CYBER-PHYSICAL SYSTEMS FOR CRITICAL INFRASTRUCTURE PROTECTION:

A Wireless Sensor Network Application for Electric Grid Monitoring

by

MARTIN SAINT

B.S. University of Redlands, 1985

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Master of Science

Interdisciplinary Telecommunications Program

2013

This thesis entitled:
Cyber Physical Systems for Critical Infrastructure Protection:
A Wireless Sensor Network Application for Electric Grid Monitoring

written by Martin Saint
has been approved for the Interdisciplinary Telecommunications Program

Sharon K. Black

Timothy X Brown

Frank S. Barnes

Date_____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Abstract

Saint, Martin (M.S., Interdisciplinary Telecommunications Program)

Cyber-physical Systems for Critical Infrastructure Protection: A Wireless Sensor Network Application for Electric Grid Monitoring

Thesis directed by Professor Sharon K. Black

Critical infrastructure includes resources which are essential to the function of society. Despite an increased focus on protecting U.S. critical infrastructure, some sectors including the electric grid are currently more vulnerable than ever. Existing critical infrastructure protection (CIP) regulations and the monitoring and control systems used to achieve them have not met expectations. This indicates that the next generation of grid control should explore new architectures, and that there is a need to prioritize improvement efforts and adapt current technologies.

The general methodology for approaching critical infrastructure protection involves 1) identifying the critical infrastructure, 2) identifying threats, vulnerabilities, and potential losses, and 3) prioritizing measures which mitigate the risks. While we focus primarily on technical measures to improve grid control, any effort which involves a change in priorities or investment practices will also require regulatory, standards, and economic support, and consideration is given to the current impact of regulation.

This thesis suggests that protection of infrastructure is a natural priority for a smarter grid, and that efforts should focus on transmission and distribution due to the potentially large

impact on reliability. We explore the question of whether a cyber-physical system in the form of wireless sensor networks can be used to improve CIP. We examine efforts by others to design a wireless sensor module for monitoring transmission and distribution lines, and note that this work includes little information about the performance of the communications subsystem. Laboratory testing of throughput, reliability, and backhaul for one example communication network are undertaken here. We also examine reliability of the short message service as one alternative for backhauling aggregated sensor data.

We found that maximum data payload throughput in this system is less than the line rate stated in the protocol, that throughput drops by approximately half for hops between each node, and that the network is robust in the face of node failure. We find that SMS is a reliable method of transmitting small amounts of sensor data. These performance metrics can be used to refine the design of the powerline sensor, and in designing the architecture of the sensor and communications networks.

Acknowledgements

This work was supported by U.S. Department of Energy grant DE-OE0000436. The chapter on short message service testing was partially funded by Intrado, Inc.

Contents

List of Tables

List of Figures

1 Background and Approach

1.1 Definition

The USA PATRIOT Act of 2001 defines the term "critical infrastructure" as:

> *... systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters* [1].

The Department of Homeland Security (DHS) identified eighteen[1] critical infrastructures and stated that:

> *Protecting and ensuring the resiliency of the critical infrastructure and key resources (CIKR) of the United States is essential to the Nation's security, public health and safety, economic vitality, and way of life. Attacks on CIKR could significantly disrupt the functioning of government and business alike and produce cascading effects far beyond the targeted sector and physical location of the incident* [2].

1.2 Critical Infrastructure Protection

It also set out a process of "critical infrastructure protection" (CIP), often presented as a framework of risk-based assessment and prioritization [2]. It can be generalized into the following steps:

- Identify critical infrastructures
- Identify potential threats to the infrastructures
- Determine infrastructure vulnerability
- Assess the risks and probability of losses
- Identify and prioritize measures which address the risks

[1] The full list includes: Agriculture and Food, Banking and Finance, Chemical, Commercial Facilities, Communications, Critical Manufacturing, Dams, Defense Industrial Base, Emergency Services, Energy, Government Facilities, Healthcare and Public Health, Information Technology, National Monuments and Icons, Nuclear Reactors and Materials and Waste, Postal and Shipping, Transportation Systems, and Water.

- Implement measures to address the risks

1.3 Purpose of the Paper and the Focus on Energy

Of the 18 critical infrastructures identified by the DHS, all are dependent on the electric power grid. Thus, the purpose of this thesis is to focus on the current level of protection of the grid in the United States, and methods to improve on that protection. A more detailed description of the structure and function of the electrical grid is provided in Chapter 3. The electric grid is usually discussed in the context of four broad functional areas: generation, transmission, distribution, and load or consumption. I focus on the electrical transmission and distribution networks as distribution is among the least automated elements and some 80% of outages occur at this level [3]. While transmission has seen more automation than the distribution system, the monitoring and control elements discussed here are readily adaptable to the transmission network, and could be used to protect it against events which have a lower frequency but a higher consequence.

Many efforts are underway to make the electrical grid smarter, particularly via the incorporation of more advanced communication and control networks. In this paper I suggest that using these networks to improve critical infrastructure protection is a worthwhile priority. Fan et al. [4] defines the Smart Grid as "an *intelligent* electricity network that integrates the actions of all users connected to it and makes use of advanced information, control, and communication technologies to save energy, reduce cost and increase reliability and transparency."

The need to improve communication and control methods is an essential element of a smarter and more reliable electrical grid. Traditional industrial control systems and equipment

are centralized and expensive, and have failed to achieve the desired level of grid protection, suggesting that the next generation of monitoring and control should explore new architectures. In this paper I propose integrating relatively simple, robust, and economical cyber-physical system components based upon wireless sensor networks with the electrical grid. As will be explored in later chapters, cyber-physical systems are those which emphasize the interaction between physical processes, computation, and communications. Wireless sensor networks are one manifestation of a widely distributed cyber physical system oriented toward monitoring and control. They are low cost, simple to configure, and robust. By providing a pervasive digital skin to the modern communications network, they enable interaction with the physical world on a scale which was previously inconceivable.

Wireless sensor networks[2] (WSNs) are networks composed of self-organizing nodes which consist of sensor(s), a processor, memory, a wireless radio, and a power source. A gateway serves to transfer information from the nodes to the access network, usually the Internet. From here data may be sent for additional processing, storage, analysis, or response. Working in reverse, control messages may be sent through the access network, gateway, and a coordinator node, eventually reaching the end nodes which have actuator capability in addition to sensing. With a combination of sensors and communication networks contributing to a smarter electrical grid, I hope to show via this thesis one method of protecting critical infrastructure by enabling better monitoring and control. The WSN system proposed, with appropriate sensors, could be applied in many other critical infrastructure protection scenarios, such as monitoring a water system or critical industrial process.

[2] These are also sometimes referred to as wireless sensor and actuator (or actor) networks (WSANs). The terms are used interchangeably here.

Despite an increased focus on protecting U.S. critical infrastructure, particularly the electrical grid, as will be shown, it is currently more vulnerable than ever. Fragmented regulation with unintended consequences, aging infrastructure, increasing demand and complexity, more capable and motivated attackers with a greater number of attack surfaces, construction challenges, and inadequate investment have all contributed to an electrical grid environment that is increasingly less, not more, secure.

Current objectives for improving critical infrastructure protection for the electric grid are well delineated, but it can be shown that the current measures have not been effective. Due to the nature of unintended regulatory consequences and sanctions, power system operators are often more focused on compliance with regulation than on solving the underlying issues regulators hoped to address.

1.4 Research Question

The following question was explored during the research for this thesis:

- Can wireless sensor networks be used to improve CIP as part of the next generation of electrical grid controls?

1.5 Significance

Cyber-physical systems and critical infrastructure protection applications are of current interest to the electric industry because vulnerabilities and failures affect national security and the function of civil society, and they exist at the intersection of several significant trends in electric distribution networks of the future. The electrical grid itself is a widely distributed cyber-physical system, so control systems with a similar architecture are a natural fit. As the application of communications technologies provides both opportunities and challenges for

modernization, the increasing need for security demands that the industry find new ways to protect the traditional electric grid and guard against attacks through the new vectors introduced by incorporating modern data networks into the power system. It is useful to examine if current critical infrastructure protection regulations and the monitoring and control systems used to achieve them have met the goals of regulations, and if not, where we should focus our improvement efforts, and why.

1.6 Scope and Assumptions

Like the electric grid, wireless sensor networks have a large number of important, but unresolved, research issues, such as optimal routing and power management, security, and quality of service. I chose to examine throughput in multi-hop mesh networks in greater depth as this is an important metric for the WSN implementation proposed. Other issues were not within the scope of my research even though they would have an impact of final system design. I use a wireless sensor network based on IEEE 802.15.4, ZigBee, and 6LoWPAN as an example, but with the intent of demonstrating that the concepts can be generalized to other WSN standards and protocols.

As it is necessary to transmit sensor data from the access network to utility operations centers, a form of backhaul will be necessary. I investigate characteristics of the short message service (SMS) as an example backhaul solution. It is widely available, inexpensive, and appropriate for small bursts of data similar to those that are transmitted by a WSN.

1.7 Organization of the Chapters

Chapter 1 provides an introduction to the topic and the organization of the thesis. Despite efforts to protect critical electric infrastructure, current techniques are not enough. I hypothesize

that monitoring and control can be improved by the use of cyber-physical systems, specifically wireless sensor networks transmitting data over an enhanced smart grid communications network. I raise the question of whether this can be demonstrated, and propose a methodology following the DHS's accepted process of identifying and evaluating critical infrastructure and related risk.

Chapter 2 reviews the literature relevant to the topics, including some perspective on critical infrastructure protection, the rationale for focusing on transmission and distribution monitoring in the electrical grid, and works specific to the sensor and communication systems that can be used to improve electrical grid control.

In Chapter 3 I cover the current structure and function of the electrical grid. The goal is to provide a basic view of how the system functions, and more importantly, to indicate parameters which are essential for stable and reliable operation. With these parameters it is possible to understand why monitoring and control are essential and how and where it is done.

Chapter 4 describes issues related to traditional grid controls, while Chapter 5 provides an outline of wireless sensor actuator networks as a manifestation of one type of cyber physical system that could integrate with grid controls to improve critical infrastructure protection. The chapter includes a description of general wireless sensor network architecture and the function of the WSN hardware and protocols.

Chapter 6 provides a description of some of the previous work completed by other researchers on using WSNs for monitoring the transmission and distribution segments of the electric grid. In Chapter 7 I expand this work with the details of laboratory testing conducted by the author concerning the use of WSNs in the electric grid, with some elements of a specific

WSN communications configuration including maximum throughput, the effect of multi-hop mesh networking on throughput, and robustness.

Sensors and actuators lie at one terminus of the network, but individual or aggregated nodes must be connected to a larger system for monitoring and control. Chapter 8 examines the reliability and capacity of one of the simplest backhaul alternatives using cellular networks and SMS, and the final chapter provides some overall analysis, and recommendations.

In the Appendix, regulatory foundations related to identifying and protecting critical infrastructure are examined particularly the rules, agencies, and standards which apply to governing the electrical grid. A brief mention is made of some of the challenges, and unintended consequences, related to regulation.

1.8 Thesis Statement and Approach

In starting my research I considered that despite regulatory efforts to make the electrical grid secure, it is actually becoming less secure and more vulnerable. I hypothesize that improved system monitoring and control will correct this. I raise the question of whether this can be demonstrated, and suggest the possibility of improving electrical grid critical infrastructure protection through the use of cyber-physical systems, specifically wireless sensor actuator networks for monitoring of transmission and distribution lines. I used the general framework for a risk-based assessment and prioritization given above to approach this question, following an accepted process of identifying and evaluating critical infrastructure and related risk. Limitations on the scope of the inquiry are also given, such as the focus on electrical transmission and distribution and the examination of selected wireless sensor network architectures and communication protocols.

1.9 Methodology

I follow the DHS approach to critical infrastructure protection as a broad theme through the thesis, 1) identifying the critical infrastructure, 2) identifying threats, vulnerabilities, and potential losses, and 3) proposing and prioritizing measures which mitigate the risks. I examined a selection of relevant policies, laws, and regulations that affect how the electrical grid is defined and protected, and how these mandates have failed to accomplish the intended level of protection.

I examined the characteristics of a system for accomplishing this goal, and completed laboratory testing to evaluate performance elements in the communications system include measures of throughput and reliability. This information is useful when considering the capabilities and limitations of the proposed system for communicating control data, along with determining the characteristics of these systems and why would we consider them. With this information we can provide some insight into system performance that might prove useful for design and evaluation purposes.

1.10 Audience

This paper is intended for those interested in the application of telecommunications and data networking to improve the operation of the electrical power grid. It may prove useful for individuals with backgrounds in power engineering, telecommunications, enterprise networking, control systems, and wireless sensor/actuator networks who are working on grid integration. It may also be of interest to those working in the field of homeland security, or with a concern for the impact of regulation on grid infrastructure.

1.11 Problematic Trends

In this section I show trends toward increasing demand, and declining investment, reserve capacity, and reliability. These trends have significant financial and socio-economic costs and leave the grid more vulnerable to disruption.

The increased use of electronics by industry, business, and consumers demands better quality power to prevent damage to sensitive devices and insure they function correctly, yet power quality is declining. In spite of conservation efforts, power demand is increasing:

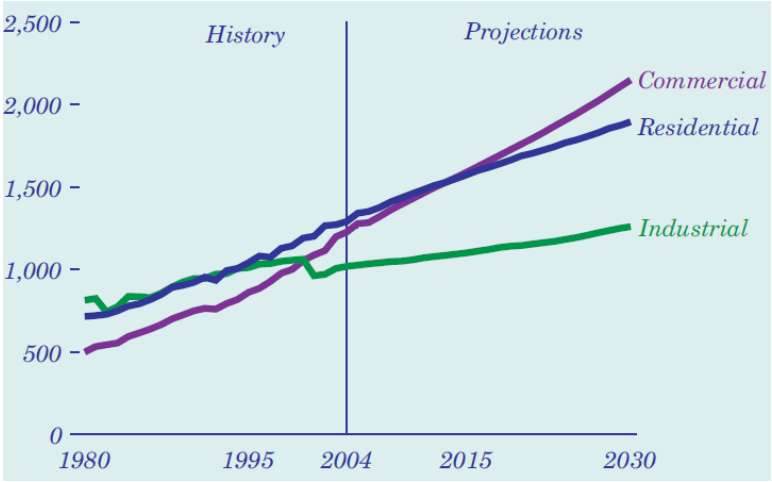


Figure 1-1 Annual electricity sales by sector 1980 to 2030 (billion kWH). [5]

Due to the difficulties in obtaining environmental and construction permits, long lead times, and high capital investment required, new sources of generation are not keeping pace with demand [6]. This is seen in the declining reserve capacity available during periods of peak demand, with the available capacity margin indicating the current excess traditional generating capacity, and the potential capacity margin taking into account alternative sources such as individual building generators:

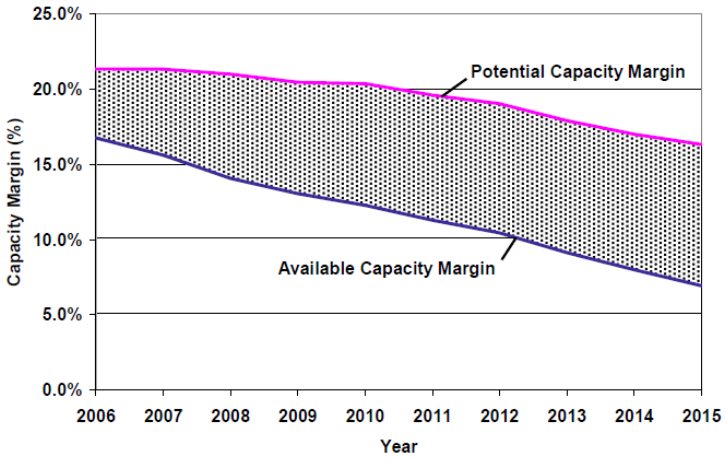


Figure 1-2 U.S. generation capacity margin. [6]

Even with an aged infrastructure and increasing demand, investment is declining:

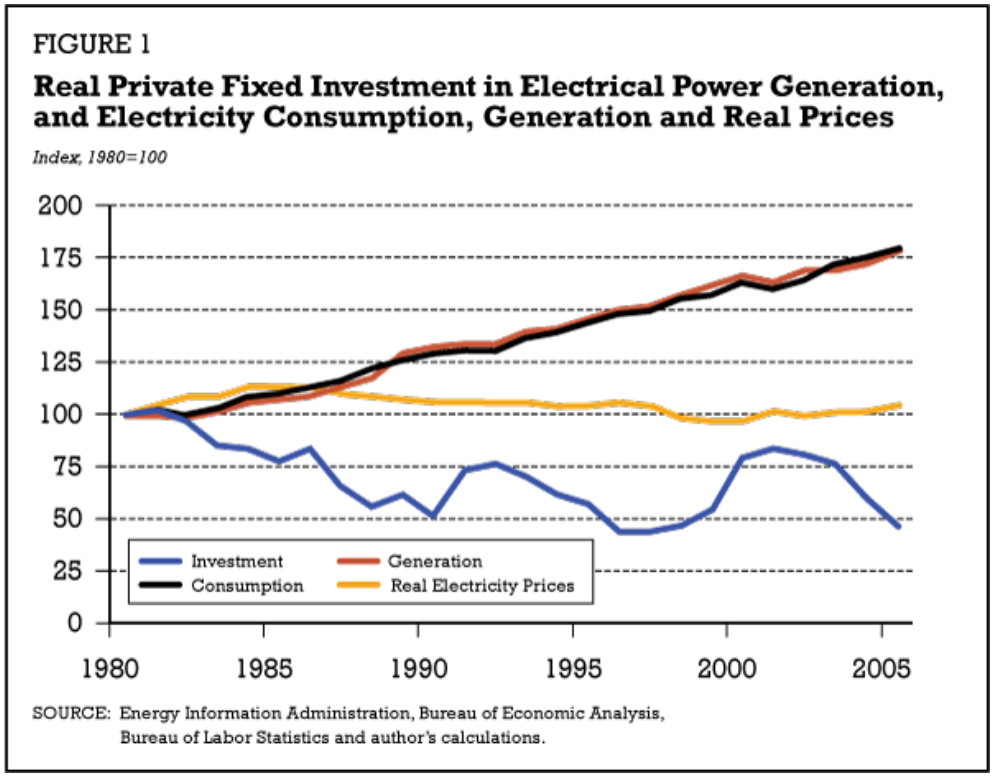


Figure 1-3 Investment in electrical power generation. [7]

Transmission resources are subject to construction constraints similar to those affecting generation, along with the difficulty of obtaining rights-of-way for new transmission lines and the issues of *wheeling* and *congestion* discussed later in the *transmission* section of this paper.

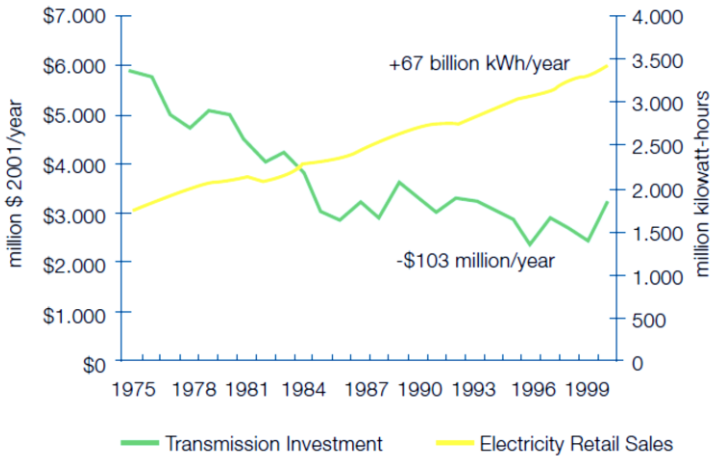


Figure 1-4 Declining transmission investment and increasing demand. [8]

The *IEEE Recommended Practice for Monitoring Electric Power Quality* Standard 1159-1995 [9] is accepted as the standard for measuring transmission and distribution reliability. It provides several indices which quantify the annual frequency and duration of power outages experienced by utility customers. Three of the indexes and their composition are as follows:

System Average Interruption Duration Index (SAIDI) The average of number of minutes of interruptions each year:

$$SAIDI = \frac{Total\ Duration\ of\ Customer\ Interruptions}{Total\ Number\ of\ Customers\ Served} = \frac{\sum r_i N_i}{N_T} \tag{1}$$

System Average Interruption Frequency Index (SAIFI) The average number of annual interruptions experienced by a customer:

$$SAIFI = \frac{Total\ Number\ of\ Customer\ Interruptions}{Total\ Number\ of\ Customers\ Served} = \frac{\sum N_i}{N_T} \quad (2)$$

Customer Average Interruption Duration Index (CAIDI) The average length of an interruption, also the average restoration time for the utility:

$$CAIDI = \frac{Total\ Duration\ of\ Customer\ Interruptions}{Total\ Number\ of\ Customer\ Interruptions} = \frac{\sum r_i N_i}{\sum N_i} = \frac{SAIDI}{SAIFI} \quad (3)$$

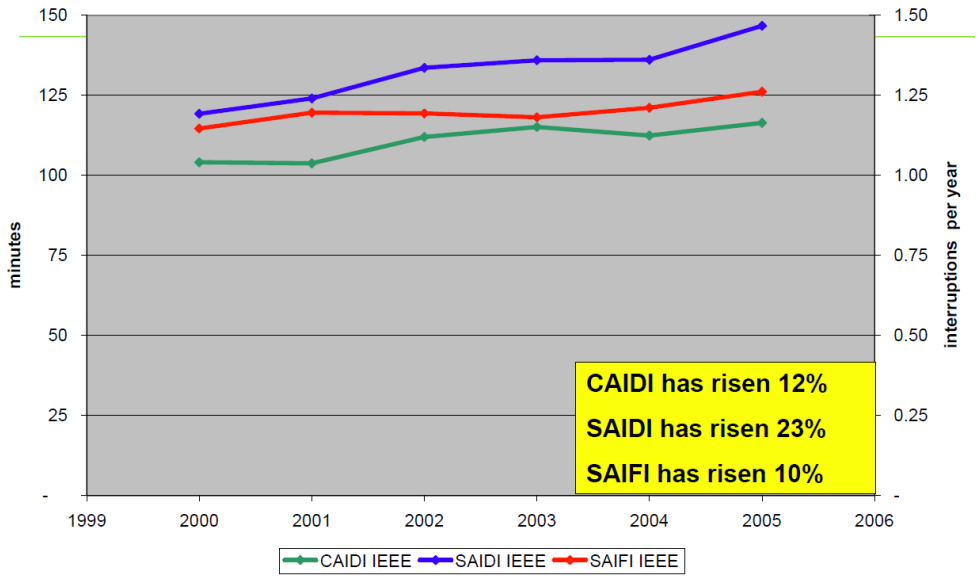


Figure 1-5 Reliability trend using IEEE metrics. [10]

With the grid under increasing pressure and declining investment, outages have increased:

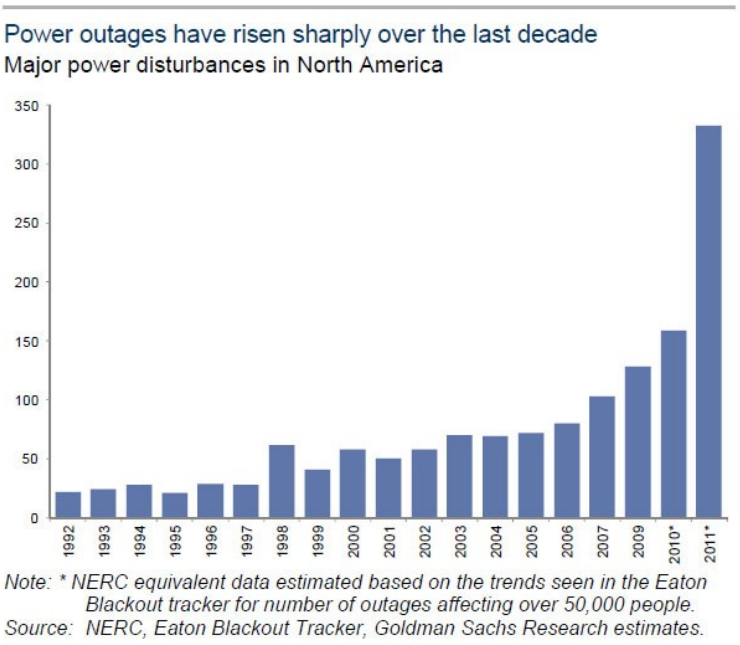


Figure 1-6 Trend of major power disturbances in North America (incidents per year). [11]

Grid disturbances cost the economy an estimated $150 billion annually [12], along with threats to security and health:

Industry	Amount ($)
Cellular Communications	41,000
Telephone Ticket Sales	72,000
Airline Reservation System	90,000
Semiconductor Manufacturer	2,000,000
Credit Card Operation	2,580,000
Brokerage Operation	6,480,000

Figure 1-7 Average cost for one hour of power interruption. [13]

Alternative generation, energy storage, and efforts aimed at reducing demand and shifting peak loads provide the potential for mitigating some of the issues noted. Better monitoring and control of elements of the power grid also have the potential to partially address the impact of the

problems noted by reducing vulnerability and disruption, assisting with recovering more quickly from disturbance, and providing the ability to utilize existing assets more efficiently.

2 Literature Review and Previous Work

Significant work has been done by other/previous researchers to date on the issues of: a) critical infrastructure protection, b) electric utility security, c) problems in the transmission and distribution portions of the electrical grid, and d) options that sensor and communication systems offer to improve electrical grid control. This chapter provides an overview of this previous work.

2.1 Critical Infrastructure Protection

The theme of threats to critical complex and interrelated systems was explored in a 1984 report titled *America's Hidden Vulnerabilities: Crisis Management in a Society of Networks* by Woolsey, Wilcox, and Garrity [14]. Critical infrastructure, the authors stated, was a national security problem. Approaches to deal with the problem converged on the process of 1) determining the systems upon which the well-being of individuals and civil society depended, 2) identifying the vulnerabilities to these systems, and the threats which could exploit the vulnerabilities, and 3) developing methods of mitigating the vulnerabilities in the system. Variations on this theme may be found in [2], Lewis [15], Willis [16], Moteff [17], and Masse [18].

Through the 1990s, the ideas related to this approach continued to be developed, encompassing both the physical and information infrastructure. In *Dits et ecrits 1954-1988,* for example, written in 1994, Foucault [19] refers to *problematization* as something that has "happened to introduce uncertainty, a loss of familiarity; that loss, that uncertainty is the result of difficulties in our previous way of understanding, acting, relating. " The new problematization of security became the need to understand the threats to *system vulnerability*, and the need to find

ways to protect vulnerable systems considered vital to the function of modern society. The 1997 Report by the Presidential Commission on Critical Infrastructure Protection entitled *Critical Foundations* [20] existed even before the terrorist activities of 9/11/2001 served as an inflection point for homeland security activities, and became a catalyst for the current focus on critical infrastructure protection.

Almost a decade later, in a 2008 article entitled *The Vulnerability of Vital Systems: How 'Critical Infrastructure' Became a Security Problem* [21], Collier and Lakoff discuss the origins and implications of critical infrastructure protection. Against the backdrop of two world wars and the perceived threat of Soviet nuclear attack during the cold war, they argue that the prevailing focus was protecting physical infrastructure from foreign threats. With the energy crisis of the 1970s, the technological failures of select facilities including the Three Mile Island and Chernobyl nuclear power plants, and the reality of growing terrorism, security experts began to consider threats which could not be prepared for or responded to with traditional strategic methods. These new threats were considered very difficult to deter, and not enough history or information existed to accurately calculate the probability of occurrence. The impact from these events could not be estimated in the same traditional terms that the probability, damages, and prevention of something like a dropped bomb could be quantified. The work of Collier and Lakoff in particular, while filled with historical lessons and speculation, stops short of providing any insight on the path forward or the likely evolution of the current state of affairs. They also make no attempt to explain the evolution, current status, or catalysts of similar efforts beyond U.S. borders. Collier and Lakoff [21] note that the DHS process of identifying and protecting critical systems is still a relatively new way of approaching national security, and stable organizations and methods for addressing the problems have not yet evolved. As a result, there

exists a variety of bureaucracies, plans, techniques, and resource allocations designed to address the problems, but not necessarily in a coordinated fashion.

2.2 Issues Specific to Electric Utility Security

When discussing energy infrastructure in particular, the Bush Administration's 2003 document entitled *The National Strategy for the Physical Protection of Critical Infrastructures and Key Assets* states:

> *Typically, these companies seek to recover the costs of new security investments through proposed rate or price increases. Under current federal law, however, there is no assurance that electricity industry participants would be allowed to recover the costs of federally mandated security measures through such rate or price increases.* [22].

Gungor and Lambert [23] stated in their 2006 article entitled *A Survey on Communication Networks for Electric System Automation,* that meeting the current and future challenges to reliable power requires integrating a data communications network with a hybrid, decentralized architecture into the power grid. They consider communications between the core network and the substation to be the "last mile," however, and make little mention of specific applications downstream of the substation except for automated metering infrastructure.

Meeus et al. [24] state in their 2010 article entitled "Smart Regulation for Smart Grids," in the *European University Institute Working Papers*, that innovations in grid technology will require addressing specific regulatory mechanisms such as an investment planning process, and social issues such as climate change or the socioeconomic impact of CIP. Since these are issues that affect civil society, the authors argue that they should: 1) be addressed by government, 2) public funding should be contributed to assist in transforming them, 3) regulators should be open to allowing utilities some freedom to experiment, test, and pilot projects which allow the utilities

to gain experience and refine grid technology, and 4) this should be done even if the pilot projects involve the use of public funding or cost recovery from utility customers.

In his 2010 article entitled "Securing the Electricity Grid, " in *The Bridge*, the quarterly publication of the U.S. National Academy of Engineering (NAE), S. Massoud Amin, formerly responsible for research and development on infrastructure security at the Electric Power Research Institute (EPRI), stated that there is no doubt that the existing power delivery system is vulnerable to both natural disasters and intentional attacks [25]. He described the existing electric grid control system as one based on strong central control, with powerful centralized control and computing facilities and dependence on very few communication links. He noted that these systems are most vulnerable during power disruptions and other stresses on the electric system, and this is the very time when they are needed most. Most importantly, Dr. Amin offered that the electric utility network could reconfigure and remain operational in the face of threats and local failures if they had distributed intelligence and control.

In the U.S. Department of Energy's 2009 *Smart Grid System Report* [26] transmission and distribution metrics related to power quality and reliability were the only measures in the report noted as declining, as shown in the following chart.

#	Metric Title	Type	Penetration/ Maturity	Trend
	Area, Regional, and National Coordination Regime			
1	**Dynamic Pricing**: fraction of customers and total load served by RTP, CPP, and TOU tariffs	build	low	moderate
2	**Real-time System Operations Data Sharing**: Total SCADA points shared and fraction of phasor measurement points shared.	build	moderate	moderate
3	**Distributed-Resource Interconnection Policy**: percentage of utilities with standard distributed-resource interconnection policies and commonality of such policies across utilities.	build	moderate	moderate
4	**Policy/Regulatory Progress**: weighted-average percentage of smart grid investment recovered through rates (respondents' input weighted based on total customer share).	build	low	moderate
	Distributed-Energy-Resource Technology			
5	**Load Participation Based on Grid Conditions**: fraction of load served by interruptible tariffs, direct load control, and consumer load control with incentives.	build	low	low
6	**Load Served by Microgrids**: the percentage total grid summer capacity.	build	nascent	low
7	**Grid-Connected Distributed Generation (renewable and non-renewable) and Storage**: percentage of distributed generation and storage.	build	low	high
8	**EVs and PHEVs**: percentage shares of on-road, light-duty vehicles comprising of EVs and PHEVs.	build	nascent	low
9	**Grid-Responsive Non-Generating Demand-Side Equipment**: total load served by smart, grid-responsive equipment.	build	nascent	low
	Delivery (T&D) Infrastructure			
10	**T&D System Reliability**: SAIDI, SAIFI, MAIFI.	value	mature	declining
11	**T&D Automation**: percentage of substations using automation.	build	moderate	high
12	**Advanced Meters**: percentage of total demand served by advanced metered (AMI) customers	build	low	high
13	**Advanced System Measurement**: percentage of substations possessing advanced measurement technology.	build	low	moderate
14	**Capacity Factors**: yearly average and peak-generation capacity factor	value	mature	flat
15	**Generation and T&D Efficiencies**: percentage of energy consumed to generate electricity that is not lost.	value	mature	improving
16	**Dynamic Line Ratings**: percentage miles of transmission circuits being operated under dynamic line ratings.	build	nascent	low
17	**Power Quality**: percentage of customer complaints related to power quality issues, excluding outages.	value	mature	declining
	Information Networks and Finance			
18	**Cyber Security**: percent of total generation capacity under companies in compliance with the NERC Critical Infrastructure Protection standards.	build	nascent	nascent
19	**Open Architecture/Standards**: Interoperability Maturity Level – the weighted average maturity level of interoperability realized among electricity system stakeholders	build	nascent	nascent
20	**Venture Capital**: total annual venture-capital funding of smart-grid startups located in the U.S.	value	nascent	high

Figure 2-1 DOE summary of smart grid metrics and status.

