Title
Interoperable and Secure Communication for Cyber Physical Systems in the Energy Grid

Author
Lee, Eun Kyu

Publication Date
2014

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Interoperable and Secure Communication for Cyber Physical Systems in the Energy Grid

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Computer Science

by

Eun Kyu Lee

2014

Abstract of the Dissertation

Interoperable and Secure Communication for Cyber Physical Systems in the Energy Grid

by

Eun Kyu Lee

Doctor of Philosophy in Computer Science

University of California, Los Angeles, 2014

Professor Mario Gerla, Chair

A smart grid system constitutes a control loop, and recent research has tried to implement the loop using a cyber-physical system (CPS) model integrating physical processes (i.e., power flow) with networked computing capabilities. More specifically, the system aims to read/control energy resources in a physical domain for the purpose of energy balance. This requires each resource to be capable of accepting control command messages as well as sending out data, making bi-directional interactions with external entities. Regarding the interactions, we pose two research questions. (1) How to enable the interactions? (2) How to secure the interoperation? To answer the questions, this dissertation designs and develops an interoperable and secure communication model for cyber-physical systems and extends it to examine the realization of smart grid interoperation, the Internet of Energy (IoE). The contribution of the dissertation is three-fold. First, we develop a generic middleware model in which an object represents various physical functions and inter-communicates with other objects in a unified manner, maximizing the interoperability. Second, we conduct an experimental study. We develop and deploy a real-world testbed of Microgrid on our campus, on which we run a variety of energy services, demonstrating feasibility of the Internet of Energy. Last, we propose several schemes securing inter-networking in the cyber-physical systems. Access control mechanisms perform authentication and authorization in a fine-grained, prioritized manner, while an anti-jam key establishment algorithm makes objects' wireless communications faster and more robust against jamming attacks.

The dissertation of Eun Kyu Lee is approved.

Jack W. Carlyle

Rajit Gadh

Gregory J. Pottie

Carlo Zaniolo

Mario Gerla, Committee Chair

University of California, Los Angeles

2014

Dedicated

to my family

who have given endless encouragement for my life.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family.

I would like to express my deepest gratitude to Professors Mario Gerla and Rajit Gadh for their valuable guidance, caring, patience, and providing me with an excellent atmosphere for doing research.

It is an honor for me to have Professors Jack Carlyle, Gregory Pottie, and Carlo Zaniolo as committee members in my doctoral study. I am very much grateful to them for their consistent support and encouragement.

I would like to thank my colleagues in the Network Research Laboratory (NRL) and the Smart Grid Energy Research Center (SMERC), who are always willing to help me and give their best suggestions to my research.

I would like to note that the testbed of the smart green building has been developed and deployed by the UCLA SMERC research team. I thank Rui Huang, Tiana Huang, Omar Sheikh, Wenbo Shi, David Yao, Dr. Peter Chu, and for their contribution to the testbed. Most part of the EV testbed has been developed by Ching-Yen Chung, Joshua Chynoweth, Charlie Qiu, Bin Wang, and Yubo Wang.

I would also like to thank Dr. Soon Oh, Joshua Joy, Jae-Han Lim, and many students in the classes of CS 218 and CS 219 for their contribution to security research in the dissertation.

I would like to thank all the institutions that financially supported my research. This dissertation has been supported in part by grant from Korea Institute of Energy Research (KIER) fund, UCLA and KIER joint research project on Smart Green Building, by the National Electric Sector Cybersecurity Organization Resource (NESCOR) project led by Electric Power Research Institute (EPRI), and by participation in the International Technology Alliance sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defense.

Vita

2000	B.E. (Electronics), Dongguk University, Korea.
2002	M.S. (Engineering), KAIST, Korea.
2002–2008	Researcher, Electronics and Telecommunications Research Inst., Korea.
2011	M.S. (Computer Science), UCLA, Los Angeles, U.S.
2014	Ph.D. (Computer Science), UCLA, Los Angeles, U.S.

Publications

R. Huang, E.-K. Lee, C.-C. Chu, R. Gadh, Integration of IEC 61850 into a Distributed Energy Resources System in a Smart Green Building, *IEEE Power & Energy Society General Meeting*, July 2014.

M. Gerla, E.-K. Lee, G. Pau, and U. Lee, Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds, *IEEE WF on Internet of Things*, March 2014.

E. Lee, E.-K. Lee , S. Oh, and M. Gerla, Vehicular Cloud Networking, *IEEE Communications Magazine*, 52(2), Feb. 2014.

E.-K. Lee, P. Chu, and R. Gadh, Fine-Grained Access to Smart Building Energy Resources, *IEEE Internet Computing*, 17(6), pp.48-56, Nov.-Dec. 2013.

E. Lee, F. Yu, S. Park, S.-H. Kim, Y. Noh, E.-K. Lee, Design and Analysis of Novel Quorum-based Sink Location Service Scheme in Wireless Sensor Networks, *Springer Wireless Networks*, 20, pp.493-509, July 2013.

E.-K. Lee, R. Gadh, and M. Gerla, Energy Service Interface: Accessing to Customer Energy Resources for Smart Grid Interoperation, *IEEE Journal on Select Areas in Communications (JSAC)*, 31(7), pp.1195-1204, July 2013.

E.-K. Lee, R. Gadh, and M. Gerla, Resource Centric Security to Protect Customer Energy Information in the Smart Grid, *IEEE Smart Grid Communications*, Nov. 2012.

E.-K. Lee, S. Oh, and M. Gerla, Physical Layer Security in Wireless Smart Grid, *IEEE Communications Magazine*, 50(8), pp.46-52, Aug. 2012.

E.-K. Lee, S. Oh, and M. Gerla, RFID Assisted Vehicle Positioning in VANETs, *Elsevier Pervasive and Mobile Computing*, 8(2), pp.167-179, April 2012.

E.-K. Lee, S. Oh, and M. Gerla, Frequency quorum rendezvous for fast and resilient key establishment under jamming attack, *ACM SIGMOBILE Mobile Computing and Communications Review (MC2R)*, 14(4), 2011.

E.-K. Lee, S. Oh, and M. Gerla, Timely and Robust Key Establishment Under Jamming Attack in Critical Wireless Networks, *IEEE MILCOM*, Nov. 2011.

E.-K. Lee and M. Gerla, Secured Bilateral Rendezvous using self interference cancellation in wireless networks, *IEEE/IFIP Mediterranean Ad Hoc Networking Workshop*, June 2011.

E.-K. Lee, S. Oh, and M. Gerla, Randomized Channel Hopping Scheme for Anti-Jamming Communication, *IFIP Wireless Days*, Oct. 2010.

S. Oh, E.-K. Lee, and M. Gerla, Adaptive Forwarding Rate Control for Network Coding In Tactical MANETs, *IEEE MILCOM*, Oct. 2010.

E.-K. Lee, S. Yang, S. Oh, and M. Gerla, RF-GPS: RFID Assisted Localization in VANETs, *IEEE MASS Workshop on Intelligent Vehicular Networks*, Oct. 2009.

E.-K. Lee, Y. Yoo, C. Park, M. Kim, and M. Gerla, Installation and Evaluation of RFID Readers on Moving Vehicles, *ACM VANET*, Sep. 2009.

CHAPTER 1

Introduction

1.1 Research Background

Internet of Things. The Internet of Things (IoT) has been envisioned as the first evolution of the Internet[1]. It aims to provide a future network infrastructure interconnecting everything - from conventional computers to purpose-built local networks (e.g., body area networks and in-vehicle networks) and to everyday objects that fill our physical environment. Figure 1.1 illustrates the evolution. Theses IoT entities autonomously collaborate to make our surroundings more intelligent. In addition, connecting objects enables us to interact directly with our physical environment as well as access unprecedented volumes of data that would be otherwise unreachable. Just as the Internet opened the World Wide Web era by interconnecting computers in the world, the IoT will lead to revolutionary applications that can dramatically improve the way people live.

The IoT has caught much attention of research communities. As IoT research is still at a very preliminary stage, what IoT models will look like in the real world completely remains unveiled. Nevertheless, we envision that at least three keywords, peer-to-peer communications, local intelligence, and security, will dominate in the emerging IoT research. First, P2P communications represent that an everyday object will communicate with peer objects. And, people holding mobile devices will access to objects directly, not through intermediate gateways. Next, local intelligence indicates that a set of objects in a local area form a small autonomous system and together perform a common task favoring the local environment and people occupying the environment. For instance, the objects sense physical phenomena,

[1]Examples include Smarter Planet (IBM), The Internet of Everything (Cisco), The City 2.0 (HP Labs), and The Internet of Things (SAP).

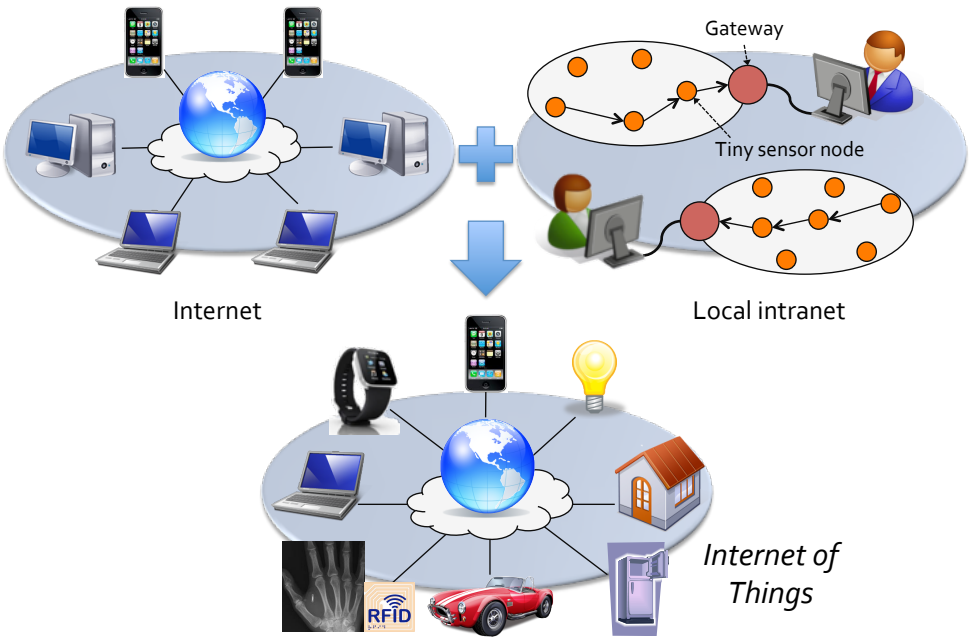


Figure 1.1: Concept of the Internet of Things aiming to interconnect everything around us.

learn local contexts in real time, and make a collaborative decision in order to self-repair regional errors. Recent emerging examples include smart building, intelligent home, and autonomous vehicle fleets. We also expect that the humans, more specifically information from online-social networking, will play an important role in the decision making process. Last, as IoT objects pervade our everyday life, accessing them may immediately impact our physical activities. Thus, security and privacy issues will become more critical than the existing IT systems processing information in a cyber domain. However, as IoT assumes a fully decentralized system, it will be extremely difficult to find out optimal solutions. Many research-leading institutes envision that the ultimate goal of the IoT is to build a smart world surrounded by intelligent environment that interacts with the existing information network so as to benefit people. We hope that our study can also contribute to accomplishing the research vision.

Smart Grid. A smart grid is a nation-wide project to modernize the century-old power infrastructure by resolving two intrinsic issues. First, the power grid is losing control to its reliability. When power demand (consumption) exceeds supply only possible response

is to shut off electric loads on the demand side, causing blackouts. Such a sudden power outage has a huge impact on our economy and individuals' activities[2]. In the past, a small number of large-sized customers such as tall buildings consumed most of energy. Thus, the supply side accurately estimated power consumption of future and planed to generate enough amount of power. Today, a large number of small-sized customers use power-hungry equipment and appliances (e.g., air conditioners). This makes the estimation complicated and inaccurate, increasing the chances of unexpected power demand exceeding supply, i.e., energy imbalance. Second, recent energy crisis stirs up governments to energy independence. As a response, they have developed national policies to improve their energy efficiency and security [5, 9]. A key feature in the policies is to stimulate the use of renewable energy sources, also responding to the threat of climate change. However, most renewables have limitations such as unpredictable, intermittent power generation due to dynamic weather changes. The existing power grid was not designed to connect to the unstable energy sources, thus incapable of resolving the limitations efficiently.

The smart grid aims to solve these problems based on the IoT model, making the existing power grid more intelligent and interoperable. All the energy objects[3] in the smart grid are equipped with embedded sensor/actuator systems and make collaborative interactions with each other. The universal interoperation enables the smart grid to monitor the health of power flow over the entire grid in real time, detect potential failures in advance, and quickly respond to risks threatening the grid's reliability. The smart grid also promotes bidirectional flow of electricity, which is compared with the existing power grid where electricity always flows from central bulk generators to end consumers. Various types of renewables will be installed at the consumers' side and supply surplus power back to the grid infrastructure. By integrating the information network and the electrical power network, a smart grid system aims to balance power demand and supply, eventually making the power grid more sustainable.

[2]It is reported that the Northeast Blackout, which hit the Northeast United States in 2002, caused a loss of approximately $6.4 billion [18].

[3]An object in the smart grid represents any energy resource involved in the energy activities of generation, transmission, consumption, and storage. Thus, for instance, it could be a bulk power plant, a utility pole, a desk lamp, or a solar panel at home.

Challenge: Understanding Cyber-Physical System as an Enabling Technology. Recent IoT research has focused on developing a cyber-physical system (CPS) as a core enabling technology, advanced from embedded systems technology. Traditional sensor nodes have been deployed to measure physical phenomena in a local area, and a centralized cloud server collects and analyzes sensory data. Likewise, RFID (Radio-Frequency Identification) tags storing identification are easily deployed and accessed in our surroundings. These embedded systems implicitly assume read-only objects deployed and focus more on computations on the data collection, allowing us to observe our environment. On the other hand, emerging objects are multi-functional. They can sense physical quantity and control their own physical operations simultaneously. The addition of a control function radically changes the way an object is connected to an IoT system - that is, the system can read and/or control these objects in a physical domain for its own purpose. The CPS captures the "control" property and promotes integration of physical processes (i.e., power flow in the smart grid) with networked computing capabilities. Consisting of three components (computation, communication, and physical process), the CPS properly realizes the control loop in the smart grid - that is, a smart grid system communicates with energy objects: reading their energy data in real time, analyzing data collection to detect the possibility of energy imbalance, and controlling these objects to increase power generation or reduce consumption. What distinguishes the smart grid system from other CPSs is that the system functions at a large scale. Since the smart grid is a nation-wide infrastructure, energy objects are deployed over geographically distributed areas. Moreover, most energy consuming objects (e.g., air conditioners and lights) are instrumented within buildings under control by different organizations having dissimilar business interests. In this setting, *communications* between objects become more complicated and challenging. Unlike other components in the CPS, however, few research has examined the communication issue in depth yet.

1.2 Contribution

To realize a cyber-physical system, an object must be capable of accepting control command messages as well as sending out data, making bi-directional interactions with external entities. Unfortunately, we have barely explored how to facilitate the bi-directional interactions with multi-functional objects. Regarding the interactions above the networking layer, thus, we pose two research questions.

- *How to enable the interactions so that the IoT objects can interoperate each other?*

  An object will interact with diverse external systems. It communicates with a data collection server and even peer objects. We also expect that various types of objects provided by different vendors are deployed ubiquitously. An existing object, however, makes a dedicated connection to a pre-designated gateway server in a proprietary manner, limiting the scope of the interconnection.

- *How to secure the interoperation?*

  Universal, unauthorized access to objects engenders new security and privacy problems that immediately threaten our everyday life. Remote hacker shutting off the power of an Electric Vehicle with the driver sitting inside is no longer a futuristic scenario.

To answer the questions, this dissertation designs and develops an interoperable and secure communication model for a cyber-physical system and extends it to examine the realization of smart grid interoperation, the Internet of Energy (IoE). The contribution of the dissertation is three-fold. First, we develop a generic middleware model in which an object represents various physical functions and inter-communicates with other objects in a unified manner. Second, we conduct an experimental demonstration of the Internet of Energy model: exploring a variety of energy services, developing a real-world testbed on our campus, and running the services on top of the testbed Last, we propose several schemes securing the cyber-physical systems' inter-networking: advanced access control systems and anti-jamming wireless communications. The following paragraphs summarize our contributions.

Interoperable communication interface for a cyber-physical system [Chapter 2]. In the smart grid, an energy object links a cyber component (information exchange) with a physical component (power flow), building a CPS. For instance, an LED light object provides its power usage data and accepts command messages turning on/off or dimming the brightness. It is obvious that new objects having a variety of physical functions will be introduced. Moreover, these objects directly interact with peer objects as well as conventional computers and mobiles. To maximize interconnection in the smart grid consisting of such heterogeneous objects, each object must operate as a standalone system and have unlimited accessibility. To enrich universal access, we need a generic cyber-physical system model in which an object represents various physical functions and inter-communicates with other objects in a unified, interoperable manner.

To resolve the issue, Chapter 2 proposes a middleware system implementing the concept of Object-Oriented REST [52, 55]. The physical component of an object is modeled (or represented) as a cyber system using Object-Oriented Programming (OOP) paradigm. Like a Class in OOP, the system maintains both its own data values (variables) and controls (methods). Then, the object communicates these state representations via a REpresentational State Transfer (REST) web service that defines access actions to the object (e.g., data reading and controlling) using HTTP methods of GET, PUT, and POST. In this way, all the objects in the IoT make collaborative interactions seamlessly. The chapter develops an Energy Service Interface (ESI), the implementation framework of the proposed middleware system. The ESI is then evaluated in a testbed where a variety of energy objects (e.g., smart plug, dimmable LED lights, smart submeters, and office appliances) are interconnected via different communications technologies such as ZigBee, WiFi, and serial connection.

Energy services - realizing the IoT model in the smart grid [Chapter 3]. Once obtaining the interconnected cyber-physical system model, following research question is about how interoperation among objects is realized in the real world. A full-fledged cyber-physical system in the smart grid interconnects physical input (power usage measurement) and output (power control). They, as essential factors, make the CPS work and thus enable

the smart grid interoperation for energy balance. Beyond the interconnection, the CPS involves intelligent data processing based on the physical inputs, e.g., learning dynamic physical contexts and making decisions for control. The full complexity of the CPS remains as an open problem. Our study especially focuses on physical domain factors making control decisions, thus directly influencing the interoperation. For instance, a building owner may not want to control his energy objects (e.g., turning off an air conditioner) unless the benefit far overweights that of uncontrolled building operations. Unfortunately, these factors are very unpredictable and not understood via a simple simulation study.

