - 3rd Quarter 2016 - Submitted to: Privacy, Security and Trust 2016 - Peer Reviewed Conference

  Abstract: The Internet of Things (IoT) is a massive group of devices containing sensors or actuators connected together over wired or wireless networks. With an estimate of over 25 billion devices connected together by 2020, IoT has been rapidly growing over the past decade. During the growth, security has been identified as one of the weakest areas in IoT. When implementing security within an IoT network, there are several challenges including heterogeneity within the system as well as the quantity of devices that need to be addressed.

  To approach the challenges in securing IoT devices, we propose using machine learning within an IoT gateway to help secure the system. We investigate using Artificial Neural Networks in a gateway to detect anomalies in the data sent from the edge devices. We are convinced that this approach can improve the security of IoT systems.

- Using Machine Learning to Monitor System Health in IoT 4th Quarter 2016 - Conference TBD

- Multi-Tiered Architecture for Internet of Things 1st Quarter 2017 - Conference TBD

- Secure and Lightweight Device-to-Gateway Authentication 2nd Quarter 2017 - Journal

- Using Machine Learning for Anti-Forensics in IoT Systems 3rd Quarter 2017 - Journal

Appendix B
Research Questions - High Level/Scoping

The following high level research questions provide the scope of study for this dissertation.

B.1 What problem are you trying to solve?

- Make an IoT architecture with tunable QoS, security and resiliency and the ability to scale with economic tradeoffs.

B.2 What will you know when you're done?

- The systematic tradeoffs of QoS, security, resiliency, and scalability with economic underpinnings.
- Identifying the achievable performance of this revised architecture.
- Understand and compare old limitations with new limitations (which should be less in several dimensions)
- Explain when "one size fits all" and where specific IoT architectures are better
- Explain if IoT architectures per service are going to be deployed, and what happens when we try to share an architecture for several "verticals" (e.g., lighting, other).
- How scalability, resiliency, and security interact and where key architectural decisions have to be made, consistent with the requirement of economic feasibility for a given design.

B.3 How will you know when you're done?

- An architecture, with prototype proofs of concept and ability to describe tradeoffs will be achieved.

B.4 How original is this?

- IoT security, scalability and resiliency are some of the most talked about areas of IoT
- Adapting HPC concepts to the system, as well as, multi-tiered gateways, is a new and innovative way of approaching this problem
- This architecture will be useful for smart factories, cities, homes, etc.
- This architecture could provide a foundation for IoT systems in a wide range of areas including smart grids, hospitals, buildings, etc.

B.5 Who will care? What's the impact?

- IoT security, scalability and resiliency are some of the most talked about areas of IoT
- Adapting HPC concepts to the system, as well as, multi-tiered gateways, is a new and innovative way of approaching this problem
- This architecture will be useful for smart factories, cities, homes, etc.
- This architecture could provide a foundation for IoT systems in a wide range of areas including smart grids, hospitals, buildings, etc.

B.6 Why did you select this problem among many?

- A strong, secure foundation is necessary for a strong infrastructure. The IoT infrastructure is new and is in need of strong security measures in an effort to create a stable and resilient infrastructure that can be implemented in factories, cities, homes, etc.

B.7 Why will this be a significant contribution to the literature in your area?

- Provide an architecture for security and resiliency within an IoT system which currently has limited documentation.
- Produce literature on machine learning in IoT systems to provide real-time system monitoring
- Produce literature on best practice for secure device to device authentication