In 2010, Locke and Gallagher, both researchers at the National Institute for Standards and Technology (NIST) identified in their "NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0," [27] that communication applications providing what they term *wide area situational awareness*, which includes wide area monitoring, control, and protection are a priority and should be protected. They mention substations and distribution grids as example systems, and also place distribution grid management and automation on the list of priorities for a smarter grid. Aligning with the priorities stated in the NIST document, and given that the electrical distribution network is among the least automated elements of the grid, and that some 80% of outages occur at the distribution level [3], it is natural to focus on improving this area of vulnerability.

While transmission has seen more automation than the distribution system, the monitoring and control elements discussed here are readily adaptable to the transmission network, and could be used to provide greater levels of monitoring and redundancy. While transmission does not suffer the number of interruptions typically experienced in the distribution network, the consequences of disruption are greater.

The preceding works provides a perspective on the history of critical infrastructure protection, demonstrating roots that date back to at least the Second World War, and serving to explain the focus on systemic, interconnected, and cascading vulnerabilities. The changing view of critical infrastructure, the relatively recent focus, and the number of sectors and agencies involved has led to fragmented, incomplete, and occasionally even contradictory regulatory efforts to define how critical infrastructure will be identified and protected.

2.3 Sensor and Communication Systems to Improve Electrical Grid Control

An option for greater electric utility security and recovery is the application of cyber-physical systems[3] using wireless sensor networks for transmission line monitoring. In 1993, Seppa reported in his article *A Practical Approach for Increasing the Thermal Capabilities of Transmission Lines* that, due to the expense, the then-current practice was to monitor a single critical span as representative of the entire system [29]. In 1996, Engelhardt and Basu stated in their *Design, Installation, and Field Experience With an Overhead Transmission Dynamic Line Rating System,* that this "single span sampling" approach was flawed due to variability in weather and exposure throughout the system [30].

In 2003 the Bush Administration released *The National Strategy for the Physical Protection of Critical Infrastructures and Key Assets* [22]. The authors of the document highlighted the need to "improve technical surveillance, monitoring and detection capabilities". They also stated that:

> *Adequate protection of our critical infrastructures and key assets requires:*
> - *Improved collection of threat information;*
> - *Comprehensive and relevant threat assessment and analysis;*
> - *Robust indications and warning processes and systems; and*
> - *Improved coordination of information sharing activities.*
>
> *Accurate, timely information is a fundamental element of our national critical infrastructure and key asset protection effort.*

Cyber-physical systems are the means through which we will be able to collect the necessary information to monitor the security of the electrical grid.

---

[3] The term cyber-physical system is attributed to Helen Gill from the National Science Foundation. This term came from the word cybernetics created by Norbert Wiener in 1948, who in turn based this term on the Greek word "kybernetes," meaning steersman or pilot. The nautical term serves as an apt description of control systems. Wiener described cybernetics as the intersection of control and communication, a closed-loop system of feedback where physical processes communicate information on their status to the control system, which then influences how the physical systems operate. Cyber-physical systems highlight the intersection of physical processes, computation, and communications. Unlike embedded systems, where the focus is primarily on computation, cyber-physical systems recognize the important role of interaction with the physical world [28].

With their presentation entitled *Power Line Sensornet – A New Concept for Power Grid Monitoring,* at the 2006 IEEE Power Engineering Society General Meeting, Yang, et al. are credited with some of the earliest work recommending wireless sensor networks for transmission and distribution monitoring [31]. Their paper noted many challenges to incorporating sensor networks into the grid, and focused heavily on the requirements for developing a sensor, with only a passing mention of the possibility for communications based on the IEEE 802.11 or 802.15.4 standards. Sensing equipment was not directly interconnected at that time, but rather data was gathered and transmitted to a central location where control systems and human operators used it to manage utility operations. Wired, or sometimes cellular wireless, sensors for the transmission and distribution network were placed on electrical lines and towers and were able to measure current, voltage, waveform, temperature and heating, sag, conductor strength, galloping[4], icing, wind speed, and contact with vegetation and animals. Multiple sensors based on both data and video technologies were required to monitor all of these parameters, however, and they are expensive; on the order of $10,000 to $50,000 per sensor [31]. They state that a "typical" utility with 25,000 km of electric lines and several thousand capacitors, transformers and breakers over 20,000 to 80,000 square kilometers could require 100,000 sensors [32]. Yang et al. set a practical target for the price of a future sensor meant for widely-distributed monitoring at under $100, and provided a block diagram of a proposed sensor [31].

[4] Gallop is oscillation of the line itself due to the effect of wind, and can cause conductors to come too close to each other or to other objects. It also places stress on poles and insulators, and can be made significantly worse if there is also a build-up of ice on the conductors.

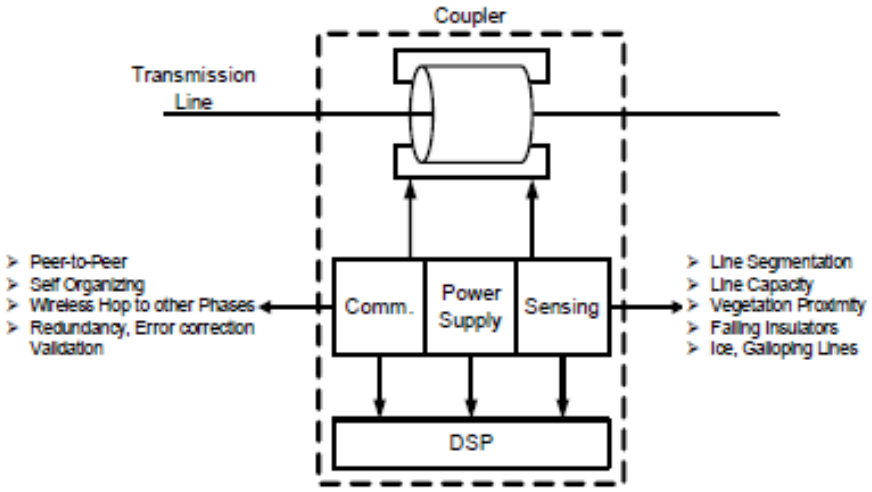


Figure 2-2 Schematic of sensor proposed by Yang, et al.

At the 2007 IEEE Power Engineering Society General Meeting, Yang et al. added a survey they conducted among transmission and development professionals to the body of knowledge. In their report entitled *A Survey on Technologies for Implementing Sensor Networks for Power Delivery Systems* [32] the authors reported that industry experts believe that while much work needs to be done to improve substation monitoring, they feel it is even more important to focus on monitoring assets outside of the substation. They also identified the following potential applications of sensor networks including:

- Overhead conductor temperature, sag and dynamic capacity
- Overhead structure integrity, reclosers, capacitors, and sectionalizers
- Underground cable and neutral conductors, temperature and capacity
- Overhead and underground faulted circuit indicators
- Pad mount and underground network transformers
- Wildlife and vegetation contact warning
- Underground network transformers, switches, vaults, manholes, switches

In *A Survey on the Communication Architectures in Smart Grid* [33]Wang, Xu, and Khanna discuss the use of communications as a means of improving power quality, efficiency, and optimization. They also mention the possibility of using sensors for transmission line monitoring, connecting them through a wireless sensor network until they reach a "measurement collection site." This is essentially the application explored in this thesis, although Wang, Xu, and Khanna do not elaborate further on the details of the system they envisioned.

In 2011, Erol-Kantarci and Mouftah, in their article entitled *Wireless Multimedia Sensor and Actor Networks for the Next Generation Power Grid* [34], discussed using wireless sensor actor networks in the power grid along with traditional sensing equipment like weather stations, sagometers[5], and power donuts[6] which are typically connected through wired communications. Other available commercial monitoring methods and products, as well as some generic sensor nodes, are surveyed in [32].

Yang et al. note many challenges to incorporating sensor networks into the grid, several of which are also discussed by Gungor et al. in [35] and [36]. These include poor link quality, low bandwidth, long latency, jitter, quality of service issues (a topic further discussed by Howitt et al. in [37]), the need to create distributed decision and control algorithms which allow for remote processing, and the ability to manage the amount of data generated by a large number of sensors, a concern also mentioned in [38]. In [35] and [36] they also mention the need for further sensor node development, the need to integrate with other existing and emerging grid

[5] A name used in devices by several manufacturers, with the trademark held by the Electric Power Research Institute. It is an instrument for monitoring conductor ground clearance on overhead transmission lines.

[6] A donut-shaped intelligent end device manufactured by USi. It is clamped around the utility line and can measure conductor temperature and inclination; data which can be used to compute line tension, sag, and clearance. Data is sent via GSM/GPRS.

communications technologies, and the need to develop standardized architectures and protocols. In another paper Yang et al. [32] note specific challenges to wireless communications, including security, electromagnetic interference, fading, bandwidth overloading, wireless protocol immaturity, and a need for more testing in the substation environment. They do note that work is underway on a new standard, IEEE P1777 *Using Wireless Data Communications in Power System Operations*, but at this time the standard is only a draft.

Isaac et al. wrote in *A Survey of Wireless Sensor Network Applications From a Power Utility's Distribution Perspective,* for AFRICON 2011 [39] that the need for further research in a number of key areas, matching the applications noted above where wired sensing is currently done. These include conductor sag, conductor temperature, thermal capacity and dynamic rating, galloping, and vegetation and animal contact monitoring. They also mention fault circuit indicators, underground cable monitoring, tower and pole monitoring, and energy harvesting. They identified four other key application areas for wireless sensor networks in the transmission and distribution system. These include the prevention of theft, including cables and lattice tower members; monitoring tower tilt and subsequent conductor sagging due to geologic, wind, soil or other conditions; detecting leakage currents which damage insulators; and fault detection and location. They note that they are not currently aware of a utility that has deployed a sensor network, and cite the lack of standards and an accepted data-management structure as significant challenges, along with a need for further work on security.

In this paper I begin to explore one of the communications issues which will affect the design of a system for power line sensing using wireless mesh networks; that of multi-hop mesh throughput. I also investigate the performance characteristics of one potential method of data backhaul using SMS. Further background literature related to these specific topics will be noted

in the appropriate later chapters. There is no other data available on these issues in this context that I am aware of. I also attempt to provide regulatory and technical foundations demonstrating the need for new methods of grid control; an integrated approach not found in other documents with a single focus.

3 The Electrical Grid

In this section I cover the current structure and function of the electrical grid. Goals of this chapter are: 1) to provide a basic view of how the system functions, 2) to indicate parameters which are essential for stable and reliable operation. 3) denote how and where essential monitoring and control it is done. As electric systems developed, certain fundamental concepts and terms also evolved that assist in understanding the function of the power grid, the requirements and opportunities for monitoring and control, and the parameters and electromechanical equipment necessary to manage the basic power grid.

3.1 Origin

Beginning in the late-1800s electrical utilities constructed individual systems to serve dedicated customers, with each company's distribution system connecting the generation station and the customer. Over time utilities began to interconnect their own varied generation assets and islands of customers to improve efficiency, capacity, and reliability, and under government pressure during World War I connected to other electrical utilities as well [40]. Systems based both on direct current (DC) and alternating current (AC) existed, but for practical reasons, primarily the lack of DC transformers, DC could be delivered at only one voltage throughout the system. Safe operation dictated that the DC systems be operated at a relatively low voltage, but as will be explained, this led to very high energy losses. AC voltage could be easily "transformed" up and down for different applications, and this flexibility was particularly important for longer-distance transmission as the transmission and distribution system loses less energy at higher voltages. In 1887 Nicola Tesla introduced a complete electrical system based on AC power, and while islands of DC existed for decades, over time the U.S. electrical grid became an AC-based system. A few specialized DC applications exist, such as high-voltage DC

transmission, but the systems are expensive to construct and have limited applications. Nationwide, utility systems continued to interconnect until they merged into three relatively independent regional electrical networks for the entire U.S., each with their own synchronization: the Eastern Interconnect, the Western Interconnect, and the Texas (ERCOT[7]) Interconnect. These systems include limited ties to electrical systems in Canada and Mexico.

3.2 Fundamental System Terminology[8]

The electric grid is usually discussed in the context of four broad functional areas: generation, transmission, distribution, and load or consumption. As electric systems developed, certain fundamental concepts and terms also evolved that assist in understanding the function of the power grid, the requirements and opportunities for monitoring and control, and the parameters and electromechanical equipment necessary to manage the basic power grid.

[7] Electric Reliability Corporation Of Texas

[8] Many of the following definitions are taken from the *Glossary of Electric Industry Terms* [41] by the Edison Electric Institute. While a detailed description of each term would be more accurate, the simplified version is intended to allow discussion of the important parameters of a functional grid without devoting excessive detail to their composition or management.

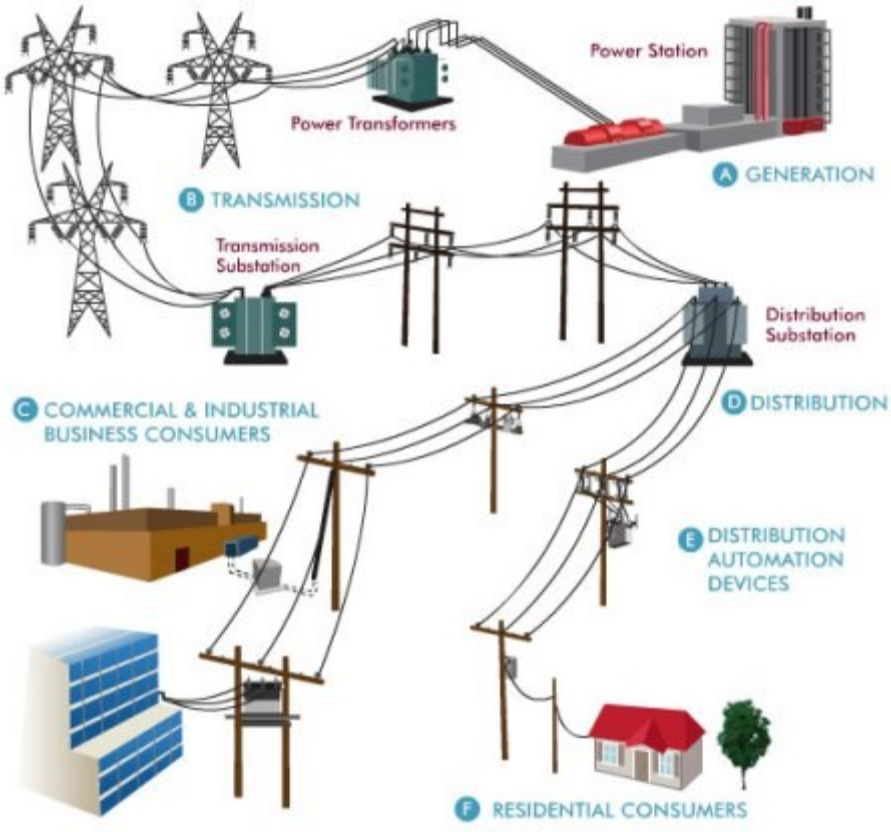


Figure 3-1 Electrical grid overview. [42]

Alternating Current (AC) An electric current that reverses its direction of flow periodically as contrasted to direct current.

Current A flow of electrons in an electrical conductor. The strength or rate of movement of the electricity is measured in amperes at a pressure measured in volts.

> **Ampere (amp)** The unit of measure of an electric current. It is proportional to the quantity of electrons flowing through a conductor past a given point in one second. It is analogous to cubic feet of water flowing per second.

Volt The unit of electromotive force or electric pressure analogous to water pressure in pounds per square inch. It is the electromotive force that, if steadily applied to a circuit having a resistance of one ohm, will produce a current of one ampere.

Ohm The unit of measurement of electrical resistance. It is that resistance through which a difference of potential, or electromotive force, of one volt will produce a current of one ampere.

Voltage (of a Circuit) The electric pressure of a circuit in an electric system measured in volts. It is generally a nominal rating based on the maximum normal effective difference of potential between any two conductors of the circuit. The voltage of the circuit supplying power to a transformer is called the primary voltage, as opposed to the output voltage or load-supply voltage that is called secondary voltage. In power supply practice the primary is almost always the high-voltage side and the secondary the low-voltage side of a transformer, except at generating stations.

Frequency The number of cycles per second through which an alternating current passes. Frequency has been generally standardized in the United States electric utility industry at 60 cycles per second (60 hertz).

Frequency Bias A value, usually given in megawatts per 0.1 Hertz (MW/0.1 Hz), associated with a Control Area that relates the difference between scheduled and actually frequency to the amount of generation required to correct the difference.

Frequency Deviation A departure from scheduled frequency.

Frequency Error The difference between actual system frequency and the scheduled system frequency.

Frequency Regulation The ability of to assist the interconnected system in maintaining scheduled frequency. This assistance can include both turbine governor response and automatic generation control.

Frequency Response (Equipment) The ability of a system or elements of the system to react or respond to a change in system frequency.

Frequency Response (System) The sum of the change in demand, plus the change in generation, divided by the change in frequency, expressed in megawatts per 0.1 Hertz (MW/0.1 Hz).

Scheduled Frequency 60.0 Hertz, except during a time of correction.

Power (Electric) The time rate of generating, transferring, or using electric energy[9], usually expressed in kilowatts (kW).

Apparent The product of the volts and amperes of a circuit. This product generally is divided by 1,000 and designated in kilovolt amperes (kVA). It comprises both real and reactive power.

Reactive Power The portion of electricity that establishes and sustains the electric and magnetic fields of alternating-current equipment. It is used to control voltage on the transmission network. Reactive power must be supplied to most types of magnetic equipment, such as motors and transformers. It also must supply the reactive losses on

[9] This can also be expressed as the product of voltage and current.

transmission facilities. Reactive power is provided by generators, synchronous condensers, or electrostatic equipment such as capacitors and directly influences electric system voltage. It is usually expressed in kilovars (kVAr) or megavars (MVAr), and is the mathematical product of voltage and current consumed by reactive loads. Examples of reactive loads include capacitors and inductors. These types of loads, when connected to an alternating current voltage source, will draw current, but because the current is 90 degrees out of phase with the applied voltage, they actually consume no real power.

Real The energy or work-producing part of Apparent Power. The rate of supply of energy, measured commercially in kilowatts. The product of real power and length of time is energy, measured by watt-hour meters and expressed in kilowatt-hours.

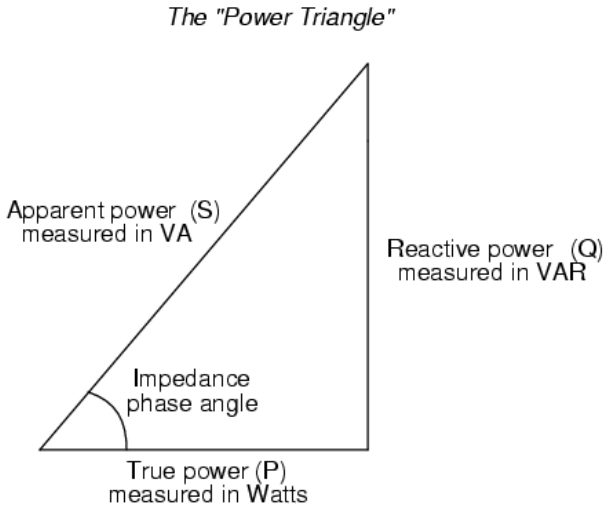


Figure 3-2 The relationship between real, reactive, and apparent power. [43]

Inductance The property of an electric circuit by virtue of which a varying current induces a voltage in that circuit or in a neighboring circuit.

Synchronism The timing of alternating current generators so that their voltage waves go through their maximum and minimum values at exactly the same rate. Alternating current generators must be in synchronism to operate on the same system.

Several other important terms are defined below, although they do not come from the Edison Electric Glossary.

Capacitance The property of an electrical circuit that resists a change in voltage. Banks of capacitors are installed in the electric system and have the effect of negating inductance and raising real voltage.

Impedance The vector result of the combination of resistance and reactance, which together act as an impediment to current flow.

Resistance In DC systems the opposition to current flow, and in AC systems the component of impedance responsible for real power loss.

Reactance The impediment to the flow of AC within a device based upon its influence on the relative timing of the current.

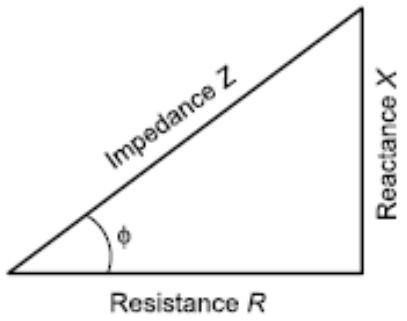


Figure 3-3 Relationship between resistance, impedance, and reactance. [44]

The relationship between voltage, current, resistance, and power is shown with the most fundamental equation in electrical engineering, Ohm's Law:

$$V = I * R \tag{5}$$

As power can be expressed as

$$P = V * I \tag{6}$$

It may also be shown as

$$P = V^2 / R \tag{7}$$

Although the relationships may be derived algebraically, they are often conveniently expressed by referring to the "power wheel." As in the following case, U is often used in Europe as the notation for volts (V).

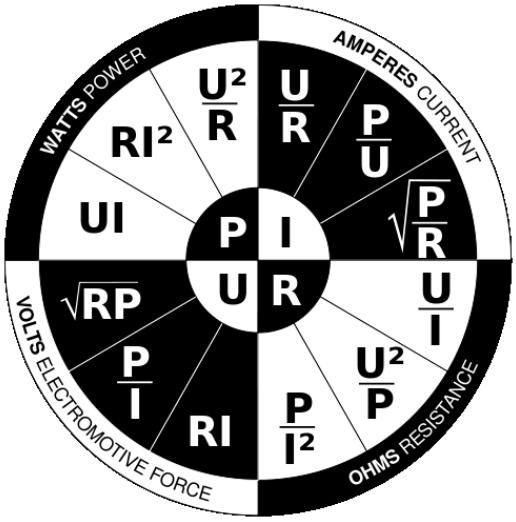


Figure 3-4 Ohm's Law Wheel. [45]

3.3 The Four Functional Areas of the Electric Grid

3.3.1 Generation

Generators work on the principle of electromagnetic induction. When a conductor moves in relation to a magnetic field, voltage is induced in the conductor. Power plant generators are three-phase, or constructed with three terminals which produce power which is 120 degrees out of phase with each other. This permits constant instantaneous power from the generator, and reduces the overall size of conductors required for transmission. In the U.S. AC generators operate at 60 Hz, and as will be explained further, it is important that all generators connected to a common system be able to coordinate their output so that the sine waveforms produced are as nearly synchronous as possible.

Generators convert mechanical energy to electrical energy. The rotating shaft of the generator is driven by a prime mover, such as a steam turbine or directly through hydropower. Most generation is thermally driven, which means that some fuel such as coal or natural gas is burned to produce the steam which drives the turbine. In the case of nuclear plants, it is the heat from the reaction which is used to create the steam for the turbine.

As it is impractical to store large quantities of electricity, at any instant the quantity of electricity generated must match the quantity demanded by the load. This leads to three types of generating plants: baseload, load-following, and peaking. Baseload generation, such as provided by coal-fired and nuclear plants, is economical to operate but cannot be quickly modulated. It is used to meet the level of relatively constant loads which are determined to historically exist on a power system. Load-following units, such as combined-cycle gas turbine driven generation, have the ability to run for long periods of time, but may be turned off, and can

vary their output more readily than baseload units. Peaking plants, such as gas turbine, may be started, stopped, and regulated quickly. They are usually only used when electricity demand is near its peak. The electricity generated from these plants is relatively expensive as they may use more expensive fuels, and building the plants involves high fixed costs even though they may be idle most of the time.

Environmental concerns, difficulty obtaining permits and approvals, and high cost all serve as impediments to constructing new traditional generation facilities. At the same time, increasing demand and aging infrastructure poses a threat to reliability and adequate capacity in the electric sector. Advances in sensing, communications, and control hold the promise of reducing and shifting demand, improving efficiency, enabling the incorporation of alternative generation, and providing for better monitoring and maintenance.

3.3.2 Transmission

Unlike the first generating stations, as a power plants grew they began to be located outside of the population centers they served, necessitating the long distance transmission of power to the local distribution networks. The transmission system is composed of transmission substations and transformers to boost the voltage from the level at which it can be safely generated to levels high enough to be efficient for long distance transmission[10]. It is also composed of high voltage transmission lines, commonly between 138 kV and 765 kV, and shares distribution substations which contain transformers which reduce the transmission voltages back down to distribution levels. Substations also contain related equipment such as switchgear or circuit breakers, which are used to protect the system and disconnect parts of the network for

[10] Power loss in an electric circuit, or transmission line, equals the resistance of the conductor multiplied by the square of the current. By increasing the voltage, current may be reduced while still transmitting the same quantity of power. $P=I^2R$ where P is the power loss, I is the current and R is the resistance, or $P=V^2/R$ where V is the voltage.

maintenance. Measurement, metering, control, and communications equipment are also housed in substations and at points along the transmission network, allowing parameters such as voltage, current, and power quality to be remotely monitored and some equipment to be remotely controlled.

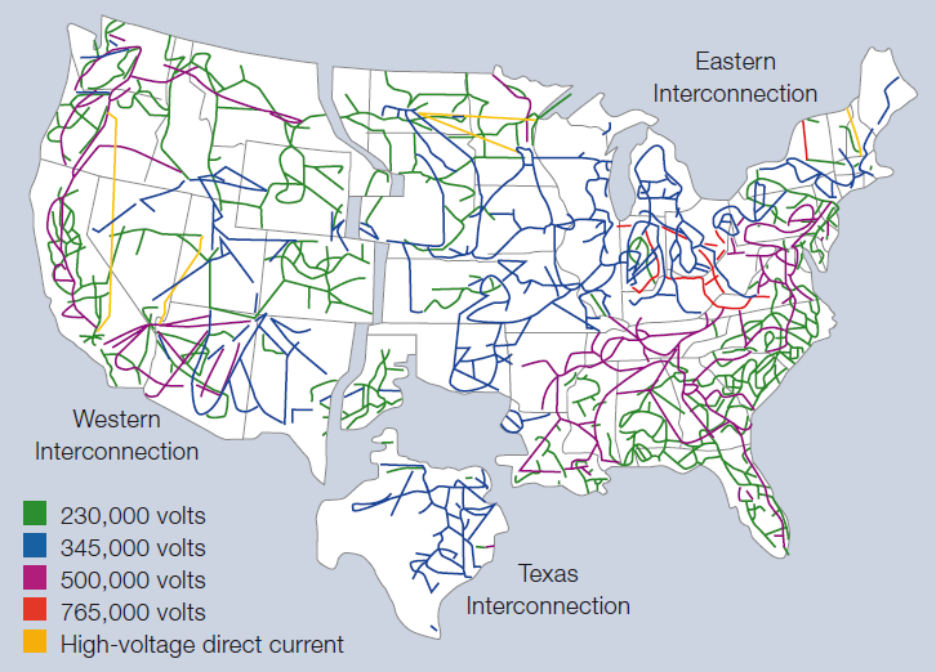


Figure 3-5 U.S. transmission grid. [8]

The amount of power that can be transmitted over a given conductor is limited by a number of factors, including the conductor's material, size, length, and distance from other conductors and potential ground elements. Transmission loads are also governed by *thermal limits*, *stability limits*, and *voltage limits*. *Thermal limits* are primarily a function of heating of the conductors due to resistance in the conductor material, leading to excessive power loss and sag in the line. While not the primary constraint, transmission thermal limits are affected by ambient temperature, and are part of the concept of *dynamic rating,* which permits the maximum carrying capacity of the line to be adjusted for factors such as ambient temperature and the

amount of time the line has been heavily loaded. The *stability limit* refers to the difficulty of keeping remote generators in synchronism with each other, particularly as feedback is required due to ever-varying loads on each generator. *Voltage limits* affect power transmission because reactance in the transmission line causes a drop in voltage over its length, and for practical reasons electrical systems generally do not permit a drop to less than 95% of the design voltage. Shorter transmission lines are usually constrained by the thermal limit, and longer lines by the stability limit. Voltage limits may also constrain longer lines, although it is possible to install equipment which helps to boost or regulate voltage.

Transmission networks are typically connected in a grid or mesh topology, rather than a point-to-point or hub-and-spoke. This creates redundancy and allows for electricity to take multiple routes, bypassing generation and transmission resources which may be accidently or intentionally taken offline. As electricity follows the path of least resistance, affected by phase, amplitude, and impedance, it is difficult to predict the exact circuit it will follow between varying generation and loads when there is more than one possible route. Small changes in voltage and impedance may be made to influence the power path, and hence the load on individual lines, but a great deal of state information and computation is required to monitor and control the system to prevent transmission line overloads. Increasingly, the U.S. electrical grid faces the issue of *congestion,* where loads cannot be matched to the most economical remote generation due to the lack of capacity and potential for overloads in the transmission system. This problem became worse following FERC Order 888 *Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities* which allowed power to be sold in competitive markets and transferred to the buyer over the electrical grid. This process, known as *wheeling,* creates additional load and stress on the grid, particularly

as distances grow large (i.e. hydro-power generated in Washington and purchased in California). NERC has created Transmission Loading Relief (TLR) procedures, and tracks the frequency at which they are implemented. Within two months of approval of FERC Order 888 TLR instances were occurring at a rate six times that of previous years [46].

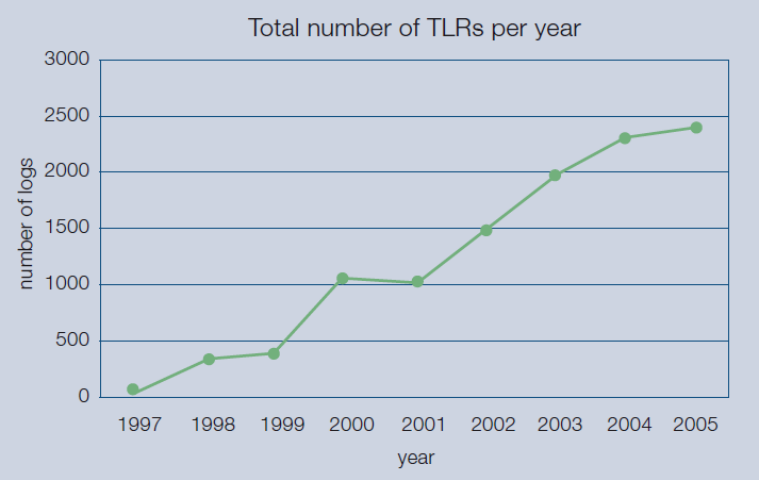


Figure 3-6 Increasing transmission loading relief incidents. [8]

Utilities are required to make their transmission lines available to third-party generators and marketers at the same rate they charge to carry power from their own generation facilities, yet the third-party generators and marketers do not have to invest in the construction of transmission lines themselves [47]. Congestion, lack of incentive, and the difficulty of obtaining rights-of-way and permits to build new transmission lines present a significant threat to critical energy infrastructure.

3.3.3 Distribution

Distribution networks begin at the distribution substation, which typically step power down from transmission voltages to the 4 kV to 35 kV range. The substation may contain equipment similar to that found in transmission substations, although currently they are often less

automated, providing an opportunity for improvement. Like transmission lines, distribution is three phase, although the conductors may be split near the load to serve individual neighborhoods. Residential and small commercial loads in the U.S. are typically served at 120 V and 240 V, requiring another small transformer near the point of consumption. Distribution networks are often arranged in a radial topology, although ring and mesh are not uncommon. Even with ring and mesh configurations, a disconnect is usually left open so that the networks are operated as point-to-point connections with the ability to close the disconnect and failover to another route if necessary. While power flows have historically been from substation to load, this is changing with the introduction of distributed generation and microgrids, which will require more sophisticated monitoring and control.

3.3.4 Load

While an important element of the electric system, beyond the meter the grid becomes the domain of the consumer more than the regulator or the utilities. While this domain is generally not considered part of regulated critical infrastructure, if the goal is to deliver end-to-end reliable electricity, this link in the chain cannot be ignored. As "smart" buildings, homes, and systems become interconnected with the electrical grid the potential for disruption of the grid increases, even if consumer systems are not directly interconnected with utility control and communications architecture.