To address the issues, Chapter 3 conducts experimental research in three major steps [53, 54]. First, it explores a variety of energy services realizing the interoperation: examining how the services evolve, analyzing factors leading the evolution, and identifying new technological demands to accommodate the factors. Second, our research team develops and deploys a real-world testbed in our campus. It contains most types of energy resources, ranging from Electric Vehicle (EV) and solar panel to smart light and smart appliances. Last, we run the services on top of the testbed to demonstrate their feasibility and evaluate performance. The experimental study also revealed several engineering and non-technical problems that must be considered to realize the cyber-physical system in practice.

Fine-grained, decentralized access control [Chapter 4.1]. In the smart grid, any entity can interact with energy objects in various ways, for example, reading data and controlling operations. Such different access actions induce different operational consequences and indicate different levels of privacy violation. Moreover, due to the property of universal accessibility in the IoT, the identity of an accessing entity is often unknown in advance. To accommodate new challenges, we need to design (1) finer-grained access control, that is, an object distinguishing the privilege of controlling from that of reading and (2) stateless, decentralized access control, such as, an object not storing the accessing user list in advance. Unfortunately, a typical access control mechanism in the Internet cannot work since it relies on a pre-registered user list and grants privileges in a coarse-grained manner.

Chapter 4.1 proposes Resource Centric Security (RCSec) that resolves the first issue by

leveraging the concept of Access Control Entry (ACE) in Unix file systems [50, 51]. Each file maintains its own ACE, say, "`rwxr-xr--`" or "`111 101 100`". The first 3 digits specify that the file owner is permitted to read (`r`), write (`w`), and execute (`x`) the file whereas the last 3 digits indicate that non-group users can only read the file. Adopting this concept, an energy object performs access control with three levels of privileges. When accessing an object, an entity is permitted to read energy data, change configuration, and/or control the objects physical operations according to its security token. To address the second issue, RCSec designs a decentralized ACE. The entity is granted a security token from a certificate authority in advance. Each token contains state information describing the entitys privilege (`r`, `w`, and/or `x`) regarding the object. Instead of storing user information and privilege rules at local, the object validates the token presented by the accessing entity and performs access control based on the included state information. This enables a security system to cope with complex interactions as well as operate in a large-scale smart grid environment.

Multi-factored authorization [Chapter 4.2]. It is intuitive that controlling operations influence our daily life more immediately than reading data. Thus, the former can be prioritized higher than the latter. This prioritization is observed everywhere in the smart grid. For instance, an energy object capable of powering off a whole city receives higher priority than the one turning off our office lights. A multi-factor authentication has been a powerful solution to addresses this issue. A user is required to present more than two authentication factors (e.g., ID and fingerprint) to access highly prioritized objects. But, the multi-factor authentication works only in authentication, requires the presence of human beings, and fails to support stateless interactions. Thus, this scheme is very limited in the IoT environment despite its effectiveness.

To overcome the limitations, Chapter 4.2 extends RCSec and proposes Multi-Factor Authentication and Authorization (MFAA) performing access control without human involvement [39, 56]. An entity is granted multiple security tokens from different certificate authorities, where each token functions as a factor. When accessing an object requiring multiple factors, the entity presents the corresponding numbers of qualified tokens. The

tokens are cryptographically correlated with each other despite being granted from different authorities. Thus, MFAA prevents collusion attacks. MFAA seeks to decrease the probability of presenting false evidence of the entity's qualification, and thus the number of factors implies protection level. The object can also set different numbers of factors for different access actions and change the number according to its own context dynamically.

Fast anti-jamming wireless communication [Chapter 4.3]. Significant portions of energy objects in the smart grid will be accessed via wireless communications. Although it is obvious that using wireless communications offers significant benefits over wired connections, the wireless technology is known to be vulnerable to physical layer security such as jamming attacks [60]. Previous research addressed the issue by proposing random, spread-spectrum based wireless communications in which two nodes having no prior knowledge of each other establish a common secret key. The key is then used to compute a common frequency-hopping sequence for following data communications. But, this key establishment suffers from severe time latency up to the order of 100 seconds.

To resolve this latency problem, Chapter 4.3 proposes Frequency Quorum Rendezvous (FQR), a fast and resilient key establishment scheme exploiting a quorum theory [57, 58]. Each node independently selects its own hopping sequence based on a public quorum during the key establishment phase. The intersection property of the quorum system guarantees that the pair nodes rendezvous within a bounded time. Thus, FQR provides two important benefits. First, it decreases time latency because nodes are able to exchange a common key faster. Second, FQR is very robust against attacks because each node independently constructs own random sequence, making jamming attacks difficult and inefficient.

Conclusion. This dissertation makes a conclusion in Chapter 5 by summarizing the proposed systems for interoperable and secure interaction in the smart grid network.

CHAPTER 2

An Interoperable Communication Interface for Cyber-Physical System

Smart grid consists of a myriad of heterogeneous systems, and their seamless interoperation is the key element for the success of our future power system. To facilitate interoperability, National Institute of Standards and Technology (NIST) presents a conceptual model consisting of seven domains [12]. Among them, the customer domain - industrial, commercial, and residential sectors - is the primary energy consumer; it consumes 72% of total energy in the U.S. [13]. Therefore, it is essential to interoperate with energy resources in the customer domain in order to achieve the goal of smart grid - balancing the power demand with supply.

The customer interoperation is generally realized via a number of customer energy services. To enable the services across the customer domain, NIST defines two gateway actors - utility meter and Energy Service Interface (ESI) [12]. They place at the boundary of the customer domain and exchange energy data between the customer domain and other external domains (e.g., distribution, operation, and market). That is, they take the bridge role for inter-domain communications. Owned and fully controlled by a utility company, the utility meter measures and collects aggregated energy usage data for the customer billing purpose. However, as the utility meter is designed to handle the aggregated data only, the ESI is expected to process most emerging energy service data. In this way, the ESI would exceptionally contribute to the customer interoperation, and there is no doubt that smart grid interoperation cannot be accomplished without its help. However, as yet, few research investigated the development of the ESI in depth.

To address the growing concern of the ESI, this paper examines its design issues and

implements and deploys a prototype that realizes the customer energy services. For the ESI design, we classify the issues into four categories. First, the ESI consumes energy services that external domains provide to the customer domain. To this end, it translates the semantics of external services into internal contexts as well as exchanges messages in a standard way. Next, the ESI provides energy services, as a service provider, to external domains. It creates a new service using internal energy data and defines communication interfaces through which external systems access the service in an efficient and interoperable way. Third, the ESI implements a security mechanism for inter-domain communications. It protects internal energy resources and prevents a security failure from being spread across domains. Last, the ESI is a logically-defined communication interface and can be implemented in any physical system. This property allows various types of ESI implementation architectures to be deployed.

This paper also builds an ESI testbed that includes an ESI, an energy management system, customer energy resources, and a demand response service server. We carefully take the design issues into consideration and implement our ESI on the energy management system. The ESI implements a demand response client module. It adopts a standard XML format to represent the customer energy data and implements an object-based web service interface for inter-domain communications. All the service objects are protected by a resource centric security mechanism. We build an additional testbed in which the ESI is deployed on a public Cloud. To verify the operation of the ESI and to evaluate its performance, we run experiments with several energy service scenarios. Throughout the experiments, we demonstrate that the ESI interacts with the demand response service (as a service consumer) and provides remote energy management services to external users (as a service provider). The experimental results are analyzed further to identify how the security mechanism and the Cloud architecture affect the service performance.

2.1 Customer Interoperation via Energy Service

The customer domain will include a myriad of energy resources (load, generation, and storage). Because the customer domain is the primary energy consumer in smart grid, interoperating with the resources is the most critical concern to balance the power demand with the supply. A couple of literature list the use cases of customer energy services that accomplish the *interoperation* [6, 7, 12]. The following subsection categorizes the customer energy services that the energy resources in the customer domain are involved.

2.1.1 Customer Energy Services

Energy Usage Collection. The energy usage collection is the simplest form of services. The customer domain periodically transmits customers' energy usage data to a utility company that processes the data for customer billing.

Efficient Energy Management. The efficient energy management service informs the customers of the details of energy usage information so that they clearly understand their usage pattern. A sub-metering system in the customer facility keeps tracking of the energy usage of individual energy loads. An Energy Service Provider (ESP)[1] may analyze the usage pattern and guide the customers to consume power more efficiently. User-friendly devices such as In-Home Display show the breakdown of the usage and corresponding costs. An Energy Management and Control System (EMCS)[2] allows the customer to control the energy loads even remotely. The energy usage data is also used to validate bills from the utility company.

Customer Feedback. Customer feedback represents the deliver of customer energy information to an ESP that exploits the information to create further services. This can eventually contribute to the reliability of power infrastructure. For instance, a demand forecast service delivers information of expected customer energy demand to a power supplier

[1] An ESP can be a third party service provider, a utility company, an operator, or a power supplier. This paper uses their meanings interchangeably.

[2] The concept of the EMCS includes those of building automation system (control) and energy management system (measurement).

so that the supplier can accurately estimate the amount of future power generation. To this end, an ESP provides power price data and weather forecast to a customer who already understands his energy usage patterns. Then, the customer examines local energy needs of the future. As an another example, the sub-metering system measures power quality, and the ESP uses it to pinpoint a potential failure spot and to maintain the health of the power supply system.

Direct Load Control. The Direct Load Control (DLC) permits external users to control internal energy loads. Currently, ESPs provide a primitive form of the service targeting at specific energy loads such as an electric water heater. They directly control the operations of the heater during the period of power supply emergencies. In the future, the ESPs will control various types of energy loads for different purposes. For instance, an ESP remotely stops charging an Electric Vehicle (EV) at a customer's garage when the power price goes beyond a contracted value, which can save the customer's electricity cost. The ESP may launch grid stabilization services such as frequency regulation and Volt-Amps Reactive (VAR) compensation. A combination of inductive and resistive loads at the customer facility can be controlled to help balance active and reactive power.

Intelligent Demand Response. Power suppliers often confront a shortage of generation capacity during peak-demand periods. Instead of constructing additional power plants to cope with the shortage, a Demand Response (DR) service tries to reduce energy consumption in the customer domain by sending DR signals to customers. A utility operator calls contracted customers who, then, manually stop their building operations by expecting financial incentives. New challenges of the *intelligent* DR service are automation and invisibility. When the customer domain receives DR signals from the utility, it automatically performs a predefined DR strategy (shedding and shifting the loads) to achieve the energy curtailment of a service contract. The strategy must be invisibile: It minimizes interference with ordinary building operations. Potential inclusion of EV, storage, and generation at the customer facility will play an important role for the invisible DR strategy.

Distributed Generation and Microgrid. Distributed Generation (DG) represents electricity generation that feeds into the distribution grid directly, not through the transmission

infrastructure. The DG can reside within a customer facility as a form of renewable sources - micro-turbines, wind-powered generators, and Photovoltaic arrays. They can be managed and controlled directly by an ESP as a part of the DLC service. A customer may lease the DG and storage, and the ESP remotely monitors the health of the energy assets and controls them directly when necessary. With the DG, the concept of microgrid enables the customer facility to rely less on bulk power sources. In this sense, the microgrid can be leveraged for the intelligent DR service. Upon receiving the DR signal, the customer prepares to operate his own DG. This can compensate some portions or all of energy needs that must be reduced by the terms of the DR service.

Energy Market - Power Trading. With sophistication of the storage and generation technologies at the customer facility, the customer will actively interact with the wholesale and/or retail energy market. He can choose to buy power from one or more power suppliers or generate power on-site for sale in the market. That is, he buys or sells electricity in more flexible ways as a "prosumer". In the trading, real time exchange of price, schedule, and location information across domains becomes of the most importance, because it is critical to dispatch power to the right place at the right time.

Environmental Monitoring. The environmental monitoring service deals with an increasing concern of environmental sustainability. The service measures the greenhouse gas emission of bulk power generation, and the information is shared among the entire smart grid. An on-site generator and building operations also report their emission contents, e.g., carbon dioxide. This way, a customer can enumerate the influence of their energy usage activities on the environment, when he consumes power sourced from various suppliers. He may choose to buy clean energy for higher prices or participate in an emission trading market with his on-site renewables.

2.1.2 Gateway Actors for Customer Interoperation

To facilitate the interoperable energy services, the smart grid conceptual model [12] defines two gateway actors - utility meter and ESI. Sitting at the boundary of the customer domain,

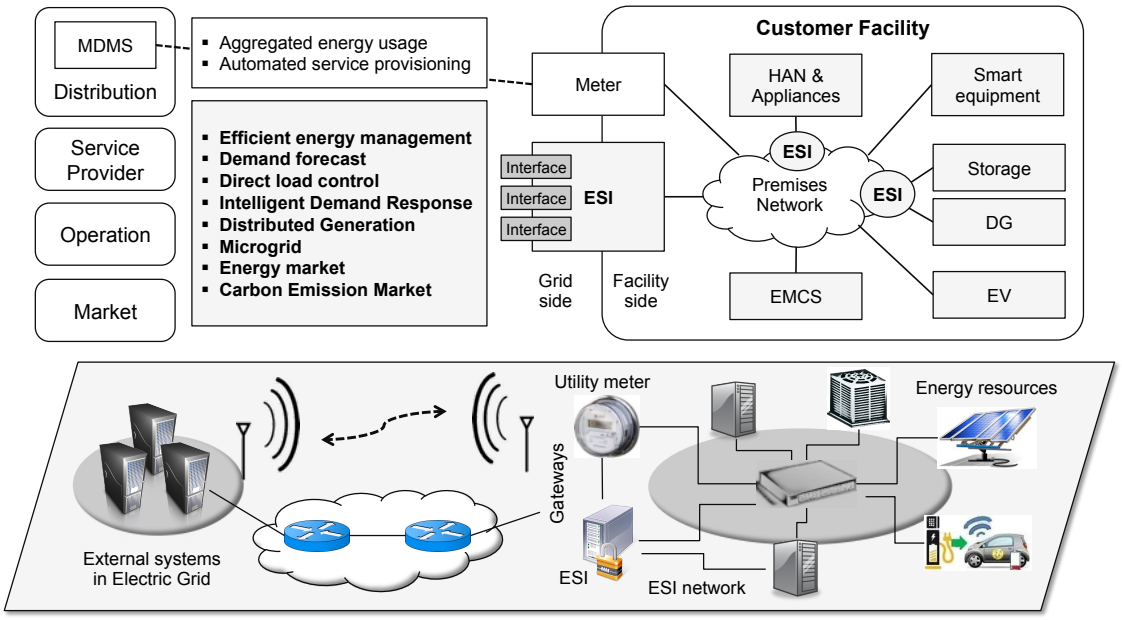


Figure 2.1: Smart grid interoperation with the customer domain. Two gateway actors are responsible for providing customer energy services bi-directionally.

they are responsible for inter-domain communications. That is, external systems (utility companies and ESPs) interconnect with the customer energy resources through them. Figure 2.1 shows an overall system architecture that represents the smart grid interoperation with the customer domain.

2.1.2.1 Utility meter

The primary function of the utility meter (a.k.a. smart meter) is to provide an accurate, remote measurement of energy usage for customer billing. To this end, a utility company installs the utility meter at a customer facility in which the meter periodically measures aggregated energy usage data. The data is then collected by an aggregator via RF communication within a neighborhood area network, and then delivered to Meter Data Management System (MDMS) in the utility via a cellular network[3]. These hardwares and communications constitute an Advanced Metering Infrastructure (AMI). As an element of the AMI, the

[3]Different utilities may use different communication channels. Our description is based on the system specification of a utility company in California.

utility meter is under full control of the utility while residing within the customer boundary. A meaningful number of utility meters already have been deployed in the U.S., and thus the AMI system architecture has been understood well. Many researches have also investigated on meter data processing and its communications.

2.1.2.2 Energy service interface

The ESI is initially designed as a communication interface to serve demand response signaling. But, because the utility meter is functionally limited[4], the ESI is expected to transmit data streams of most customer energy services. When expecting that a great amount of energy resources will be involved in the energy services, the ESI becomes the most important entity in the customer domain in terms of smart grid interoperability. Unlike the utility meter, however, few literature of smart grid has, as yet, investigated more than its conceptual idea. As of today, a couple of public reports simply discuss the definition of the ESI - "The ESI serves as the information management gateway through which the customer domain interacts with ESPs' [4, 12, 36, 41]. The discussions share common characteristics summarized below.

The ESI represents a logical boundary of the customer domain. The ESI logically distinguishes the operation of the customer facility from the outside world. To this end, the ESI consists of two sides: the facility side and the grid side. The facility side communicates with internal energy resources directly, or is connected to data storage and processing modules in an EMCS. The grid side interacts with external systems such as ESPs and operators.

The ESI represents a bi-directional service interface. The ESI serves as a platform that provides energy services to both the facility side and the grid side. It exposes raw data and a set of energy services that the customer domain provides to external domains - e.g., direct load controls and demand forecasts. We call this function a *facility service interface*. In the same way, the ESI delivers energy services and data from external systems to the customer

[4]The utility meter has a very long lifetime and is cost sensitive so that it is barely upgraded. Moreover, a utility company is unwilling to collect customer energy data other than from the utility meter due to privacy issue.

domain - *grid service interface*.

The ESI does not represent a physical device. The ESI is a logical interface in the form of software component and thus can be implement on various physical devices such as EMCS and an energy gateway. Or, it can be realized in a standalone server system. It is generally assumed that the ESI is implemented in a customer's EMCS that communicates with external systems via the Internet.

Development of the ESI is still at an enfant stage, and there are many issues to be considered before its real world deployment. This paper tackles those issues - we discuss their designs and eventually implement a prototype.

2.2 ESI Design Issues

Reviewing the customer energy services helps us understand the functional requirements of the ESI. This section investigates how to design an ESI system.

2.2.1 Grid Service Interface - Interconnecting with External Energy Services

At the core of the ESI are communication and interoperation with the smart grid infrastructure. The ESI must be able to interoperate with external energy services that other domains in smart grid provide to the customer domain. For instance, a Demand Response Automation Server (DRAS) provides a DR service by generating and delivering DR event signals to the customer. Then, the ESI accepts and understands the service signals; delivers event information to internal energy resources; and sends a DR event report back to the DRAS. Many standardization efforts have defined communication protocols of such grid services. Examples include Open Automated Demand Response (OpenADR) [75], Energy Market Information Exchange (EMIX)[5], Energy Interoperation (EI)[6], IEC Common Information Model (CIM) family of standards, and Weather Information Exchange Model (WXXM) [8].

To support the grid service interface, it is essential that the ESI translates the informa-

[5]https://www.oasis-open.org/committees/emix/
[6]https://www.oasis-open.org/committees/energyinterop/

tion delivered from external domains into the semantics and protocols that are used within the customer facility. Data mapping is especially critical, because most existing customer systems have used proprietary protocols that were not designed for interoperation. The ESI also translates contexts for security. Since it sits at the boundary between two independent domains, the ESI seamlessly interconnects dissimilar security mechanisms so as to balance their security policies. Misconfigured mapping of security contexts may create vulnerable points and increase the possibility to leak unauthorized private information.

2.2.2 Facility Service Interface - Serving Energy Services to External Domain

The ESI provides energy services to external domains, which primarily delivers data that the customer domain generates. In a customer feedback service, for instance, it transmits customer energy information to ESPs. It also accepts command messages that eventually control the customer equipment - e.g., remote energy management service and direct load control. From the viewpoint of the entire smart grid, these interactions make the customer domain appear as a service provider. Thus, the ESI needs to define what and how specific service data is transferred through itself. This subsection discusses five topics.

Interface abstraction. The ESI must consider interface abstraction that determines the appropriate level of internal details that the interface exposes to external domains. High level of abstraction exposes internal business logics with less details, while low level of abstraction exposes more details of internal operations. DLC and event-based DR show two extreme examples with respect to control capability. In DLC, the ESI allows external systems to control energy loads directly, which realizes the lowest level of abstraction. On the other hand, it receives a DR event signal from a DRAS, from which the facility determines its own control strategy without revealing the list of controlled energy loads. The abstraction level must match to the requirements of energy services and applications [2, 41]. A well-designed abstraction transmits only the necessary service data to external domains, while shielding them from changes that occur within the customer domain. In this way, the ESI maintains consistent views of internal energy services, which is critical to the interoperability goal of

smart grid.

Separation of concern. An ESI consists of several service interfaces, each of which is responsible for one functionality of energy services as shown in Figure 2.1. The separation of concerns indicates that an interface is functionally independent of other interfaces so that any changes on the interface hardly affect the others. For instance, an interface for an event-based DR service must be distinguished from a bidding DR service. In the same way, sub-functions within an interface must be separated in an appropriate way.

Data representation. The ESI must provide energy data in a standardized format. Various customer equipment generates different formats of data that must be semantically understandable over smart grid. To cope with the heterogeneity, smart grid takes a Canonical Data Model (CDM) approach [3]. A data producer transforms its output to a standardized information model, and then a consumer transforms it back to own terminology. This approach is especially crucial in the customer facility, because there are still many legacy systems generating data in the proprietary format. The ESI must be able to handle and transform such data into a standard form. Standardization efforts[7] touching this issue include Facility Smart Grid Information Model (FSGIM), ISO/IEC 15045, Open Building Information eXchange (oBIX), Building Automation and Control networks (BACnet), OPC Unified Architecture, and ZigBee Smart Energy Profile (SEP).

Service interaction model. The ESI must support an efficient interaction model that determines how it communicates with external systems. Traditionally, a middleware has performed this task. In order to fulfill the interoperability requirement, recent standardization efforts consider web services. For instance, OpenADR specifies two types of bindings (SOAP and HTTP) and two types of message exchange patterns (PUSH and PULL). However, some energy services may require a constraint of high performance. In this case, we may consider an API based asynchronous messaging system. Another consideration is a content-centric middleware that fully supports contextually-driven associations amongst energy resources [47].

[7]FSGIM - http://spc201.ashraepcs.org/standards.html; oBIX - http://www.obix.org;
BACnet - http://www.bacnet.org; OPC/UA - http://www.opcfoundation.org;
SEP - http://www.zigbee.org/Standards/ZigBeeSmartEnergy

Extensibility. The extensibility and flexibility are important requirements for the ESI, because the customer-engaged energy services evolve over time. New applications will be introduced, and corresponding service interfaces are newly added into the ESI. Some of them may reuse existing interfaces as sub-functions, and existing interfaces become deprecated or remain for backward compatibility. Implementation of internal functions also keeps changing with technological advancements. The design of the ESI must be extensible enough to allow such innovations of service interfaces, while ensuring the level of abstraction to be consistent.

2.2.3 System and Architecture

2.2.3.1 Cyber Security

Cyber security is one of two cross-cutting issues in smart grid, and it is well known that the relative importance of Confidentiality, Integrity, and Availability (CIA) is reversed in smart grid [12]. This is because it is more fundamental that authorized messages must be delivered to the right place at the right time. Confidentiality is also critical, but its importance varies according to the provided services. Thus, this article omits discussion on it. This section discusses three important topics of availability, access control, and integrity.

Availability. Customer energy data must be consistently available for the successful provisioning of energy services. And such availability is primarily related to the network connection at the ESI. The connection can be threatened by external attacks such as Denial of Service (DoS). Or, data may be unavailable due to internal reasons such as the ESI failure. One interesting solution is to move the ESI to the Cloud that is generally assumed to be more accessible and reliable. An external system contacts the Cloud to access customers's energy resources, without knowing the location of the customer facility. The availability is also affected by traffic congestion in an emerging converged network, where the energy data is transmitted over the same IP network that has delivered residential data. A strong logical separation of traffic and proper QoS mechanisms can minimize the impact of non-critical traffic on the energy data [86].

Access control. An access control permits only authorized users to read customer data

and to control energy resources. To this end, access control verifies the identity of an accessing user and grants him access permission to specific energy resources. An emerging issue of the access control to the resources is the increasing complexity of an access control rule. The customer domain will include a myriad of equipment, each of which is equipped with an embedded system. This enables external systems to access individual resources. Moreover, an access for data reading must be distinguished from that for resource control. The access control rule must take the new condition into account, which makes it complicated. In any scenario, the rule must be carefully designed, because any abuse of privilege makes the smart grid system unreliable potentially as well as violates privacy policy. An ideal access control realizes the principle of least privilege. Say, an ESP is permitted to perform a DLC to an energy load. The granted privilege must prohibit it from taking other actions and must be revoked immediately after its use.

Integrity. The ESI takes care of both message integrity and system integrity. It must be able to detect any tampered messages and to quarantine compromised interactions. A forged DR signal may misinform a large group of customers of an urgent DR event, which can lead to the failure of their reducing energy consumption during on-peak period. This can potentially cause a serious blackout. A proper usage of a message integrity code can mitigate the risk. System integrity represents the capability of preventing the cascade effect of a security failure. Say, a service interaction is compromised. This must not affect other interfaces and functionalities. A well-designed separation of concerns can mitigate the impact by removing an unforeseen chain of events among interfaces.

2.2.3.2 Architectural Consideration

The ESI is an interface specification for inter-domain communications, and its physical location is of less importance. This property introduces several new design issues as below.

The ESI is generally implemented on an EMCS, but its functionality can be realized by a coordination of several subsystems. Or, a customer facility may have multiple EMCSs, and thus multiple ESIs. For instance, the owner of a multi-tenant office building may want

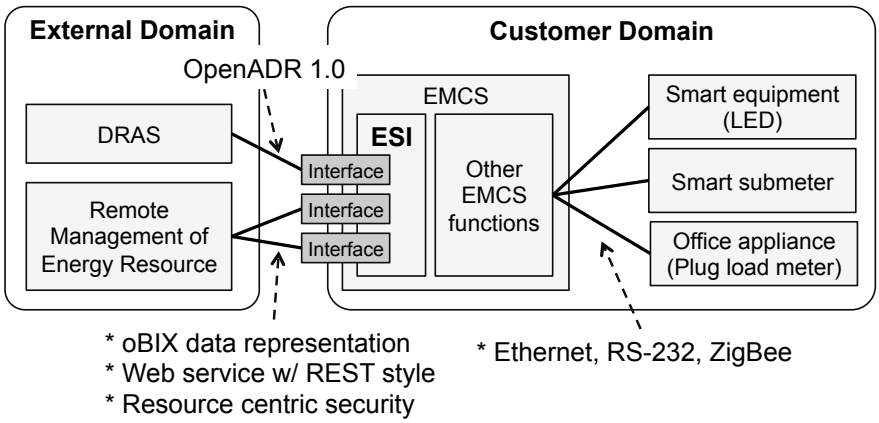
to distinguish sub-domains, each of which is managed through internal EMCSs. Moreover, solar panels on the roof and EV charging stations may be under control of another EMCS. There is still a central system running an ESI responsible for inter-domain communications, and the internal systems organize a network in a hierarchy.

The ESI can be deployed in a remote site. Especially, with the advancement of the Cloud technologies, there are many drivers to move the ESI to the Cloud. First, a local ESI may not be universally accessible, because many customer systems reside behind firewalls and Network Address Translations (NATs). This limits external systems' access to the customer data. Next, there are still many small facilities that cannot afford a local EMCS system for the ESI. Even if there are some, the systems may not store and process the increasing volume of energy data. Moreover, it is very hard to expect that all the security patches are applied to all the systems in a timely manner. Last, the Cloud enables efficient data accesses. Customer systems are located over geographically distributed areas, and some of their connections are quiet slow. When an ESP collects energy information from multiple customers and generates a new benchmark, it can do the task more quickly and efficiently once all the data is already on the Cloud.

2.3 Implementation of the ESI Prototype

2.3.1 ESI Testbed

The integral part in developing and deploying the ESI is to realize two service interfaces in an interoperable and secure manner. In the grid service interface, the ESI implements the counterpart module of an external energy service that can understand the service's context as well as translate the context to the terminology that the customer domain uses. Implementing the facility service interface is more challenging: it requires prior deployment of energy resources and implementation of EMCS in the customer facility. Then, the ESI models resources as service objects, abstracts their communication interfaces according to the separation of concern principle. To ensure interoperation, the objects and the interfaces



(a) ESI testbed consists of ESI, EMCS, energy loads, and energy services.

* EMCS – the front and rear side views

* Dimmable LED light

* Plug load meter

(b) Pictures of the EMCS and two types of energy loads.

Figure 2.2: ESI testbeds.

must be implemented in a standardized way. A security mechanism being deployed must take into account the abstraction level.

Accommodating the issues above, this section develops and deploys an ESI testbed in a small customer facility where one ESI is implemented on an EMCS. In addition to the ESI functionalities, the EMCS performs data collection; management of energy resources; database management; and service data generation. We deploy three types of energy loads in our office. Office appliances such servers are plugged into plug load meters that measure energy usage and turn on/off the input power. A smart submeter is attached to each circuit within a circuit break panel. An LED light represents a smart equipment that can adjust own operations beyond a simple on/off status. The LED operates with 8 steps of brightness

and temperature that affect its power consumption directly. The energy loads are connected to the EMCS via Ethernet, RS-232 and ZigBee. Figure 2.2a illustrates the testbed system that includes the ESI, the EMCS, energy loads, and an external energy service. Figure 2.2b shows pictures of our testbed - the front and rear side views of the implemented EMCS; the dimmable LED light; and the plug load meter.

2.3.2 Energy Services

2.3.2.1 Open Automated Demand Response

The ESI implements an OpenADR client as a grid service interface that realizes an automated DR service [75]. OpenADR is expected to be the first grid service based on real-time power price that will appear in the near future. Thus, examining it helps us understand how a customer domain interoperates with external energy services. To illustrate the interoperation, we implements and deploys a DRAS server system in our testbed [11]. The OpenADR client accepts and interprets DR signals from the DRAS. At the moment, the client module handles event programs only, not the bidding program.

A DR event initiates when the DRAS issues a message, *EventInfo*, and then, generates an *EventState* signal[8] that represents the DR event. The DRAS supports connections from both smart and simple clients. The smart client is capable of dealing with the *EventInfo* information within the *EventState* signal. Included in *SmartClientDREventData* entity, it contains event details. For instance, the *eventInfoTypeID* identifies the type of event information and takes one of values of PRICE_ ABSOLUTE, PRICE_RELATIVE, LOAD_AMOUNT, etc. For the simple OpenADR client, the DRAS translates the *EventInfo* information into a simpler form, named *SimpleClientEventData*. In the entity, two variables describes the event state. The *EventStatus* element indicates the temporal state of the event (FAR, NEAR, or ACTIVE), whereas *OperationModeValue* denotes the operational state of the energy loads in the event (NORMAL, MODERATE, or HIGH). Our ESI implements both smart and simple

[8]All the data in OpenADR is represented as XML form, and their schemas are available at http://openadr.lbl.gov/src/1.

clients and periodically "pulls" the *EventState* message from the DRAS. This PULL mode is typically used, since the OpenADR client has more control over the communications, e.g., firewalls and private networks using NAT. A PUSH mode can be considered in scenarios where very low latency is required.

To address the security issue of OpenADR, in particular the message integrity, we implement a Message Authentication Code (MAC) on top of the existing OpenADR system. Following the NISTIR 7628 guideline [83], our testbed takes a hash-based MAC (HMAC) with SHA-256.

2.3.2.2 Remote Management of Customer Energy Resource

To show the ESI serving as a service provider, we consider a scenario of "remote management of customer energy resources". It includes various types of services - an external user is able to obtain energy usage data of individual energy loads; retrieve historical data; and control the loads directly. The ESI leverages two technologies of oBIX and web services to implement the facility service interface supporting the services above. This subsection describes how they can realize the services while fulfilling the design issues discussed earlier.

Open building information exchange. The ESI takes the oBIX specification as a standard data representation because of several outstanding advantages.

First, it supports an Object-Oriented (OO) design pattern with low-level abstraction. This helps us model information of electrical and mechanical systems in a facility as an object. Like the OO programming, each object is modeled by a set of *value* objects like "int" and "str" and a set of *op* objects that defines an operation with input and output objects. The object model allows inheritance so as to model complex energy data by means of a contract mechanism. Realized by *is* object, it establishes the conventional "is a" relationship with various overriding rules. In this way, an object cannot only represent a physical unit directly, but also a particular functionality as a collection of sub-objects. This capability allows us to abstract service interfaces in a more flexible manner according to the service requirements.

Next, oBIX exploits XML to express its underlying object model, which maximizes the interoperation property of standard data format. To this end, it specifies four syntaxes - each object type maps to one XML element; an object's children are mapped as children elements; the XML element name maps to the predefined primitive object type; and every other objects are expressed as XML attributes. Last, the OO design of oBIX is also beneficial to the design issues of extensibility and separation of concern. The reusability of the object model helps make the ESI more flexible to accomodate future innovation of the services. The inheritance property simplifies the development of complex energy services, and each service and interface implementation can be easily separated from others.

Leveraging the advantages of oBIX, we implement data and service models for our scenario. The box below illustrates a *History* object, a historical archive of an energy usage data over time. The *is* attribute in the *obj* element indicates that it is extended from a standard oBIX object *obix:History*. The example also shows that the *query* operation to read history records takes an argument whose object type is defined as *psxml:HistoryFilterEx* and returns history records in the object format of *obix:HistoryQueryOut*.

```
<obj href="http://myESI/History/" is="obix:History">
  <int name="count" val="541"/>
  <abstime name="start" val="2012-01-02T00:00:00.000-08:00"/>
  <abstime name="end" val="2013-01-06T00:00:00.000-08:00"/>
  <op name="query" href="query" in="psxml:HistoryFilterEx" out="
     obix:HistoryQueryOut"/>
  <feed name="feed" href="feed" in="obix:HistoryFilter" of="
     obix:HistoryRecord"/>
  <op name="rollup" href="rollup" in="obix:HistoryRollupIn" out="
     obix:HistoryRollupOut"/>
</obj>
```

In a similar way, the box below represents a *power* object that contains information of power draw data, power factor, and communication quality regarding plug1 (a plug load

meter) connected to the EMCS via ZigBee. As shown in the *href* attribute, the object is a sub object of a *plug1* and *zigbee*.

```
<obj href="http://myESI/Points/zigbee/plug1/power/">
  <real name="voltage" val="120.2" unit="obix:units/volt"/>
  <real name="current" val="1.21" unit="obix:units/ampere"/>
  <real name="power" val="160.03" unit="obix:units/watt"/>
  <real name="powerFactor" val="0.996" unit="obix:units/power_factor
     "/>
  <int name="rssi" val="-72" unit="obix:units/decibel"/>
</obj>
```

Web service. Our data representation model enables to design data in an object oriented way, and then each service is mapped into an object having both values and method signatures. The ESI takes Web Service (WS) as a service interaction model, because it accommodates this property as well as exposes service interfaces in an interoperable manner. In terms of WS, each energy object is accessed via a URI, and its data is passed around as an oBIX document. To realize this access activity, we take HTTP binding and implement the ESI in the REpresentational State Transfer (REST) style. Supporting a resource centric access, instead of method centric one, REST utilizes a small set of verbs to transfer an object's state via XML [38]. Its resource-oriented access mechanism best fits to our data model. Similarly, The RESTful paradigm has been used in recent literatures of smart grid [35].

More specifically, three request types in oBIX are mapped into HTTP methods - Read, Write, and Invoke. Read request uses GET for any object having *href* attribute and returns information of an object as an oBIX document. Write is targeting at any object whose *writable* attribute is set to true, and is implemented with PUT. Invoke supports operations of any object by using the POST method. An oBIX document is passed to a server as an input in both Write and Invoke. Below oBIX document represents a smart plug object, "plug1". The root element indicates its access URI, and the *ref* tells the link to the associated

sub-object, "power". The document also shows that the load is controllable via two ways: Write and Invoke. To turn it on or off, an external system sends directly a PUT request targeting at the *connectLoad* object within the same URI or sends a POST request to the hyperlinked URI targeting at the operation object, *controlLoad*.

```
<obj href="http://myESI/Points/zigbee/plug1/">
  <str name="deviceName" val="BSPE12SOYZM43001"/>
  <str name="version" val="C2V5.57"/>
  <bool name="connectLoad" writable="true" val="true"/>
  <op name="controlLoad" href="control" in="obix:WritePointIn" out="
     obix:Point"/>
  <ref name="power" href="power"/>
</obj>
```

2.3.3 Decentralized Access Control Entry

A security mechanism satisfies the security requirements discussed earlier. To address the integrity issue, we implement HMAC with SHA-256. We address the availability issue with a cloud technology in the following subsection. Access control is more challenging, because our interaction model, WS, exposes the values and operations of each energy resource via three different operations: Read, Write, and Invoke. This property is interpreted as "fine granularity of data access and load controls". Fine granularity introduces a new challenge, because different access actions induce different operation consequences and indicate different levels of privacy violation. Say, you might be okay that ESPs read the energy usage of your air conditioning system, but you do not allow them to turn it off on a hot summer day.

To resolve the fine granularity problem, authors in [10, 51] proposed a new access control concept that performs authorization on the action level (e.g., Read, Write, and Invoke). We adopt Resource Centric Security (RCSec) [51] that implements both access control and encryption in a distributed manner. To address the fine granularity issue, RCSec leverages the concept of a filesystem Access Control Entry (ACE). That is, each object maps to an

attribute with three-digit privilege level. For instance, the object "plug1" in the previous example is related to an attribute "plug1=111". The first digit indicates permission of Read, and the following two digits indicate the rights of Write and Invoke, respectively. Each user maintains his own set of attributes in his private key. When the user tries to access the object with his key having an attribute of "plug1=100", he is permitted to read energy usage data, but not to change any values or control the energy load. In this way, RCSec performs authorization based on what the end user has, and this does not require the ESI to maintain any information for access control such as user id-password pairs and user-privilege mapping rules.