Electrical loads are the reason the electrical grid exists, and have important engineering implications for the way the grid functions. A charging plug-in electric vehicle (PEV) may draw as much power as an entire house, so the appearance of several within a neighborhood may easily exceed the original design for the neighborhood's distribution system unless they can be

monitored and controlled. It is also important to monitor the load on the total system, both to plan for peak demand and to meet instantaneous requirements.

The type of load also has important implications for the way the grid functions, based upon its *impedance*. Different types of loads have characteristic *resistance* and *inductive* or *capacitive reactance*. Baseboard heaters and incandescent light bulbs are examples of purely resistive loads, which only consume *real power*. Many loads, such as motors, are a combination of resistive and inductive impedance, which effectively draw *reactive power*. While loads with capacitive impedance are effectively able to supply reactive power, the number of motors connected to the grid means that it is heavily skewed towards needing generators to supply both real and reactive power. It is possible to connect capacitor banks, which produce reactive power, to cancel inductive loads and reduce the demands on generation and transport, but this introduces the need for monitoring and control of power quality in the system.

4 Traditional Grid Control Systems

The electrical grid has been called the most important engineering achievement of the 20th Century, and the transmission and distribution system is the largest machine ever made [48]. However, it still uses technologies that have changed little in 100 years, and that are poised for a revolution in their physical, organizational and conceptual structure. This section will focus on one element of that change, the control and communications networks. These are an important component of the current power grid, and their evolution is one of the very elements Ericsson [49] claims makes possible and serves to define the future "smart grid."

Broad challenges to communications and control network evolution include the regulatory environment, and standards which are nonexistent, inadequate, or lacking integration. The design of the existing communications network is also a significant issue, with limited reach and capability, low bit rates, inability to collect data in real-time, minimal control functionality, and the inability to connect to other communications and control networks or future smart grid elements. There is a reliance on proprietary hardware, software, protocols and support. Security is also significantly lacking, with inadequate identification and physical security of critical assets, a poor authentication and authorization capability, and the lack of secure protocols and software.

4.1 The Evolving Electrical Grid

In order to consider the challenges to current electrical grid communications, it is useful to see how the electrical grid itself is evolving. The traditional electrical grid relies on bulk generation, often located some distance from the majority of users. A transmission network is used to move bulk electricity closer to the end users. Substations and distribution networks

reduce the transmission voltages and supply power to end users. Electricity only flows in one direction, from generator to consumer. Operations rely heavily on manual control with the assistance of relatively isolated and limited supervisory control and data acquisition (SCADA) systems, transmitting mostly low bit rate data, and with SCADA systems seldom reaching closer to the end user than the substation. Investments in infrastructure have not kept pace with demand, and it is generally accepted that an aging infrastructure and workforce have cast the future reliability and security of the grid into doubt.

Fan et al. [4] defines the Smart Grid as "an *intelligent* electricity network that integrates the actions of all users connected to it and makes use of advanced information, control, and communication technologies to save energy, reduce cost and increase reliability and transparency." While this is an excellent high-level definition, it abstracts some of the very practical details which make the future grid very different from the present. For example, consideration will need to be given to distributed generation and storage, where end users of electricity may also become islanded generators of power, or store power locally for their own use, and there will be times when this power flows back into the grid. The future grid also contemplates a much greater use of renewable energy, such as wind and solar, which are also highly variable sources of generation. To address this issue and others, large-scale storage such as pumped hydro and compressed air will be incorporated into the grid infrastructure. The distributed, variable, stored and bidirectional flow of electricity will be a significant change from present grid operations.

There will also be a significant increase in the connectivity and intelligence of the data and control layer of the grid. The largest change will be the involvement of the end user in the data network, even beyond the connection of smart meters. Providing real-time information and

pricing to the consumer has several goals: changing behavior, conserving energy, shifting load to off-peak periods, better reflecting the actual cost of variable demand, and providing credit for energy generation. Another significant change will be the ability to identify and control end user devices. For instance, it will become possible to cycle an air conditioner compressor during periods of peak energy use, program a dryer to run when demand is lowest, and bill a consumer for recharging their electric vehicle regardless of where they plug it in. Further changes will take place in the communications architecture itself because there will be a need for synchronous real time communications among many more devices, and there is a need to increase the interconnection, reliability and security between the different elements of the grid.

4.2 The Evolution of Electrical Grid Communications

Early electric grid controls were electro-mechanical and primarily concerned with regulating generation. Communications, if they existed at all, were completely separate systems. The telephone system could be used by a customer to call in a problem, or to communicate between a utility plant and office. Later, radios were used to dispatch service vehicles or for voice communications with field employees. More recently the grid has incorporated communications technology for SCADA systems, or distributed control systems (DCS). SCADA systems typically consist of process sensors which are connected to remote terminal units (RTU) or simpler programmable logic controllers (PLC) which are responsible for converting the sensor output to a digital signal and sending it via a communications architecture. The communications architecture delivers the signals to a computer system which can monitor, store and manipulate the data or send control commands. The computer system may do this automatically, or through a human-machine interface (HMI) with a human operator. The functionality of the overall system is determined in large part by the software which runs on the

computer system. SCADA systems were traditionally supervision-oriented and DCS systems were control-oriented, but the differences have largely disappeared, and both are often referred to now using the more generic term "industrial control systems."

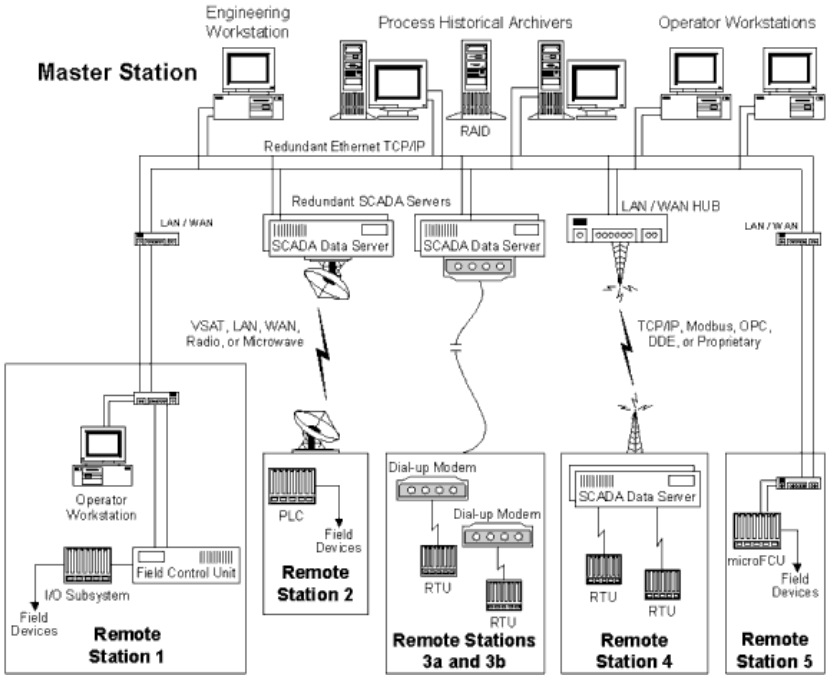


Figure 4-1 Example SCADA system architecture. [50]

Early control systems reflected the thinking and technology of the time. Distributed sensors were connected to a central computer, virtually all elements of the system were proprietary to one manufacturer, and data flowed only from the sensors to the computer, reflecting monitoring but not control capability. As the systems evolved, two-way communication became more common, although generally at very low bit rates. The central computer was replaced with multiple servers. Standard communication protocols, such as RS-232 serial were sometimes used throughout one system, although there was no single standard used by all vendors or all systems. Further evolution permitted systems to be networked, and for the software to run on standard personal computer architecture and operating systems such as

Microsoft Windows. Some systems have adopted open standards, allowing devices from many different manufacturers to interconnect, and to use common networking protocols such as the Transmission Control Protocol/Internet Protocol (TCP/IP) to run over a variety of media, from Ethernet to wireless. Despite the changes, industrial control systems still use a primarily hub and spoke or star connection topology. Electrical grid communications have been evolving away from serial and toward IP, and in the future will be called upon to carry data for a much wider variety of applications.

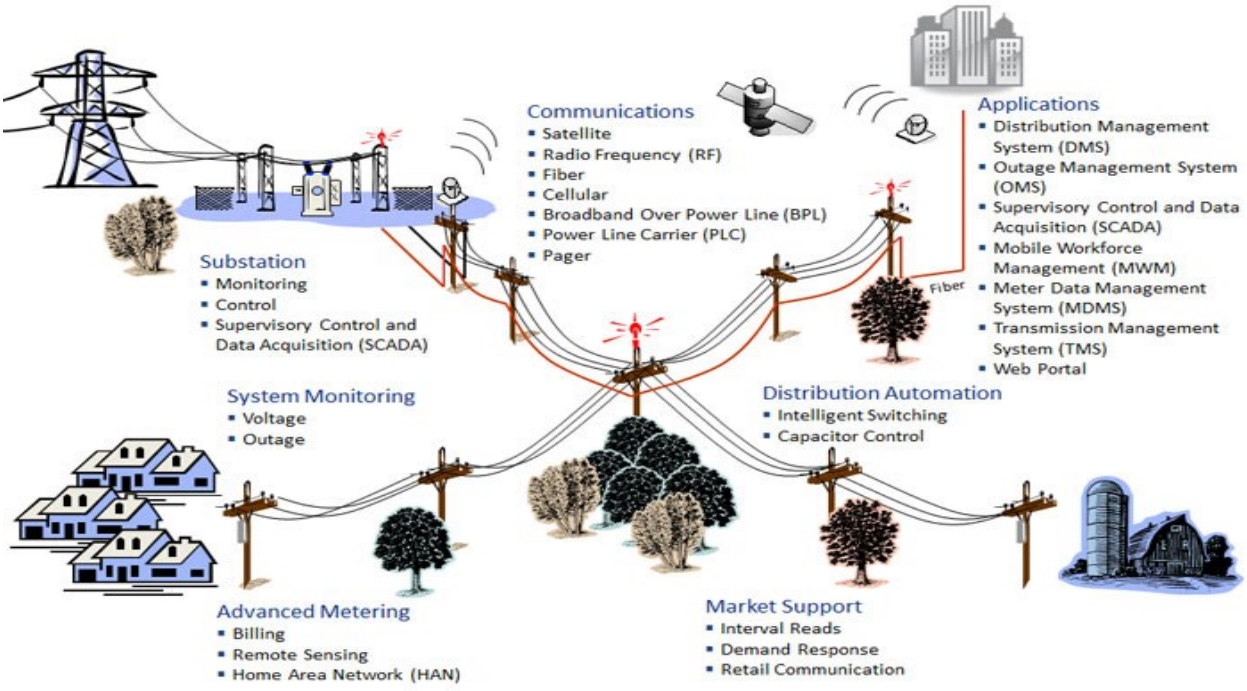


Figure 4-2 Example electrical grid communication uses and carriers. [51]

While the graphic above provides some examples of the classes of things which can be monitored, controlled, or communicated, the actual number of specific potential data points is much larger. The graphic below shows some of the information that could be monitored on a single transformer:

	Communication			
	Units	Alarm	Local	Remote
Transformer Parameters				
Tank pressure	psi	Y	Y	Y
Tank vacuum	psi	Y	Y	Y
Oil temperature	degrees	Y	Y	Y
Winding temperature	degrees	Y	Y	Y
Pressure relief device operation	on/off	Y	Y	Y
Sudden pressure relay operation	on/off	Y	Y	Y
Liquid level	on/off	Y	Y	Y
Hydrogen gas	%	N	N	Y
Water content in oil	%	N	N	Y
System Parameters				
Fans on	on/off	Y	Y	Y
Loss of control power	on/off	Y	Y	Y
Ambient temperature	degrees			Y
Input current	amps	N	N	Y
Input voltage	volts	N	N	Y
Output current	amps	N	N	Y
Output voltage	volts	N	N	Y

Figure 4-3 Potential monitoring parameters for an electrical transformer. [52]

To carry the example further, the Common Information Model (CIM) was developed by the Electric Power Research Institute (EPRI) and is now maintained by the International Electrotechnical Commission (IEC) [53] [54]. It is in widely used in North America and by many utilities throughout the world. It provides an extensible model based upon the Unified Modeling Language (UML) for describing components of a power system and a common Extensible Markup Language (XML) format for exchanging data between software applications and other utilities. It provides a hierarchical class structure similar to object oriented programming. The following data points, or more, could be assigned to the windings on the transformer in the above example, for instance:

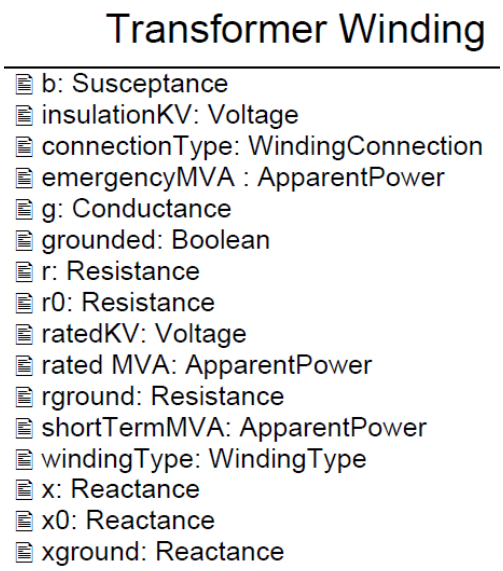


Figure 4-4 Example transformer winding attributes. [55]

Of course it is also possible to forego monitoring anything about the transformer directly. If the transformer fails power will go out, customers will call the electric utility, and a crew can be dispatched to investigate the problem. While this is the historical scenario, it is neither proactive nor efficient, demonstrating the past and future of electrical grid communications. It also demonstrates that it is not possible to say with certainty the things that are, or could be, monitored in the electrical grid, although each grid entity will know what it currently does monitor on its own system.

4.3 Limits to Legacy Electrical Grid Control Systems

As has been noted, a number of challenges to grid communications are not strictly limited to single hardware, software or protocol issues. There are also regulatory and mindset influences, and the complex interplay of different systems and goals that makes for multi-faceted, controversial and complicated future potential directions. This section will examine some of the issues.

Considering the evolution of electrical grid communications, and in light of the issues, it is less than surprising that many control systems have evolved slowly from their initial implementation, or that advanced control functionality is limited in certain organizations. Older systems largely reflect proprietary hardware, software and protocols with little standardization, low bit rates, little real-time reporting or control capability, and limited reach and networking. These systems are not well suited for the evolving smart grid, yet utilities and regulatory authorities may not have the ability or even desire to advance their capabilities. Future systems, however, will need to expand far beyond the traditional definition of SCADA and become complex data networks with control capability and more, such as the ability to connect devices never intended to run on an industrial communications network.

Despite the fragmentation in location, structure or regulatory agency, electric utilities have been united by a mentality of "the lights must stay on." This focus on reliability and the prevention of failure predisposes the industry to certain choices which make it difficult to implement desirable features of advanced control and communications. Hardware tends to be expensive because it must be highly reliable, robust, and long-lived by design. Proprietary solutions increase costs and reduce the ability to interchange components and interoperate systems, while limiting utilities to the features and pace of innovation offered only by the proprietary vendor. Proprietary solutions can be desirable to the utility, however, because they offer systems which can be certified, insured, and supported with a single point of contact and are reasonably certain to work from end-to-end. Proprietary solutions, including networks and protocols, are desirable for the vendor because they increase margins for products and services sold to the utility, and make it difficult to switch systems or components with another vendor in any sort of incremental fashion.

Increasing the reach and capabilities of the system will also increase complexity, maintenance requirements and costs and add additional points of failure. The need for real-time, two-way communications further increases the complexity of the system and generates a tremendous amount of data. This data comes with the need to manipulate, store and retain the information it contains, as well as further regulatory, policy, privacy and legal challenges. It creates the need for more complex human-machine interfaces, and for the first time, the need for an HMI on both the utility and customer ends of the system. The customer end is particularly problematic due to the need to manage the perceptions of the customer, and the widely varying technical capabilities and desired levels of involvement of each customer. In addition to interfacing with the customer directly, the future smart grid will require systems to interface with home area networks (HANs) and smart devices ranging from electric vehicles to hot water heaters.

4.3.1 Security

Security of both legacy and evolutionary control systems is one of the top concerns for the electrical grid. Physical security, while an important concern, will not be addressed here beyond its passing relation to securing the communications components, and to note that physical security is generally considered the first step in system security. Most existing networks were designed for simplicity, reliability and ease of use. Because security stands in opposition to these goals, and adds cost, it is often minimal when contemplated at all. The historical lack of standards or requirements and the slow pace of upgrades to existing systems often means that little has been done to address modern security concerns. Another drawback of existing architectures is their star topology, which creates, at the core, the possibility of a single point (or area) of failure.

Certain elements of the existing utility communications infrastructure, while a barrier to future development, have actually assisted with a relative measure of security. One-way low bit rate applications offer fewer opportunities and lesser incentive to adversely affect the network. The limited number of monitoring and control points present few points of entry onto the network. Running a limited-scale network used only for local control also limits the access and compromise opportunities for potential attackers. Proprietary software and communications protocols require specialized knowledge with limited applications, so while these systems may not be hard to crack they offered a certain level of "security through obscurity." This relative obscurity is of no assistance when the attacker is someone with inside access to the network, such as a disgruntled employee, and with the Internet providing easy access to even obscure information obscurity is no longer considered a deterrent. Strong authentication, authorization and accountability schemes, at a minimum, are one way to reduce this attack vector, but many existing systems have little, if any, capability to do so.

Watts [56] indicates that the requirements of the future smart grid communications infrastructure will also create significant additional challenges for control system cyber security. High speed, two-way communication will create more attack possibilities. The expansion of the network and more monitoring and control points will also increase the attack potential, as will the addition of customer interfaces. The interconnection of networks will present a greater number of vulnerabilities, particularly if communications are standardized over the Internet. The use of wireless devices and protocols will open new avenues for access, presenting challenges because wireless spectrum cannot be physically secured [56]. The control system trend is toward "open" software or protocols, but this also makes the source code and necessary knowledge to exploit these systems readily available. In addition to the increase in vulnerability, expanding

the scope of systems and interconnection also increases the potential scale of any damage should a system be compromised.

4.4 Looking Forward

Recognizing that wholesale changes to the electric grid infrastructure, including the related communications component, will only come about slowly without impetus, Congress has passed specific legislation such as the Energy Independence and Security Act of 2007. This act requires, among other things, that the Secretary of Energy study and regularly report to Congress the regulatory framework which affects deployment of smart grid initiatives. Similar regulatory reviews are being conducted in many states. Recognizing the role of regulation should result in incremental changes which remove barriers to smart grid development and encourage initiatives which address existing challenges such as improvements in grid control and critical infrastructure protection. New standards, while not without their own challenges, continue to be adopted and refined. While the industry lacks a unifying standard, by making steady progress toward defining necessary functionality and interoperability, a core set of standards supported by market forces may emerge in a fashion similar to the evolution of the Internet.

The aging electric grid infrastructure brings each element closer to replacement. Networks of the future are moving toward being IP-based, and increasingly use open, rather than proprietary, devices, protocols and platforms. Devices from different manufacturers are becoming more interchangeable, and open protocols allow these devices and networks to be interoperable. As component replacements are effected, it is possible to incorporate devices designed for improved control and communications. Even if this capability is not immediately utilized, or is not fully capable of supporting all features desired in the future, it still permits an incremental increase in the intelligence and capability of the communication network. Ericsson

[49] notes that by taking a cue from existing data networks, expensive proprietary equipment designed for high reliability is being replaced by redundant, cheap, commercial off-the-shelf (COTS) equipment. Although not without security issues of its own, this provides a different path to an equivalent level of reliability. Lacking an integrated network, it is possible to run complementary networks with each responsible for distinct functions. Select information from a legacy SCADA network may be converted to data which is accessible by an end user over the public Internet, for instance, or new sensors on a network may be connected to the Internet via a wireless connection and virtual private network (VPN) to report data to an existing SCADA system. The problem of a lack of unique identifiers for devices on the system can by handled by providing media access control (MAC) and Internet protocol (IP) addresses for new equipment similar to the current addressing scheme on the Internet, although many legacy sensors are too simple, inexpensive or resource-limited to include MAC/IP addressing.

5 Cyber-Physical Systems

The chapter includes a description of general cyber-physical system characteristics, wireless sensor actuator network (WSAN) architecture, and some specific examples of how the hardware, protocols, and networks function. These are essentially machine-to-machine communications, although they may involve human monitoring and control as well. I focus on wireless sensor and actuator networks as an example of cyber-physical systems, although it will be seen that WSANs are designed to be highly constrained, with limitations on power consumption, cost, transmission range, and processing capability. When applications allow these constraints to be relaxed a greater variety of device and network architectures also become suitable for electrical grid control. The future impact of these systems is expected to rival that of the Internet:

> *Cyber-physical systems will transform how we interact with the physical world just like the Internet transformed how we interact with one another.* [57]

Of the earth's 7 billion people [58] approximately 5.9 billion have mobile phones [59]. The total population places something of an upper bound on the maximum number of person-to-person (P2P) communications devices, while machine-to-machine (M2M) communication devices are estimated to grow to over 50 billion in the next decade [60]. For Internet Protocol based communications, the IPv6 address space permits 3.4×10^{38} addresses, enough for 4.8×10^{28} per person, or approximately 1,500 addresses per square foot of the earth's surface, even if conservatively allocated [61]. M2M communication is essential for critical infrastructure protection, and for development of the smart grid, with networked monitoring and control incorporated into generation, transmission, substation, distribution, and consumption activities. Each of these components of the electrical grid have unique communication requirements from a

network design perspective, with a need for different throughputs, bandwidth, latency, reliability, quality of service, cost, energy consumption, and other metrics unique to the application.

Wireless sensor networks have applications throughout all functional areas of the electrical grid, although in their typical implementation they are not necessarily suited for all tasks. These networks also have application beyond the electrical grid, with the potential to monitor or control virtually any state or process for which a sensor or actuator can be constructed, although the discussion of such potential outside of the electrical grid will not be considered here. WSANs have found early adoption in the smart grid ecosystem as part of advanced metering infrastructure (AMI) installations. Commercial and industrial buildings and processes have used sensor and actuator networks, often wired, for decades for everything from fire alarm systems to energy management, with varying levels of integration among different systems at each facility. Advances in wireless communications, protocols, standards, processing power, cost reduction, and a variety of other factors are permitting an explosion of WSAN applications and interoperation which will dwarf all previous efforts. Desirable smart grid objectives such as demand reduction and response, load shifting and shedding, and consumer empowerment can all be enabled through the use of WSANs, although a focus on critical infrastructure protection is arguably a higher priority and more readily achievable in the near term. The growing number of embedded and networked consumer devices can also be included in these networks, and functionality equal to and exceeding that long utilized in commercial and industrial applications is becoming available on a smaller scale for so-called building area networks and home area networks (HANs).

A variety of networking technologies, both wired and wireless, are suitable for use throughout the electrical grid. Among the wireless options, microwave technologies have long

been used for intermediate to long-range applications. Medium range options include WiMAX (IEEE 802.16), and cellular communications using general packet radio service (GPRS), short message service (SMS), and 3G and LTE voice and data. Local area networks using WiFi (IEEE 802.11) have entered the consumer lexicon, and personal area networks (PANs) using Bluetooth (IEEE 802.15.1) and low-rate PANs using IEEE 802.15.4 and ZigBee are becoming more common, although there are other competing standards as well. The future is likely to include body area networks which have an even smaller range and their own specialized functions.

While originally termed Personal Area Networks due to their relatively short range, devices using protocols such as IEEE 802.15.4 and ZigBee have found application in fields as diverse as health care and building automation. Because of this, some authors have suggested the term Premise Area Networks would be more accurate [62]. This document highlights an application of ZigBee-based WSANs for transmission and distribution line monitoring, another example of a use well outside the personal networking space.

5.1 Wireless Sensor Actuator Networks

WSANs are networks composed of self-organizing nodes which consist of sensor(s), a processor, memory, a wireless radio, and a power source. These nodes are often referred to as *motes*. Each node or mote is intended to serve as a bridge between the physical and electronic worlds. Node communications are generally organized as an ad-hoc mesh. One of the nodes, operating on permanent power, is responsible for additional functionality in the network, such as coordination or processing and storing messages from the end nodes. A gateway serves to transfer information from the mesh nodes to the access network, usually the Internet. From here data may be sent for additional processing, storage, analysis, or response. Working in reverse, control messages may be sent through the access network, gateway, and coordinator node,

eventually reaching the end nodes which have actuator capability in addition to sensing. Motes are designed to consume very little power, usually by spending most of their time asleep. This allows them to be powered for years at a time by just batteries. Due to the extreme power constraints, WSANs are only capable of sending data over short distances and at low bit rates. Motes using the ZigBee standard, for instance, are limited to about 100 meters and 250 kbps under the best of conditions. A more complete description of the components appears below.

5.1.1 Generic Mote Architecture

Motes are relatively small; the size of a shoebox on the large end and targeted to be as small as a speck of dust on the small end, although no fully functional motes have achieved this nano-scale yet. As motes are typically designed to be deployed in large numbers, it is important that the cost per unit be small.[11] Due to the need for small size, low cost, and most importantly low power consumption, the capabilities of the components are extremely constrained. The physical architecture may depend on the application. A single-chip solution may contain a system-on-a-chip radio and a microcontroller with processor and memory in one device, offering a highly-optimized solution which minimizes cost, power requirements, and size. A two-chip solution could contain a separate microcontroller and transceiver, offering more flexibility or performance. Sensors may be built-in or coupled separately. Minimizing power consumption is an active area of research, and includes not only the physical components but also networking and routing protocols. The mote components form five functional areas:

Processor. This is the core element of the mote, responsible for processing the sensor or actuator data. The processer spends most of its time asleep, but may be active when sending,

[11] During a presentation at the International Symposium on Advanced Radio Technologies (ISART) [63] a target figure of $2 to $4 per mote was quoted, although it is unlikely that any devices in this price range currently exist. For the electric grid, the sensor or associated equipment such as transformers which permit monitoring of line voltage levels could far exceed the cost of the remaining mote components.

receiving, or processing data, or idle when awake but not processing. WSAN processors have very limited functionality.

Sensor(s). Motes may have one or more sensors, and the sensors available obviously determine the application for which the device may be used and are subject to the constraints noted. Sending video, for instance, is not practical due to the low bandwidth available. The "pins" used for sensor input usually may also be configured for output, so that control signals may be delivered to a component instead of simply receiving sensor data.

Memory. Used to store program instructions and data from the sensors or processor. Again due to power, cost, and size constraints, memory capacity is very limited.

Transceiver. A low power, low data rate, short range radio. Like the processor, it must spend most of its time asleep to conserve power as receiving, or worse yet sending, data is among the most power-intensive tasks performed by the mote.

Power. Usually a battery which must last for a long period, up to several years. Other sources of power may also be used, such as solar panels or more novel methods such as energy harvesting from the surrounding environment [36].

5.1.2 Gateways and Beyond

Most current implementations of WSANs are not based on the Internet Protocol. This creates certain challenges because the tools used to manage enterprise networks cannot be used to manage WSANs, and standard enterprise networking equipment like switches and routers cannot be directly used in the WSAN. WSANs are designed to form their own mesh networks, although this does not preclude other architectures such as a star topology or point-to-point link. Eventually all electrical grid communications will need to send their data to an access or

transport network such as the Internet, a cellular service, or an enterprise network. In IP networking the device which translates between different networks is usually called a router. Similar functionality is required for a WSAN, although the term gateway is often used because the device must not only translate between different networks, but usually also different protocols in a manner which goes beyond simply encapsulating a lower-layer protocol with networking or transport protocols.

Gateways are a problematic device because they are designed to work with only a certain set of protocols, or with a manufacturer's proprietary equipment. They must be permanently powered, and unlike low-cost motes, may cost several hundred dollars. They are also a single point of failure, and require learning another skill set to manage. With the evolution toward IPv6-based WSANs such as detailed in the Internet Engineering Task Force's (IETFs) 6LoWPAN standard, or the forthcoming ZigBee IP, gateways will function more like standard enterprise wireless access points or routers. Even with the move toward IPv6-based WSANs, however, some form of wireless gateway or router will be required to route traffic between the WSAN and the access or transport network.

Once data has left the WSAN and passed through the gateway there is still a lot that must be done for it to be useful, but these are well-defined domains that will not be discussed in detail here. The data may need to move locally on an Ethernet network, for instance, or be sent over the public Internet via TCP/IP. An application must present the data in a format where it can be read and interpreted by human or automated operators, or an application must exist which can send a control or actuator signal back to the mote. These processes are not unique to the electrical grid or critical infrastructure protection, although the issues of incorporating WSAN

data into a utility operations center or automated control system are neither trivial nor beyond the scope of work that would be needed to develop a complete end-to-end system.

5.2 Current Options for Constructing a Wireless Sensor Actuator Network

The general mote architecture previously described could be realized via a wide variety of possible hardware options, and there exist a number of standards distinguished primarily by the method of radio communication. Other characteristics addressed by a standard could include the operating environment, temperature, humidity, electromagnetic compatibility, interoperability, and coexistence with other systems [64]. DASH7, Z-Wave, and Wireless HART are common standards which differ in their intended end use and radio frequency band, although all common WSAN standards typically use unlicensed industrial, scientific, and medical (ISM) frequencies. The most common standard currently used is ZigBee, which, along with Wireless HART, ISA100, and 6LoWPAN, is used on devices which also depend upon the IEEE 802.15.4 standard. As this section is intended to demonstrate the capabilities and function of a WSAN, rather than serve as a comprehensive comparison of protocols, the following explanations focus on IEEE 802.15.4, ZigBee, and 6LoWPAN as examples.

5.2.1 IEEE 802.15.4

The IEEE first released the 802.15.4 standard for low rate wireless personal area networks in 2003, although it is used for many applications beyond what the name personal area network might indicate. The latest version is IEEE 802.15.4 – 2011 [65]. The standard defines the link-layer (MAC) and physical (PHY) layers for the radio and has the following characteristics.

TABLE 5-1. IEEE 802.15.4 STANDARD RADIO CHARACTERISTICS

Region	Frequency	Channel Numbers	Bit Rate	Modulation
Europe	868 MHz	0	20 kbps	BPSK
United States	902 to 928 MHz	1 to 10	40 kbps	BPSK
World	2400 to 2483.5 MHz	11 to 26	250 kbps	O-QPSK

Channel access is controlled using carrier sense multiple access with collision avoidance (CSMA-CA) in a beaconless mode. A beacon-enabled mode is available as an option which can reserve time slots for high priority data, but it is not commonly used. For the lower frequencies binary phase shift keying (BPSK) is used as the modulation scheme, and for the 2.4 GHz frequencies offset quadrature phase shift keying (O-QPSK) is used.[12] Communication may be via a broadcast message, which is sent once, or via unicast, where three retries are allowed after the initial attempt if no acknowledgement (ACK) is received. The use of ACKs is optional, and many of the higher-layer protocols which run over IEEE 802.15.4 implement their own acknowledgement schemes. IEEE 802.15.4 networks can use a globally-unique 64-bit address (which includes the hardware MAC address) for each node, or a shorter 16-bit address which is valid only within one personal area network (PAN) to reduce overhead in smaller networks. There are four frame types: data frames, acknowledgement frames, MAC layer command frames, and beacon frames, although the latter three are optional. An example of the frame format is shown in Figure 5-1. The physical frames may be up to 127 bytes in size, with the actual payload varying between 72 and 116 bytes depending upon the addressing and security options enabled. Optional Layer 2 security is provided via 128-bit AES encryption.

[12] Understanding the different modulation schemes is not important to the theme of this paper, but a nice explanation is available in Appendix D from this report [66], and helps to explain the gains in throughput at the higher frequencies.