2.3.4 Separating Read Operation from Control Operation

Deploying the ESI on the Cloud is one convincing option for the real world deployment. This is also expected to help resolve the availability issue in the security context by utilizing more reliable and secured computing resource on the Cloud. Despite these benefits, however, we must solve the privacy problem of the Cloud technology, because people do not want the Cloud to read their private data or to control internal energy resources.

Our fundamental approach to address the issue is to extend RCSec so that it distinguishes the Read operation from the Control operation. Then, we move the Read service to the Cloud. To this end, we implement two service interfaces of the ESI and a Database on a Cloud system hosted by Amazon, while control and security functionalities remain in a lightweight mini-EMCS. Figure 2.3 represents the new system architecture. More specifically, all the energy data is first encrypted at the mini-EMCS and then transmitted to the Cloud, and the Database in the Cloud stores encrypted energy data with meta information. So, external users retrieve historical data by contacting service interfaces on the Cloud directly while the Cloud does not access the plaintext of the energy data. The Cloud, on the other hand, forwards the users requests for control operations to the mini-EMCS that performs access control and verifies message integrity.

Separating and moving the Read operation to the Cloud is feasible mainly because RCSec

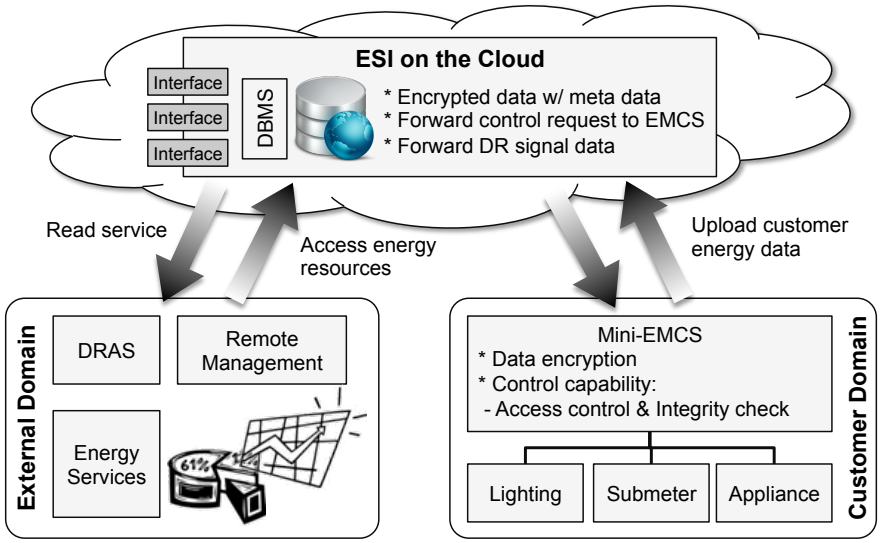


Figure 2.3: Moving the ESI to the Cloud. The energy data is stored on the Cloud, while the customer still has full control over data encryption and access control. The external user contacts only the Cloud to access the customer energy resources.

is implemented based on Attribute-Based Encryption (ABE) algorithm [40] that realizes a secret sharing scheme using a pairing-based cryptography. That is, the mini-EMCS encrypts energy data using a list of attributes, not using specific user' public key or specific symmetric key. Since RCSec decouples encryption from users' identity, the Cloud does not authenticate and authorize the accessing users. Instead, any users can access encrypted energy data via the Cloud, but only qualified ones having the matched set of attributes (protected via the encryption algorithm) can decrypt the ciphertext.

2.4 Experiment and Discussion

2.4.1 Automated Demand Response

In this automated DR experiment, the DRAS server generates a DR event of a Real Time Pricing (RTP) program. The box below shows the event signal - the *SmartClientDREvent-Data* entity in the *EventState* message. The event starts at 1pm and lasts until 4pm. During the event, the unit power price changes every hour; it becomes 2 times, 3 times, and 2 times

more expensive than a normal price. The event data is generated one hour before the event - an hour-ahead DR program.

```
<p:drEventData>
  <p:notificationTime>2012-08-20T12:00:00.000-07:00</
     p:notificationTime>
  <p:startTime>2012-08-20T13:00:00.000-07:00</p:startTime>
  <p:endTime>2012-08-20T16:00:00.000-07:00</p:endTime>
  <p:eventInfoInstances>
    <p:eventInfoTypeID>PRICE_MULTIPLE</p:eventInfoTypeID>
    <p:eventInfoName>price</p:eventInfoName>
    <p:eventInfoValues>
      <p:value>2.0</p:value>
      <p:timeOffset>0</p:timeOffset>
    </p:eventInfoValues>
    <p:eventInfoValues>
      <p:value>3.0</p:value>
      <p:timeOffset>3600</p:timeOffset>
    </p:eventInfoValues>
    <p:eventInfoValues>
      <p:value>2.0</p:value>
      <p:timeOffset>7200</p:timeOffset>
    </p:eventInfoValues>
  </p:eventInfoInstances>
</p:drEventData>
```

The ESI pulls the *EventState* message from the DRAS every 15 minutes. Once the ESI notices that the price goes up, it performs a predefined DR strategy. In this experiment, we register one LED light to our strategy so that the price change is seen through the brightness of the LED. As the price goes up, the LED gets dimmed proportionally. Since the power draw of the LED is proportional to the brightness level, we can observe the change of energy

usage during the DR event as shown in Figure 2.4a.

We note that the experimental results demonstrate an automated DR, but not address the invisibility that the intelligent DR requires. The invisibility still remains a big research challenge that has not been studied much yet. One plausible approach is to prioritize the energy loads so that lower-prioritized loads are shed first upon DR events [15]. The prioritization can be automated by the combination of many factors - residents' load-usage profile, time, space, and environmental context.

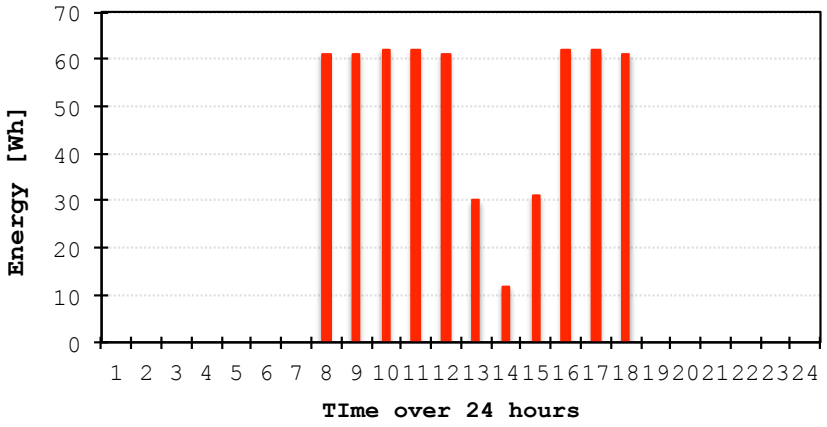
2.4.2 Energy Management Service

Next experiment demonstrates that the ESI provides customer energy services. A user retrieves a historical energy usage data from the ESI. To this end, he prepares an request message as shown in the box below and contacts the service object via a URI to trigger the Invoke operation. The URI in the box implies that the service provides an aggregated energy usage information. The user would retrieve daily usage data of smart submeters and plug load meters for a month.

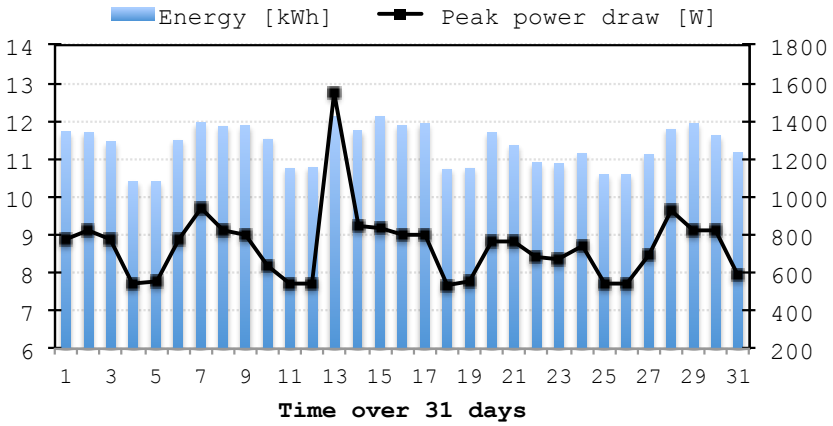
```
http://myESI/History/aggregated/rollup/

<obj xmlns="http://obix.org/ns/schema/1.0">
  <reltime name="interval" val="P1D"/>   // daily
  <int name="limit" val=""/>
  <abstime name="start" val="2012-08-01T00:00:00"/>
  <abstime name="end" val="2012-08-31T23:00:00"/>
  <str name="orderby" val="ttime"/>
  <bool name="reverse" val="false"/>
</obj>
```

The retrieved energy data is drawn in Figure 2.4b. The bars show that energy consumption decreases on weekends, but the difference is slight. This is mainly attributed to the type of deployed energy loads. In our testbed, several servers and switches running 24 hours are



(a) DR event affects the operation of LED. The bars indicate energy usage of the LED over 24 hours. The operation of the LED is also scheduled so that it turns on at 8am and off at 7pm every week day.



(b) Retrieve aggregated energy usage over 31 days. The maximum peak power demand is hit on 13th of the month.

Figure 2.4: Two energy services of automated DR and remote energy management, and their experimental results.

Data size [KB]	10	200	400	600	800	1000
Encrypt [ms]	1383	1369	1372	1380	1389	1399
Decrypt [ms]	225	224	225	229	230	240

Table 2.1: Data encryption and decryption time [ms]

plugged into the plug load meters. And few such loads that is affected by human activities are deployed. We note that a preliminary analysis on the types of energy loads and their operating characteristics is fundamental to establishing the right DR strategy. The figure also draw a curve of peak power draw (maximum instantaneous power consumption), ranging from 500W to 1600W. The maximum peak was hit on Aug. 13 on which we ran another experiment with power-hungry loads such as a dryer and a refrigerator. Regardless of the high value, however, the peak draw hardly affects the energy usage on that day, because the experiment lasted short. However, this high peak matters in practice, since most utility companies charge customers based on their peak values. Finding the energy usage patterns and encouraging to change them to save the energy cost is the primary goal of the energy management service. In the future, the energy data can be further analyzed for advanced energy services. Recently, data mining technology and network optimization algorithms have been leveraged to analyze the time series of energy usage data, which then performs energy efficiency benchmark [63], fault detection and diagnosis, and anomaly detection [22].

2.4.3 Running the ESI on the Cloud

In this experiment, we measure the time overhead of RCSec, once it is applied to our ESI testbed on the Cloud architecture. The mini-EMCS encrypts data using 10 attributes on average. The experiment changes the data size and measure the encryption time. The encrypted data is then uploaded to the Cloud that provides the Read service to end users. We also measure the decryption time at a user. We implement both the mini-EMCS and the user at a conventional laptop computer, running with 2.2 GHz Intel Core 2 Duo processor and 2 GB memory, that performs encryption and decryption, respectively. The box below

Num. attribute	1	5	10	15	20
Overhead [ms]	900	1648	2357	3097	3823

Table 2.2: Overhead of authorization with varying numbers of attributes.

illustrates a sample of energy data that the Cloud provides. It shows that 5 attributes are used in the header and encrypted data is attached in the data element.

```
<obj href="http://Cloud/Points/zigbee/plug1/power">
  <obj name="header">
    <str name="attributes">(ucla_esi AND expire>=1362978960563) AND
       (power>=110 OR (BSPE12SOYZM43001>=100 AND power>=100))</str>
  </obj>
  <obj name="data">FJl31tjE3yDjXdIgNtH715DVuBRUXalfkredN9OmxNyN48...
     </obj>
</obj>
```

Table 2.1 shows that the data size hardly affects the performance. It also indicates that the encryption overhead is comparatively greater than the decryption. This is mainly attributed to difference of mathematical complexity within the encryption and decryption. The encryption algorithm constructs an elliptic curve and performs bilinear maps (pairing) based on the curve's elements. For each attribute, it constructs two bilinear group elements, which requires two exponentiations. The pairings and exponentiations dominate the encryption overhead. In the decryption algorithm, on the other hand, bilinear map operations are partially replaced by exponentiations, which is more lightweight, in a recursive manner. This optimization reduces the processing overhead. We refer [76] for detailed analysis. The observation that the decryption overhead is comparatively low benefits the Cloud architecture. The customer domain encrypts data in advance, and users do not experience it. Thus, the decryption overhead only contributes to the overall service performance. In this sense, the Cloud architecture considerably compensates the overhead of the security mechanism.

Next experiment measures the overhead of the authorization protocol for the control operation that travels to the EMCS via the Cloud. As more customer equipments are connected to the EMCS, the authorization rule would also get complex. RCSec handles such complexity by applying varying numbers of attributes in the authorization protocol. This experiment, therefore, changes the number of attributes and measures the protocol latency. The result, shown in Table 2.2, indicates that the latency overhead is non-trivial in numbers. In particular, the processing overhead overwhelms the communication latency (89% vs. 11% on average, not shown in the table). However, when considering that the control operation occurs infrequently and does not require real time performance, the overhead is acceptable. Moreover, the overhead is compensated by benefits coming from the decentralization property of the protocol. In the authorization protocol, an end user does not register to the EMCS. Instead, he obtains his own private key from a certificate authority. Complex authorization rules are embedded into exchanged messages, instead of being managed by the EMCS. Then, authorization is carried out on-the-fly using the private key. This property makes the EMCS much simpler and more lightweight. Since the protocol completely separates the EMCS from the user, it easily scales well in a distributed environment such as smart grid.

CHAPTER 3

Realizing Smart Grid Interoperation - Microgrid Platform

In the future smart grid, the customer domain will also generate and store energy with potential inclusion of Electric Vehicles (EV), batteries, and Distributed Energy Resources (DER). With such increasing energy capabilities, the domain technologically advances and the smart grid expects it to operate as a *Microgrid* [49]. Up to date, however, the EMS concept is simple and its core functions are too limited to support microgrid operations. The microgrid must be able to *manage the new types of energy resources efficiently* so that they interoperate with traditional energy loads. In addition, it must interoperate with the grid infrastructure to achieve the energy balance over the entire smart grid - that is, it must *be smart gridable*.

To address the issues, we enhance the EMS model and propose *Microgrid Platform (MP)*. We design and develop a MP system that takes into account all the three demands (including fine-grained management). Moreover, we deploy the system in a small customer site, where MP interoperates with a real-world retail energy market. The contributions of this chapter are three fold:

Designing energy management system with microgrid perspective. We examine the design issues of MP with two topics. An autonomous operation discusses DER management; energy forecast; and data analysis. A smart grid interoperation classifies energy services that MP must support into two types and discusses each of them in detail. We also review previous works in each topic.

Implementing and evaluating a MP prototype in a laboratory environment. We

implement a building-level testbed in which various types of energy resources are deployed. MP implements many protocols and communication technologies to connect to them and develops advanced data processing and control-decision mechanisms for microgrid operations. The testbed also realizes a few energy services that realize the smart grid interoperation. On top of the MP testbed, we conduct experiments to demonstrate the capability of our microgrid platform.

Deploying and running MP in the real-world energy market. We deploy MP in a small customer site having two 4-stories buildings. With MP, the customer participates in an Automated Demand Response (ADR) service that a local utility company offers. In the ADR service, the customer automatically curtails his energy consumption when MP receives DR signals from the utility, expecting financial incentives. In this paper, we share our experience and discuss lessons learned from the real-world energy service.

3.1 Energy Management in Building Facility

Energy management system. Traditionally, a Building Automation System (BAS) controls facility equipment simply for the purpose of occupants' comfort and optimal business operations. For instance, as an integral part of the Heating, Ventilation, and Air Conditioning (HVAC) system, the BAS accepts inputs from temperature and humidity sensors and controls the heater and chiller to optimize the working environment. Today, the introduction of the smart grid changes the customers' awareness and expectation on their energy management. They want to see breakdowns of their energy usage and to take actions to reduce energy costs.

To meet the emerging customer needs, recent research on the customer facility has developed an EMS that replaces the unintelligent BAS system. It performs fine-grained energy measurements and controls, say, at individual home/office appliance level. It optionally analyzes the collected data and controls equipment in a way to maximize the energy efficiency.

Although the EMS model makes the customer facility more intelligent, however, it still misses two important functional demands: *autonomous building operation* and *smart grid*

interoperation. The autonomous operation denotes the capability of enabling various types of energy resources to interconnect with each other so as to manage the overall energy balance at the facility level. These days, more building owners are interested in instrumenting new types of energy resources like a solar panel or a battery within the facilities. Leveraging the resources, for instance, the EMS must be able to decide when to purchase power from the grid to charge a battery and when to use the battery to power building equipment while not sacrificing the occupants' comfort. The smart grid interoperation is that the facility contributes to maintaining the energy balance at the entire grid level. With increasing capabilities of energy consumption, generation, and storage, the facility can respond to the grid's signals by reducing power consumption or by feeding power back to the grid in real time. The EMS must also play a communication gateway role to enable the interoperation across domains.

Microgrid model. These new requirements and corresponding behaviors in the facility are explained by a Microgrid model [49]. Coined in the field of DER, a microgrid is defined as "a localized grouping of electricity generation, energy storage, and loads that normally operates connected to a traditional centralized grid"[1]. It is able to separate and isolate itself from the grid seamlessly with little or no disruption to the internal loads during a grid disturbance. We envision that the microgrid is the future model of a customer's building facility or at least a group of facilities. The existing EMS model, therefore, must be enhanced to support two additional functions that are essential to realize the microgrid. To address this issue, we propose a *Microgrid Platform (MP)*, the implementation framework of an advanced EMS system for microgrid operations.

3.2 Microgrid Platform

This section discusses the design issues of MP to support two functional demands. For better understanding, we illustrate a system architecture where MP interconnects with both customer energy resources and external systems in Figure 3.1.

[1]http://en.wikipedia.org/wiki/Distributed_generation

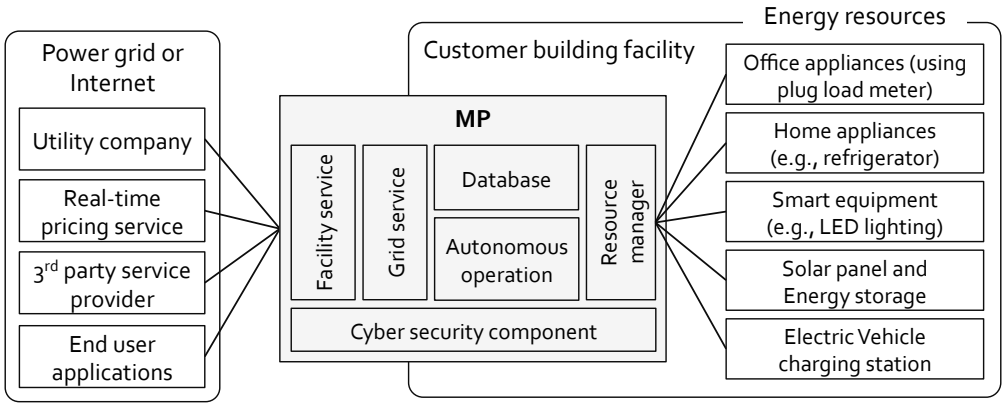


Figure 3.1: An information system architecture around the customer domain, including MP, energy resources, external systems, and energy services.

3.2.1 Autonomous Operation

A unique feature in the microgrid, comparing with conventional building facilities, is its enhanced energy capabilities: consumption, generation (say, using DERs), and storage. The operational goal of MP is to manage the emerging energy resources efficiently and thus to make the microgrid operate intelligently. To these ends, we discuss three issues.

3.2.1.1 Managing distributed energy resources

A DER represents electricity generation that feeds into the distribution domain directly. In terms of the customer facility, it generally refers to on-site renewable sources: Photovoltaic arrays and thermal energy storage. It benefits the microgrid in ways of enhancing local reliability, reducing power losses in transmission, and supporting local voltages. The DER also enables the microgrid to respond in real time to the grid's signals without interfering with normal building operations. Despite such potential advantages, however, one of the biggest obstacles is on contextual heterogeneity. Because a number of vendors are still using proprietary protocols, the DERs barely interoperate with each other and with other energy resources seamlessly.

To resolve the problem, a few standardization efforts have proposed specifications: IEC

61850[2], IEEE 1574[3], OPC Unified Architecture[4], open Building Information eXchange (oBIX)[5], Facility Smart Grid Information Model (FSGIM)[6]. In academia, researches have developed programmable APIs that allow us to access customer energy resources in a unified manner [16, 48]. Dawson-Haggerty *et al.* leverage SensorML [25] and define a data model that fits to tiny sensor devices embedded into energy resources [35].

3.2.1.2 Forecasting energy activities

As the energy activities in the facility become more dynamic and complex, it is critical to predict them accurately for the purpose of the energy balance at both the facility and the grid levels. First, MP performs a demand forecast. An increasing number of power hungry equipment such as air conditioning and EV has been installed even within small houses. This makes the energy consumption unpredictable. MP must be able to estimate future energy usage accurately, which is of the most important to both the microgrid and the smart grid. Second, MP performs a generation forecast. Although the renewables represent clean energy sources, they still suffer from being variable or intermittent. As the smart grid relies on them, such unpredictability may threaten the energy balance. Accurate generation forecast can mitigate the risk. Last, MP performs a storage forecast. It can be computed from the above demand and generation forecast today. If the storage accounts for the EV battery in the future, the forecast computation will become more complicated as EVs move around.

To make an accurate forecasting, MP uses various types of data sources - from historical data to mathematical model, weather data, and other societal data. Zhu *et al.* run demand forecast and solar generation forecast from history data, and then develop a battery (dis)charging scheduling algorithm [87]. Huang *et al.* propose a hybrid mathematical model that takes weather forecast and history data to improve the prediction accuracy of a solar panel over both short-term and mid-term periods [42].

[2]http://www.iec.ch/smartgrid/standards/
[3]http://grouper.ieee.org/groups/scc21/1547/1547_index.html
[4]https://www.opcfoundation.org/UA
[5]http://www.obix.org
[6]http://spc201.ashraepcs.org/standards.html

3.2.1.3 Making analysis and control decision

MP collects a huge amount of data from the energy resources in real time. The data collection is further analyzed so as to estimate energy cost; to understand the building's energy usage pattern; to compare it with neighbors' through energy benchmarking tools; to pinpoint where the energy is being wasted more accurately; and to make a control decision for efficient energy management. For instance, Bellala *et al.* analyze time series data of energy usage in a commercial campus [21]. Then, they detect anomalous usage periods representing unusual power consumption. Detecting and correcting the anomaly can save the electricity bill. In [66, 67], given EV owners' charging profiles and real-time power price, the authors develop a Vehicle-to-Grid (V2G) scheduling algorithm that works at a large scale EV charging structure. The Monitoring-Based Commissioning (MBCx) project exploits the measurement data and diagnostic tools in order to perform commissioning[7] on 24 non-residential buildings throughout the state of California [68].

3.2.2 Smart Grid Interoperation

The microgrid interacts with external systems, and energy services instantiate such interoperation. A couple of literatures present use cases of energy services [7, 41] and we classify the services into two categories: facility service and grid service. In the facility service, the customer facility provides service data to external systems in the grid infrastructure, whereas it receives and consumes service data delivered from the grid in the grid service. MP must be able to support both services.

3.2.2.1 Facility energy service

The facility serves as a service provider to external systems. MP provides a *data service* that transmits data related to energy resources, which includes:

- History data is a collection of historical data of energy measurement per each energy

[7] Commissioning is a process of verifying energy performance and design intent for non-residential buildings and correcting deficiencies.

resource.

- Current measurement includes information of status and power on energy resource. "Resource status" represents current status of an energy resource. In addition to simple on/off equipment, an LED light has two status types of brightness and temperature, whereas the status of a power storage is indicated by State of Charge (SOC) [%]. "Resource power" represents the instantaneous capability of power consumption, generation, and storage. In addition to current, voltage, power, and energy values, we include power quality data (e.g., reactive power) in this category.

- Future forecast is an estimation of future energy activity on each energy resource.

In addition, MP provides a *control service* in which it accepts command messages from external systems that eventually control internal resources or to change their configuration.

MP provides/accepts service data to/from external systems via service interfaces. To this end, public reports introduce an *Energy Service Interface (ESI)*, a communication gateway software through which the microgrid interacts with the grid infrastructure [12, 36, 41]. Authors in [52] discuss three design issues to develop the communication interface. First, interface abstraction determines the appropriate level of internal details that it exposes to external systems. High level of abstraction exposes internal business logics with less details, while low level of abstraction exposes more details of internal operations. A well-designed abstraction transmits only the necessary service data to external systems, while shielding them from changes occurring within the customer facility. Next, MP must represent service data in a standardized format. To address the issue, the smart grid takes a Canonical Data Model (CDM) approach. A data producer transforms its output to a standardized information model, and then a consumer transforms it back to own terminology. Last, MP must support efficient interaction models that determine how MP communicates with external systems. Recently, a Web Service (WS) has been paid attention as a candidate due to its interoperability and scalability.

3.2.2.2 Grid energy service

The grid provides energy services to the facilities in the smart grid. A few services have been introduced, which can be classified into two groups.

Direct load control. Direct Load Control (DLC) permits external systems to control internal energy resources. Thus, it corresponds to the control service that the facility provides. Currently, utility companies provide the primitive form of services targeting at specific energy loads such as an electric water heater. In the future, they will control various types of energy resources for different purposes. For instance, an energy service provider remotely stops charging an EV at a customer's garage when the power price goes beyond a contracted value, which saves the customer's electricity cost.

Facility-centric load control. In Facility-centric Load Control (FLC), the grid transmits power-related signals (not control messages) to the facility that, then, takes own control actions based on information in the signals.

An *Automated Demand Response (ADR)* is a well-known FLC service. When power suppliers confront a shortage of generation capacity during peak-demand periods, a Demand Response Automation Server (DRAS) generates and transmits a DR event signal to customers to encourages them to reduce their energy consumption during the event period. Upon receiving the signals, thus, MP executes load curtailment (shedding and shifting the energy loads) of a service contract. The customer expects financial benefits for responding to the event signals. Lawrence Berkeley National Laboratory (LBNL) published an interoperable communication specification for the ADR, Open Automated Demand Response (OpenADR) 1.0 [75]. The next version 2.0 has been officially standardized by the OpenADR alliance[8].

A *Real-Time Pricing (RTP)* is another FLC service that the grid will provide in the near future. While the DR signal may contain real-time price, we see the RTP service as continuous changes on the power price, not event-based changes. Moreover, it has a different strategy to encourage customers to cooperate for the energy balance. The ADR service aims

[8]http://www.openadr.org

at incentivizing participants responding to intermittent DR events while the RTP service expects energy customers to take actions to reduce their own energy bills. With this new approach, we expect the customers (i.e., MP) to interact with the service more actively. As of today, the service model has been under discussion in parts of Energy Interoperation (EI)[9].

The next FLC service will be a *Transactive Energy (TE)*. In the future, we will have active and dynamic retail energy markets. There, the microgrid having multifarious energy capabilities will interact with the market more frequently: it may choose to buy power from one or more power suppliers and/or generate power on-site for sale in the market. The customers buy or sell electricity in more flexible ways as "prosumers". Say, one may sell the power in his/her energy storage directly to neighbors in the community. A collaborative group from both private and public sectors is leading the discussion on the TE markets[10]. The EI has defined the communication specifications for the TE services, which leverages Energy Market Information Exchange (EMIX)[11] and Web Service Calendar (WS-Calendar)[12].

3.3 Implementation of Microgrid Testbed

To demonstrate the feasibility of MP, we deploy a testbed by extending our previous system [52]. We additionally instrument new types of energy resources and develop the Microgrid Platform running on top of a Linux distribution. For experiments, we also develop several external energy services.

3.3.1 Energy Resources

The energy resources considered in the testbed are smart submeter, office/home appliances, smart equipment, EV charing station, and solar panel. Figure 3.2 pictures some of them.

Smart submeter. Unlike a conventional smart meter that measures aggregated energy

[9]https://www.oasis-open.org/committees/energyinterop
[10]http://www.gridwiseac.org/about/transactive_energy.aspx
[11]https://www.oasis-open.org/committees/emix
[12]https://www.oasis-open.org/committees/ws-calendar

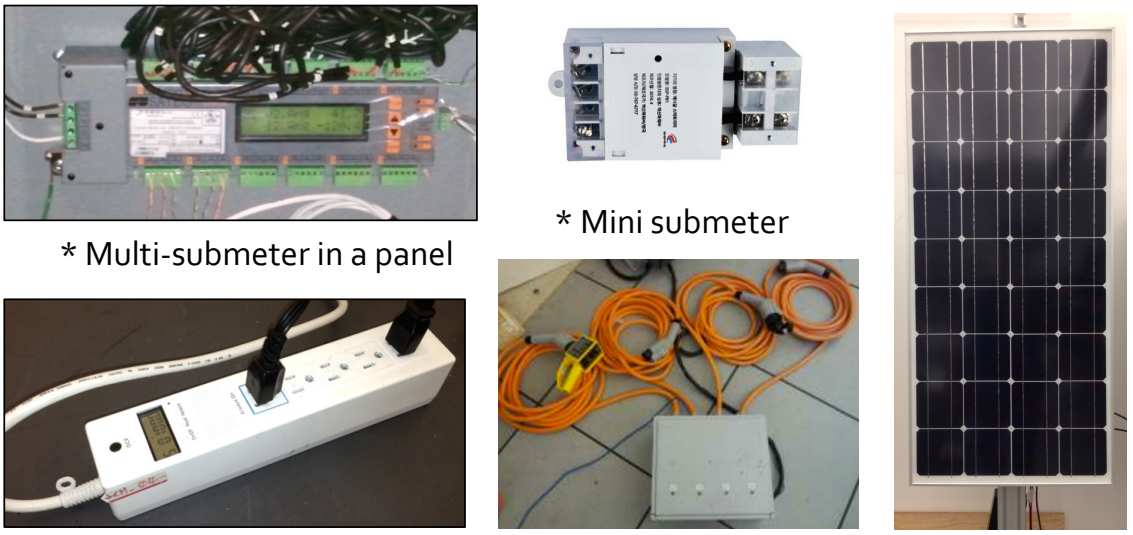


Figure 3.2: Energy resources in the Microgrid testbed.

usage, a smart submeter provides fine-grained measurement and control. Our testbed deploys two types of submeters. We instrument a panel-level multi-submeter that simultaneously connects up to 36 single phase circuits within a panel[13]. Using it, we monitor two groups of energy loads - the lightings and power outlets at an office. We also install mini submeters that are instrumented to single power lines[14]. For instance, it can directly connect to a light switch that turns on/off a set of fluorescent lights. These submeters use Current Transformer (CT) to convert current to voltage, and an embedded microcontroller calculate the real, reactive, and apparent powers and energy usage. They are with relays, and the microcontroller switches the power upon requests.

Office appliance with plug-load meter. As the plug-loads including all the office appliances account for more than one third of the total power consumption in a building [17], it is necessary to mange them carefully. To this end, we deploy two types of plug-load meters: smart plugs and smart power strips. Office appliances are plugged into them: computers, monitors, desk lamps, and network switches. The plug-load meter is functionally same to the submeter, i.e., energy measurement and control. It communicates with MP using a ZigBee

[13]http://www.satec-global.com/eng/products.aspx?product=42
[14]http://www.bspower.co.kr/en/smartmeter.do

module.

Smart equipment. Smart equipment represents such energy resources that must be accessed directly. Recent programmable thermostat and LED lights fall into this category. Each equipment has its own operation cycles beyond a simple on/off control and is able to adjust the operations upon external requests. Our testbed deploys dimmable LED panel lights that adjust their brightness and color temperature in 8 steps. Each light uses a ZigBee module to transmit its status and to accept control commands to/from MP. For scalable experiments, we additionally develop a light emulator that creates 200 LEDs, each of which operates exactly same as the real device (brightness, temperature, and energy consumption).

Smart home appliance. The home appliances are functionally same to the smart equipment. Each manages own operation cycles and must be accessed directly. MP connects to two types of appliances via the Ethernet: a clothes dryer and a refrigerator[15]. It is able to change the strength of the heat (high, low, or no heat) as well as turn on/off the operation by sending signals to the dryer. The refrigerator adjusts the operating cycles of compressor, defrost, and fan. To measure energy usage, the mini submeters are instrumented to their input power cables.

EV charing station. UCLA has instrumented a number of charging stations at campus parking structures [28]. Each station powers several EVs via J1772 connectors simultaneously and supports multiple charging levels in a fair and safe manner [29,32]. It is capable of measuring charging capacity as well as charging rate. Each station sends the charging data in real time to a management server in our laboratory that controls the stations based on subscribers' profiles and preference [30, 31, 82]. MP commuicates with the stations via the server. Because of low penetration of EVs, however, we could not collect enough data for experiments. As a complementary work, we simulate charging activities based on measurements and obtain ample amount of data.

Solar panel. MP also connects to a Photovoltaic (PV) solar panel. We are currently installing a new one on the roof, and, yet, this version of testbed implements a virtual panel

[15]http://smartgrid.ucla.edu/projects_adr.html

Notation	Description
C	capacity, $C = 5$ kWh
N	number of cells in a series to form 1 kWh of array, $N = 8$
A	area of 1 kWh panel, $A = 9.952\ m^2$
ϵ	solar generation efficiency, $\epsilon = 16.5\ \%$
V_{max}	maximum voltage on each solar cell, $V_{max} = 41$ V
I_0	saturation current, $I_0 = 1.919 \times 10^{-40}$ A
I_L	light-generated current, $I_L = 5.754$ A
q	elementary charge, $q = 1.6 \times 10^{-19}$ Columbus
κ	Boltmann constant, $\kappa = 1.38 \times 10^{-23}$ J/K

Figure 3.3: Notation used in our virtual PV resource.

that follows the same hardware specification of the real device. Our panel resource obtains the real-time solar radiation data (Global Horizontal Irradiation (GHI), Direct Normal Irradiation (DNI), and temperature $T\ [^{\circ}C]$) from Solar Resources and Meteorological Assessment Project (SOLRMAP)[16] and then computes power ($\mathbb{P}$), voltage ($\mathbb{V}$), and current ($\mathbb{I}$) values using the following equations. Fig. 3.3 gives the description of notation used in our scheme. Solar power is

$$\mathbb{P} = \frac{GHI+DNI}{2} \times A \times \epsilon \times C$$

For voltage $\mathbb{V}$, we compute P_i for V_i running from 0 to V_{max} every 0.01 step as following.

$$P_i = V_i \times (I_L - I_0 \times (exp(\frac{qV_i}{\kappa(T+273.15)}) - 1))$$

Then, we pick V_j such that $|P_j - P_c|$ is minimum, where $P_c = \frac{\mathbb{P}}{N \times C}$. The voltage at this point, j, represents the most accurate value that we can obtain. Then, we compute $\mathbb{V} = V_j \times N$ and $\mathbb{I} = \mathbb{P}/\mathbb{V}$. In this way, we update energy data of two virtual PV resources every 30 min.

[16] http://www.nrel.gov/midc/

3.3.2 Microgrid Management

MP communicates with the energy resources via Ethernet, RS-485 serial, and IEEE 802.15.4. It supports various application protocols such as Modbus, IEC DLMS (Device Language Message Specification), SEP 1.0, and several proprietary protocols. MP collects and stores both power-related measurement and status information from the energy resources every 5 minutes on average. In addition, it maintains meaningful meta data regarding each resource. For instance, each mini submeter is managed with a load type, location, and the load's priority. A resource owner configures the meta data, and thus the data keeps reflecting physical characteristics of the plugged load and user contexts. MP also develops a control strategy. It prioritizes energy resources based on resource type, current status, criticality, and capacity, and then determines groups of resources and corresponding control sets. Upon receiving a DR signal, for instance, it finds a predefined strategy corresponding to the DR event and executes the control set. MP provides a scheduling function through which a user pre-schedules the operations of energy resources. The dimmable LED lights are now reserved to be ON only during office hours, while a user can still turn them on/off any time.

3.3.3 Energy Services from the Grid

In addition to basic DLC services, our testbed implements two FLC type of services in which the microgrid is interested most - the ADR service and the RTP service.

Open Automated Demand Response. We deploy an OpenADR 1.0 server that provides the ADR service by exploiting the open source [11]. The server generates an *EventState* signal to trigger a new DR event. It supports connections from both smart and simple clients. The smart client is capable of interpreting the *EventInfo* information within the *EventState* signal. Included in *SmartClientDREventData* entity, it contains event details. For instance, the *eventInfoTypeID* denotes an event type and takes one of values of PRICE_ABSOLUTE, PRICE_RELATIVE, LOAD_AMOUNT, etc. For the simple OpenADR client, the server translates the *EventInfo* information into a simpler form, named *SimpleClientEventData*. In the entity, two variables describe the event state. The *EventStatus* element indicates the

temporal state of the event (FAR, NEAR, or ACTIVE), whereas *OperationModeValue* denotes the operational state of the energy loads in the event (NORMAL, MODERATE, or HIGH). To address the security issue of the OpenADR, in particular the message integrity, we implement a Message Authentication Code (MAC) on top of the existing OpenADR. Following the NISTIR 7628 guideline [83], our testbed takes a hash-based MAC (HMAC) with SHA-256.