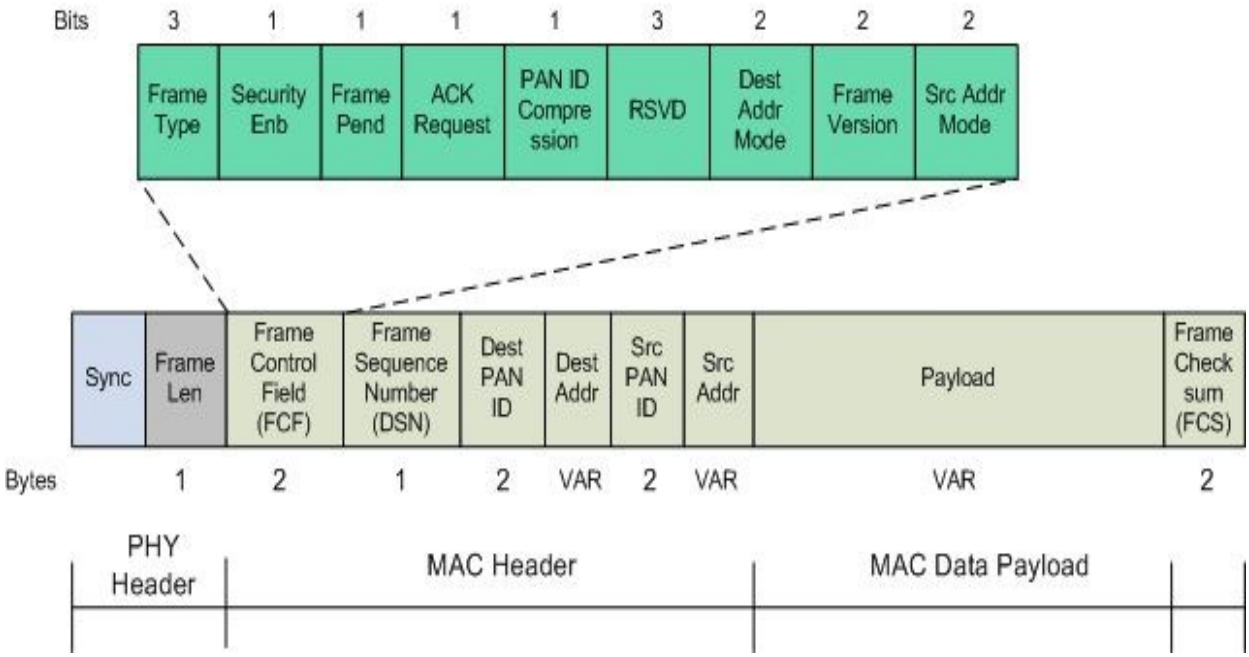


Figure 5-1 General IEEE 802.15.4 frame format. [67]

5.2.2 ZigBee

ZigBee is a proprietary protocol [68] owned by the ZigBee Alliance, and is intended for use over IEEE 802.15.4 based radios. While the ZigBee protocol stack resembles the layers in the OSI model, the ZigBee protocol is designed to work with the MAC and PHY layers of IEEE 802.15.4 exclusively, with the upper layers of the protocol stack relying on features such as addressing from the IEEE 802.15.4 standard. While ZigBee uses the MAC and PHY layer from IEEE 802.15.4, it does not permit the option for beacon mode, and therefore relies exclusively on CSMA/CA.

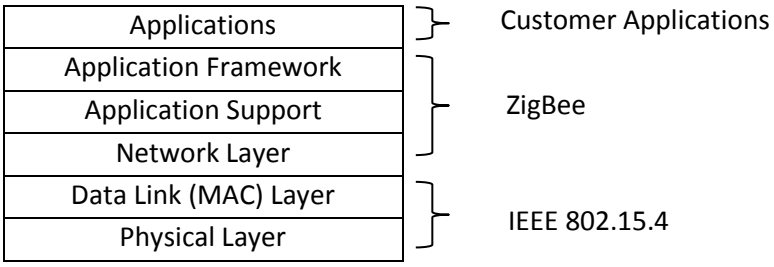


Figure 5-2 ZigBee protocol stack.

Applications. Vendor-specific applications may be written which take advantage of specific application profiles and the lower layers of the protocol stack.

Application Framework. The ZigBee protocol allows for profiles, which are capable of supporting specific services such as Smart Energy or Home Automation. Profiles are managed by the ZigBee Alliance, with public profiles being interoperable across all vendors, and vendor-specific profiles being available only on the devices of that manufacturer. The application framework looks for a registered end point identifier in the packet which determines the application profile to which the packet should be passed. If no registered identifier is found, the packet is dropped.

Application Support. The APS is a layer or sublayer that roughly corresponds to Layer 4 in the OSI model, and provides transport services such as filtering duplicate packets and providing acknowledgements.

Network Layer. The network layer is similar to layer 3 in the OSI model, and is responsible for addressing and routing. Packets may be delivered via broadcast or unicast and there is a field which permits the maximum hop count to be specified, although the protocol limits packets to a maximum of 30 hops. Routing is accomplished via the Ad hoc On-demand Distance Vector (AODV) standard [69], a reactive routing protocol which does not determine the node locations

or establish the route until a packet must be sent. At that time a routing request packet is broadcast to all nodes on the network, and each node records the address of the originating node and the node from which it received the routing request packet. When the routing request packet reaches the intended destination, the destination node sends a unicast route reply, and the intermediate nodes record the address of the original destination node and the node from which they received the route reply. When the route reply packet reaches the originating node the route is established and regular communications commence.

The ZigBee devices serve as wireless radios or modems to create a wireless PAN with point-to-point, tree, star, or mesh topology potential.

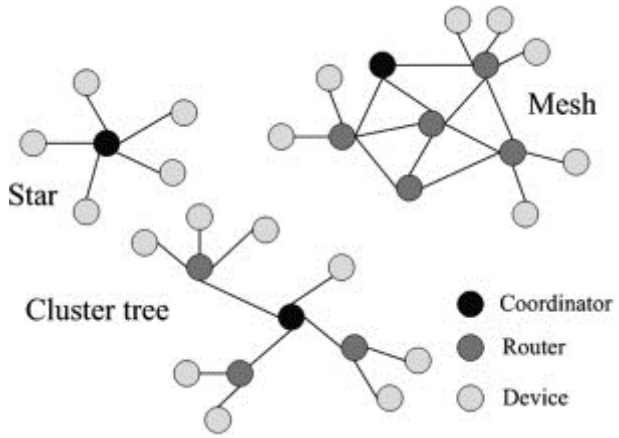


Figure 5-3 Example ZigBee device topologies. [70]

As shown in the topology diagram above, ZigBee networks are composed of three kinds of ZigBee devices. End devices are designed to sleep as often as possible, cannot route data, cannot be used by other devices to join the network, must join a PAN before sending or receiving any data, and must have a parent device to store and route messages for it. ZigBee routers can route data, can allow other end devices and routers to join the network, must join a PAN before they can send, receive, or route data, and can buffer messages for sleeping child devices. Coordinator

devices serve similar functions to routers, but is responsible for selecting the appropriate RF channel and PAN ID for the network [71]. Each network must have exactly one coordinator, and can be configured with all routers and no end devices, or up to ten end devices per router or coordinator. Each coordinator can be responsible for a network of up to 65,535 nodes.

5.2.2.1 ZigBee IP

ZigBee addressing is based upon the Layer 2 MAC address and does not map directly to the IP protocol stack. Because of this it is necessary to use a gateway to translate the ZigBee protocol to IP-based protocols. In addition to the previously mentioned issues related to gateways, one drawback of this arrangement is that traditional IP network management tools based upon the Internet Control Message Protocol (ICMP) do not work in the ZigBee network. Some individual manufacturers have introduced their own remote network management suites, but these involve using a model very different from the open standards of an IP network.

In order to more readily integrate with IP based networks and take advantage of the tools and solutions which are available, there is a trend toward IP-based sensor networks. While not yet available, the ZigBee Alliance is currently working on ZigBee IP.

5.2.3 6LoWPAN

6LoWPAN is an open standard [72] created by the IETF to enable IPv6 addressing on IEEE 802.15.4 devices in highly constrained networks. It is used instead of ZigBee, and while some devices which use the IEEE 802.15.4/6LoWPAN combination are currently available, they are more complicated and expensive than ZigBee devices. This is likely to change as demand increases. Because the protocol runs on IEEE 802.15.4 hardware it is subject to the limits on power, throughput, and the 127 byte maximum transmission unit (MTU) specified in the IEEE

standard. The IPv6 standard specifies a 40 byte header, and requires the ability for all links to carry 1280 byte packets. Because the data payload in sensor networks is often small, or there is the potential for significant fragmentation if a full-size IPv6 packet must be transmitted, there was a desire to create a standard in which the header overhead does not occupy such a significant portion of the available payload. This led to the creation of the 6LoWPAN standard, in which an adaptation layer effectively compresses the IPv6 header. This is done by removing any unnecessary or redundant information from the IP header and inferring certain information from the link header. The link local address, for instance, may be created using the MAC address derived from the 802.15.4 header. Following the underlying IPv6 protocol, the header scales to include only the information needed, and unnecessary fields do not require any space as they are simply eliminated. The IPv6-required support of 1280 byte packets is done through fragmentation and reassembly, with a much smaller header overhead per fragment. Unlike ZigBee devices, each node functions as both an end point and a router, with the transmitting node sending "wake up" packets to a sleeping receiver node prior to transmission.

The figures below, from [73], depict an example of header compression. In this example, the version field is omitted because it is always 6 for IPv6, the traffic class and flow value are omitted because they are empty, the length field is omitted because the length can be inferred by subtracting the length of the IPv6 header from the length of the IEEE 802.15.4 length field, the next header is UDP or TCP so this 8-bit field can be reduced to 2-bits, and the IPv6 address can be determined based upon the 64 bit address used in IEEE 802.15.4 [74].

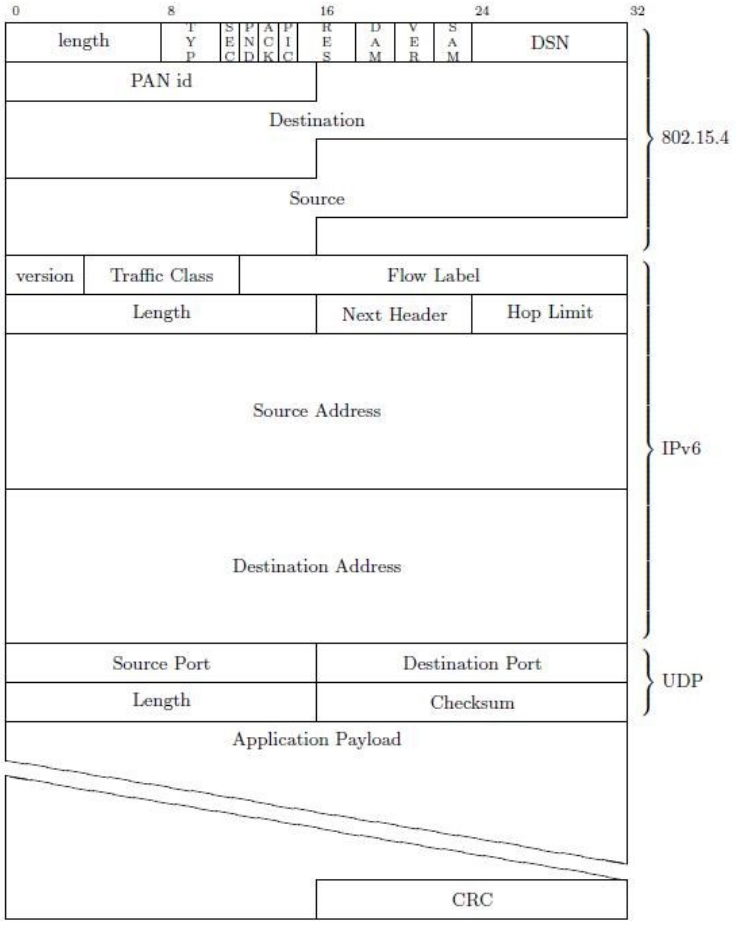


Figure 5-4 Uncompressed IEEE 802.15.4, IPv6, and UDP headers [73].

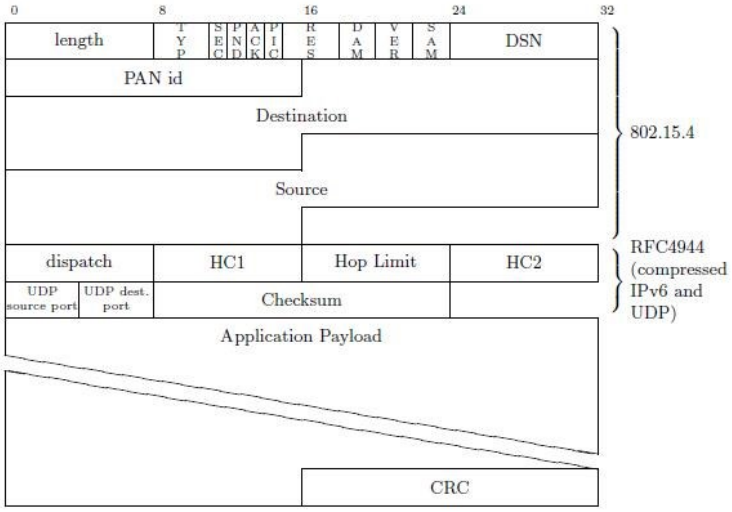


Figure 5-5 Compressed IEEE 802.15.4, IPv6, and UDP headers [73].

5.2.4 Connecting Sensors and Actuators

Virtually anything for which a sensor or actuator can be built can be controlled via a WSAN. Sensors permit mapping of physical characteristics of the environment to quantitative measurements [23]. The exact mechanics of how the sensor or actuator is coupled to the system is more of a vendor issue than a conceptual one, but a general description is given here. A sensor is any device which reads some state, quantity, or event, such as breaker opening, and provides an output signal that is converted to a digital quantity, read into memory and processed, and has an appropriate output passed to the transceiver. The reverse process may also take place, with the output from the processor being sent to an actuator, which may be operate a switch, motor speed control, etc. When the sensor reads data it is converted to a voltage. In the case of sensors which work with the Arduino logic board and XBee brand ZigBee devices which were used in the experiment which follows, the input must be between the range of 0 and 1.2 volts. For many devices this will require building a voltage divider circuit which reduces the input voltage to the appropriate range. Inputs and outputs may be Boolean digital (on/off or high/low), analog-to-

digital quantized (one of 256 discrete levels with the equipment noted), or pulse-width modulated. Pulse width modulation changes the amount of time a digital pulse is on or off, simulating an analog value which permits speed control of a motor or dimming of lights. Outputs in the test setup which follows are limited to 5 volts, which again may require transformation to control devices which take higher inputs. While the sensor may send data continuously, only certain events, levels, or sampling intervals need be passed to the transceiver.

Data given to and sent by the transceiver is received by a gateway, either directly or after traversing other WSAN nodes. From here it may go over a corporate network or the Internet until it becomes an input for a display or application. The example below, for instance, shows power output for a wind turbine sent to the free online data logging and display service Cosm [75]. For electric utility applications the data would ultimately be incorporated into the utility network operations center.

Figure 5-6 Power output from a wind turbine displayed by Cosm.

5.2.5 Relaxing the Constraints

The systems detailed above represent the current mainstream for highly constrained devices, but a variety of other alternatives could be considered in situations where one or more of the limitations do not apply. Where permanent power is available, for instance, sensors could be coupled to an IEEE 802.11 device. While consuming approximately one hundred times the power of an IEEE 802.15.4 device, considerable increases in throughput and even range are possible. Non-traditional sources of power, such as the scavenging of electromagnetic energy from high voltage lines could permit the use of permanently-powered devices. Another alternative is low-power WiFi, such as in systems made by GainSpan Corporation, which use only approximately one tenth the power of traditional WiFi. While more expensive than IEEE 802.15.4 devices, cost may not be a significant factor when taken together with the cost of the sensor, or weighed against the cost of the equipment being protected or the cost of an outage or successful sabotage. WSANs are also conceived to be self-organizing and deployable in large numbers, but if large numbers of devices are not required than nodes may be custom-configured, or outfitted with directional antennas to increase range or reduce interference. Microcontrollers may be upgraded to embedded microcomputers, which are still relatively low cost and low power, with the advantages of additional processing power or memory, which allows pre-processing of sensor data and puts greater intelligence at the edge, rather than the core, of the network.

6 Concept for a WSAN Application for Transmission and Distribution

In [76] Yang et al. demonstrate construction of a prototype sensor module for transmission and distribution line monitoring using off-the-shelf components. This module represents the creation of an actual device of the type proposed conceptually in their earlier paper [31]. The sensor is capable of monitoring line current, ambient temperature, and conductor temperature, and is intended to provide parameters which allow dynamic line rating. No cost for the prototype unit is given, but the author's state they believe the units can be produced to sell in the $100 to $200 range. The unit is powered via energy scavenging. Expanding on the block diagram from their earlier paper, the authors provide the following schematic, as well as photos:

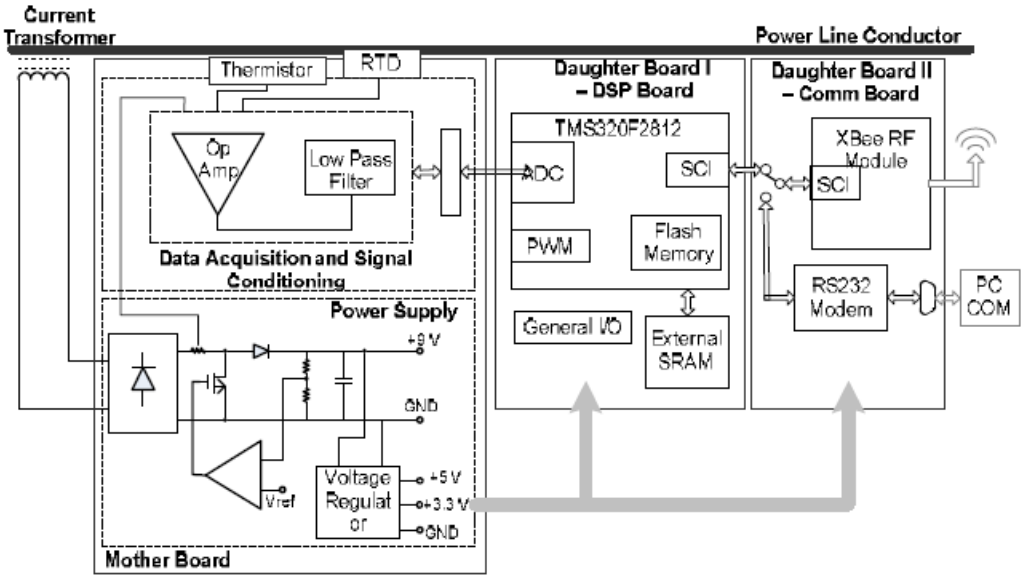


Figure 6-1 Power line sensor module schematic from Yang et al [76].

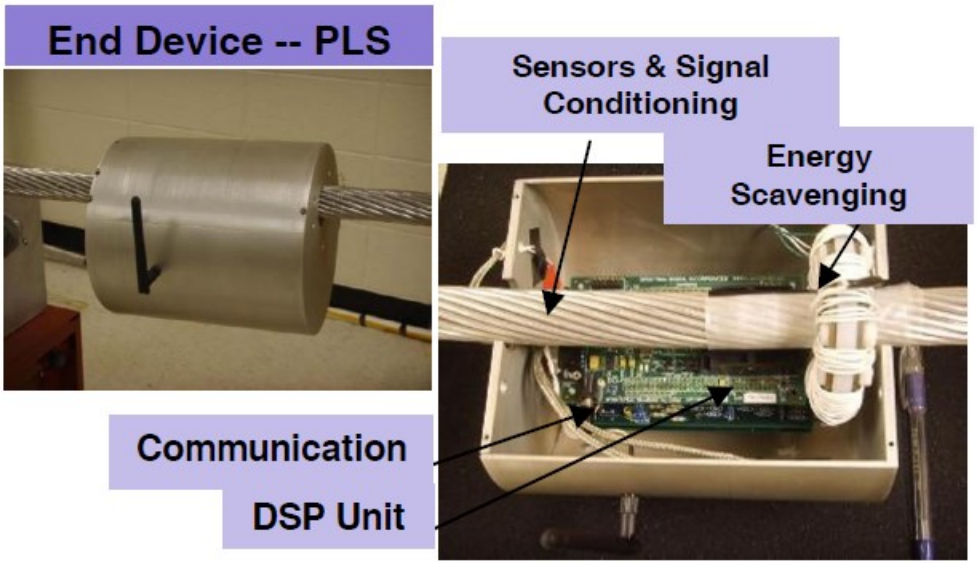


Figure 6-2 Prototype power line sensor module from Yang et al.

As can be seen from the schematic, communications are via ZigBee using Digi International's XBee Pro module. The authors reproduce the manufacturer's specification for the device, but as in the previous paper, they focus primarily on the sensor, not the communications system. The authors state they performed a laboratory experiment where they sent a data stream from a PC, looped it through the ZigBee device in the power line sensing module and received it back at the computer. They performed an indoor test with the unit attached to a test conductor flowing 500 and then 1000 amperes, and outdoors without attaching the sensor node to an electrical conductor. They measured received signal strength (RSSI) and percentage of successful receiving (PSR). The results were as follows, and indicate that they were not influenced by the amount of current in the conductor, but were impacted dramatically by the obstructions and environment.

TABLE 6-1. ZIGBEE PERFORMANCE TEST (INDOOR) FROM YANG ET AL.

Current (A)	Range (m)	RSSI (dBm)	PSR
1000	50	-70 to ~ -75	100%
	100	~ -92	~ 45%
500	10	-45 to ~ -50	100%
	50	~ -73	100%
	100	~ -93	~ 45%

TABLE 6-2. ZIGBEE PERFORMANCE TEST (OUTDOOR) FROM YANG ET AL.

Range (m)	RSSI (dBm)	PSR	Conditions
200	~ -76	~ 95%	Close to line of sight
400	~ -83	~ 80%	Trees
500	-92	~ 35%	Trees & Buildings

While considerable work was obviously done on the design and construction of the sensing module, a great deal more could be undertaken to characterize the performance of the communications system. There are also no performance requirements stated for power line monitoring against which communications performance can currently be compared, such as metrics for a sampling interval, size of the data set, throughput, or tolerance to latency and quality of service issues. The work of Yang et al. does, however, demonstrate the feasibility of constructing a sensor which allows better monitoring of transmission and distribution lines, and permits higher and more efficient power flows via the ability to do dynamic line rating. The lack of a practical, affordable sensor previously served as the primary obstacle to performing this level of monitoring.

7 Laboratory Testing of ZigBee Robustness and Throughput

In this chapter we document throughput testing and a "failover" test using ZigBee communications. While these do not characterize all performance attributes of the system, they do provide insight into what the capabilities for reliable communication are, and could serve as a basis for comparing what the system is capable of versus what is required for the power line monitoring application.

7.1 Objective

The remainder of this chapter discusses the use of empirical testing to determine characteristics of ZigBee networks. Throughput was tested in a network with one, two, and three hops using the default AODV mesh routing protocol. Additionally, a "failover" test was performed to observe the effects of losing a node in the established route. Testing is considered within the context of applications which might be employed within smart energy networks.

7.2 Scope

For purposes of this project, I will investigate an implementation of a PAN using IEEE 802.15.4 and ZigBee. IEEE 802.15.4 is a physical layer and media access control standard [77]. ZigBee is an extension to this protocol which adds networking options, security, and an application framework [68]. Two specific features of the ZigBee protocol significant to the tests conducted are the ability to construct a self-healing ad hoc mesh network, and routing protocols. For purposes of simplicity, the combination of the IEEE 802.15.4 and ZigBee protocols will simply be referred to as ZigBee protocols for the remainder of this chapter unless it is necessary to make a specific distinction.

The ZigBee protocol is capable of implementing security features, and practices for securely incorporating these devices into control systems have been documented [78]. No security measures were enabled for this exercise, but it should be noted that they add additional overhead and therefore would reduce throughput below the values found as part of this lab. The ability for end devices to sleep as a power conserving measure is an important feature of most ZigBee networks, and devices are designed to wake quickly, on the order of 2 ms, which assists in power conservation. As all devices in the test network were required to perform routing functions, none were configured as end devices, and therefore the sleep feature was not enabled as part of the tests conducted. ZigBee devices are also capable of using a number of predefined "profiles." The ZigBee Smart Energy profile, for instance, adds support for features like metering support, and demand response and load control support [79] [80]. The standard ZigBee device profile was used for all testing, with no attempt to implement specialized profiles or features.

ZigBee, in conjunction with other devices, protocols, and applications, forms the basis for some of the most popular current implementations of HANs and advanced metering infrastructure (AMI) mesh networks, as well as seeing usage in a variety of utility, commercial, building, and industrial WSAN applications with smart grid integration potential. New standards continue to evolve, such as DASH7 [81] and Z-Wave [82], which may ultimately prove more popular than ZigBee, however performance testing of these alternatives was not undertaken.

A typical WSAN implementation involving sensors or actuators, microcontrollers, ZigBee devices, a gateway, and standard LAN and Internet technologies is shown in the figure below.

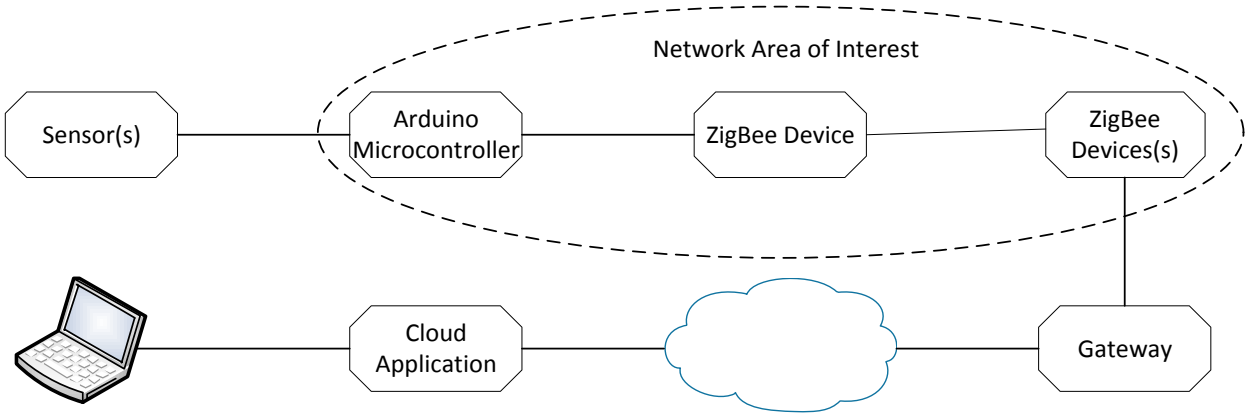


Figure 7-1 Example WSN network with ZigBee devices.

In this project I will examine primarily the ZigBee portion of the sample network shown, and based upon published specifications and laboratory testing attempt to characterize elements of the ZigBee protocol. While devices were physically arranged in a point-to-point fashion for our throughput testing, default mesh protocols were used. It is useful to think of the testbed as a mesh network in all respects, except to say that non-essential nodes were removed and those remaining simply represented the shortest path route. No attempt was made to conduct tests using other topologies.

Our testing was limited to four nodes by the equipment available in the lab facility, but as is shown by Conti and Giordano [83], the decay in throughput places a limit on the number of hops over which data can be practically transmitted, so the goal of designing a network with no more than three hops is not necessarily unrealistic. They conclude that only small to moderate scale ad hoc networks can be implemented efficiently. Devices in the 2.4 GHz frequency were used for all tests, and testing was conducted in an environment with good signal strength. Device proximity and power output were adjusted until error free frames were consistently sent,

and no attempt was made to investigate throughput as a function of low signal strength or radio frequency interference.

7.3 Previous Work and the Goals of This Experiment

In the classic work *Computer Networks: A Systems Approach* Peterson and Davie state that the two principal metrics of networking are throughput and delay [61]. ZigBee is designed as a low-power, intermittent low-data rate standard capable of operating in a variety of network topologies. It has a much lower latency than Bluetooth when recovering from the sleep state of the duty cycle, but is not designed as a real-time protocol. This makes the standard suitable for applications such as PANs, but unsuitable for real-time, high-throughput, low-latency uses such as synchrophasor[13] monitoring.

Previous work has characterized elements of ZigBee networks such as received signal strength indication (RSSI) and packet delay [84], and in [85] the authors attempt to use a combination of calculations, network simulation, and practical testing to determine throughput in a one-hop ZigBee network. I am not aware of any studies detailing throughput as a function of message size, one of the tests conducted here, or with ZigBee in a multi-hop mesh configuration using actual devices. Throughput is an important parameter because ZigBee end nodes spend up to 99% of their duty cycle sleeping, and in this interval a sensor and microcontroller memory may accumulate a large buffer of data. Alternatively, the ZigBee device may be configured for very short sleep intervals, or none at all, if a continuous stream of small messages must be relayed. Different traffic characteristics require a suitable protocol, and the traffic type and protocol may have large impacts on network performance and reliability. In a study titled

[13] A distribution automation device which takes real-time, synchronized power quality measurements from multiple remote points on the grid.

Communications Requirements of Smart Grid Technologies [86], the U.S. Department of Energy provides the following guidelines for required bandwidth for various smart grid applications:

Application	Network Requirements				
	Bandwidth	**Latency**	**Reliability**	**Security**	**Backup Power**
AMI	10-100 kbps/node, 500 kbps for backhaul	2-15 sec	99-99.99%	High	Not necessary
Demand Response	14kbps- 100 kbps per node/device	500 ms-several minutes	99-99.99%	High	Not necessary
Wide Area Situational Awareness	600-1500 kbps	20 ms-200 ms	99.999-99.9999%	High	24 hour supply
Distribution Energy Resources and Storage	9.6-56 kbps	20 ms-15 sec	99-99.99%	High	1 hour
Electric Transportation	9.6-56 kbps, 100 kbps is a good target	2 sec-5 min	99-99.99%	Relatively high	Not necessary
Distribution Grid Management	9.6-100 kbps	100 ms-2 sec	99-99.999%	High	24-72 hours

Figure 7-2 Communications requirements for smart grid.

Yang et al. do not provide any guidelines for the amount of data produced by their prototype powerline sensor, and in the general case the amount of data would depend on the type and quantity of sensors, the sampling interval, the amount of data preprocessing which could be done, and the monitoring application requirements at the network operations center. In the DOE chart above many of the throughput values fall in the range of 10 to 100 kbps per node. This indicates that a wide variety of useful, discrete sensor and actuator information can be sent at data rates between these bounds, although powerline monitoring is not directly addressed. By determining the performance of the communications system it would be possible to tune other parameters of the powerline sensor module and the network architecture to fit within the communication system capabilities. To reduce throughput, for instance, the sampling interval

could be increased, the number of data points reduced, or the system could be designed to send only values which are outside of a certain range, rather than all measurements.

ZigBee is designed to be low-throughput, with the protocol stating a maximum line rate of 250 kbps, which may give the impression that throughput is not an important consideration in ZigBee networks. The low bandwidth is in reality a function of power constraints designed to maximize battery life, and throughput becomes an increasingly important metric as the number of network nodes and congestion increase. Appropriate network design depends, in part, on knowing how the protocol will behave under different conditions.

In a widely cited paper titled *The Capacity of Wireless Networks* [87] the authors examine the effects of multiple hops on throughput in a wireless mesh network using the IEEE 802.11 protocol. As a transceiver operating on a single channel cannot receive and transmit at the same time, it can be seen that the transceiver will spend approximately half of its time receiving and half sending, thus intuitively the throughput can be at most half of the one-hop throughput. In the best-case scenario the nodes in a multi-hop mesh network are arranged in a linear fashion, and each radio can only hear the transmissions of its adjacent neighbor on either side. The authors expect throughput to be

$$T = c / \sqrt{n} \tag{5}$$

where T is the throughput, c is the maximum rate at which one node can send or receive data, and n is the total number of nodes. Where all nodes are co-located, the worst-case scenario requiring simple time division multiplexing, the authors expect the throughput to be

$$T = c / n \tag{6}$$

In [88] the same authors test their hypotheses experimentally in a variety of configurations, finding actual throughput to be

$$T = c / n^{1.68} \quad (7)$$

a result worse than either of the expected outcomes. All messages were one kilobyte in size, although the authors state it would be interesting to run the experiments with messages of different sizes, something I do in this experiment.

Throughput of the payload data, rather than raw throughput of the total number of bits, is also affected by the amount of protocol overhead. A ZigBee data frame is shown below. The data payload is variable, depending upon factors such as type of addressing used and security features enabled. The maximum data payload per frame for the configuration used in this lab was 84 bytes according to the manufacturer's manual [71]. This maximum size was confirmed by empirical testing, which showed that data payloads over 84 bytes were fragmented into multiple packets.

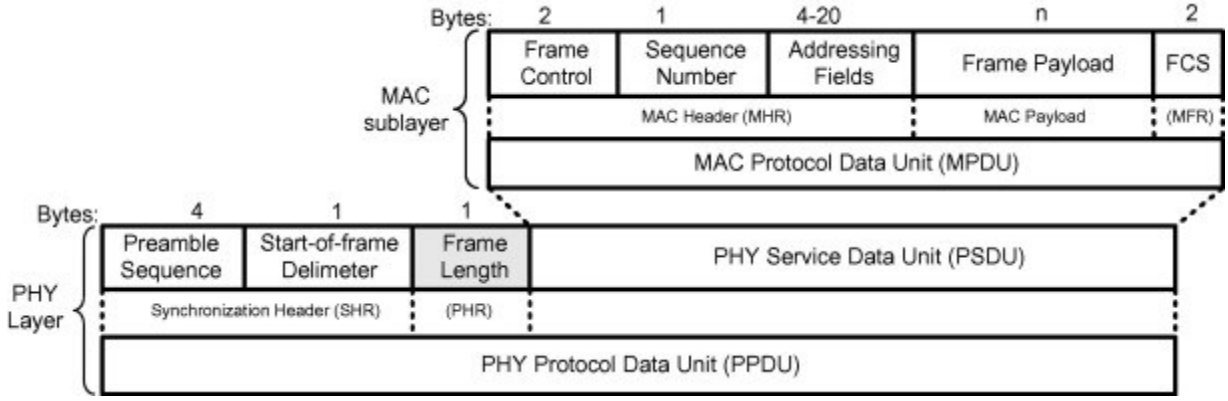


Figure 7-3 ZigBee data frame. [89]

The factors noted result in an expected throughput of the data payload which is substantially less than the stated protocol line rate of 250 kbps.

7.4 Hypothesis and Research Questions as the Basis for Testing

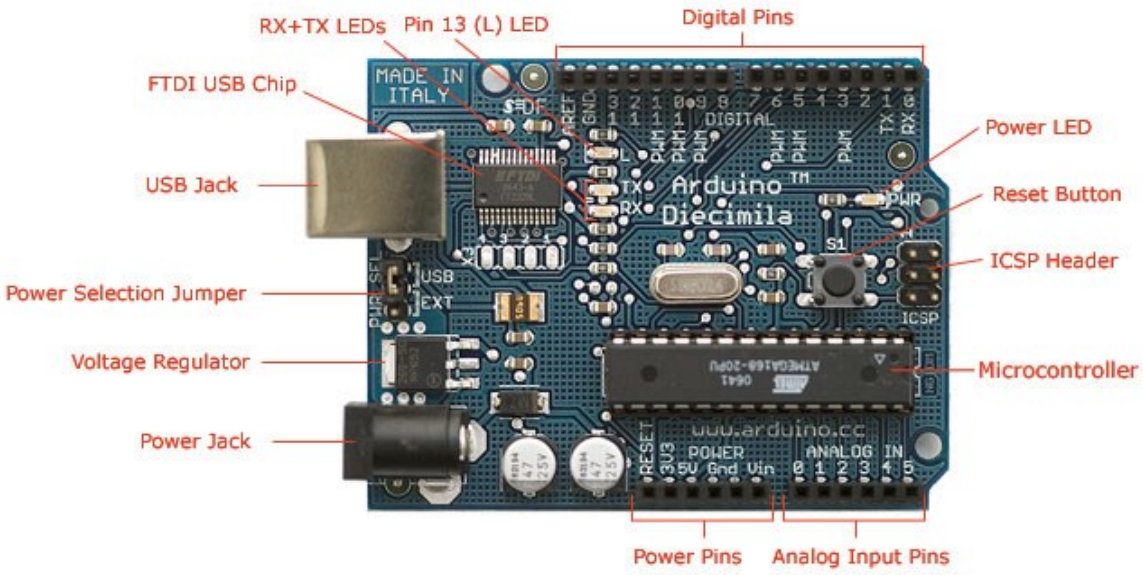
It is expected that due to overhead, the payload or data throughput will be less than the stated line rate, and I wish to discover a baseline value for maximum payload throughput. Due to the necessary overhead associated with each message, I anticipate throughput performance will be worse for data which is transmitted as a series of small messages, and that where possible it would be advantageous to buffer data so that it is sent in larger packets at longer intervals. I also anticipate that in multi-hop networks, throughput per hop will decline by a factor which lies somewhere between the limits noted previously. As a self-healing, ad hoc mesh protocol I also anticipate that the network will be robust in the event of disruption, a factor which would support using many inexpensive redundant devices instead of a single highly-reliable device when designing the network.

In the project I ask the following specific questions:

- For a single hop, what is the maximum tested, or practical, data payload throughput, and how does this compare with the published specification? How does this change as a function of message size?
- For a frame with a large data payload, expected to be near the theoretically highest throughput, how does throughput change as a function of the number of hops in the network?
- What are the effects of removing a device in the routing path actively relaying data when another device not currently in the routing path is available for failover?

7.5 Methodology

Small networks were constructed with open source Arduino microcontrollers [90] and Digi International's XBee ZigBee devices using mesh protocols.



Photograph by SparkFun Electronics. Used under the Creative Commons Attribution Share-Alike 3.0 license.

Figure 7-4 Arduino microcontroller. [91]

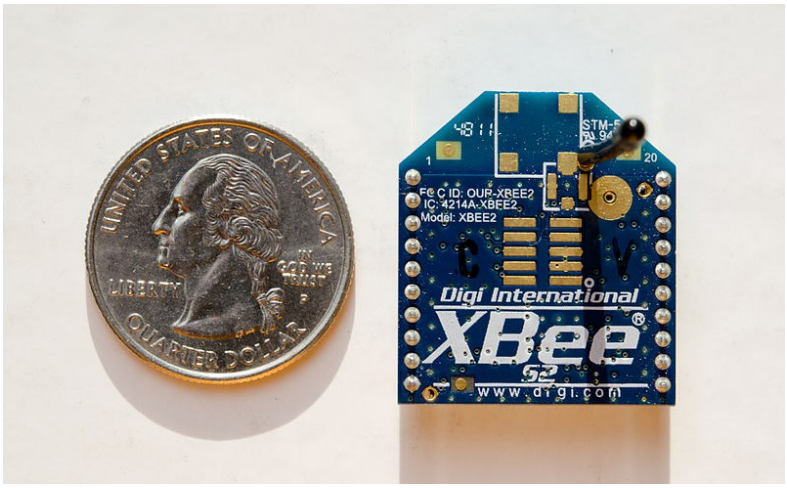


Figure 7-5 ZigBee module. [71]

As a hardware-intensive project it was of limited scale, with a maximum of four Arduino and ZigBee devices used to make a network with a maximum of three hops. The microcontrollers contained the processor and memory, and were used to send messages through the ZigBee network. These were received by a ZigBee device interfaced to a laptop.

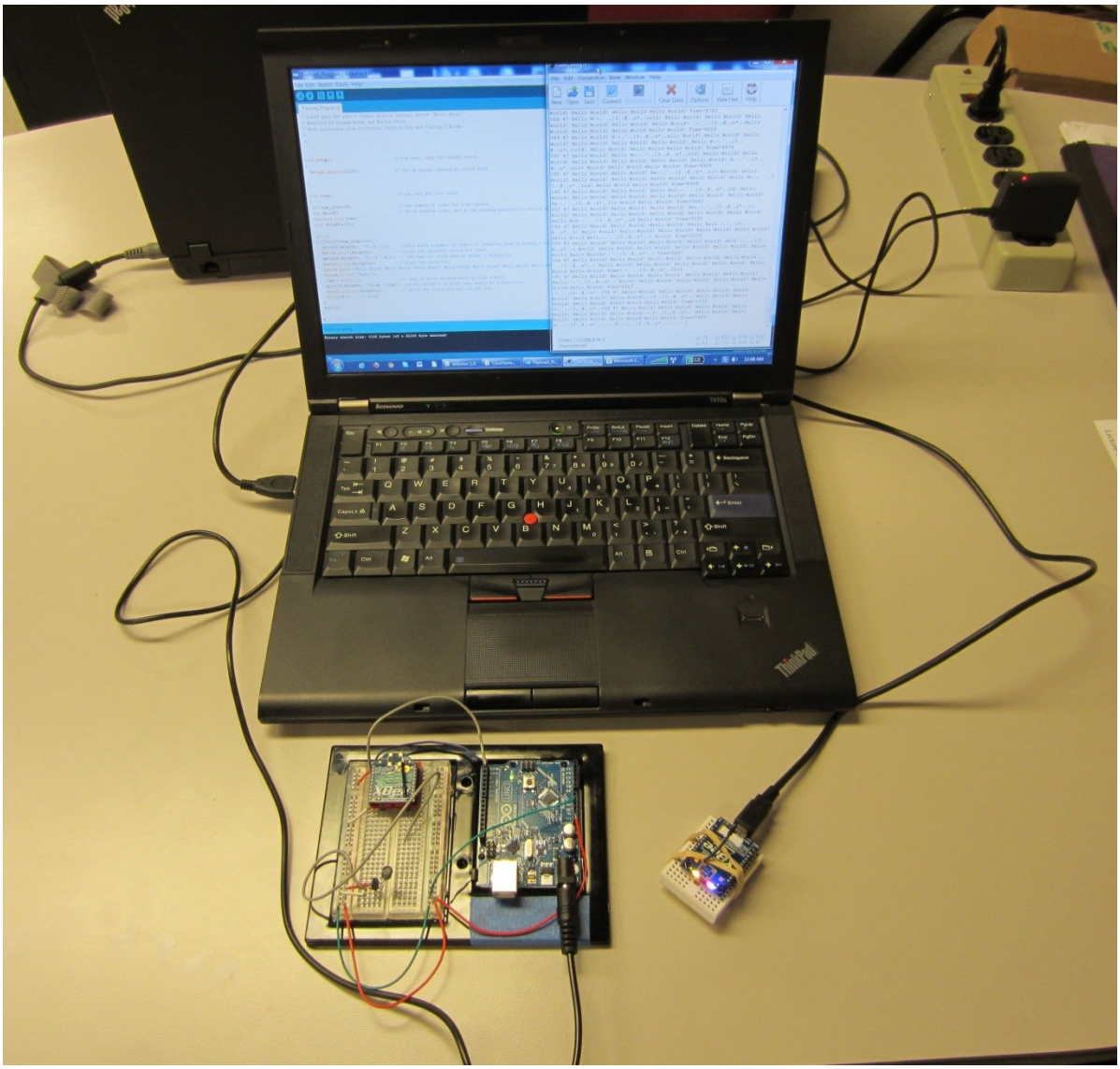


Figure 7-6 Sending node, and receiving node interfaced to laptop.

Sensors were not used, instead, a program running on the microcontroller sent data payload packets of variable, controlled size, which included a count of the number of packets and the

time sent. Each loop of the program included a delay which represented the processing and insertion time, and the delay was set to the minimum value possible to permit consistently reliable receipt of the data. See Appendix A for an example of uncorrupted received data, and Appendix B for an example of corrupted data received in the terminal. The minimum time was determined by testing a variety of values for a given amount of data and determining the lowest reliable value. This value was then used in a calculation which scaled the delay to other packet sizes, and was used as a starting point for empirical testing which determined the lowest reliable delay for the data payload being sent. Data was viewed in a terminal program on the laptop. Arduino "sketch" programs were written using the open source Arduino programming language based on the Wiring programming language, which is in turn based on the C programing language, and the Arduino development environment based on the Processing software language. The ZigBee devices were configured with appropriate firmware and parameters using the manufacturers X-CTU software interface (see Appendix C), Hayes AT modem commands from the terminal program, and a hardware interface module connected to the laptop via a USB cable.

For the first test a one hop network with an Arduino and ZigBee sending node and a ZigBee, interface module, and laptop receiving station were constructed. Maximum throughput was determined by sending a data payload of known size via a loop in the program which executed 99 times. The data received was examined for completeness and reliability, and throughput was determined by simply multiplying the payload size by the loop count and dividing by the total amount of time required to send the data. This test was repeated for payloads of 4, 8, 16, 32, 64, 84, and 160 bytes.

For the second test two intermediate router nodes were added to the above configuration. They were located close enough that throughput was not reduced due to the effects of low signal

strength, but were sufficiently far from each other that a node could not be "skipped" to complete the transmission of data. Locations were determined by varying the transmit power and using field adjustments to work with convenient physical obstructions. Throughput tests were conducted in a similar manner, but only using 64 byte data payloads.

The third test was conducted in a manner similar to the second, but the two intermediate nodes were placed adjacent to each other. It was determined which of the two adjacent nodes was originally routing the data, and this node was disconnected while relaying frames, with observations made about the co-located failover node and effects on the data transmission. No attempts to measure the rate of throughput were made during this test.

7.6 Equipment and Lab Setup

A complete list of lab equipment is provided in Appendix D, and significant components were noted above. Design and assembly work was required to interface the significant components and insure correct connections, configurations, and power requirements for the pins on each type of device. Small components such as LEDs were used as troubleshooting and activity indicators. Certain other small components such as pushbuttons and sensors were used to set up, test, and troubleshoot the connections in preparation for the tests conducted, and provide options for future additional work. In the interest of brevity, further details are not provided here.

7.6.1 First Setup – Maximum Throughput, One Hop

A diagram for the first setup, which tested one-hop throughput for a variety of packet sizes, is shown below. While the ZigBee devices communicated wirelessly, they were placed within a meter of each other for easy operation and to insure good signal characteristics.

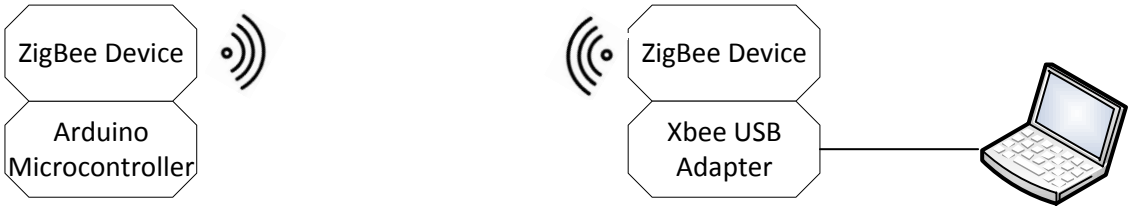


Figure 7-7 One hop throughput test setup diagram.

7.6.2 Second Setup – Maximum Throughput, Multiple Hops

The setup for the second series of tests was per the diagram above, with the addition of two intermediate ZigBee router nodes. The Receiving and Router Nodes were placed approximately 15 meters apart, with the sending node approximately half that distance away from its closest router as shown in the diagram below. As I only had four motes available, testing beyond three hops was not conducted.

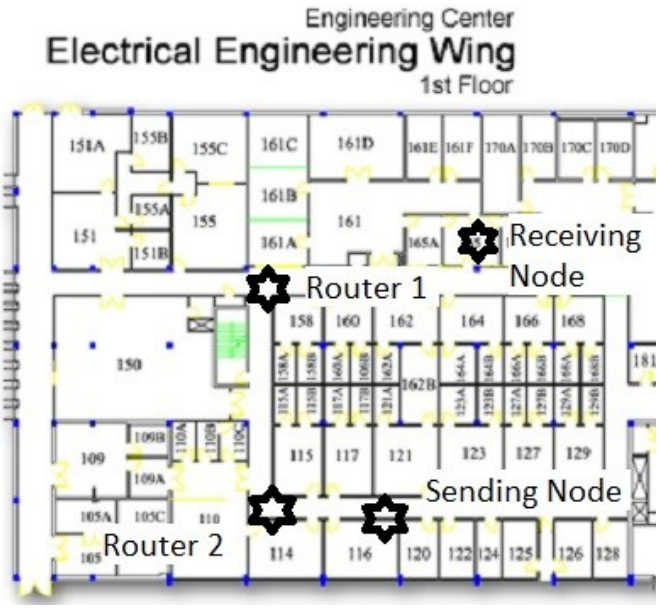


Figure 7-8 Three hop throughput location diagram.

7.6.3 Third Setup - Robustness

For the third test, a measure of robustness in the self-healing ad hoc mesh protocol, equipment was arranged as indicated below:

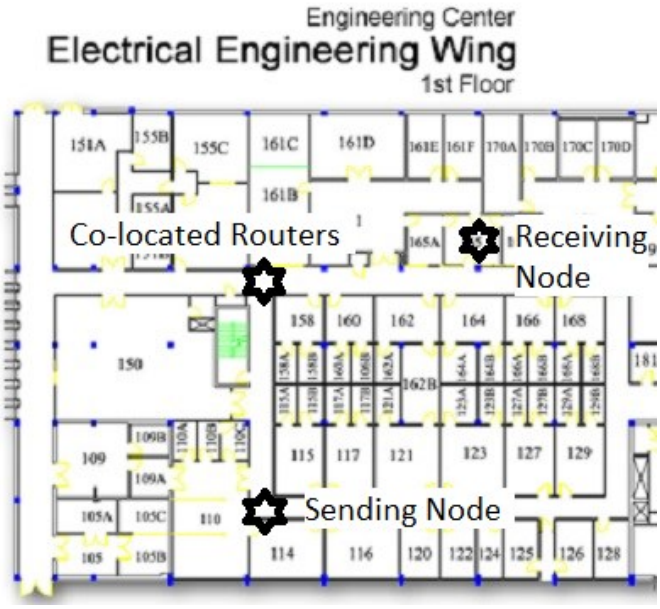


Figure 7-9 Robustness testing location diagram.

7.7 Results

7.7.1 First Test - Maximum Throughput, One Hop

For the test of maximum throughput at different data payload sizes the results are as below. The first test was done over one hundred times with the 64 byte payload and showed no variation in the total sending time. Subsequent tests at different payload sizes were performed ten times each, again with no variation in the total sending time. As such, no range bars are indicated for each point on the graph.

TABLE 7-1. THROUGHPUT BY PAYLOAD FOR ONE HOP

Payload per Loop (B)	Loop Delay (ms)	Total Payload (B)	Total Sending Time, 99 Frames (ms)	Throughput (kbps)
4	1	396	135	23.23
8	3	792	348	18.02
16	5	1584	578	21.70
32	10	3168	1120	22.40
64	19	6336	2099	23.91
84	25	8316	2737	24.06
160	47	15840	5428	23.11

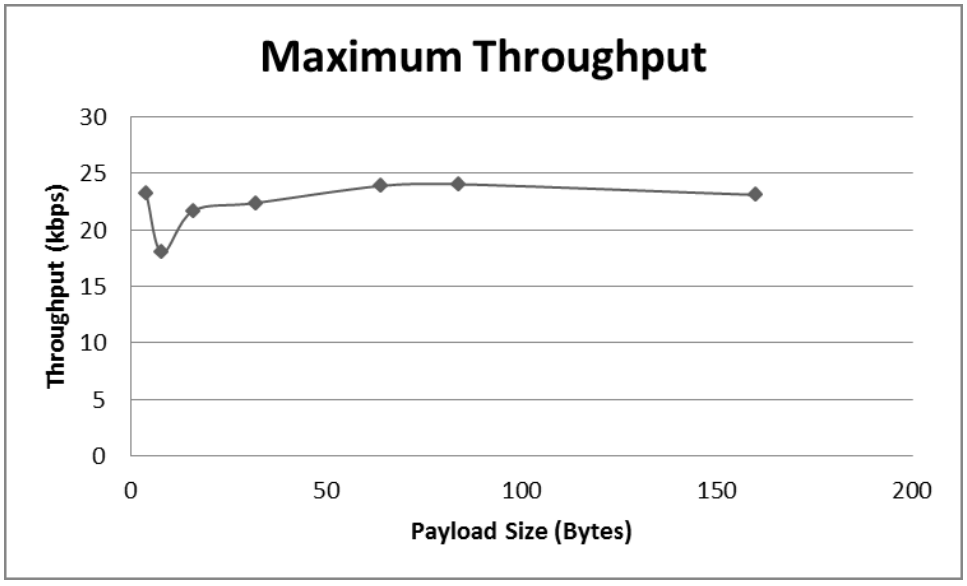


Figure 7-10. Graph of maximum throughput.

7.7.2 Second Test - Maximum Throughput, Multiple Hops

For the second test, the results are indicated below. Each test was repeated ten times, with no variation in the total sending time.

TABLE 7-2. THROUGHPUT FOR MULTIPLE HOPS

# Hops	Payload per Loop (B)	Loop Delay (ms)	Total Payload (B)	Total Sending Time, 99 Frames (ms)	Throughput (kbps)
2	64	40	6336	4169	12.04
3	64	72	6336	7305	6.87

7.7.3 Third Test - Robustness

For the third test, no throughput measurement as done. Instead, the behavior of the network and effect on received data was observed when the co-located intermediate node routing data was disconnected. As soon as the node was disconnected, I observed the following behavior:

- Five empty ZigBee frames are received at the terminal, followed by four data frames which continue the sequence transmitted by the sender. Following this I observed a drop of somewhere between 112 and 120 sequential frames.
- The co-located failover router node begins transmitting the data frames.
- The elapsed time during which the 112 to 120 frames are lost is between 4.7 to 5.6 seconds. Once the terminal begins receiving data again all subsequent frames are received reliably in correct sequential order.

7.8 Analysis

Yang et al. propose a prototype power line sensing module containing a ZigBee transceiver, but little information is provided about the performance of the communications system or the level of performance needed to support the application. The laboratory testing shown here provides some insight into what is possible.

7.8.1 First Test

In the results for the first test throughput for the 4 byte payload appears uncharacteristically high. This is explained by the "frame assembly timeout" configuration option for the XBee module. In the case of 4 byte payloads, even with the frame assembly timeout set to the minimum, three program "loops" worth of data were carried by a single

protocol frame. This effectively reduced the amount of overhead for the payload, and therefore permitted an increase in payload throughput.

Beginning with the 8 byte payload, throughput increases up until the maximum permissible frame payload of 84 bytes. This is expected as the amount of overhead per frame remains fixed, but the amount of payload per frame increases, so the effective payload throughput increases. Beyond 84 bytes the payload must be fragmented into multiple frames and reassembled at its destination, reducing the effective payload throughput.

In the first test I establish the maximum throughput for the data payload, a value found to be just over 24 kbps with the maximum payload possible without fragmentation and reassembly. This value is less than 10% of the IEEE 802.15.4 stated line rate of 250 kbps, an important consideration in designing the communications system for the sensor application. It may be necessary to aggregate fewer devices and provide more gateways, for instance, do more preprocessing of the data in the logic board such that only exceptions are reported, or design an application that only requires intermittent data rather than a continuous stream.

When compared to the bandwidth requirements given in the smart grid communications chart, 24 kbps falls within the range of 10 to 100 kbps shown for many applications, but indicates that significant optimization could be required for effective performance of multiple devices on the system.

7.8.2 Second Test

The second test indicates that throughput falls by close to half at each transceiver node. This is consistent with conventional thinking on the topic, and most closely matches the $T = c / n$ model proposed by the authors in [87] and [88]. In this case c is the maximum payload

throughput of 24 kbps, not 250 kbps, so the implications noted above apply here as well. If the x-axis is drawn logarithmically and a straight line is fit to the data points than the slope (s) of the line may be used to estimate throughput for additional hops, so:

$$T \approx c / n^{s} \tag{8}$$

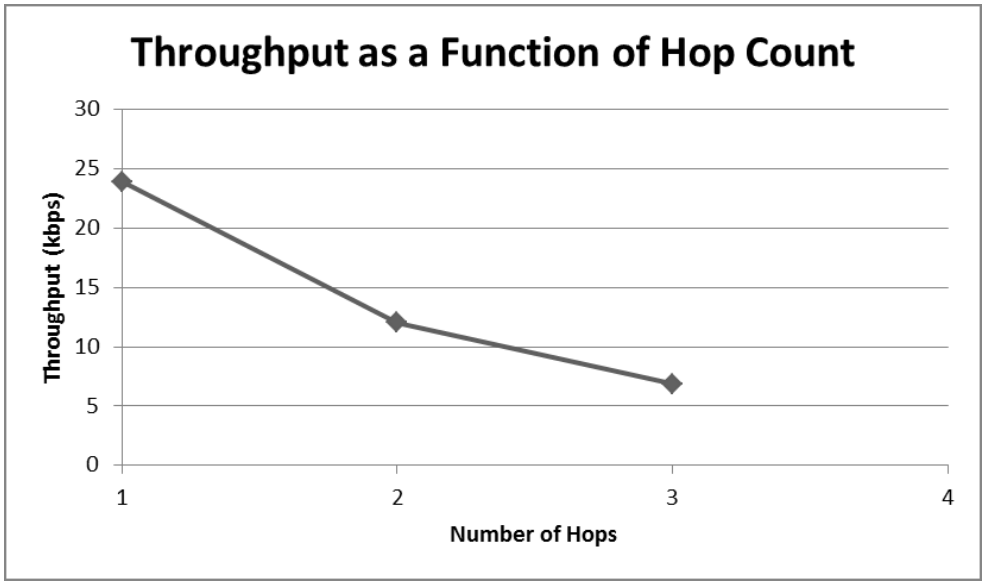


Figure 7-11. Throughput as a function of hop count.

Network design needs to account for the implications of multiple hops, and architecture should include as few as possible. While it would have been desirable to test a network with a larger number of nodes if more devices had been available, it is clear that transmission beyond a few hops becomes impractical due to the low throughput.

7.8.3 Third Test

As expected, the protocol is robust and automatically fails over to a co-located node. Some frames held in the buffer of the failed device, however, are lost forever.

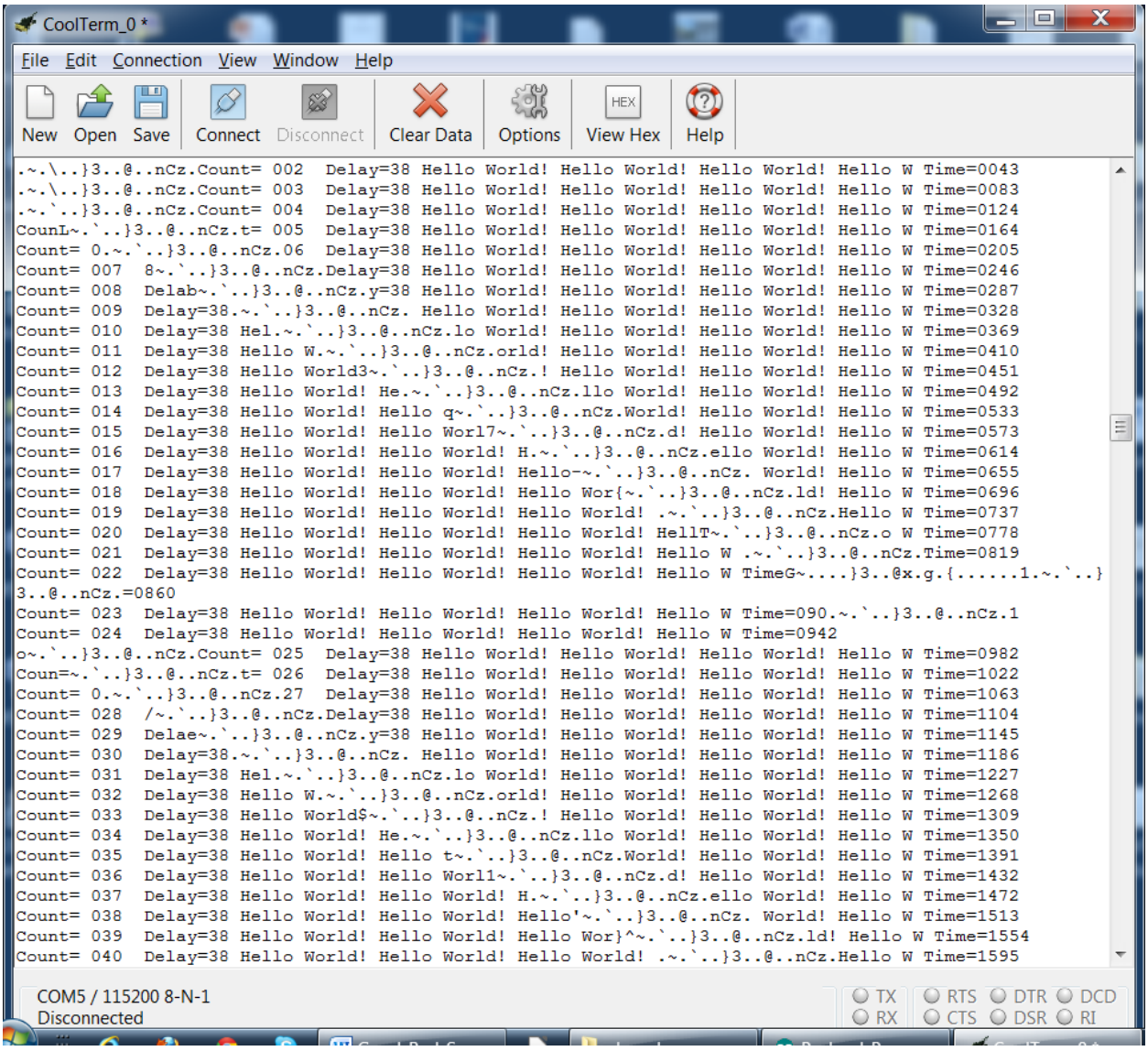
Taken together, the testing suggests some considerations for design of the powerline monitoring device and data collection network. As the transceivers are compact, inexpensive,

and have abundant power it is worthwhile considering adding a second co-located transceiver in the same PAN to each monitoring device, which could take over in the event of failure of the primary transceiver. It also suggests that additional transceivers communicating on different channels could possibly be used to send data from different sensors located in the device as a method of increasing effective throughput in the system.

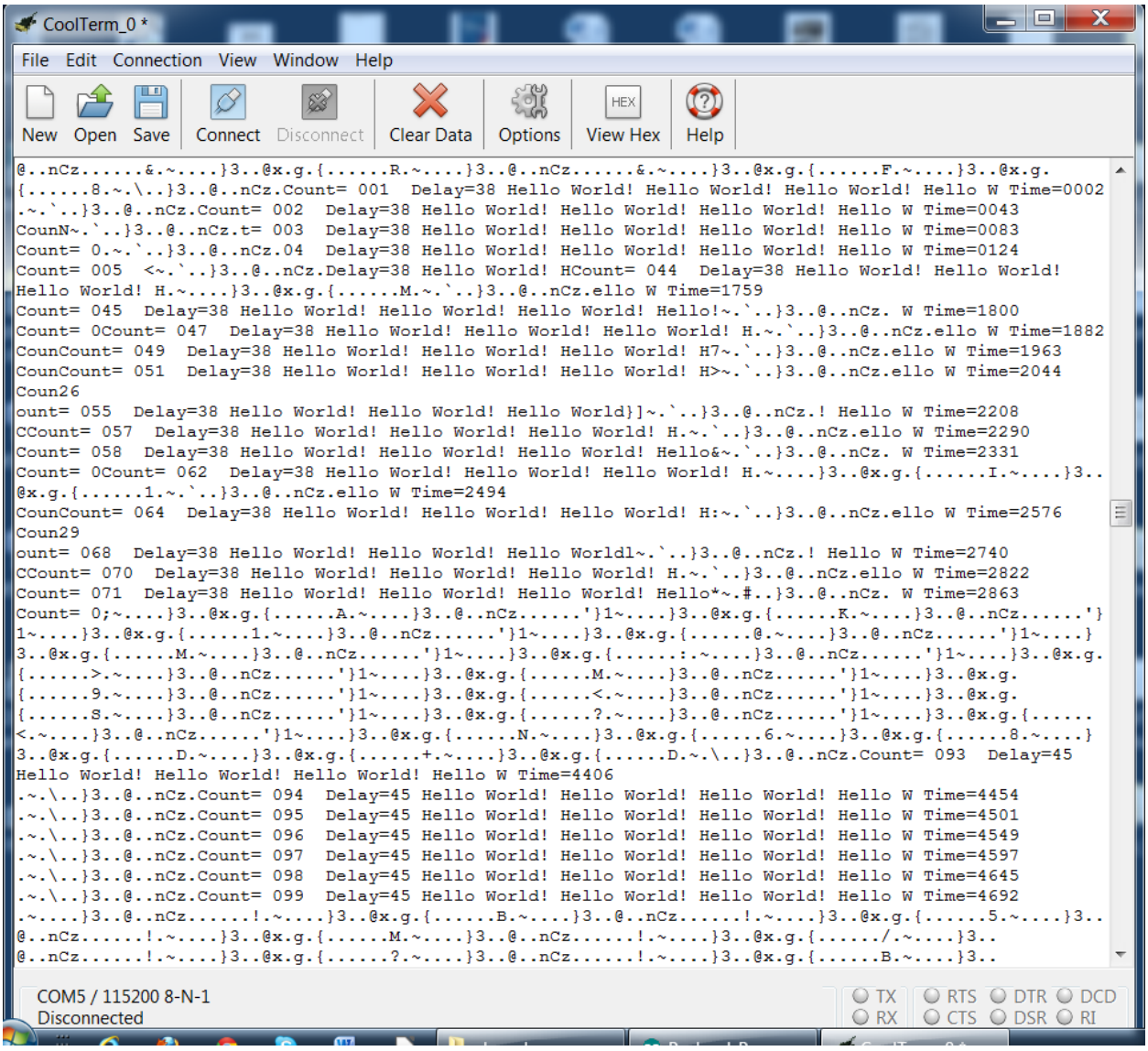
7.9 Chapter Summary

In this chapter I discuss an application of WSANs for transmission and distribution monitoring. This chapter also details laboratory testing of some elements of a specific WSAN configuration, including maximum throughput, the effect of multi-hop mesh networking on throughput, and robustness. Testing demonstrates that throughput of the actual maximum data payload is significantly less than the 250 kbps rate stated in the underlying protocol, an important consideration when designing the powerline monitoring application. Throughput is also significantly affected by the number of hops in the mesh network, falling by approximately half at each node. As expected in an ad hoc mesh network, the communication nodes are able to automatically reconfigure to route around the loss of a node, an important consideration for maintaining a robust mesh network.