Real-time pricing for retail energy market. To assess the feasibility of the RTP service, our testbed implements an RTP server that provides price forecast for a retail energy market. The server, in the absence of an RTP model in the real world, exploits the wholesale market price provided by California Independent System Operator (CAISO)[17]. More specifically, it obtains three types of price forecast from CAISO - Day-Ahead Market (DAM); Hour-Ahead Scheduling Process (HASP); and Real-Time Market (RTM). The DAM provides an estimated power price of every hour for 24-hour ahead. The HASP and RTM provide an hour-ahead/10-min-ahead price estimation of every 12/5 minutes, respectively. Since CAISO does not provide the price forecast for the location of our campus, the server takes the price value for the city of Long Beach. The RTP server also takes inputs of demand forecast and weather forecast from CAISO, and then eventually determines three types of price forecast (DAM, HASP, and RTM) for the retail energy market.

Consuming the service data. MP implements communication counterparts of the above two energy services for interoperations. With respect to the ADR service, it implements both smart and simple clients that periodically "pull" the *EventState* message from the server. This PULL mode is often preferred over a PUSH mode since the OpenADR client has more control over the communications, e.g., firewalls. It, then, identifies when the event starts and ends and other event contexts. MP also pulls the price forecast from the RTP server periodically. Different applications may use three types of forecast differently. Our testbed primarily fetch the HASP and RTM forecasts every hour and 10 minutes and executes scheduling algorithms according to the price changes.

[17]http://oasis.caiso.com/mrioasis/

3.3.4 Energy Service Provider

MP provides energy services to the grid, which makes the microgrid play an energy service provider role in the smart grid. In addition to basic energy services, it realizes the facility-side forecasting that helps the grid understand the facility's energy behaviors accurately.

Basic energy services. MP provides fundamental data services that most EMSs can do. These include (1) historical energy data for individual resource as well as for the aggregated one; and (2) real-time measurement on resources' status, their energy activities (consumption, generation, and storage), and power quality. (3) MP also accepts command messages from the grid that eventually control the internal energy resources. This corresponds to the DLC service on the grid side.

Demand forecasting. MP performs a demand forecast on individual resource level that estimates future energy consumption every hour ahead. To this end, it takes a persistence model in which the future demand is estimated fully based on historical data. This is highly effective in very short-term prediction, i.e., an hour-ahead. Due to its simplicity, the model has been widely used by public sectors. In particular, we extend an existing customer baseline load calculation that has been used by local utilities to calculate customers' curtailment on DR events. That is, given historical data over last 5 weeks, we put weights on the day and latest data and estimate the demand of next hour.

Generation forecasting. MP also performs generation forecasting on our solar panels. To this end, it implements a hybrid model that integrates a persistence model and a statistical model to ensure high accuracy over both short-term and mid-term predictions. More specifically, our model takes Auto-Regressive Moving Average (ARMA) [85] as the statistical model, and the following equation describes the process to predict solar generation $S(t)$ at time t.

$$S(t) = \sum_{i=1}^{p} \alpha_i S(t-i) + \sum_{j=1}^{q} \beta_j e(t-j),$$

where the first term, the Auto-Regressive (AR) part, includes the order p of the AR process and the AR coefficient α_i, and the second term, the Moving Average (MA) part, includes

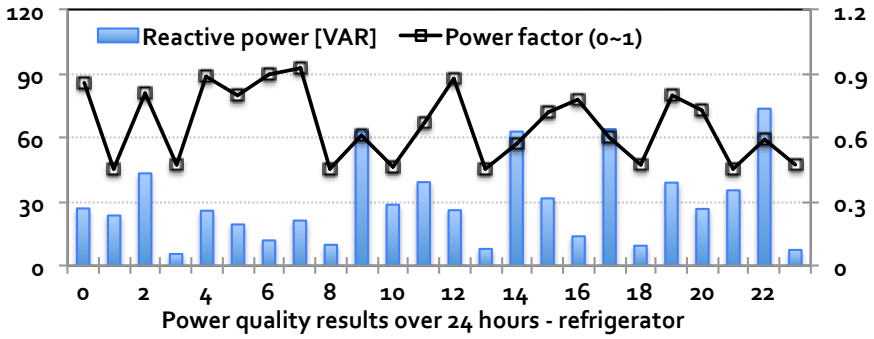


Figure 3.4: Reactive power and corresponding power factor on a refrigerator over 24 hours.

the order q of the MA process, the MA coefficient β_j, and the white noise $e(t)$.

Storage forecasting. Given both the demand and generation forecasting, MP performs a storage forecasting. To this end, we run a Battery Management System (BMS) that stores energy, generated from the solar panels, into a virtual energy storage whose maximum capacity is 25 KWh. The forecast value represents the surplus or shortage of power within the microgrid at a moment and thus indicates the possibility of microgrid islanding.

Energy service interface. MP develops the ESI using the existing implementation model [52]. That is, the service data is represented via the oBIX specification [33] and is then exchanged via the Web Service model with REST (REpresentational State Transfer) style [38]. The RCSec (Resource Centric Security) carries out access control on action levels (i.e., Read, Write, and Invoke) [51]. In addition to the oBIX, we extend the IEC 61850 specification [44] to represent data from our solar panels and energy battery.

3.4 Experiments - Microgrid Operations

This paper focuses on the microgrid operations on our laboratory-level testbed[18]. To this end, we run the following experiments - measuring power quality; computing energy forecast; responding to the real-time price; and leveraging the DERs and energy storage.

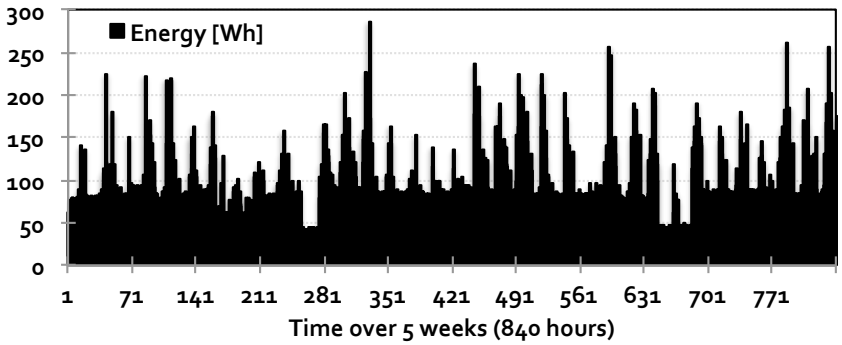
[18]We note that this paper omits basic experimental results, i.e., measurement of energy usage and resource control.

3.4.1 Power Quality Measurement

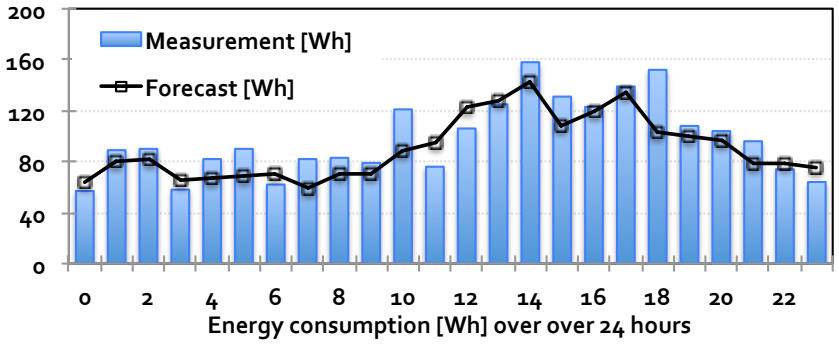
Our experiment measures the power quality on a refrigerator. The reactive power and the power factor are computed by the phase difference (θ) between the input voltage (V) and current (A), and together represent the amount of energy being wasted. For instance, the power factor of 0.7 means that the 100 W device requires 143 VA (Voltage-Ampere) apparent power (=100/0.7) to operate, and 43 VA is wasted. The experimental results in Figure 3.4 show the power factor of 0.66 (=$\cos\theta$) on average and the accumulated reactive power of 714.8 VAR (VA Reactive) that is compared with the consumed energy of 726.8 Wh over 24 hours. The θ value in our experiment is mainly attributed to inductance and capacitance in the electric circuit of the refrigerator, and can be usually compensated by using a synchronous condenser. A combination of inductive and resistive loads at the customer facility can be used in grid stabilization services such as frequency regulation and VAR compensation. An appropriate control on the loads can help balance active and reactive power in the smart grid.

3.4.2 Energy Forecasting

Demand forecast. In the demand forecast experiment, we train our persistence model with energy usage data of a smart power strip. A couple of office appliances are plugged into the strip. And, two students come and go to use it to power their laptops. So, the usage patten is more like irregular as shown in Figure 3.5a. It consumes 95 Wh on average with max of 286 Wh and min of 26 Wh. The strip consumes energy more than the average for 294 hours out of 840 hours, which computes around 35% in the whole measurement. The standard deviation of the usage values is 40.45. Based on the training, MP performs a demand forecast for the next 24 hours as shown in Figure 3.5b. The figure also draws bars representing the real measurement in order to verify the accuracy of the forecasting. Our model forecasts energy demand of 90.2 Wh a day, but the strip consumes 97.8 Wh in real, which computes around 92.2 % of accuracy. The highest accuracy appears at hour 13 with 97.6 %, while the lowest one is 67.6 % at hour 18.



(a) Energy usage data over last 5 weeks for training.



(b) Demand forecasting over 24 hours and testbed measurement.

Figure 3.5: Demand forecast on a set of energy loads.

Month \ Hour	0-7	8	9	10	11	12	13	14	15	16	17	18	19-23
Jan.	0	0	0.35	1.2	2.1	2.2	2.0	1.9	1.8	1.2	0.43	0	0
Aug.	0	0.15	0.42	1.7	2.5	2.8	2.7	2.6	2.3	1.8	1.2	0.54	0

Table 3.1: Power generation [KW] on a solar panel in summer and winter.

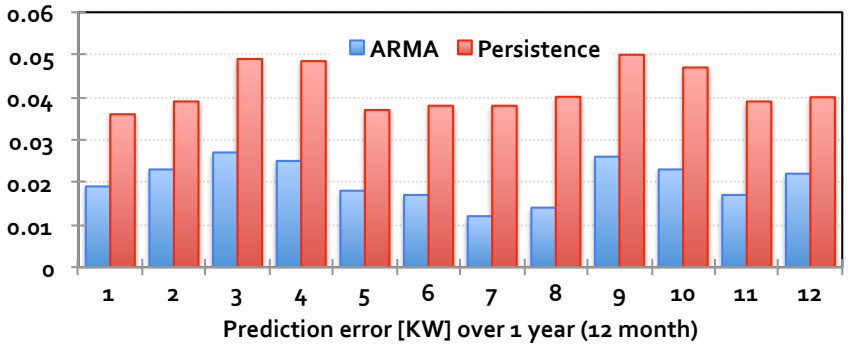


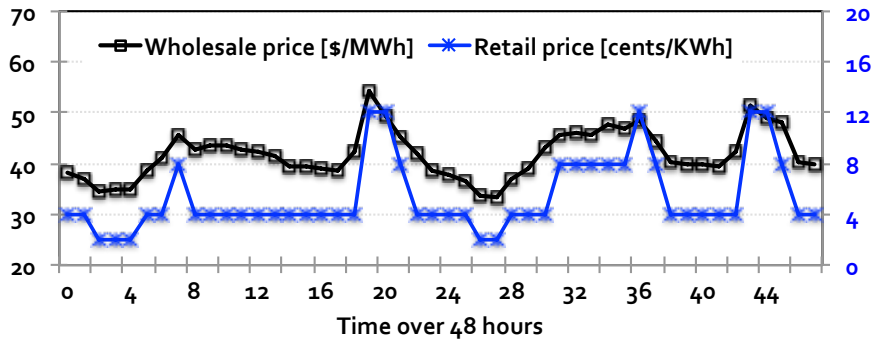
Figure 3.6: Prediction error [KW] with one hour-ahead solar generation forecast using MSE over one year.

Solar generation forecast. Our solar panel is fixed with an angle of 20 degree, and its maximum capacity is 5 KW. The relationship among power, voltage, and current is as follows: maximum point voltage is 328 V; maximum point current is 10.58 Amp; open circuit voltage is 381 V; and short circuit current is 11.5 Amp. Table 3.1 shows power generation of our solar panel in August 2012 and January 2013. During a winter day, the panel generates 13.2 KW of power for 10 hours with efficiency of 15.3%. In winter, it generates 18.7 KW of power for 12 hours with efficiency of 15.6%.

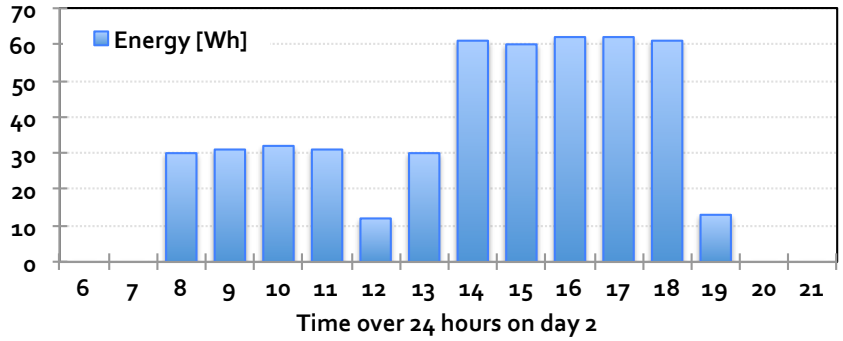
We measure the prediction accuracy through Mean Squared Error (MSE). To this end, we take solar forecast data over one year from the ARMA model and obtain real measurement data. For comparison, we also implement a simple persistence model, run prediction, and compute MSE values. Figure 3.6 illustrates the error values [KW] over one year. Taking January as an example, the ARMA model shows better performance than the persistence model by 44.38%. The error in the ARMA model shows 0.0206 KW on average with max of 0.028 KW in March and min of 0.0121 KW in July.

3.4.3 Responding to Real-Time Pricing

The next experiment runs the RTP service. The RTP server acquires a power price forecast from a wholesale market in California. Figure 3.7a draws a curve of day-ahead prices (DAM) over 48 hours. The wholesale market price is 41.7 [$/MWh] on average with max of 54.27



(a) Wholesale market price (from CAISO) and corresponding retail market price over 48 hours.



(b) The adjustment of brightness on LED affects the energy consumption.

Figure 3.7: ADR server generates retail market price, and MP responds to it by adjusting the brightness of LED light.

and min of 33.37 during the period. The server, then, determines a retail market price based on the wholesale market price and other factors. Say, a unit price is 4 [cents/KWh]. The star-marked line in the figure shows the changes of power price in the retail market.

MP periodically fetches the retail market price from the server and is informed of the price increasing. Taking the price values between 7AM and 2PM on day 2 in Figure 3.7a, for instance, the price becomes 2 times, 3 times, and 2 times more expensive than the unit price. MP responds to this change by performing a predefined control strategy. In this experiment, we register one LED light to our strategy. As the price goes up, the LED gets dimmed proportionally. Since the power draw of the LED is proportional to the brightness level, we easily measure corresponding changes of energy usage as shown in Figure 3.7b. We note that the LED is also scheduled to turn on during office hours (8AM to 7PM) only. Taking

Power capability	Energy resource	Specification
Generation	solar panel 1	5KWh, fixed angle of 20 degree
	solar panel 2	5KWh, tracking angle
Consumption	EV 1	110V, peak demand of 1500W
	EV 2	240V, peak demand of 3500W
	LED light	peak demand of 60W
	plug-load meter	peak demand of 200W
	refrigerator	peak demand of 500W
Storage	battery	25KWh of capacity

Table 3.2: Configuration for DER experiments. We measure data every 15 min.

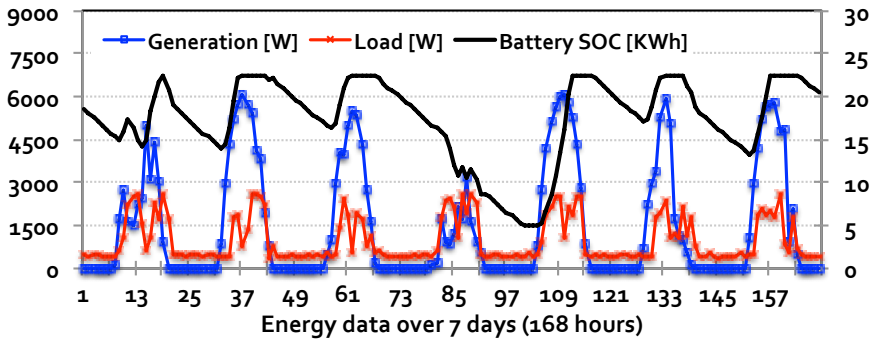


Figure 3.8: Microgrid operation with power generation, load, and energy storage.

the initial result, we run another experiment with 200 LED lights. They are deployed in 4 different offices, and occupants set priorities on lights individually. We also take into account potential inputs from occupancy sensors and brightness sensors. In this way, all the lights are ranked instantaneously, and MP adjusts their brightness according to the price changes.

3.4.4 Utilizing Distributed Energy Resource

Given the capabilities of both power generation and consumption, this experiment runs the battery management system to show that MP manages the energy storage to achieve autonomous microgrid operations. To this end, we take our experimental data as shown in Table 3.2. Both the charging and discharging efficiency on the battery are 80%. That is, we lose 20% of power in the procedure of charging and discharging. The maximum State

of Charge (SOC) is 90%, i.e., 22.5 KWh, while the minimum SOC is set to 20% (5 KWh). Whenever the solar panels generate power, MP stores it to the battery. The energy loads draw power from the battery first. They use another source of power (e.g., from the grid) when the battery is fully discharged.

Figure 3.8 illustrates the experimental results over 7 days (168 hours) in 2013. The square mark curve (blue) represents aggregated power generation by two solar panels - max of 6,066.8 W and min of 0 W. The star mark curve (red) represents aggregated power consumption - max of 2,596.9 W and min of 351.3 W. The solid line (black) represents the SOC of the battery that both the generation and consumption influence directly. Since the generation is usually greater than the consumption, the SOC often reaches to the maximum of 22.5 KWh. The panels generate 67.3 KWh of surplus power for 30 hours in total. On the other hand, the demand exceeds the power capacity of generation and storage from hour 100 to hour 103. This is mainly attributed to low power generation on day 4, which exemplifies the unpredictability of renewables.

3.5 Field Study - Automated Demand Response

In this section, we deploy our Microgrid Platform in a small customer facility, helping the building owner participate in the ADR service.