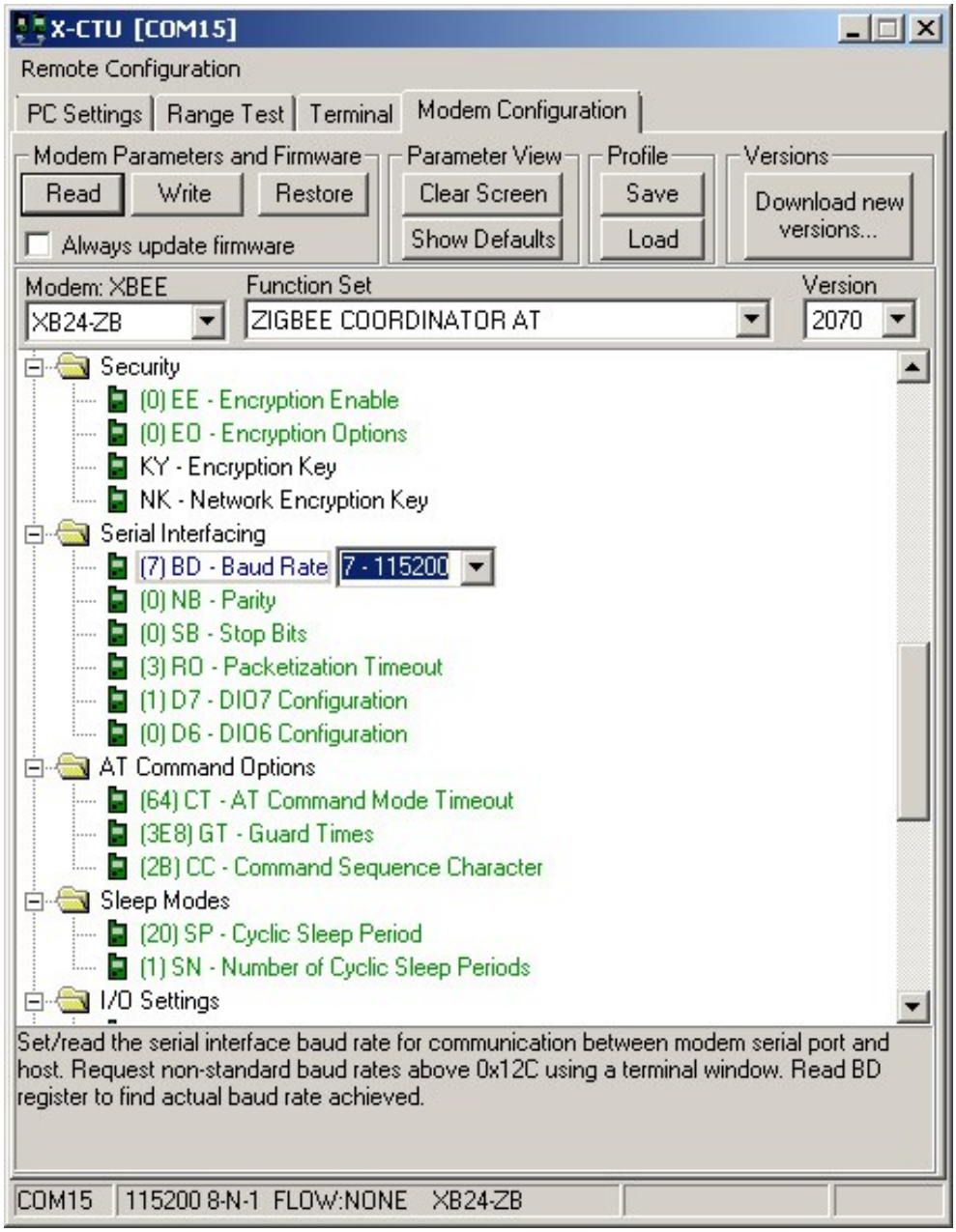
7.10 Exhibit A – Uncorrupted Received Data



7.11 Exhibit B – Corrupted Received Data



7.12 Exhibit C – X-CTU Firmware Configuration



7.13 Exhibit D – Lab Equipment List

- Laptop running Windows 7 and terminal program
- Arduino Uno SMDs
- XBee 2mW Wire Antenna - Series 2 (ZigBee Mesh)
- USB cables
- Parallax XBee USB Adapter Board
- Jumper Wire
- Voltage Regulators - 3.3V
- Common BJT Transistors - NPN 2N3904
- Electrolytic Decoupling Capacitors - 10uF/25V
- Wall Adapter Power Supplies - 9VDC 650mA
- Mini Photocell
- Breadboards, Clear Self-Adhesive
- Break Away Headers - Straight
- 2mm 10pin XBee Sockets
- LED - Assorted
- TMP36 - Temperature Sensors
- Solar Cell Small - 0.45W
- 9 Volt Alkaline Batteries
- Breakout Boards for XBee Module
- Resistors 10k Ohm 1/6th Watt PTH

- Resistors 330 Ohm 1/6th Watt PTH
- XBee Shields
- Arduino Stackable Headers - 6 Pin
- Arduino Stackable Headers - 8 Pin
- DC Barrel Jack Adapters - Breadboard Compatible
- 9V to Barrel Jack Adapter
- Hook-up Wire - Gray
- Momentary Push Button Switch - 12mm Square
- Buzzer - PC Mount 12mm 2.048kHz
- Humidity and Temperature Sensor - RHT03
- Variety of small hand and electrical tools, multimeter, and soldering equipment.

8 The Issue of Backhaul

Sensors and actuators lie at one terminus of the network, but individual or aggregated nodes must be connected to a larger system for monitoring and control. Gateways may be connected to virtually any type of public or private wide area communications structure, including general packet radio service (GPRS), the public switched telephone network (PSTN), fiber backbone, or satellite. This chapter examines the reliability and capacity of one backhaul alternative using cellular networks and the short message service (SMS). It is chosen because it is relatively available, simple, and inexpensive. It is also among the least capable of all of the alternatives, making it a worthy candidate for further study to determine the potential limitations and suitability of the service. Power lines often traverse largely rural areas, and rural cellular services are frequently characterized by challenging terrain and long distances between cell sites. These areas are subject to weak signals and intermittent connection due to fading. Transmission lines are placed on high towers which may help improve coverage in remote areas, however we expect there to be a complementary role for mesh networks as well. Because SMS transmits messages as short bursts of text, it is well suited for the intermittent small bursts of data characteristically sent by sensors if it can be shown to be adequately reliable under unfavorable conditions.

> *Recently, Short Message Service (SMS) functionality of the digital cellular network has been applied in order to remotely control and monitor substations... these communication technologies are suited to the applications that send a small amount of data… [23]*

8.1 SMS Laboratory Testing[14]

This chapter presents a study of SMS air interface delay in strong and weak signal environments. The performance of SMS in unfavorable conditions, in this case weak signal conditions, is relevant for areas with poor coverage or a weak signal environment.

8.2 SMS Background

The short message service is one of the most popular mobile data services. The International Telecommunications Union (ITU) recently reported: "The total number of SMS sent globally tripled between 2007 and 2010, from an estimated 1.8 trillion to a staggering 6.1 trillion. In other words, close to 200,000 text messages are sent every second." [92]

This chapter studies SMS communication in a laboratory setting and presents the results of measurements of the time it takes to transmit an SMS message over the air interface. This time is termed the *air interface delay*. The chapter is organized as follows: First is a description of the measurement setup. Second is a report of data for a variety of conditions; these conditions vary the size of the SMS, the signal strength, and whether the receiver is fixed or mobile. Third is a report on a delay model and technique for estimating the model parameters; with these parameters it is possible to characterize the distribution of the delays.

---

[14] Portions of the work of this chapter were completed with Suzana Brown of the University of Colorado Boulder and submitted as part of the *Comments of the University of Colorado Interdisciplinary Telecommunications Program* Before the Federal Communications Commission In the Matter of *Facilitating the Deployment of Text-to-911 and Other Next Generation 911 Applications* PS Docket No. 11-153 and *Framework for Next Generation 911 Deployment* PS Docket No. 10-255.

The work of this chapter was presented at the International Awareness Conference on Sustainable Wireless Solutions for Environmental Monitoring, 23 February 2012 as part of the ICTP-ITU/BDT School on Sustainable Wireless ICT Solutions in Trieste, Italy.

Results and business implications are also scheduled to be presented at the IST-Africa 2013 Conference on May 29-31, 2013 in Nairobi, Kenya.

SMS has been proposed for a number of time-sensitive applications such as infrastructure monitoring [93], sensor data collection [94], medical patient monitoring [95], and 911 emergency communications [96]. SMS is designed for text messaging, an application which is asynchronous and delay tolerant. In the case of [96], the FCC has proposed SMS as an additional method of contacting emergency services via 9-1-1.

A better characterization of SMS delay distribution would provide a method to assess the suitability of SMS for these and other applications. For the purpose of my analysis the performance of SMS in unfavorable conditions, as found at the edge of coverage, present support for using SMS as a mode of communication in conditions ranging from ideal to marginal. For the purpose of this testing I use GSM phones because of the worldwide popularity of the GSM standard and simplicity of accessing the network facilitated by SIM cards. While the model developed is specific to GSM, similar models can be developed for other standards such as CDMA 2000.

8.2.1 SMS Architecture

In a cellular system, the location of a mobile station (MS) is determined by a registration and paging process over control channels. SMS messages are transported via those control channels. A voice call is set up using a control channel to communicate with the tower, but as soon as the base station establishes that the request is for a voice call it transfers the call to a traffic channel. The mobile switching center (MSC) controls multiple base transceiver stations (BTS). The MSC also handles functions such as registration, authentication, location updating, handovers and routing to roaming subscribers. It has access to at least three databases: the Home Location Register (HLR) that keeps detailed records of each subscriber in the MSC network; the Visitor Location Register (VLR), which keeps a record of subscribers who have

roamed into the jurisdiction of the MSC from other networks; and the Authentication Center (AuC), which authenticates each subscriber that attempts to access the network.

SMS is based on the capability of a digital cellular terminal to send and receive alphanumeric messages [97]. These short messages can be up to 160 characters in length and can be sent concurrently with voice traffic. When a subscriber is not on a voice call SMS utilizes the stand-alone dedicated control channel (SDCCH), but while the subscriber is on a call it uses the slow associated control channel (SACCH) [98].

The payload length of SMS is limited by the constraints of the signaling protocol, precisely 140 octets. Short messages can be encoded using a variety of alphabets; the default in GSM is a 7-bit alphabet. This leads to the maximum individual short message sizes of 160 7-bit characters [99].

The path of a voice call from one MS to another MS differs from a path SMS takes. Figure 8-1 is a diagram of a standard mobile voice call. The main difference is that the MSC sends the SMS message to a SMS center (SMSC) and if the recipient is in the same network it is delivered directly. If the recipient is in a different network it goes via a gateway as pictured in Figure 8-2.

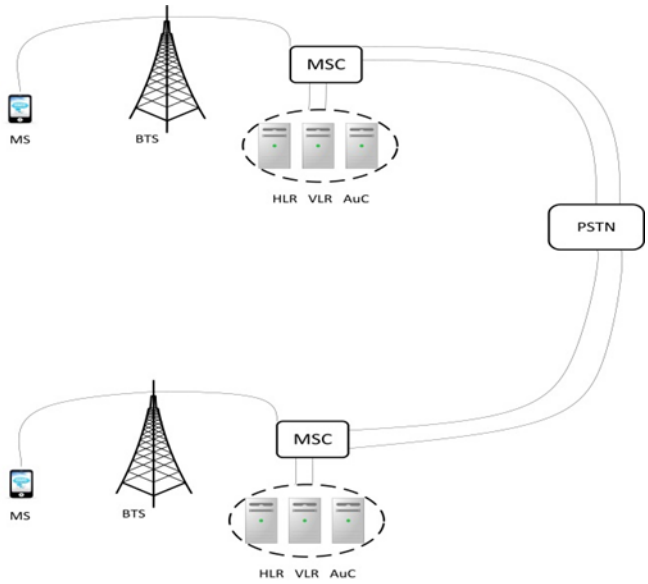


Figure 8-1 Path of a mobile voice call.

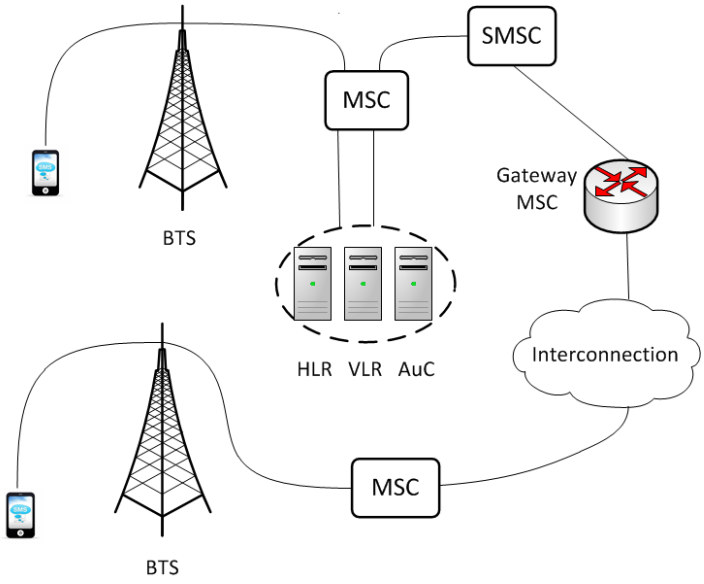


Figure 8-2 Path of an SMS message.

8.3 SMS Prior Work

The air interface delay can be computed from the first principles by examining the GSM protocol [100]. There are, however, considerable vendor and operator specific parameters which may also be configured. As will be shown, channel errors also introduce significant variability to the delay, and it is precisely this variability which will be characterized using a measurement approach.

Hung et al. derive the distribution of SMS transmission delay based on 40,000 data points obtained from commercial operations [101]. The authors consider a variety of standard distributions to which to fit the data. They find that the transmission delay distributions are not symmetrical, have heavy tails, and cannot be approximated completely by any one distribution. The conclusion is that the distributions can be used to roughly model transmission delays but are not useful for precise estimation.

In another study, Tseng et al. apply GSM-SMS for rural agricultural data acquisition [99]. The authors use SMS because of its low power requirements, widespread GSM coverage, the capability to save messages, and the group broadcast functions of the system. The authors conducted several delay measurements. One sent 60-character SMS messages through the SDCCH channel and measured an average 3.2 seconds for a message from the MS to reach the SMSC. They then measured the end-to-end one way delay from a field monitoring platform to a remote host platform; this took 10 to 15 seconds. In another measurement, the sending time for an SMS message to any receiving endpoint ranged from 10 to 20 seconds. The accuracy of data transmission via SMS was 100%; the retransmission rate was 2.73%, and the data loss rate 0.66%.

Collesei et al. explain that network delay depends on the length of the SMS message [98]. A message of 60-characters, for example, sent via either the SACCH or SDCCH, in average radio propagation conditions, requires 3.2 seconds before the message reaches a SMSC and can be routed to its destination. If the propagation conditions are extremely favorable, this may require as little as 2.9 seconds.

To our knowledge no other study directly compared the air interface delay for good signals with those at the edge of coverage. Further, we develop a model incorporating the GSM protocol that provides a good fit to the data.

8.4 Methodology

The experimental setup and test methodology were as follows:

8.4.1 Test Setup

Testing took place inside the Pervasive Communication Laboratory inside the Engineering Center at the University of Colorado. The basic setup is shown in Figure 8-3. The mobile phone used for testing is a Telit EVK2 GSM Development Board with a SIM card for the AT&T/Cingular network. This is connected to a Larsen Special remote mobile antenna through an S.M. Electronics SA3550S manual step attenuator (0-3000 MHz, 0-50 dB in 1 dB steps). An Anritsu Spectrum Master MS2721B spectrum analyzer is connected to the mobile phone and antenna through a 20 dB attenuator and a Mini-Circuits 15542 Splitter.

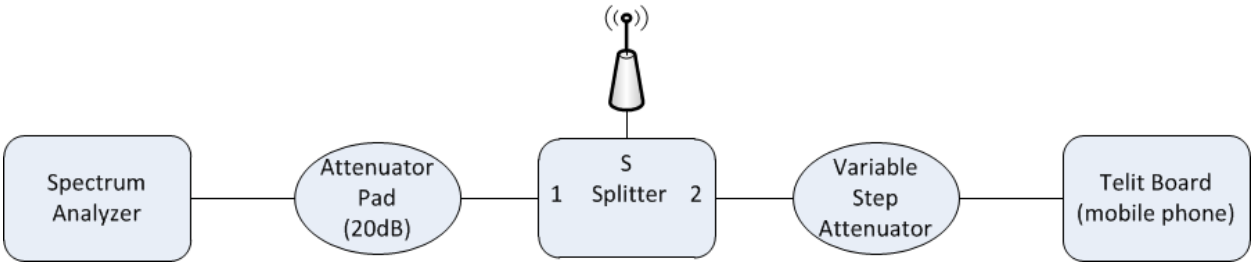


Figure 8-3 Equipment configuration for all tests.

The Federal Communications Commission (FCC) Spectrum Dashboard website was used to determine the GSM spectrum available and licensees for the test site in Boulder, Colorado. The SIM card used for testing carried the Corr Wireless brand, but they were not listed as having licensed spectrum in the area, which indicated they were reselling service from another carrier. As GSM in the United States was known to operate in the cellular bands around 850 MHz and in the PCS band around 1900 MHz, the Spectrum Dashboard was used to determine the exact frequencies licensed in Boulder. A series of test calls were placed and the spectrum analyzer was used to detect activity in these GSM bands. Multiple calls were used to verify that spectrum in the range of 824 – 835 MHz and 845 – 846.5 MHz were used for the reverse link (phone to BTS) and 869 – 880 MHz and 890 – 891.5 MHz were used for the forward link (BTS to phone), licensed to AT&T/Cingular wireless. Unless noted otherwise, all measurements were taken on the reverse link.

The Telit board was instructed to send text messages, and then queried for the received signal strength (RSSI). Tracking the RSSI enabled the path loss to be tuned to specific repeatable received power levels.

The 20 dB attenuator reduced the signal from the mobile phone board to within the dynamic range of the spectrum analyzer. This path was tested and calibrated using a signal

generator and the spectrum analyzer. The path was found to have approximately 30 ± 1 dB of overall attenuation with the step attenuator connected but set to zero attenuation. At this setting, the base station signal strength was strong enough for good reception at the mobile. The variable attenuator allowed the signal level at the phone to be varied between this strong signal level and a level sufficiently weak that communication was not possible. The spectrum analyzer was placed in zero span mode so that the packet transmissions from the MS could be recorded over time.

8.4.2 Defining and Measuring the Air Interface Delay

The call connection setup for SMS, when viewed on a spectrum analyzer, appears as a series of spikes, one for each packet in the exchange. Air interface delay for sending SMS is measured on the spectrum analyzer from the beginning of the first spike (initial RACH message) to the end of the last spike, as pictured on Figure 8-4. This does not measure the initial random delay before the RACH message is sent, but this is at most one multiframe period (235.38 ms).

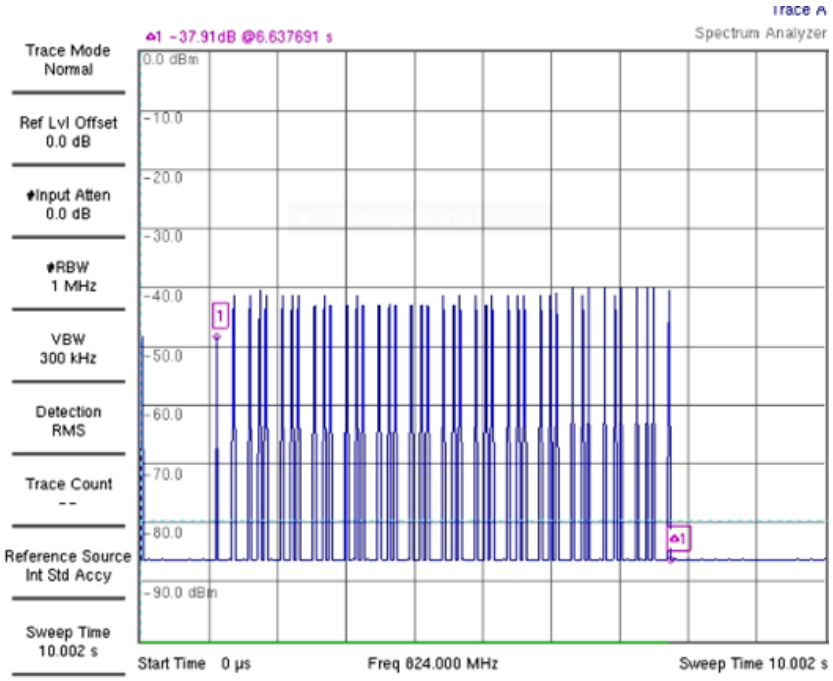


Figure 8-4 Spectrum analyzer screenshot showing time to send an SMS message.

Performance was measured at both strong and weak signal levels. The level of cut-off attenuation where an SMS call could not be successfully placed was found empirically. Once this level was established, reducing the attenuation by just 1 dB was enough to enable a successful SMS connection. The measured signal level at this attenuation was -109 dBm and denoted a *weak* signal. A *strong* signal was 12 dB higher (-97 dBm).

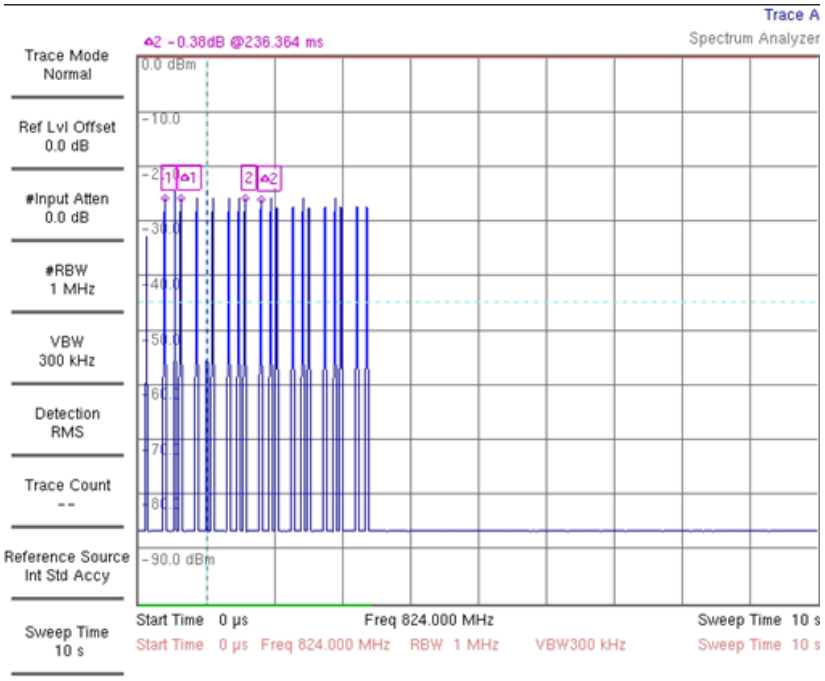


Figure 8-5 Spectrum analyzer screenshot showing measurement of a Superframe.

8.4.3 First Test Setup - Various Message Sizes

The first test was intended to establish a baseline for the time delay of SMS messages and evaluate messages of different sizes. The message sizes used were 1, 60, and 160 characters, covering the extremes of the range of permissible SMS message sizes. They were all measured with a strong signal.

8.4.4 Second Test Setup – 60 and 160-character Messages at Different Signal Levels

This test evaluated the effects of low signal strength, such as when a mobile phone signal is obstructed or at the edge of coverage, on the ability to connect a call. Measurements were made under weak signal conditions and compared to the previous results made with a strong signal.

8.4.5 Third Test Setup - Choreographed Mobility

The final set of tests was designed to evaluate the effect of mobility-induced fast fading when the signal is weak. A 60-character message was transmitted while moving the antenna in a choreographed manner as described in [102], with one experimenter pacing back and forth on a path four meters long.

8.5 Results

Figure 8-4 is a screen shot of the spectrum analyzer showing the individual frames and the spacing of a GSM multiframe. This multiframe structure is characteristic for GSM, and each grouping of three spikes represents two SDCCH with one SACCH in the middle. The standard multiframe length is 235.38 ms as described by Redl et al. [100]. The measured value is 236 ms and is typical, thus our measurements are accurate to within a few milliseconds. Reported results are derived from at least 30 observations.

8.5.1 Various Message Sizes

For this test the delay time of different size messages was compared. Three message sizes were chosen, 1-character, 60-characters, and 160-characters. The results are shown in Table 8-1.

TABLE 8-1. AIR INTERFACE DELAY

Message Size	Average Delay (sec)	Standard Deviation
1 character	3.28	0.29
60 characters	4.06	0.38
160 characters	5.59	0.38

The longest message delay is only 1.7 times larger than the smallest message as most of the air interface delay consists of connection setup overhead. This indicates that message size does not increase delay considerably.

8.5.2 Comparison of Time Delay of 60 and 160-character Messages at Different Signal Strengths

This test compared messages of 60 and 160-characters at two different signals strengths. The strong signal is set at -97 dBm and the message is transmitted with 100% reliability. The weak signal strength is set at -109 dBm and the message is transmitted 82% of the time. If a packet was transmitted (as reported by the Telit board), in each case it was successfully delivered to the destination. The recorded time is that of the successfully transmitted 60-character messages only, and the results are presented in Table 8-2. Table 8-3 shows the time delay for the 160-character message.

TABLE 8-2. DELAY FOR 60-CHARACTER MESSAGES AT TWO SIGNAL STRENGTHS

Signal Level (dBm)	Average Delay (sec.)	Standard Deviation
- 97 (strong)	4.06	0.38
- 109 (weak)	4.49	0.76

TABLE 8-3. DELAY FOR 160-CHARACTER MESSAGES AT TWO SIGNAL STRENGTHS

Signal Level (dBm)	Average Delay (sec.)	Standard Deviation
- 97 (strong)	5.59	0.38
- 109 (weak)	6.19	0.93

It is apparent that at the edge of coverage it takes longer to send a message of the same size, and more importantly that the standard deviation approximately doubles for both the 60-character message and the 160-character message compared to the strong signal environment.

This extended delay is explained by the time of retries to establish a connection, and layer two retransmissions.

8.5.3 Choreographed Mobility-induced Fading

The final test was a choreographed mobility experiment using a methodology similar to that described by Rensfelt et al. [103]. The MS antenna was moved in a choreographed manner to simulate mobility and concurrently sent a series of 60-character messages. The test was started at a low signal strength of -109 dBm, and a path of approximately eight meters was traced in a random fashion. The result is shown in Table 8-4.

TABLE 8-4. DELAYS FOR 60-CHARACTER MESSAGE, STATIONARY AND MOBILE

60-character SMS	Average (sec.)	Standard Deviation
Stationary	4.09	0.38
Mobile	5.24	1.66

The mean of the air interface delay in the mobile case in-creases by more than a second as compared to stationary, but the standard deviation triples. The large standard deviation implies that in realistic situations (when the sending device is moving) the time delay varies considerably.

8.6 Air Interface Delay Probability Model

The empirical cumulative distribution function for the 60-character SMS air interface delay is shown in Figure 8-6 for both the strong and weak signal case. Figure 8-7 shows the empirical cumulative distribution function for the case of the mobile user. With a probability of 0.9 and strong signal, the delay of SMS messages will be under 5 seconds. With a weak signal this delay will be less than 8 seconds, and in the mobile case the delay will be less than 6.5 seconds.

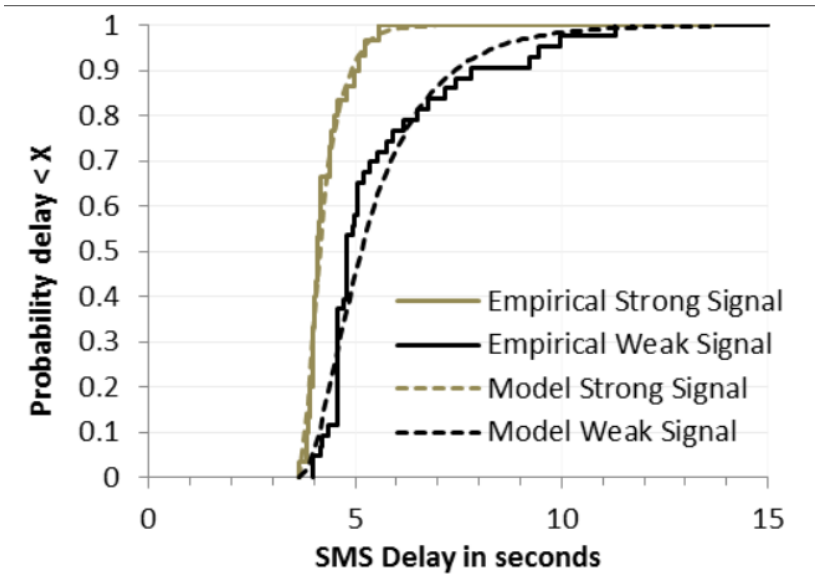


Figure 8-6 Cumulative distribution function for the air interface delay of a 60-character message.

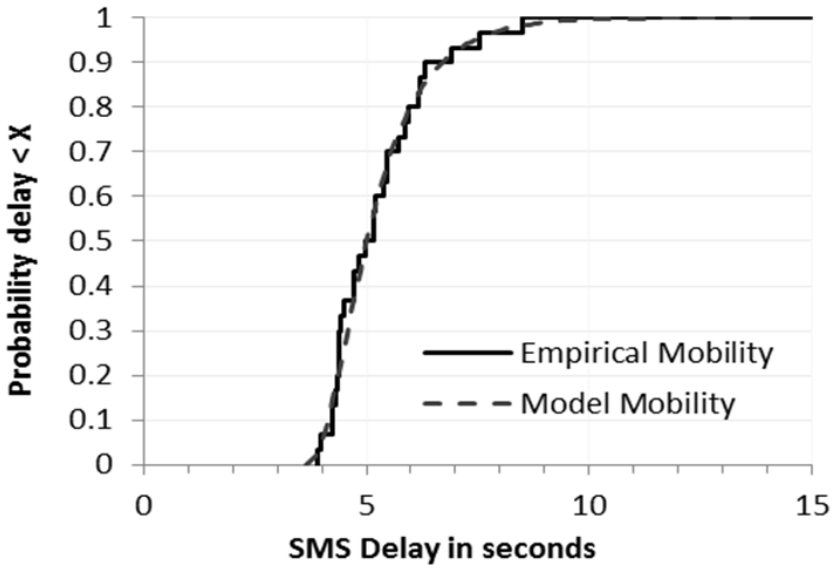


Figure 8-7 Cumulative distribution function for 60-character air interface delay.

In [99] it is shown that the distribution of the delay is neither normal nor well described by other common distributions. It is possible to take an approach based on knowledge of the SMS connection process instead. Delay is composed of four components. The first is a deterministic minimum time to deliver the message in the best case, $t_{\min}$. The second is a uniform distribution between 0 and t_s, i.e. $U \sim \boldsymbol{U}(0, t_s)$, where t_s = 235.38 ms is the multiframe duration. This accounts for the random time between when the first RACH is sent and the assignment of the SDCCH. The RACH can fail due to collisions or not being received and this occurs with unknown probability p_r. If a retry occurs, it happens after retry timeout tr = 1 second and this process can be repeated until successful. Thus, this third component is geometric, $G \sim \boldsymbol{G}(p_r)$. The last component is a function of the number of frames sent. The SMS message consists of n frames, and each time a frame is sent the transmission can be in error and require the frame to be resent with probability p_s. Each of the n frames has to be sent and resent until they are successful. The number of attempts to send each frame is geometrically distributed and the total number of attempts to send all n frames is distributed as a negative binomial, $B \sim \boldsymbol{NB}(p_s, n)$. Thus, by letting T be the random variable for the air interface delay:

$$T = t_{min} + U + t_r G + t_s B \tag{8}$$

It is possible to estimate $t_{\min}$ from the minimum time to send a message under strong signal conditions, and n can be similarly determined. The quantities t_s and t_r are known from the SMS protocol. This leaves p_r and p_s as the quantities remaining to be determined. For a given set of k air interface delay measurements, $\{t_1, t_2, \ldots, t_k\}$, it is possible to compute the maximum likelihood values of the probabilities p_r and p_s using a grid search over possible values.

This approach was applied to the three data sets for the 60-character SMS air interface delay. The parameters of the model are shown in Table 8-5, and these parameters were used to plot the model cumulative distribution functions in Figure 8-6 and Figure 8-7. This model provides a good fit to the empirical cumulative distributions. The fit for the strong signal and mobile scenarios is very good. The fit for the weak signal has more deviation, perhaps because a significant fraction of the transmission attempts (18%) were unsuccessful and not included in the empirical data.

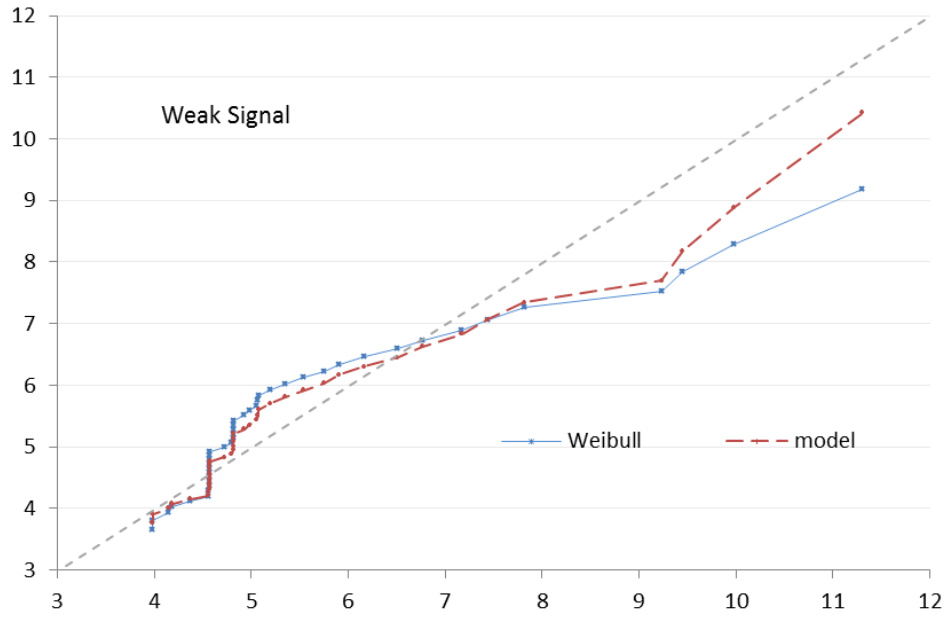


Figure 8-8 Weak signal fit (seconds – seconds).

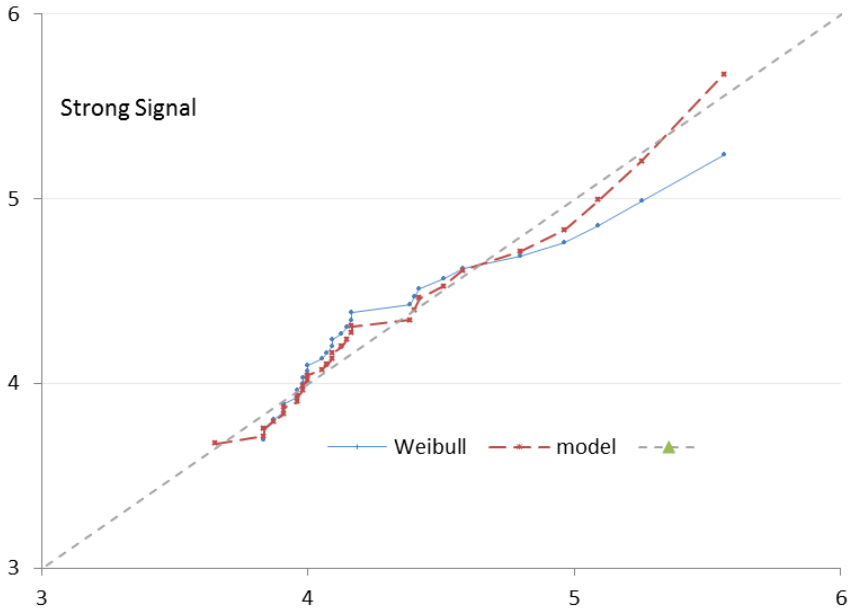


Figure 8-9 Strong signal fit (seconds – seconds).

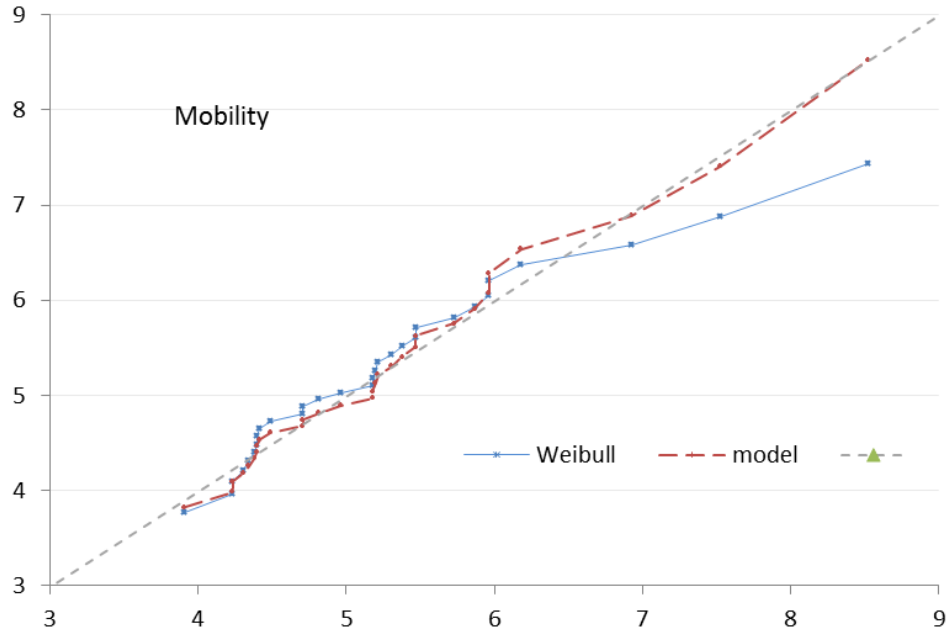


Figure 8-10 Mobility fit (seconds – seconds).

TABLE 8-5 GSM MODEL PARAMETERS DETERMINED FROM DATA.

60-character SMS	t_{min} (sec.)	n	p_r	p_s
Strong Signal	3.654	13	0.10	0.11
Weak Signal	3.654	13	0.49	0.20
Mobile	3.654	13	0.38	0.22

For comparison, consider a Weibull-based model distribution. The Weibull-based model is

$$T = t_{min} + W \tag{9}$$

where $W \sim \boldsymbol{W}(\lambda, k)$ is a Weibull distribution with scale and shape parameters λ and k. The Weibull is a general distribution that can help identify if the data follows some standard distributions. The Weibull is fit to the excess delay beyond $t_{min} = 3.5$ *sec.* This is smaller than the value in Table 8-5. Delays as small as 3.654 sec were observed. With (9), since the probability of a *0* excess delay is *0*, a small gap was necessary to get a valid fit. We note that such heuristics are not necessary in using (8). The resulting parameters are shown in Table 8-6. The shape parameter close to $k = 2$ suggest that the Rayleigh distribution is the best fit.

TABLE 8-6 WEIBULL MODEL PARAMETERS DETERMINED FROM DATA

60-character SMS	t_{min} (sec.)	λ	k
Strong Signal	3.5	0.86	2.01
Weak Signal	3.5	2.31	1.66
Mobile	3.5	1.97	2.04

The next exploration is the goodness of fit between these models and the empirical distribution. To that goal, Q-Q plots were devised comparing the GSM-based model distribution with the Weibull-based distribution. The result for weak signal, strong signal, and mobility are shown in Figure 8-8, Figure 8-9, and Figure 8-10, respectively. In all three cases (strong and weak signals, and the mobility case) the model is closer to the experimental data than the Weibull as determined by its closeness to the 45 degree line. Note in particular that the fit is

significantly better in the tail of the distribution, which is important for modeling extreme excess delays.

8.7 Analysis

The purpose of this study was to characterize the SMS air interface delay. Three different scenarios were considered.

The first scenario, measuring the time delay of three SMS messages of different sizes, shows that the overhead of an SMS message is large and the time delay ratio of the smallest message (1-character) to the largest message (160-character) is 1:1.7. This test also established the longest time of delay expected for the longest message at 6.8 seconds in an environment with a strong signal.

The above measurements are conducted with a strong signal, defined as a reliable service, in this case -97 dBm. The next test was of 60 and 160-character messages at a weak signal, defined as unreliable service, in this case -109 dBm. For both message sizes the time delay in a weak signal is longer primarily due to access retries and layer two retransmissions. Both messages take an average of 10% longer in the weak signal scenario. More importantly, the standard deviation doubles in both cases, and now, with a probability of 0.9, the longest message takes 7.8 seconds to transmit.

In the case of mobility the situation changes slightly. Because the standard deviation increases considerably, it will take at most 8.6 seconds to send a 60-character message with a probability of 0.9. A mobility scenario in a weak signal environment is the most realistic but the most complicated case.

The modeling of the data in Table 8-5 shows that at the edge of coverage the retry probability is four to five times greater than when at good coverage, and the layer two retry probability is twice as large. With these parameters we can estimate outlier events such as the likelihood of packet failure due to retries.

It is noteworthy that all SMS messages that were successfully sent were received by the recipient, so the loss rate is 0%. In each test at least 30 text messages were sent, and two different scenarios were conducted for each test. We performed three tests, so the total number of text messages sent was approximately 300. This is not as large as the field study discussed by Hung et al. [101] which had 40,000 messages but the results of these lab tests are compatible with the larger field studies.

8.8 Capacity and Congestion in the Cellular Network

In the U.S., GSM spectrum is allocated in 25MHz blocks per market. These blocks are frequency divided among two or more carriers, but for this exercise I do not attempt to determine the capacity on a per-carrier basis. The 25 MHz block is divided into GSM frequency channels (frequency division multiple access, FDMA) of 200 kHz each, so 25 MHz / 200 kHz = 125 frequency channels. One channel is used as a guard band, so 124 frequency channels are useable for traffic. If the reuse factor is 4, and the number of sectors per cell site is 3, then each sector has a capacity of 124 / 12 = 20 frequency channels, and each cell site has a total of 60 frequency channels available.

Each of these frequency channels is time-divided into 8 time division multiple access (TDMA) channels, so 60 * 8 = 480 total channels are available for voice calls on each cell site. Within these divisions, “virtual” channels exist, so it is possible for control, traffic, and other

channels to exist in the same FDMA and TDMA slot. SMS messages are sent over the Standalone Dedicated Control Channel (SDCCH), which can subdivide each slot into an additional 8 divisions, so there are 8 times as many SMS channels as voice channels available, or a total of 480 * 8 = 3,840 SMS channels per cell site.

8.8.1 Grade of Service

Grade of service is a common telecommunications engineering term, and is one of the inputs to the traffic theory model. The National Emergency Number Association (NENA) gives the following definition: ***Grade of Service*** *The probability (P), expressed as a decimal fraction, of a telephone call being blocked. P.01 is the grade of service reflecting the probability that one call out of one hundred during the average busy hour will be blocked.* [104]. The grade of service to which a mobile telephone system is designed is determined by the mobile network operator, and is typically in the range of P.02 to P.03. I will use P.02 in the examples.

8.8.2 SMS

The maximum number of SMS users per minute or hour for a given grade of service for each cell site may be determined. The following assumptions are used for SMS: 3,840 channels and P.02 grade of service, as determined above. The average air interface duration of an SMS message is approximately 8 seconds as determined previously. Using an Erlang B table or calculator, 3,840 channels and P.02 grade of service gives us 3,880E. 3,880E * 36 CCS per Erlang = 139,680 CCS * 100 seconds = 13,968,000 call seconds. 13,968,000 / 60 seconds per call = 232,800 users per hour / 60 minutes per hour = 3,880 users per minute.

8.8.3 Other Applications

Using the same formulas, it is possible to determine the percentage of calls blocked for a user group of any given size. Similarly, it is possible to determine the aggregate capacity of part or all of a cellular system, given the number of cell sites, or for a different cellular infrastructure, such as code division multiple access (CDMA). From this it is possible to determine the call capacity or blocking probability based upon a certain percentage of the infrastructure becoming damaged or unavailable, such as during a natural disaster.

8.8.4 Other Network Elements

The above calculations determine the capacity of the cellular sites only. It is possible that congestion could exist at other points in the network, such as backhaul from the base transceiver station (BTS) to the base station controller (BSC), the BSC to the mobile switching center (MSC), in the MSC itself, or at the short message service center (SMSC). While not all of the capacity calculated above would be routed to a NOC, it is also possible that while a large volume of calls or messages may be successfully offered to the NOC, equipment available may be insufficient to handle the call volume.

8.9 Chapter Summary

This chapter studied the air interface time delay of GSM SMS messages, and considered congestion in the network. The length of the message does not increase the delay more than a factor of two. On the edge of coverage, where the signal is weak, the delay increases by an average of 0.5 seconds for both 60 and 160-character messages. The standard deviations also grow much larger, implying less reliability on the edge of coverage. The GSM protocol-based delay distribution model proved to fit the empirical data well. Future work could undertake testing using other standards for mobile communication and compare the results to the findings

using GSM. The data and models presented here will assist in analyzing the performance of SMS in time or reliability-sensitive applications, and should prove useful in considering options for backhauling sensor data collected from monitoring the electrical grid. SMS is particularly useful in rural and developing areas due to its availability, low cost, and robustness, and is well suited to systems and applications which are mindful of the benefits and limitations of the protocol.

9 Overall Summary and Analysis

The policies and entities responsible for improving critical infrastructure protection for the electric grid are reasonably well defined, but it can be shown that current measures have not been adequately effective. Despite an increased focus on protecting U.S. critical infrastructure, especially since the events of 9/11/2001, the electrical grid in particular is more vulnerable than ever. Fragmented regulation with unintended consequences, aging infrastructure, increasing demand and complexity, more capable and motivated attackers with a greater number of attack surfaces, construction challenges, and inadequate investment have all contributed to an electrical grid environment that is increasingly less, not more, reliable and secure.

Many efforts are underway to make a smarter, more modern, and more capable electrical grid. In this paper I suggest that focusing on the aspect of protecting critical infrastructure is a natural starting point for modernizing the grid. The consequences of the outstanding problems, the fact that unprecedented focus has failed to improve overall security and reliability, and the realization that improving critical infrastructure protection need not require the invention of significant new technology, large long-term investments, or changes in consumer behavior make this approach a candidate for consideration as a strategic priority.

Due to the nature of unintended regulatory consequences and sanctions, power system operators are often more focused on compliance with regulation than on solving the underlying issues policy makers hoped to address. While this thesis focuses on one area of vulnerability and proposes a technical solution, these measures are not likely to see widespread adoption without further regulatory, standards, and economic support.

Traditional industrial control systems and equipment are centralized and expensive, and have failed to achieve the desired level of grid protection, suggesting that the next generation of monitoring and control should explore new architectures. As the application of communications technologies provides both opportunities and challenges for modernization, the increasing need for security demands that the industry find new ways to protect the traditional electric grid and guard against attacks through the new vectors introduced by incorporating modern data networks.

In this paper I proposed integrating relatively simple, robust, and economical cyber-physical components based upon wireless sensor networks with the electrical grid. The electrical grid itself is a widely distributed cyber-physical system, so control systems with a similar architecture are a natural fit. Wireless sensor networks are one manifestation of a widely distributed cyber physical system oriented toward monitoring and control.

I explored an application which incorporates these sensor networks with electrical transmission and distribution equipment. I focus on the electrical distribution network as it is among the least automated elements of the grid and some 80% of outages occur at the distribution level, although distribution monitoring is also readily applicable to transmission. I demonstrated a prototype WSAN node for powerline monitoring, but virtually no work has been done to characterize the communications performance or requirements of this monitoring system. By discussing requirements for grid monitoring and the functional elements of WSANs I provide background for performance testing of select characteristics of the communications network.

Using a network based on Arduino microcontrollers and ZigBee transceivers I provide some empirical data on maximum throughput, the effect of multiple hop transmissions in mesh

architecture, and the capability of the system to respond to node failures. For most applications the throughput available in one or two hop arrangements will likely meet bandwidth needs, but arrangements of three hops or more may not. These performance metrics can be used to refine the design of the powerline sensor, and in designing the network architecture.

Data from the sensing network must be relayed to a network operations center for further analysis, and to assist in guiding the response to abnormal conditions. I demonstrate testing of one simple method of backhauling data via SMS. This option is attractive due to its relative availability, reliability, and low cost. Because SMS is designed for a series of small text messages it is a good fit for the small packets of sensor data typically sent by WSANs through their gateway.

While a great deal more work may be done on everything from critical infrastructure protection regulation to WSAN testing, this thesis demonstrates the process of identifying critical infrastructure, and prioritizing mitigation measures for vulnerabilities. It also provides performance data to assist in designing a system for vulnerability mitigation in the electrical grid. It was determined that:

- Current critical infrastructure protection regulations and the monitoring and control systems used to achieve them have not met the goals of the regulations.
- Improvement efforts should focus on transmission and distribution due to the relatively low use of automation and the large impact on reliability. Technical solutions will require regulatory, standards, and economic support, however.

- Wireless sensor networks, a cyber-physical technology originally designed as a "personal" area network could be used to improve CIP as part of the next generation of electrical grid controls.
- Characteristics of these systems are such that we should consider using them to improve critical infrastructure protection for the electrical grid
- System performance requires testing of a number of important metrics, and these are essential for designing the communications and monitoring systems.

9.1 Overall Future/Further Work

Advances in technology, and to some extent regulation, along with new threats and opportunities, are bringing about a revolution in the way electricity is generated, distributed, and consumed. Better monitoring, control, and communications are at the forefront of this revolution, and additional research is needed in every area from integrating renewable energy to studying consumer motivation.

In the area of transmission and distribution, additional work could be done on applications for underground conductors, transformer monitoring, or other distribution automation applications like sectionalizers, capacitors, or reclosers.

This thesis demonstrates one new application for power line monitoring, but as was shown, there are presently no performance requirements for the system. It would be useful to determine the bandwidth required, the quality of service and security desired, and the effects of latency, jitter, and interference, fading and link quality. These requirements could be tested with the ZigBee communication system proposed, and the performance of the communication system

could be compared against the performance requirements of the application. This should point to a desired network architecture, including suitable backhaul options.

While I focus primarily on the monitoring application and communications system, there is also a need to integrate monitoring with utility operations. This could involve optimizing preprocessing algorithms, an application for integrating the sensor data into the network operations center and dealing with the large amounts of data created, and the mechanism for using monitoring data to influence control of power distribution.

> *Experience with modern networks such as the Internet has shown that high system reliability and availability can be realized through redundancy, low-cost communications and distributed solutions. [31].*

10 References

[1] *USA PATRIOT Act*. 2001.

[2] Department of Homeland Security, "National infrastructure protection plan." Department of Homeland Security, 2009.

[3] V. Aravinthan, B. Karimi, V. Namboodiri, and W. Jewell, "Wireless communication for smart grid applications at distribution level -- feasibility and requirements," presented at the 2011 IEEE Power and Energy Society General Meeting, San Diego, CA, 2011, pp. 1–8.

[4] Z. Fan, G. Kalogridis, C. Efthymiou, M. Sooriyabandara, M. Serizawa, and J. McGeehan, "The new frontier of communications research: smart grid and smart metering," *ACM e-Energy*, p. 115, 2010.

[5] U.S. Energy Information Administration, "Annual energy outlook 2006 with projections to 2030." U.S. Department of Energy, Feb-2006.

[6] North American Electric Reliability Council (NERC), "2006 long-term reliability assessment: the reliability of the bulk power systems in North America." North American Electric Reliability Council (NERC), Oct-2006.

[7] K. Kliesen, "Electricity: the next energy jolt," *The Regional Economist, a publication of the Federal Reserve Bank of St. Louis*, pp. 4–9, Oct-2006.

[8] E. Santacana, T. Zucco, X. Feng, J. Pan, M. Mousavi, and L. Tang, "Power to be efficient: transmission and distribution technologies are the key to increased energy efficiency," *ABB Review, The Corporate Technical Journal of the ABB Group*, pp. 14–21, Feb-2007.

[9] IEEE Standards Coordinating Committee 22 on Power Quality, Institute of Electrical and Electronics Engineers, and IEEE Standards Board, *IEEE Recommended Practice for Monitoring Electric Power Quality*. New York, N.Y., USA: Institute of Electrical and Electronics Engineers, 1995.

[10] C. A. Warren, "IEEE benchmarking 2005 results," IEEE Power Engineering Society Working Group on Distribution Reliability, Montreal, Canada, Jul. 2006.

[11] W. Lusvardi, "Will blackouts darken California this summer?," *CalWatchDog*, 01-May-2012. [Online]. Available: http://www.calwatchdog.com/2012/05/01/will-blackouts-darken-calif-this-summer/. [Accessed: 25-Oct-2012].

[12] W. Pentland, "Clean energy tops agenda in Connecticut," *Forbes.com*, 09-Nov-2010. [Online]. Available: http://www.forbes.com/sites/williampentland/2010/11/09/microgrids/. [Accessed: 28-Oct-2012].

[13] U.S. Department of Energy, "The smart grid: an introduction." U.S. Department of Energy, 2008.

[14] R. J. Woolsey, R. H. Wilcox, and P. J. Garrity, *America's Hidden Vulnerabilities: Crisis Management in a Society of Networks: a Report of the Panel on Crisis Management of*

the CSIS Science and Technology Committee. Center for Strategic and International Studies, Georgetown University, 1984.

[15] T. G. Lewis, *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation*. Hoboken, N.J.: Wiley-Interscience, 2006.

[16] H. H. Willis, T. LaTourrette, T. Kelly, S. Hickey, and S. Neill, *Terrorism Risk Modeling for Intelligence Analysis and Infrastructure Protection. TR-386-DHS.* Santa Monica, CA: RAND Corporation, 2007.

[17] J. Moteff, "Risk management and critical infrastructure protection: assessing, integrating, and managing threats, vulnerabilities and consequences," Congressional Research Service, Washington, DC, RL32561, Feb. 2005.

[18] T. Masse, S. O'Neil, and J. Rollins, "The Department of Homeland Security's risk assessment methodology: evolution, issues, and options for Congress," Congressional Research Service, Washington, DC, RL33858, Feb. 2007.

[19] M. Foucault, "Dits et ecrits 1954-1988," *Gallimard*, 1994.

[20] President's Commission on Critical Infrastructure, "Critical foundations: protecting America's infrastructures, the report of the President's Commission on Critical Infrastructure Protection," *National Defense University*, Oct. 1997.

[21] S. Collier and A. Lakoff, "The vulnerability of vital systems: How 'critical infrastructure' became a security problem," *The Politics of Securing the Homeland: Critical Infrastructure, Risk and Securitisation*, pp. 40–62, 2008.

[22] G. W. Bush, "The national strategy for the physical protection of critical infrastructures and key assets." United States. White House Office., Feb-2003.

[23] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," *Computer Networks*, vol. 50, no. 7, pp. 877–897, May 2006.

[24] L. Meeus, M. Saguan, J. M. Glachant, and R. Belmans, "Smart regulation for smart grids," *European University Institute Working Papers*, vol. RSCAS, no. 2010/45, pp. 1–23, 2010.

[25] S. M. Amin, "Securing the electricity grid," *The Bridge, quarterly publication of the US National Academy of Engineering*, vol. 40, no. 1, pp. 13–20, 2010.

[26] U.S. Department of Energy, "Smart grid system report," Washington, DC, Jul. 2009.

[27] G. Locke and P. D. Gallagher, "NIST framework and roadmap for smart grid interoperability standards, release 1.0," *National Institute of Standards and Technology*, p. 33, 2010.

[28] E. A. Lee and S. A. Seshia, *Introduction to Embedded Systems: A Cyber-Physical Systems Approach*. Morrisville, NC: Lulu Enterprises, 2011.

[29] J. S. Engelhardt and S. P. Basu, "Design, installation, and field experience with an overhead transmission dynamic line rating system," in *Proceedings of the IEEE Transmission and Distribution Conference*, 1996, pp. 366–370.

[30] T. O. Seppa, "A practical approach for increasing the thermal capabilities of transmission lines," *Power Delivery, IEEE Transactions on*, vol. 8, no. 3, pp. 1536–1550, 1993.

[31] Y. Yang, D. Divan, R. G. Harley, and T. G. Habetler, “Power line sensornet - a new concept for power grid monitoring,” in *IEEE Power Engineering Society General Meeting*, 2006, p. 8.

[32] Y. Yang, F. Lambert, and D. Divan, “A survey on technologies for implementing sensor networks for power delivery systems,” in *IEEE Power Engineering Society General Meeting*, 2007, pp. 1–8.

[33] W. Wang, Y. Xu, and M. Khanna, “A survey on the communication architectures in smart grid,” *Computer Networks*, vol. 55, no. 15, pp. 3604–3629, 2011.

[34] M. Erol-Kantarci and H. T. Mouftah, “Wireless multimedia sensor and actor networks for the next generation power grid,” *Ad Hoc Networks*, vol. 9, no. 4, pp. 542–551, 2011.

[35] V. C. Gungor, B. Lu, and G. P. Hancke, “Opportunities and challenges of wireless sensor networks in smart grid,” *Industrial Electronics, IEEE Transactions on*, vol. 57, no. 10, pp. 3557–3564, 2010.

[36] V. C. Gungor and G. P. Hancke, “Industrial wireless sensor networks: Challenges, design principles, and technical approaches,” *Industrial Electronics, IEEE Transactions on*, vol. 56, no. 10, pp. 4258–4265, 2009.

[37] I. Howitt, W. W. Manges, P. T. Kuruganti, G. Allgood, J. A. Gutierrez, and J. M. Conrad, “Wireless industrial sensor networks: Framework for QoS assessment and QoS management,” *ISA Transactions*, vol. 45, no. 3, pp. 347–359, 2006.

[38] R. V. Kulkarni, A. Forster, and G. K. Venayagamoorthy, “Computational intelligence in wireless sensor networks: A survey,” *Communications Surveys & Tutorials, IEEE*, vol. 13, no. 1, pp. 68–96, 2011.

[39] S. J. Isaac, G. P. Hancke, H. Madhoo, and A. Khatri, “A survey of wireless sensor network applications from a power utility’s distribution perspective,” in *AFRICON 2011*, pp. 1–5.

[40] T. P. Hughes, *Networks of Power: Electrification in Western Society, 1880-1930*. Johns Hopkins University Press, 1993.

[41] Edison Electric Institute, “Glossary of electric industry terms.” Edison Electric Institute, Apr-2005.

[42] Oncor Electric Delivery Company LLC, “Power restoration,” *Oncor.com*, n.d. [Online]. Available: http://www.oncor.com/EN/Pages/Power-Restoration.aspx. [Accessed: 20-Sep-2012].

[43] T. Kuphaldt, *Lessons in Electric Circuits Volume II - AC*, 6th ed., vol. II, VI vols. Open Book Project, 2012.

[44] A. von Meier, *Electric Power Systems: A Conceptual Introduction*. Hoboken, N.J.: IEEE Press : Wiley-Interscience, 2006.

[45] H. Henriksen, “Ohms law wheel,” *Wikimedia Commons*, 17-Jan-2010. [Online]. Available: http://commons.wikimedia.org/wiki/File:Ohms_law_wheel_PURI.svg. [Accessed: 25-Sep-2012].

[46] E. Lerner, “What’s wrong with the electric grid?,” *The Industrial Physicist*, vol. 9, no. 5, pp. 10–11, Nov-2003.

[47] A. Abel, “Electric transmission: approaches for energizing a sagging industry,” Congressional Research Service, Washington, DC, RL33875, Jan. 2008.

[48] G. Constable and B. Somerville, *A Century of Innovation: The Engineering That Transformed Our Lives*. Joseph Henry Press, 2003.

[49] G. N. Ericsson, “Cyber security and power system communication—essential parts of a smart grid infrastructure,” *Power Delivery, IEEE Transactions on*, vol. 25, no. 3, pp. 1501–1507, 2010.

[50] Automation, Plc, “All about scada,” *Automation*. 11-Jun-2011.

[51] Oncor Electric Delivery Company LLC, “Smart grid technology,” *Oncor.com*, n.d. [Online]. Available: http://www.oncor.com/EN/Pages/Smart-Grid-Technology.aspx. [Accessed: 31-Oct-2012].

[52] M. Dickinson, “Are your transformers ready for the smart grid?,” *Electric Light & Power*, vol. 16, no. 2, 01-Feb-2011.

[53] International Electrotechnical Commission, *IEC 61970-301 Energy Management System Application Program Interface (EMS-API), Part 301: Common Information Model (CIM) Base*, 2.0 ed. Geneva, Switzerland: International Electrotechnical Commission, 2009.

[54] International Electrotechnical Commission, *IEC 61968 Application Integration at Electric Utilities–System Interfaces for Distribution Management Part 11: Common Information Model (CIM)*. Geneva, Switzerland: International Electrotechnical Commission, 2003.

[55] R. Mackiewicz, T. Saxton, and R. Rhodes, “Introduction to CIM and its role in the utility enterprise: data preparation, exchange, integration, and enterprise information management,” UCA International CIM User Group Meeting, Austin, TX, Oct. 2007.

[56] D. Watts, “Security & vulnerability in electric power systems,” in *Proceedings of 35th North American Power Symposium*, University of Missouri-Rolla in Rolla, Missouri, 2003, pp. 559–566.

[57] National Science Foundation, “NSF workshop on cyber-physical systems,” *NSF Research Initiative on Cyber-Physical Systems*, 16-Oct-2006. [Online]. Available: http://varma.ece.cmu.edu/cps/. [Accessed: 24-May-2012].

[58] Central Intelligence Agency, “CIA - the world factbook,” *CIA.gov*, 30-Mar-2012. [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html. [Accessed: 05-Apr-2012].

[59] ICT Data and Statistics Division, Telecommunication Development Bureau, International Telecommunication Union, “The world in 2011 - ICT facts and figures.” International Telecommunications Union (ITU), 2011.

[60] Ericsson, “More than 50 billion connected devices.” White Paper. Telefonaktiebolaget LM Ericsson, Feb-2011.

[61] L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, 4th ed. Amsterdam; Boston: Morgan Kaufmann, 2007.

[62] Z. Shelby and C. Bormann, *6LoWPAN: The Wireless Embedded Internet*. Wiley, 2011.

[63] P. Marshall, "Fresh approaches to spectrum sharing," Presentation. International Symposium on Advanced Radio Technologies (ISART), Boulder, Colorado, Jul. 2012.

[64] International Society of Automation ISA100 Committee, "ISA100, wireless systems for automation, ISA 100 committee scope," *ISA.org*, n.d. [Online]. Available: http://www.isa.org//MSTemplate.cfm?MicrositeID=1134&CommitteeID=6891. [Accessed: 27-Sep-2012].

[65] Institute of Electrical and Electronics Engineers (IEEE), "IEEE Standard 802.15.4-2011 Wireless medium access control (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (WPANs)." IEEE, 05-Sep-2011.

[66] ERA Technology Ltd., University of Surrey, and University of Hull, "Electromagnetic compatibility of radio-based mobile telecommunications systems," LINK Personal Communications Programme, Surrey, England, 1999.

[67] C. "Akiba" Wang, "IEEE 802.15.4 in the context of ZigBee - part 2 - MAC layer," *Freaklabs - Open Source Wireless*, 15-Dec-2008. [Online]. Available: http://freaklabs.org/index.php/Blog/Zigbee/IEEE-802.15.4-in-the-Context-of-Zigbee-Part-2-MAC-Layer.html. [Accessed: 07-Nov-2012].

[68] ZigBee Standards Organization, "ZigBee-2007 r17 specification." ZigBee Standards Organization, 19-Oct-2007.

[69] S. R. Das, E. M. Belding-Royer, and C. E. Perkins, "RFC 3561. Ad hoc on-demand distance vector (AODV) routing.," *Internet Engineering Task Force (IETF)*, Jul. 2003.

[70] B. Tian, S. Han, L. Liu, S. Khadem, and S. Parvin, "Towards enhanced key management in multi-phase ZigBee network architecture," *Computer Communications*, vol. 35, no. 5, pp. 579–588, Mar. 2012.

[71] Digi International, Inc., "Product Manual: XBee®/XBee-PRO® ZB RF Modules." Digi International, Inc., Mar-2012.