3.5.1 Preliminary

3.5.1.1 Existing ADR service in energy market

Utility companies offer conventional tariffs such as Critical Peak Pricing (CPP) to customers in the ADR service. That is, customers are charged their energy bills based on the contracted tariffs, while an additional incentive is granted based on their participation on DR events. More specifically, a customer subscribes to the ADR with "curtailment rate", the amount of power that the customer must reduce upon receiving a DR signal from the utility. When succeeding in reduction, she receives a monetary incentive based on the rate. Because small-

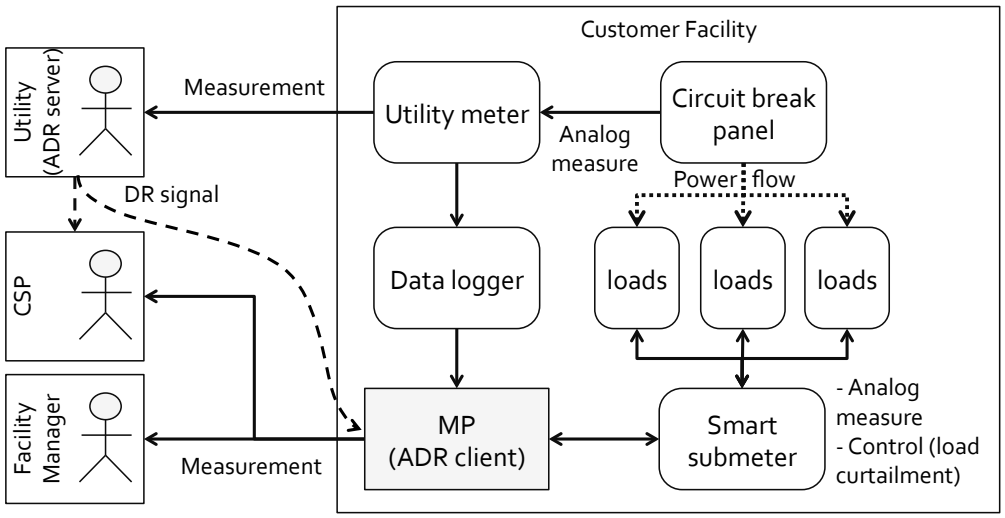


Figure 3.9: A system architecture for the ADR service in the field - ADR server, ADR client, MP, energy loads, three stakeholders, and information flows amongst elements.

sized customers can reduce few amount of power (say, a couple of hundreds of KW), they often subscribe to the service with a qualified Curtailment Service Provider (CSP). That is, a service contract is made amongst three stakeholders - customer, utility, and CSP. The customer agrees on her rate with the CSP who then agrees on an aggregated curtailment rate (from all the customers enrolled in the CSP) with the utility. Figure 3.9 shows the service architecture including three stakeholders. The ADR server transmits a DR signal to MP in the customer facility and to the CSP. MP, then, controls energy loads to reduce power consumption via smart submeters. Such curtailment is measured in three ways. First, the utility meter measures and communicates an aggregated usage data directly with the utility. Next, the CSP deploys its own data logger that obtains measurement from the utility meter. Last, MP measures the usage from the smart submeters. The building owner and the CSP share their measurement from the logger and MP.

3.5.1.2 Customer building

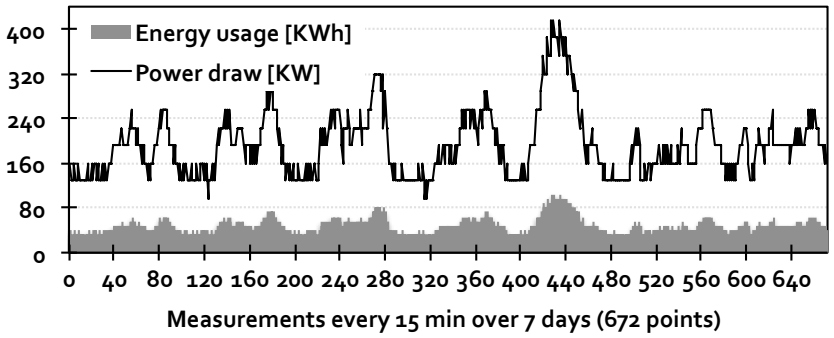
Two 4-stories buildings in the customer facility participate in the ADR. Especially, there are few occupants in the first building, and thus most energy loads can be easily curtailed. They understand and agree on potential inconvenience due to load curtailment that usually

occurs less than 5 times a year. The facility was not instrumented with any energy management system, smart submeters, or energy storage. We deploy one MP system and 6 smart submeters in two circuit breaker panels. Three submeters are installed for measurements only, while MP turns on/off energy loads via the other three submeters. Such controllable energy loads mainly are lights and office appliances. We does not control the HVAC system upon the owner's request.

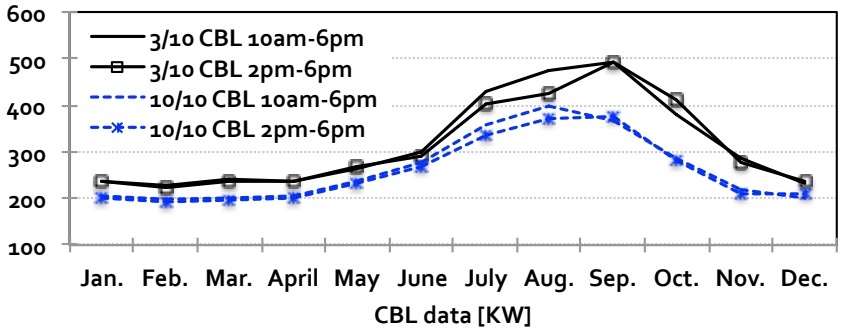
3.5.2 Curtailment Rate

Analysis on historical data. We look into the historical data of energy usage. To this end, we access a web page through which the utility provides the owner with the power [KW] and energy [KWh] usage data every 15 minute. At the same time, we access the measurement from the CSP's data logger, which provides the usage data at a finer resolution. Connected to the utility meter directly, the logger is updated instantaneous power draw every minute. It, then, transmits the raw data to our MP every 15 minutes. The raw data is aggregated and computed together so as to represent the power and energy usage every 15 minutes as shown in Figure 3.10a. The customer buildings consume 32,296 KWh of energy in total over 7 days with average of 48.06 KWh; maximum of 104 KWh at 11AM; and minimum of 24 KWh at 7AM. The figure also shows that the maximum instantaneous power draw in the buildings is 416 KW and occurs in the range of hour 427 and 434, while the average power draw is 192.24 KW.

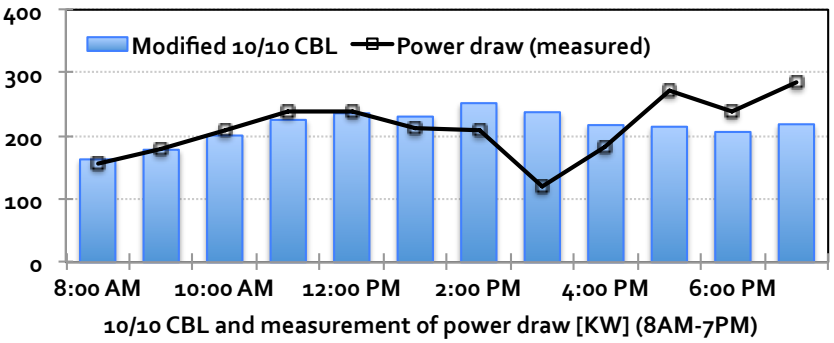
Customer baseline load. Using the historical data, the owner and the CSP calculate a Customer Baseline Load (CBL). The CBL [KW] is a reference point on which the CSP determines the amount of power consumption the customer facility reduces. Say, the calculated CBL value at 2PM today is P_c KW, and the maximum instantaneous power draw measured by the meter at the same time is P_m KW. Then, the customer officially reduces power usage by $(P_c - P_m)$ KW. If the difference is greater than a predetermined curtailment rate, then the CSP concludes the success of the load curtailment on the customer side. While many literatures investigate the baseline models [27, 34, 37], two calculations are widely used for



(a) Energy usage and power draw every 15 min over 7 days in March.



(b) Calculation of 3/10 and 10/10 CBL over 12 months.



(c) Measurement of peak power draw (KW) during a simulated one-hour DR event (at 3PM).

Figure 3.10: CBL calculation and measurement for rate calculation.

their simple computation - 3/10 CBL and 10/10 CBL. Algorithm 1 shows a pseudocode to calculate $k/10$ CBL. Fundamentally, it takes historical data of power usage over the last 10 days and then picks and averages the k largest values in order to get a CBL data of today. More specifically, t in the code indicates time, say, 11AM, and d represents date, e.g., March 20, 2013. The value h represents the frequency of power measurement per an hour. In our experiment, the power data is measured every 15 min, and thus h is set to 4. In this sense, the data m_{18}^3 represents the instantaneous power draw measured at 10:45AM on March 18.

Algorithm 1 $k/10$ CBL calculation at time t on date d

Require: Power measurement m_i^j at time t, where $d-11 \leq i \leq d-1$ and $1 \leq j \leq h$.

1: /* select maximum power draw within an hour range */
2: **for** $i = d-11$ to $d-1$ **do**
3: $\quad M_i \leftarrow \max_{1 \leq j \leq h} m_i^j$
4: **end for**
5: /* select k largest values from $M = \{M_i\}$ */
6: **for** $n = 1$ to k **do**
7: $\quad P_n \leftarrow \max_{1 \leq n \leq 10} (M_n \setminus \{P_1, P_2, ..., P_{n-1}\})$
8: **end for**
9: /* compute an average $\bar{P}$ from $P = \{P_1, P_2, ..., P_k\}$ */
10: **return** $k/10$ CBL $\leftarrow \bar{P}$

Using the calculation, Figure 3.10b illustrates both 3/10 CBL and 10/10 CBL for two time windows of 10AM-6PM and 2PM-6PM over 12 months. The 3/10 CBL values are always greater than those in the 10/10 CBL since the 3/10 CBL takes three largest values only among the measurements. The results in the figure also show that there is few differences between two time windows in both calculations. This indicates that the building consumes almost same amount of power during the time window of 8 hours.

Rate calculation. Given the CBL calculation, the curtailment rate is determined based on additional measurements. To this end, the owner and the CSP run an one-hour DR event during which the buildings turn off all the energy loads participating in the ADR. In our

experiment, the MP shuts off the power of the submeters. Then, the load drop is measured and compared with CBL data. We use a modified 10/10 CBL that the utility offers. It takes the 10/10 CBL data and adjusts it by putting more weights on 4 values closest to the moment of the DR event. The experimental results are illustrated in Figure 3.10c. The bars represent the modified 10/10 CBL from 8AM to 7PM that are averaged over last 10 days, and the line shows measurements of power draw at the event day. During the event (at 3PM), the CBL is 237.6 KW while the power draw is 120 KW. Then, we determine 100 KW of curtailment rate, after considering a buffer to the actual load drop (= 117.6 KW). Note that the building owner is incentivized for 100 KW.

3.5.3 Running ADR Service

The existing ADR server in the utility follows the OpenADR 1.0 specification and only supports the simple client mode in which data is recorded in the *SimpleClientEventData* entity of the *EventState* message. A sample signal is shown at the box below. Note that the *simpleDRModeData* represents the *SimpleClientEventData* entity. Of the three operational states for the simple client in the specification, the existing ADR server uses two states - NORMAL or (MODERATE or HIGH). That is, MP is informed that a DR event occurs only when it receives a DR message containing either MODERATE or HIGH state. The server transmits the message via SOAP communications over the Internet that is protected using the customer ID and password.

```
<p:simpleDRModeData>
  <p:EventStatus>ACTIVE</p:EventStatus>
  <p:OperationModeValue>MODERATE</p:OperationModeValue>
  <p:currentTime>354.638</p:currentTime>
  <p:operationModeSchedule>
    <p:modeSlot>
      <p:OperationModeValue>MODERATE</p:OperationModeValue>
      <p:modeTimeSlot>0</p:modeTimeSlot>
```

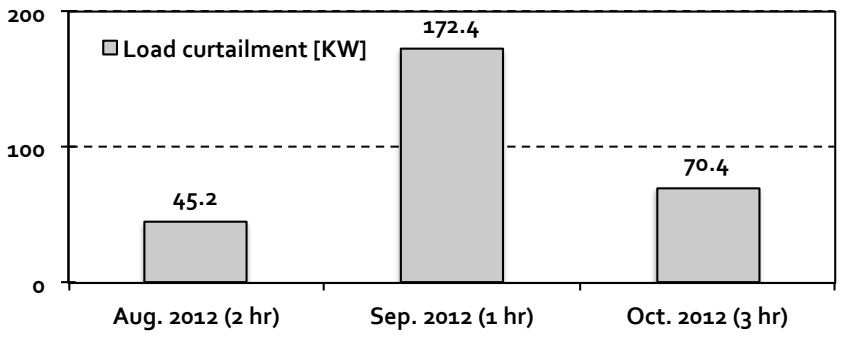


Figure 3.11: Three ADR events occur and last for 2, 1, and 3 hours, respectively. The load curtailment is the calculation of $(P_c - P_m)$. Recall that the curtailment rate is 100 KW.

```
    </p:modeSlot>
    <p:modeSlot>
      <p:OperationModeValue>HIGH</p:OperationModeValue>
      <p:modeTimeSlot>3600</p:modeTimeSlot>
    </p:modeSlot>
  </p:operationModeSchedule>
</p:simpleDRModeData>
```

During our experiments, the ADR server sent three DR event messages to MP for testing purpose, and then MP turned off the smart submeters. The results are summarized in Figure 3.11, showing that the customer building achieves 45%, 172%, and 70% of performance. The huge gap between the achievement is mainly attributed to the simple control strategy. MP simply stops building operations by turning off the smart submeters participating in the ADR so that the amount of reduction entirely relies on power consumption of energy loads currently connected to the submeters. With such inconsistent performance, the utility cannot ensure that the customer is able to reduce required amount of power consumption on emergency, failing to satisfy the grid needs.

To mitigate the risk, especially in small customer facilities, the CSP aggregates power reduction from all the enrolled customers and makes the performance more reliable. Its mediator role also benefits the customers. We note that the performance less than 100%

does not necessarily mean that the customer gets penalized. If other customers under the same CSP reduce more than 100% and thus the aggregated reduction is greater than the contracted rate between the CSP and the utility, the customer still receives an incentive even though she achieves only 45%. Likewise, her performance of 172% benefits other customers at the second event. We refer [26] for more discussion on the CSP's role.

CHAPTER 4

Securing the Interoperation

4.1 Fine-Grained Access Control

One interesting property in the Energy Service Interface (ESI) is a clear distinction between reading data and invoking operations within an object, i.e., "fine granularity of object access". The property is especially highlighted in the smart grid context because two accesses affect the operation of the customer domain differently. For instance, leakage of customer data violates privacy policy while an abuse of control capability can harm the customer directly. Therefore, the security mechanism at the ESI must be able to understand the difference. Unfortunately, however, few security schemes under the standardization efforts address this issue. Furthermore, as more energy objects are added to the customer domain and the domain interconnects with various external systems, the interactions become extremely complicated. But, existing security schemes, relying prior knowledge of user list and database, cannot handle the complexity in an efficient way: "They do not scale".

To solve the problems, we propose a new security mechanism, *Resource Centric Security (RCSec)*, that provides an access control and data encryption. To address the fine granularity issue, RCSec leverages the concept of a filesystem Access Control List (ACL), in which each file (object) maintains an entry that predefines three classes of *user*, *group*, and *others* and determines which privileges (*read*, *write*, and *execute*) are assigned to each class. In RCSec, we do access control reversely. That is, we define three privileges within an energy resource, and then assign a set of attributes[1] to each privilege. Unlike the ACL, RCSec does not predefine the classes in advance. Instead, each accessing user must show a matched

[1] An attribute is a property that represents the resource. For instance, an LED at an office 127 may have two attributes "LED" and "OF-127".

set of attributes to obtain the privilege. In this way, the user may receive permission to read data of a resource but not to invoke operations. To address scalability, we implement RCSec by exploiting Attribute Based Encryption (ABE). Given a set of attributes assigned to each resource, associated data is ABE encrypted using this set. Each external user maintains a private key consisting of his own set of attributes, and is able to decrypt the ciphertext only if his attribute set matches the resource set. Unlike the ACL, each user in RCSec is responsible for managing his own attributes. Thus, the resource (or the ESI) is free from maintaining an access control entry to perform the authorization process. This enables a security system to cope with complex interactions as well as to operate in a large-scale smart grid environment. We implement the proposed RCSec on top of our testbed of Energy Management and Control System (EMCS) and evaluate its performance in terms of overhead. Experimental results discover that the performance of the proposed scheme relies on underlying encryption algorithm.

4.1.1 Security Challenge in Communication Interface

Standardization efforts to realize the energy interoperation show three interesting characteristics. First, eXtensible Markup Language (XML) is used for data representation. Data format in BACnet and oBIX is initially defined with XML, and existing KNX is also mapped to XML data format [70]. Especially, the extensible nature of the user-defined XML schema is well leveraged when mapping analog data set to digital representation. Most service models also exploit XML technology. OpenADR and EMIX define their specifications in the XML format.

Second, the way of defining customers' energy data follows an Object-Oriented (OO) design pattern. While some protocols show this property in their schema design and UML (Unified Modeling Language) representation, explicit examples can be found in oBIX and BACnet. Given predefined primitive objects such as "int" and "list", an object in oBIX is modeled with data types (values) and operations (method signatures). An object can represent a physical device directly or represent a collection of information related to a

particular function.

Last, a web service technology is exploited for the inter-domain transportation of the customer data. The OpenADR specification defines two web service connections, SOAP (Simple Object Access Protocol) and REST (Representational State Transfer). EI utilizes web service technologies such as WS-Calendar and WS-Addressing for additional functionality. This sounds natural because the data is encoded in the XML format. But, it is noted that the ESI benefits the most from web services' capability of machine-to-machine communications in a distributed environment. This support is essential to the automation of the smart grid system. For instance, dynamic pricing programs requiring minimized human intervention can be achieved effectively only by automation technology.

When looking at three characteristics, one notices that they all together maximize the interoperability and automation. That is, each energy resource distinguishes its values and operations, and web services expose them by implementing three actions to the resource: Read, Write, and Invoke. These efforts are summarized as "fine granularity of data access and load controls". For instance, a user only reads energy usage of an air conditioning system, while another user may turn off the system to reduce power consumption during the on-peak period.

Fine granularity on the communication interface. Fine granularity is a new challenge of a security mechanism at the ESI, because different actions induce different operation consequences and indicate different levels of privacy penetration. For instance, energy usage data is read to calculate bills, but the read privilege can be misused to infer private activities of the residents. The energy assets can be configured to perform an automated demand response strategy. However, if they are controlled unfavorably, the actions would have a detrimental impact on the activities. If an adversary abuses the control privilege, and sets to maximize energy loads during the peak, the reliability of smart grid would be seriously threatened. To avoid these potential problems, the owner of the energy assets wants to permit utilities to read parts of data and service providers to control contracted energy loads only, while he has a full control over his assets.