[72] N. Kushalnagar, G. Montenegro, D. E. Culler, and J. W. Hui, "RFC 4944. Transmission of IPv6 packets over IEEE 802.15. 4 networks.," *Internet Engineering Task Force (IETF) Network Working Group*, Sep. 2007.

[73] University of California Berkeley OpenWSN Project, "OpenLowPan – Berkeley's openWSN project," *OpenWSN: Implementing the Internet of Things*, 10-Mar-2012. [Online]. Available: http://openwsn.berkeley.edu/wiki/OpenLowPan#a6LoWPAN. [Accessed: 08-Nov-2012].

[74] M. Crawford, "RFC 2464. Transmission of IPv6 packets over Ethernet networks.," *Internet Engineering Task Force (IETF) Network Working Group*, Dec. 1998.

[75] Cosm, Ltd., "Cosm - Internet of things platform connecting devices and apps for real-time control and data storage," *Cosm.com*, n.d. [Online]. Available: https://cosm.com/. [Accessed: 08-Nov-2012].

[76] Y. Yang, D. Divan, R. G. Harley, and T. G. Habetler, “Design and implementation of power line sensornet for overhead transmission lines,” in *IEEE Power & Energy Society General Meeting*, 2009, pp. 1–8.

[77] IEEE Computer Society. LAN/MAN Standards Committee, Institute of Electrical and Electronics Engineers, and IEEE-SA Standards Board, *IEEE Standard for Information Technology Telecommunications and Information Exchange Between Systems--Local and Metropolitan Area Networks--Specific Requirements. Part 15.4, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*. New York, NY: Institute of Electrical and Electronics Engineers, 2006.

[78] K. Masica, “Recommended practices guide for securing ZigBee wireless networks in process control system environments.” Lawrence Livermore National Laboratory for the DHS US CERT Control Systems Security Program, Apr-2007.

[79] ZigBee Alliance, “ZigBee smart energy features.” ZigBee Alliance, n.d.

[80] ZigBee Alliance, “ZigBee smart energy FAQ,” *ZigBee.org*, n.d. [Online]. Available: http://www.zigbee.org/Standards/ZigBeeSmartEnergy/FAQ.aspx. [Accessed: 05-Apr-2012].

[81] DASH7 Alliance, Inc., “DASH7 Alliance home page,” *DASH7.org*, n.d. [Online]. Available: http://www.dash7.org/. [Accessed: 05-Apr-2012].

[82] Z-Wave Alliance, “Z-Wave Alliance home page,” *z-wavealliance.org*, n.d. [Online]. Available: http://www.z-wavealliance.org/. [Accessed: 05-Apr-2012].

[83] M. Conti and S. Giordano, “Multihop ad hoc networking: The reality,” *Communications Magazine, IEEE*, vol. 45, no. 4, pp. 88–95, 2007.

[84] V. Mayalarp, N. Limpaswadpaisarn, T. Poombansao, and S. Kittipiyakul, “Wireless mesh networking with XBee,” presented at the 2nd ECTI-Conference on Application Research and Development (ECTI-CARD 2010), Pattay, Chonburi, Thailand, 2010.

[85] T. R. Burchfield, S. Venkatesan, and D. Weiner, “Maximizing throughput in ZigBee wireless networks through analysis, simulations and implementations,” in *Proceedings of the International Workshop on Localized Algorithms and Protocols for Wireless Sensor Networks*, Santa Fe, New Mexico, 2007, pp. 15–29.

[86] U.S. Department of Energy, “Communications requirements of smart grid technologies.” U.S. Department of Energy, 05-Oct-2010.

[87] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[88] P. Gupta, R. Gray, and P. R. Kumar, “An experimental scaling law for ad hoc networks.” University of Illinois at Urbana-Champaign, 16-May-2001.

[89] I. Gragopoulos, I. Tsetsinas, E. Karapistoli, and F.-N. Pavlidou, “FP-MAC: A distributed MAC algorithm for 802.15.4-like wireless sensor networks,” *Ad Hoc Networks*, vol. 6, no. 6, pp. 953–969, Aug. 2008.

[90] Arduino, “Arduino home page,” *Arduino.cc*, 06-Feb-2012. [Online]. Available: http://www.arduino.cc/. [Accessed: 07-May-2012].

[91] Arduino, “Arduino - ArduinoBoardDiecimila,” *Arduino.cc*, 05-Feb-2010. [Online]. Available: http://www.arduino.cc/en/Main/arduinoBoardDiecimila. [Accessed: 12-Nov-2012].

[92] Market Information and Statistics Division, Telecommunication Development Bureau (ITU-D), “The world in 2010: ICT facts and figures,” International Telecommunications Union (ITU), Geneva, Switzerland, 2010.

[93] B. Ciubotaru-Petrescu, D. Chiciudean, R. Cioarga, and D. Stanescu, “Wireless solutions for telemetry in civil equipment and infrastructure monitoring,” in *3rd Romanian-Hungarian Joint Symposium on Applied Computational Intelligence (SACI) May*, 2006, pp. 25–26.

[94] E. Paulos, R. J. Honicky, and E. Goodman, “Sensing atmosphere,” *Human-Computer Interaction Institute*, p. 203, 2007.

[95] H. Cole-Lewis and T. Kershaw, “Text messaging as a tool for behavior change in disease prevention and management,” *Epidemiologic Reviews*, vol. 32, no. 1, pp. 56–69, 2010.

[96] Federal Communications Commission (FCC), “Facilitating the deployment of text-to-911 and other next generation 911 applications PS docket no. 11-153, framework for next generation 911 deployment PS Docket no. 10-255, notice of proposed rulemaking FCC 11-134.” Federal Communications Commission, 2011.

[97] G. Peersman, P. Griffiths, H. Spear, S. Cvetkovic, and C. Smythe, “A tutorial overview of the short message service within GSM,” *Computing & Control Engineering Journal*, vol. 11, no. 2, pp. 79–89, 2000.

[98] S. Collesei, P. Di Tria, and G. Morena, “Short message service based applications in the GSM network,” in *Personal, Indoor and Mobile Radio Communications, 1994. Wireless Networks-Catching the Mobile Future., 5th IEEE International Symposium on*, 1994, vol. 3, pp. 939–943.

[99] C.-L. Tseng, J.-A. Jiang, R.-G. Lee, F.-M. Lu, C.-S. Ouyang, Y.-S. Chen, and C.-H. Chang, “Feasibility study on application of GSM–SMS technology to field data acquisition,” *Computers and Electronics in Agriculture*, vol. 53, no. 1, pp. 45–59, Aug. 2006.

[100] S. M. Redl, M. K. Weber, and M. W. Oliphant, *An Introduction to GSM.* Boston, Mass.: Artech House, 1995.

[101] H.-N. Hung, Y.-B. Lin, M.-K. Lu, and N.-F. Peng, “A statistic approach for deriving the short message transmission delay distributions,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2345–2352, Nov. 2004.

[102] M. Werner, K. Kamps, U. Tuisel, J. G. Beerends, and P. Vary, “Parameter-based speech quality measures for GSM,” in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, 2003, vol. 3, pp. 2611–2615.

[103] O. Rensfelt, F. Hermans, L.-A. Larzon, and P. Gunningberg, “Sensei-UU: a relocatable sensor network testbed,” in *Proceedings of the Fifth ACM International Workshop on*

Wireless Network Testbeds, Experimental Evaluation and Characterization, 2010, pp. 63–70.

[104] National Emergency Number Association, "NENA master glossary of 9-1-1 terminology, v. 16." National Emergency Number Association, 22-Aug-2011.

[105] C. W. Johnson, "Public policy and the failure of national infrastructures," *International Journal of Emergency Management*, vol. 4, no. 1, p. 18, 2007.

[106] A. Abel, "Electric reliability: options for electric transmission infrastructure improvements," Congressional Research Service, Washington, DC, RL32075, Dec. 2004.

[107] American Society of Civil Engineers, *2009 Report Card for America's Infrastructure.* Reston, Va.: American Society of Civil Engineers, 2009.

[108] J. Fershee, "Misguided energy: why recent legislative, regulatory, and market initiatives are insufficient to improve the U.S. energy infrastructure," *Harvard Journal on Legislation*, vol. 44, no. 2, pp. 328–330, Summer 2007.

[109] The Clinton Administration, "Presidential decision directive 63 (PDD-63): policy on critical infrastructure protection." US Office of the Federal Register, 22-May-1998.

[110] North American Reliability Corporation, "Reliability standards for the bulk electric systems of North America." North American Reliability Corporation, 21-Aug-2012.

[111] U.S. Energy Information Administration, "Electric power industry overview 2007," *EIA.gov*. [Online]. Available: http://www.eia.gov/cneaf/electricity/page/prim2/toc2.html. [Accessed: 13-Sep-2012].

[112] Homeland Security Council, "National strategy for homeland security." U.S. Government, Oct-2007.

[113] P. Behr, "Md.'s veto of advanced meter deployment stuns smart grid advocates," *The New York Times*, New York, NY, 23-Jun-2010.

[114] S. Neumann, "Overview of smart grid standards and dependencies." National Institute of Standards and Technology, 27-Jul-2009.

[115] M. Holt and C. Glover, "Energy policy act of 2005: summary and analysis of enacted provisions," Congressional Research Service, Washington, DC, RL33302, Mar. 2006.

[116] J. Morrison and D. Broad, "Wind interconnection: bridging the divide," *Electric Light & Power*, vol. 84, no. 3, pp. 32–33, May-2006.

[117] J. Weiss, "Electric power 2008 – is NERC CIP compliance a game?," *ControlGlobal.com*, 09-May-2008. [Online]. Available: http://community.controlglobal.com/content/electric-power-2008%E2%80%93-nerc-cip-compliance-game. [Accessed: 28-Oct-2012].

[118] R. Anderson and S. Fuloria, "Security economics and critical national infrastructure," in *Economics of Information Security and Privacy*, T. Moore, D. Pym, and C. Ioannidis, Eds. Springer US, 2010, pp. 55–66.

11 Appendix - The Law and Regulation of Critical Infrastructure Protection of the Electrical Grid

All industries are subject to certain laws and regulations that guide the actions they can take. In order to discuss the primary entities responsible for regulating the grid and viable solutions to improvements in the security of the U.S. power grid, this chapter provides a review of 1) the key Federal statutes (laws) that apply to the electrical power industry, 2) Presidential Directives, 3) Regulations issued by the Federal Energy Regulatory Commission (FERC), and 4) U.S. and international standards. It is followed by: 5) a discussion of some of the challenges, and unintended consequences, related to regulation, and 6) the current state of legislative efforts to create consistent, relevant standards.

11.1 Key Federal Statutes

A sampling of the relevant statutes that govern critical infrastructure protection (CIP) for the U.S. electrical system are summarized below, although some apply to other sectors as well. .

11.1.1 Critical Infrastructure Protection

The process of critical infrastructure protection (CIP) is often presented as a framework of risk-based assessment and prioritization [2] [15] [16] [17] [18]. It can be generalized into steps similar to these:

- Identify critical infrastructures
- Identify potential threats to the infrastructures
- Determine infrastructure vulnerability
- Assess the risks and probability of losses
- Identify and prioritize measures which address the risks

- Implement measures to address the risks

The Department of Homeland Security currently expresses risk assessment as [2]:

$$R = f(C, V, T) \tag{4}$$

where risk (R) is a function (f) of consequence (C), vulnerability (V), and threat (T). The composition and weighting of the risk formula elements has evolved over time [5].

The process must also consider the potential objectives of an attacker. Critical infrastructure may be destroyed, or perhaps simply incapacitated in a manner that creates chaos or catastrophe without permanently damaging the infrastructure itself, such as with a denial of service attack. It is also possible that infrastructure may be exploited to multiply or further the efforts of an attacker without otherwise affecting its function [22]. The notions of continuity and interconnected systems are essential to understanding the larger picture of critical infrastructure protection, as is the concept of proactive assessment and protection which goes beyond measures which are simply defensive.

The electrical grid is vulnerable to more than motivated and malicious attackers. Many components of the grid are nearing the end of their designed service life, and a disproportionately large portion of the workforce that maintains the grid is approaching retirement. At the same time the demand for electricity is rising, and new demands such as bidirectional energy flows and the integration of renewables are being placed on the grid. Lightning strikes and other weather-related threats are unavoidable, and even the slow and natural process of vegetation encroaching on hundreds of thousands of miles of power line requires constant vigilance.

11.1.2 Federal Power Act (FPA) of 1920, as amended

This act created the Federal Power Commission, which later became the Federal Energy Regulatory Commission (FERC), the agency responsible for regulating all interstate electricity transmission. Together with later amendments such as the *Public Utility Regulatory Policies Act (PURPA) of 1978* and the *Energy Policy Act of 1992*, the Secretary of Energy is given authority over reliability of the interstate electric transmission system. FERC is given the authority to create reliability standards, and the Department of Energy (DOE) is given the authority to make reliability and security recommendations and gather reliability data. In times of war or emergency the Secretary of Energy is authorized to require interconnection of any facilities or assets necessary to address the state of emergency.

11.1.3 Federal Power Act, 16 U.S.C. §§ 791a-825r; Public Utility Regulatory Policies Act, 16 U.S.C. § 2705; DOE Organization Act, 42 U.S.C. §§ 7101-7352; 18 C.F.R. Parts 4, 12, and 16; MOU between FERC, Army Corps of Engineers and Bureau of Reclamation

These documents designate FERC as the agency responsible for overseeing non-federal hydropower projects, but recognize the variety of interests and agencies which have a role in the permitting, construction, and operation of hydropower facilities.

11.1.4 Communications Act of 1934, as amended and Executive Order 13618 – Assignment of National Security and Emergency Preparedness Communications Functions, July 6, 2012

Executive Order 12472 (1984) stated that through the DOE, energy providers may request priority access to, or installation or repair of, telecommunications services during an emergency. *EO 12472* was repealed by President Obama's EO 13618 of July 6, 2012, which

transferred responsibility for all non-military communications and critical infrastructure protection networks to the Executive Branch and the Secretary of Homeland Security. While altering the responsible agency, similar priority access to telecommunications could be justified by the energy sector under the new EO.

11.1.5 Defense Production Act (DPA) of 1950, as amended

The Defense Production Act authorizes the Secretaries of Energy and Commerce to take measures necessary for national defense, emergency preparedness and relief, and critical infrastructure protection and assurance, including maximizing domestic energy supplies.

11.1.6 Department of Energy Organization Act, 1977

The Department of Energy Organization Act reorganized oversight of energy activities by a variety of agencies into a single Department of Energy. One example of their authority is the ability to require reporting of actual or potential incidents or emergencies that could affect electric system reliability or national security.

11.1.7 Robert T. Stafford Disaster Relief and Emergency Assistance Act, as amended, 1988

Robert T. Stafford was both a Representative and a Senator from Vermont and, while in Congress, helped to pass a law known as the Robert T. Stafford Disaster Relief and Emergency Assistance Act to coordinate federal natural disaster assistance. It promotes the creation of disaster preparedness plans and intergovernmental cooperation related to disaster planning and response, including making the DOE responsible for assisting in restoring energy systems.

11.1.8 Energy Policy Act of 1992

Widely known as an energy conservation act, the Energy Policy Act also amended the *Public Utility Holding Company Act*, the *Federal Power Act*, and the *Public Utility Regulatory Policies Act* which had important consequences for utility competition and deregulation. As a consequence, it is argued that utility companies were forced to reduce infrastructure spending in order to remain competitive [105]. Many experts currently agree that utilities are not investing adequately in electric infrastructure and security [106][107] [108].

11.1.9 Protected Critical Infrastructure Information (PCII) Program of the Critical Infrastructure Information Act of 2002

The Protected Critical Infrastructure Information (PCII) Program of the Critical Infrastructure Information Act is a program that allows the private sector to submit information regarding critical infrastructure, such as the location of key assets, to the Department of Homeland Security and provides guidelines for insuring the submitted information will remain confidential.

11.1.10 Energy Policy Act of 2005, Public Law 109-58, Title XII: Electricity, Subtitle A: Reliability Standards, Section 1211: Electric Reliability Standards; Electricity Modernization Act of 2005

Section 1211 of Title XII of the U.S. Code authorizes federal jurisdiction over the reliability of the bulk power systems and directs the Federal Energy Regulatory Commission (FERC) to designate a national Electric Reliability Organization (ERO). The ERO will create and enforce mandatory standards for reliability, subject to review and approval by FERC.

11.1.11 Energy Independence and Security Act of 2007

While primarily a law promoting clean energy standards, the Energy Independence and Security Act 02 2007 included provisions addressing the need to create standards and improve security in the move toward a smarter electrical grid, and created a Smart Grid Task Force to make recommendations on everything from policy to reliability.

11.2 Presidential Directives

Among the first federal documents to define a coordinated national policy toward modern critical infrastructure protection was the *1998 Presidential Decision Directive 63* [109], which originally identified eight areas of critical infrastructure, along with a federal agency responsible for leading protection efforts in each area. Being designated the lead agency does not necessarily imply that the agency has direct regulatory authority over the sector, however, and efforts may require the cooperation of a number of other regulatory agencies. While the DHS *National Infrastructure Protection Plan* places responsibility for protecting private sector critical infrastructure on the entities which own or operate it [2], it is not always clear how this will be enforced. This list of critical infrastructure sectors and agencies has subsequently seen deletions, additions, and modifications such that there are currently eighteen sectors [2].

TABLE 11-1. CRITICAL INFRASTRUCTURE SECTORS AND LEAD AGENCIES

Sector	Lead Agency
Agriculture and Food	Department of Agriculture/Department of Health and Human Services
Banking and Finance	Department of the Treasury
Chemical	Department of Homeland Security
Commercial Facilities	Department of Homeland Security
Communications	Department of Homeland Security
Critical Manufacturing	Department of Homeland Security
Dams	Department of Homeland Security
Defense Industrial Base	Department of Defense
Emergency Services	Department of Homeland Security
Energy	Department of Energy
Government Facilities	Department of Homeland Security
Healthcare and Public Health	Department of Health and Human Services
Information Technology	Department of Homeland Security
National Monuments and Icons	Department of the Interior
Nuclear Reactors, Materials and Waste	Department of Homeland Security
Postal and Shipping	Department of Homeland Security
Transportation Systems	Department of Homeland Security
Water	Environmental Protection Agency

11.2.1 Homeland Security Presidential Directive 5 (HSPD-5), February 28 2003

Issued under George W. Bush, the Homeland Security Presidential Directive 5 established the National Incident Management System (NIMS), a uniform system for managing domestic disasters and emergencies. It recognized: 1) the role of local governments and non-governmental organizations in preparing for and mitigating disasters, and 2) the need for partnerships in planning for both.

11.2.2 Homeland Security Presidential Directive 7 (HSPD-7), December 17, 2003

Issued under George W. Bush, the Homeland security Presidential Directive 7 directed all federal agencies to: 1) identify and prioritize critical infrastructures, 2) protect the information, 3) coordinate their actions to protect critical infrastructures, and 4) work with all state and local governments and private sector entities to carry out the Directive.

11.3 Orders of the Federal Energy Regulatory Commission (FERC)

11.3.1 FERC Orders 630 (February 21, 2003) and 630a (July 23, 2003), Critical Energy Infrastructure Information

FERC Order 630 and 630a define critical infrastructure and critical energy infrastructure information, and provide guidelines for protecting and accessing this information.

11.3.2 FERC Order Issued in Docket No. RR06-1-000, Certifying NERC as the Electric Reliability Organization, July 20, 2006

This FERC order established the North American Electric Reliability Corporation (NERC), a self-regulatory non-governmental organization, as the national Electric Reliability Organization (ERO).

11.3.3 FERC Order 693 Mandatory Reliability Standards for the Bulk-Power System, March 16, 2007

In this order FERC approved 83 of the 107 reliability standards proposed by NERC. As of this writing, the remaining 24 standards are still under review.

11.3.4 FERC Rulemaking RM 12-6-000 and RM 12-7-000 Revisions to Electric Reliability Organization Definition of Bulk Electric System and Rules of Procedure, June 22, 2012

This rule establishes a definition of the bulk electrical system as all facilities that are operated at or above 100 kV.

11.4 North American Electric Reliability Corporation (NERC) and Local Regulation

NERC was created by the electric utility industry in 1968 with the goal of promoting reliable operation of bulk power electric transmission. It is now an independent non-profit agency and the Electric Reliability Organization (ERO) for the bulk power system as designated by FERC. NERC focuses on reliability standards, compliance, enforcement, assessment, event analysis, utility communication and cooperation, infrastructure security, establishing benchmarks, education, and certification.

NERC reliability standards focus on two areas: adequacy and security. Adequacy is the ability of the power grid to meet all demand at all times, although they cannot order the construction of new generation or transmission facilities. Security is defined as the ability of the grid to continue functioning after the loss of any component, or after the most severe single contingency, generally known as the "N minus 1 reliability requirement."

It should be noted that the "bulk power system" is federally defined as 100 kV and above, so the NERC standards apply only to the generation and transmission segments of the electric grid. They do not apply to the distribution or consumer load segments. The NERC publishes and enforces its standards that address the following fourteen areas of grid operation reliability [110]:

- Resource and Demand Balancing
- Communications

- Critical Infrastructure Protection[15]
- Emergency Preparedness and Operations
- Facilities Design, Connections, and Maintenance
- Interchange Scheduling and Coordination
- Interconnection Reliability Operations and Coordination
- Modeling, Data, and Analysis
- Nuclear
- Personnel Performance, Training, and Qualifications
- Protection and Control
- Transmission Operations
- Transmission Planning
- Voltage and Reactive

The electrical system is in many respects a natural monopoly in each local area. Due to economies of scope and scale it is more efficient to build a single large power generation plant than two side-by-side competitors. Historically, with large coal, hydro or nuclear generation plants, it made more sense to build a single transmission and distribution network than to provide multiple sets of wires between the producers and consumers of power. To further complicate the issue, instead of a single large national monopoly there are instead a large number of local monopolies. These may exist under different organizational structures, including various combinations of public and private ownership. According to the U.S. Energy Information

[15] NERC terminology does not align well with what is otherwise considered critical infrastructure. NERC defines "Reliability Standards" in fourteen separate areas, with a failure in any one of them posing a threat to the reliability of the bulk power system, a system widely regarded as "critical infrastructure." One of the fourteen areas is termed "Critical Infrastructure Protection," yet it deals exclusively with cyber security, a domain far too narrow to address the full scope of CIP. I am unable to find an explanation for the choice of nomenclature.

Administration (EIA) [111], in 2007 there were 2009 non-profit, publicly-owned electric utilities, 883 member-owned co-operatives, 210 investor-owned utilities (IOU), and 9 federal electric utilities such as the Tennessee Valley Authority. Overall, some 85% of the national critical infrastructure is privately owned or operated [112].

Because of their monopoly position and critical infrastructure status, these entities must be, and are, regulated. While there are common regulatory entities, such as the U.S. Department of Energy (DOE) and the Federal Energy Regulatory Commission (FERC), even at the federal level these quickly fragment. Most generation and transmission entities, as they are above the 100 kV threshold for "bulk power" and operate interstate are subject to the regulations of the North American Electric Reliability Corporation (NERC), however nuclear power plants are primarily under the jurisdiction of the Nuclear Regulatory Commission (NRC). When discussing energy infrastructure the Bush Administration document titled *The National Strategy for the Physical Protection of Critical Infrastructures and Key Assets* states:

> *Typically, these companies seek to recover the costs of new security investments through proposed rate or price increases. Under current federal law, however, there is no assurance that electricity industry participants would be allowed to recover the costs of federally mandated security measures through such rate or price increases.* [22].

Smaller electric utilities engaged in local distribution are primarily subject to the requirements of state, local, and tribal Public Utility Commissions (PUCs), which may impose widely-varying regulations and rate calculations upon individual utilities. It is not uncommon for PUCs to focus on keeping the rate base as small as possible in order to minimize charges to the end user. This often involves amortizing assets over a longer period, which discourages upgrades and replacements with newer, more effective, or more capable technology which could improve operations or infrastructure protection. Behr [113] reports a case where the Maryland Public Service Commission forced Baltimore Gas & Electric to reject a $200 million grant for

advanced metering infrastructure (AMI) from the DOE for fear that it might increase rates for some customers.

11.5 The Role of Standards

Organizations such as the Institute of Electrical and Electronics Engineers (IEEE), the National Institute of Standards and Technology (NIST) and the International Society of Automation (ISA) all publish recommendations and standards that relate to the electric grid. Standards cover equipment specifications, communications protocols, cyber security and other topics that vary widely in their scope and focus. While standards are not in themselves regulations, they may be incorporated "by reference" such that compliance with the standard becomes a required part of the regulation.

While intended to improve grid reliability and functionality in ways far beyond just critical infrastructure protection, standards themselves have posed problems for the electric grid. When control systems were first being implemented, there were no or few required standards, and manufacturers desired to implement their own proprietary designs. Regulation or market forces have moved certain elements of control systems closer to integration, but this is not necessarily reflected in older systems. It has also been noted that the future smart grid will require a communications infrastructure more complex than any current control system, and one which may incorporate new or existing standards for other systems. The current relationship of some of these standards is depicted by Neumann [114] in Figure 11-1. This figure demonstrates that we have gone from a lack of standards to a complicated web[16] of arguably too many, with more under development, as there are many areas of smart grid communications that remain to

[16] Utilities do not necessarily fare any better on the regulatory side. The Energy Policy Act of 2005 added 550 pages of new regulations. Preparing a summary of the major points reportedly took a Congressional Research Service staff of 19 and required 152 pages [115].

be defined. There is no standard which defines how to connect to the electrical grid, for instance, an issue of increasing relevance as greater quantities of distributed and alternative generation are introduced [116]. Also notable is the lack of a unifying standard which provides direction on how and when to apply the disparate options for a common and interoperable architecture. While the diagram depicts efforts which arguably reach beyond just critical infrastructure protection, they all ultimately form the same ecosystem.

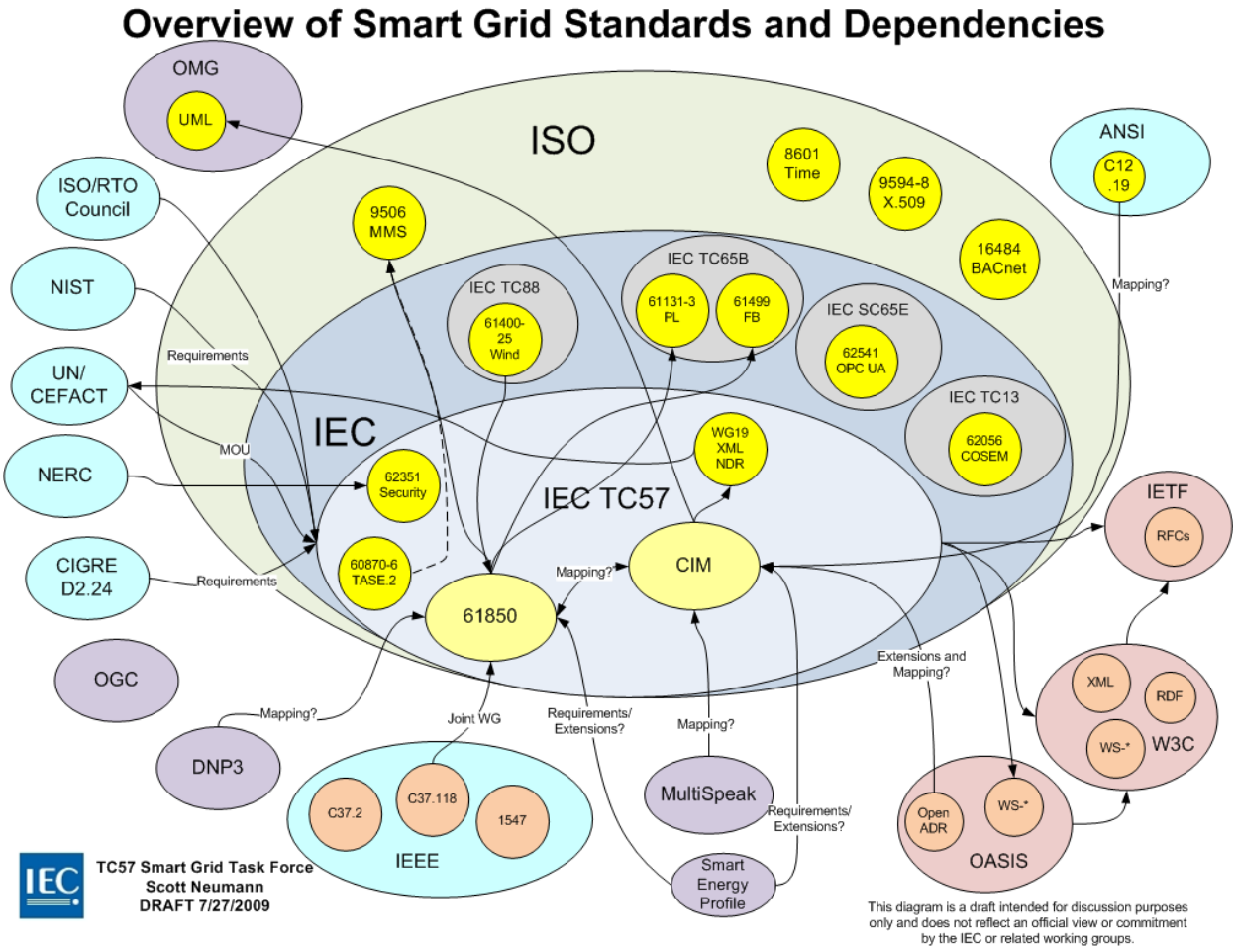


Figure 11-1 Overview of smart grid standards and dependencies [114].

11.6 Unintended Consequences of Regulation

Due to the nature of regulatory sanctions, power system operators are often more focused on compliance with regulation than on solving the underlying infrastructure protection issues regulators hoped to address. Two examples of this are the elimination of black start capability, and the maintenance of or return to insecure serial communication protocols in an effort to avoid regulation on IP-based communication.

11.6.1.1 Elimination of Black Start Capability

A black start is required when an offline generator is to be returned to operation without relying on the external power grid. This may occur under a number of conditions, including a widespread grid outage where power from an external tie-line may not be available. Independent power may also be required for related processes, such as feeding and preparing coal for combustion or running boiler pumps at a thermal generation facility. A black start normally begins by starting an on-site diesel generator using a battery, which is then used to re-establish the generating field in a small generator or hydroelectric plant. This additional generation can then be used to bring larger generators back online, perhaps even remotely via electrical tie-lines between generating facilities. Not all generators or plants are capable of a black start. One author notes the case of plant managers removing black start capability from their generators to avoid the cost and potential liability of compliance with NERC regulations related to the maintenance of this capability [117].

11.6.1.2 Secure Communication Protocols

Many control systems run over older unencrypted serial communication protocols, as do remote communication capabilities using dial-up modems. The opportunity exists to convert many of these communications to encrypted IP-based protocols. As currently written, Internet

Protocol (IP) based communications are required to comply with NERC standards, but serial communications are exempt from compliance requirements. As a consequence, utility operators are in many cases foregoing the opportunity to migrate to a more secure and capable protocol to avoid compliance activities and potential liability. One author even notes cases of IP-based communications being converted back to serial in order to avoid having to comply with regulations [118]. The same publication notes that some have called NERC CIP a "giant exercise in avoidance."

11.7 Regulatory Conclusions

As can be seen from the foregoing, regulation concerning critical infrastructure protection underscores its importance, is ever-evolving, and is not limited to a single agency or entity. The disparate regulations also serve to highlight the difficulties of determining who is responsible for critical infrastructure, and even in determining what infrastructure is critical. Beyond these two basic questions lie the issues of prioritizing the importance of CI, and determining how it should be protected. While power systems around the world may face a less fragmented regulatory environment, they all face similar issues of identification, prioritization, and protection.

Due to the nature of regulatory sanctions, power system operators are often more focused on compliance with regulation than on solving the underlying infrastructure protection issues regulators hoped to address. The evolving nature of threats to the power system include aging infrastructure, increasing demand, greater complexity, more informed and motivated attackers, and a larger number of attack vectors.

It is hard not to note the similarities between the monopoly years of the U.S. telecommunications industry and the current electric industry, particularly from a regulatory

standpoint. Much as telecommunications faced revolutionary changes following the breakup of the monopoly, it seems that the electric power industry is poised to undergo a transformation unlike anything it has seen in the last century. Clearly this will need to include a radical change in the way communication and control has traditionally been handled, but the lessons learned from restructuring the telecommunications industry and the rise of the Internet may at least provide a model as to how that change will take place. While this paper will attempt to show that the addition of wireless sensor actuator networks (WSANs) is a technically feasible and relatively simple and inexpensive method of enhancing electrical grid protection, it is unlikely that any additional improvements in this area will come about without a comprehensive regulatory and standards effort, possibly combined with investment incentives.