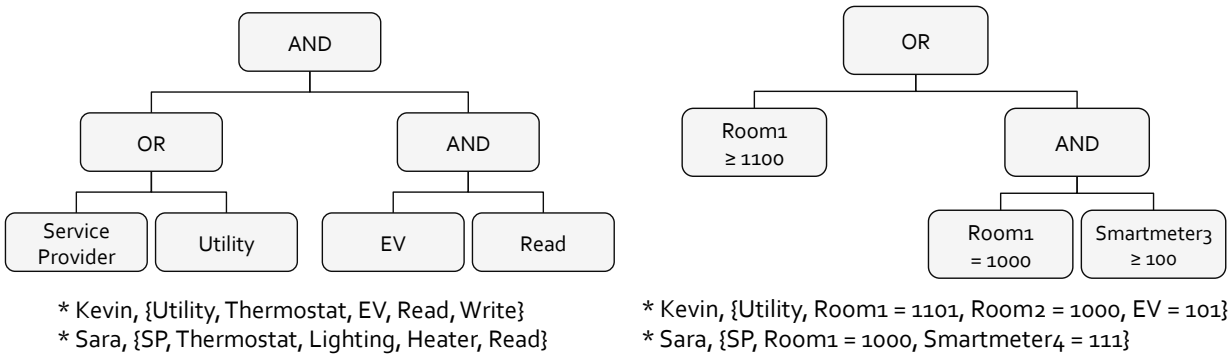
Current security schemes being considered under standardization efforts, however, cannot support the same level of granularity efficiently. The ESI or the EMCS authenticates users using its own database of ID-password sets and authorizes them with a coarse-grained rule. This way, a user would have a full control over a group of energy resources at once. However, the owner may not want the user to have excessive rights to access the resources. He would like to apply the principle of least privilege so that the user is given minimum permission that are essential to that user's work. Given the requirement of fine granularity of new data and service models in the customer domain, these schemes cannot realize the principle in an effective way. Moreover, the requirement gets greatly complicated, as more energy resources are added to the customer domain, and more external users are connected the domain. To cope with the increasing complexity of interactions, existing schemes must manage a volume of user lists and authorization rules, and develop corresponding enforcement mechanisms, which causes additional overhead to them. Thus, they cannot easily scale up in a large-scale, distributed smart grid network. To overcome the challenges, we propose a *Resource Centric Security (RCSec)* approach that takes the concept of data and service model into consideration.

4.1.2 Resource Centric Security

Based on our observation of abstraction level, the proposed RCSec realizes a fine-grained, scalable security mechanism through encryption, privilege assignment, and authorization.

4.1.2.1 Encryption

To achieve confidentiality, RCSec leverages the concept and implementation of Ciphertext-Policy Attribute-Based Encryption (CP-ABE) that encrypts data using user attributes [23]. CP-ABE realizes the secret sharing scheme [81] using bilinear map based Pairing-Based Cryptography (PBC). More specifically, each user is assigned a set of shares (attributes), and a data sender encrypts data using an open key and an arbitrary set of attributes. The encryptor creates an access policy tree, representing a Boolean formula defining the



(a) An access policy tree is created when Alice (owner) encrypts data. Two users, Kevin and Sara, have own sets of attributes.

(b) Each attribute in the access policy tree is assigned privilege. Kevin, having 4 attributes, only satisfies the tree and accesses the object Smartmeter3.

Figure 4.1: Constructing access policy tree with a set of attributes.

combination of attributes in the ciphertext. If a user presents a proper credential, i.e., any combination of his attributes satisfies the tree, he recovers the secret and is authorized to access the data. To make the tree secret, CP-ABE exploits a polynomial interpolation technique that guarantees information theoretic security. To prevent collusion attacks, an authority assigns a random number to each user whose attributes are also tied with the number.

Figure 4.1a illustrates an example of an access policy tree that consists of two types of Boolean logic gates and four attributes at the leaf position. Two users, Kevin and Sara, have 5 attributes in their private keys. A decryption process begins from the leaves by matching their attributes, and each gate returns true to its parent if children satisfy the logic. If the root returns true, then the user recovers data successfully. In this way, Kevin accesses Alice's data, but Sara cannot.

4.1.2.2 Privilege Assignment

Each attribute represents a state of permission and does not relate to other attributes. Suppose that Kevin in Fig. 4.1a is permitted to read and write the thermostat data, but only to read the EV data. But, his attributes indicate his right to write the EV data. This happens mainly due to the discrepancy between the concept of attribute and an object model.

To exploit the attributes in the resource centric model at the ESI, we apply a filesystem ACL that has been used in modern operating systems. In ACL, each object (e.g., a file in a Unix system) maintains own Access Control Entry (ACE), and a 3-digit code represents privilege to access the object - a user can read from, write to, or execute the object.

In RCSec, each object maps to an attribute with 3-digit privilege level. For instance, an object Smartmeter can be represented as an attribute "Smartmeter = 111". The first digit indicates permission of Read, and the following two digits indicate the rights of Write and Invoke, respectively. Thus, when a user has an attribute "Smartmeter = 100" and tries to access the object, it is permitted to read energy usage information but cannot turn on/off the device. In our implementation, the ESI encrypts an object data with attributes in which appropriate privileges are assigned. The privilege assignment allows inequality, e.g., "Smartmeter $\geq$ 100", in the access policy tree, whereas this is not used in user attributes. The inequality expression significantly simplifies the assignment rule. For instance, if a user has an attribute "Smartmeter = 111", he is still able to satisfy the inequality condition and to read data. In this way, the expression is capable of testing multiple privileged attributes at once.

The proposed assignment rule also supports hierarchical types of objects. Say, a building has more than one room, and several Smartmeters are deployed in each room. In this hierarchy, each room is also identified as an object that does not provide energy information directly. Instead, the corresponding XML document provides meta data about the object and access information to the sub objects of Smartmeters. For such resources, we assign privilege with 4-digit code. The first digit represents permission to access the object itself. An example is "Room1 = 1000". The last three digits have the same semantics to the 3-digit code, but with different scope of permission. For instance, "Room1 = 1100" implies that the user can read all the data produced in Room 1. This implies that the "Room1" attribute is more inclusive than the "Smartmeter" attribute and provides higher level of privilege. This rule makes it much easier to manage attributes. When there are 10 Smartmeters in the room, a user can use one attribute instead of ten attributes to access them. Fig. 4.1b depicts an access policy tree in which privilege is assigned according to the proposed rule.

4.1.2.3 Authorization Protocol

A user accesses an object in three ways: Read, Write, and Invoke. Authorization for the Read is performed at a user side with his own private key. Both Write and Invoke, on the other hand, occur at the customer domain upon receiving requests from the user. The ESI is not allowed to have the user's private key that is required for the authorization[2]. Thus, we design an authorization protocol leveraging our encryption and privilege assignment.

The authorization for Write and Invoke follows almost the same procedure, and the followings describe 4 steps of procedure for the Invoke operation. We use below notations.

- u, v are two end systems. In our example, they are an user and an ESI, respectively.
- $u \rightarrow v : M$ denotes that u sends a message M to v
- $M_1|M_2$ is the concatenation of messages M_1 and M_2
- $H(M)$ is hash of M (e.g., SHA-1).
- T_x is an access policy tree containing an attribute x.
- $\{M\}_T$ is attribute based encryption with a tree T.
- $[M]_K$ is symmetric key encryption with a key K (e.g., AES).
- N is a random nonce value.

Invoke Request (IR). A user u generates and concatenates a request message M_{req} and a nonce N_u. M_{req} includes information about an operation that u requests v to execute. The concatenated data is then encrypted with an access policy tree T_{bruin}, where *bruin* denotes the name of v, i.e., the ESI. Note that T_{bruin} does not imply that the tree has only one attribute. We assume that Certificate Authority (CA) assigns the *bruin* attribute only to v, and the attribute is not forged or stolen. Thus, v is only able to satisfy the tree. The ciphertext M_{IR} is then delivered to v as follows.

[2]In addition, RCSec pursues a pure distributed system in which the ESI never maintains any database of user list, whereas conventional authorization schemes rely on ACL stored in them for authorization.

$$u \rightarrow v : M_{IR},\ M_{IR} = \{N_u | M_{req}\}_{T_{bruin}}$$

Authorization Request (AR). Once v receives and decrypts M_{IR}, it obtains N_u and M_{req}. And it generates a nonce N_v. As mentioned, v does not manage users list or store any state information of external requests to achieve a lightweight and stateless distributed system. To this end, it creates and sends u a message M_{op} that stores the state information. M_{op} is not intended to expose to u but expected to return back to execute the operation later. Thus, v encrypts the message with a key K_{uv} and creates a new message M_x as below. The key is created using both v's private pseudonym PS_v and N_u. PS_v saves v's burden to remember u, while K_{uv} is still related to u through N_u.

$$M_x = [M_{op}]_{K_{uv}}$$
$$M_{op} = (N_v | M_{req} | T_i),\ K_{uv} = H(N_u) \oplus H(PS_v)$$

where $\oplus$ denotes an XOR operation. T_i represents the time when M_{op} is generated and is used to protect communication against replay attacks. v also encrypts N_v with an access policy tree T_{invoke} and creates a message, $M_y = \{N_v\}_{T_{invoke}}$. This is to challenge u if it is qualified or not. Two messages together are now encrypted with a key $(N_u + 1)$ and delivered to u as follows.

$$v \rightarrow u : M_{AR},\ M_{AR} = [M_x | M_y]_{(N_u+1)}$$

Authorization Ack (AA). Upon receiving and decrypting M_{AR} with own nonce N_u, u obtains M_x and M_y. It cannot decrypt M_x, instead returns back to v. u recovers a nonce from M_y only if it satisfies the tree T_{invoke}. Let N'_v be the recovered nonce. u, then, collects three data, encrypts the collection with T_{bruin}, and transmits the ciphertext M_{AA} to v as follows.

$$u \rightarrow v : M_{AA},\ M_{AA} = \{M_x | (N_u + 2) | (N'_v + 2)\}_{T_{bruin}}$$

Invoke Ack (IA). After receiving and decrypting M_{AA}, v obtains M_x, N_u, and N'_v. Using N_u, it decrypts M_x and obtains N_v, T_i, and M_{req}. Then, the authorization confirms the followings - (1) $N'_v = N_v$; and (2) $T_c - T_i \leq \Delta$, where T_c is the current time at v, and Δ denotes a timeout threshold. Once confirmed, v executes the request, M_{req}. An acknowledge message M_{ack} after the operation is encrypted and delivered to u as follows.

$$v \rightarrow u : M_{IA},\ M_{IA} = [M_{ack}]_{(N_v+3)}$$

4.1.2.4 Advantage

The RCSec scheme provides a few advantages. First, it supports fine granularity. An object is treated separately with distinguished attributes, and three request types for the object are given different privileges. Furthermore, these concepts are effectively implemented within the conceptual boundary of an attribute. Second, RCSec is developed based on an encryption algorithm. Thus, it inherently supports confidentiality and helps guarantee customer privacy. Last, RCSec is scalable. It does not require the ESI to maintain user information. Instead, each user manages own privilege in the form of attributes. Thus, RCSec can scale well even in a distributed system environment.

4.1.3 Performance Evaluation

To validate the proposed RCSec, we implement it within our EMCS testbed. Developed on a server system running the Eeebuntu distribution, it gathers energy data periodically from various energy loads. Collected data is stored and managed in the oBIX format. The EMCS also implements the ESI that realizes HTTP-based web service communications, and RCSec runs in the ESI. We omit the analysis of underlying encryption algorithm due to space limitation. Instead, we refer [76] for interested readers.

4.1.3.1 Implementation

oBIX. Following the OO paradigm, each object in oBIX is modeled by a set of *value* objects like "str" and "bool" and a set of *op* objects that define operations with input and output objects. The object model also allows inheritance to model complicated oBIX resources by means of a contract mechanism. Realized by *is* object, it establishes the classic "is a" relationship with overriding rules. oBIX supports lower level of abstraction, which gives a huge flexibility. This benefits the customer domain in that heterogeneous data formats must be effectively represented in a common data model.

HTTP REST. In order to expose data to external domains, the ESI implements the web service of the oBIX specification - HTTP binding in the REST style having a small set of verbs to transfer an object's state. oBIX maps three oBIX requests to HTTP methods: Read with GET, Write with PUT, and Invoke with POST. Below oBIX document represents a smart plug object, "plug1" in which the energy load is controllable via two ways: Write and Invoke. To turn it on or off, an external user sends directly a PUT request targeting at the *connectLoad* object within the same URI or sends a POST request to the hyperlinked URI targeting at the operation object, *controlLoad*.

```
<obj href="http://myPAS/zigbee/plug1/">
  <str name="deviceName" val="BSPE12SOYZM43001"/>
  <bool name="connectLoad" writable="true" val="true"/>
  <op name="controlLoad" href="control" in="obix:WritePointIn" out="
     obix:Point"/>
  <ref name="power" href="power"/>
</obj>
```

Access control and encryption. The proposed RCSec is implemented in the ESI. In particular, it encrypts oBIX data and the authorization works with three oBIX requests, i.e., in the application layer. To implement details of RCSec, we take algorithms approved

in the NISTIR 7628 guideline [83]. More specifically, we use AES-256 for the symmetric key encryption, SHA-256 for the hash function, and SHA-1 based random number generator to generate the nonce value. In addition, we leverage CP-ABE for attribute based encryption. Base64 encoding transforms the encrypted bytes to printable ASCII strings so that data is delivered in the XML form over the web. In order to minimize the overhead of encryption processing and to improve data access time, popular data is encrypted in advance and stored in cache, while MySQL based database management system manages all the data. Therefore, Read request is handled quickly with cached data. Whereas, both Write and Invoke requests require the authorization process, which introduces processing delay.

4.1.3.2 Experiments

Most security algorithms protect user information at the cost of additional overhead and latency. This is inevitable because stronger algorithms usually come with expensive data processing. To evaluate the proposed scheme, therefore, we examine the level of overhead in details. All the experiments are conducted by two computers that run with 2.2 GHz Intel Core 2 Duo processor and 2 GB memory.

Encryption and decryption. The first experiment assesses performance of the attribute based encryption which is the most expensive portion in RCSec. In this experiment, we vary the number of attributes (N_A) used in the algorithm, and then measure processing time to encrypt and decrypt data that is 200 KB in size. As illustrated in Figure 4.2, the processing time grows in proportional to N_A. Encryption is much more sensitive to N_A than decryption. When N_A=30, encryption is around 8 times slower than decryption. This is mainly attributed to difference of mathematical complexity within the encryption and decryption. On the other hand, another experiment reveals that the data size V_D barely affects the processing time although the results are omitted due to space limitation. From the results, we reason that the encryption part dominates the processing overhead of the encryption, and N_A used in encryption primarily influences the overhead.

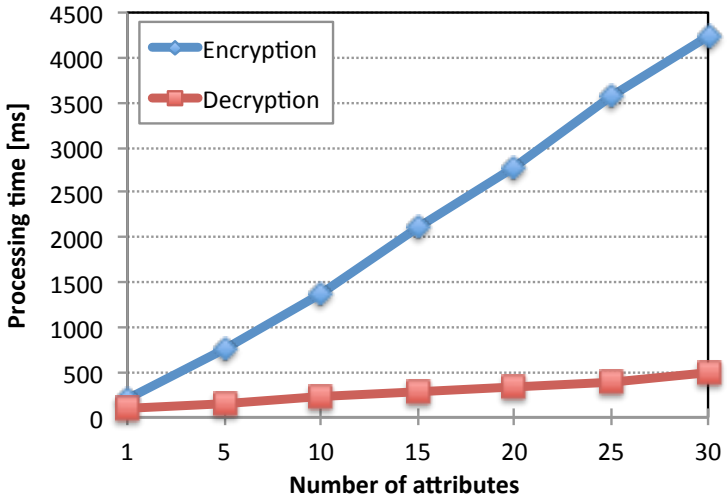


Figure 4.2: Data processing time [ms] for encryption and decryption in the attribute based encryption.

Num. attribute (N_A)	1	5	10	15	20	25	30
Msg. overhead [KB]	2	8	17	26	35	43	52

Table 4.1: Message overhead with varying numbers of attributes.

Message overhead. Having the impact of the attributes on the processing time in mind, next experiments investigate how the attributes affect the message volume. In addition to an encrypted data, a message to be transmitted contains meta information about the access policy tree and corresponding mathematical expressions. Table 4.1 shows extra volumes increased with varying numbers of attributes. We set V_D=200 KB. The result indicates that adding one attribute in encryption expands the entire message size by 1.75 KB on average. This way, the message overhead becomes 52 KB when N_A=30.

However, we note that the message overhead does not relate to V_D. That is, the overhead remains 17 KB with 10 attributes even when the algorithm encrypts data of 1000 KB. This property allows us to calculate the overall message overhead. For instance, Figure 4.3 draws a breakdown of a message in terms of volume, where we vary V_D while fixing N_A=10. The figure also shows the overhead due to the Base64 encoding, which accounts for 25% all the time. Because of the fixed size of 17 KB, the encryption overhead accounts for 47.2% when

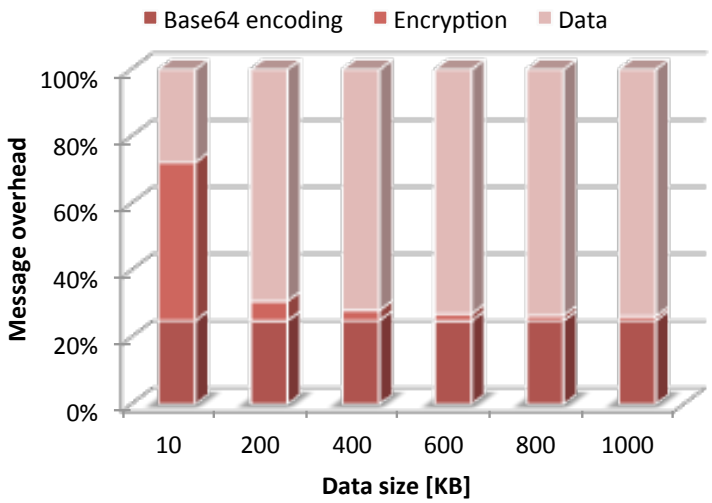


Figure 4.3: Message overhead - a breakdown of a message in terms of volume.

V_D=10KB. As V_D increases, its portion decreases dramatically: 3.1% for 400 KB and 1.6% for 800 KB. The results again reveal noticeable influence of the attributes on the proposed security scheme.

Latency on authorization. The proposed authorization protocol comprises of 4 steps (IR, AR, AA, and IA), each of which involves different computational operations. This experiment investigates the performance of the protocol by showing the breakdown of the processing time of each step. In our scenario, a user (client) requests a list of energy usage data for specified period to our EMCS (server) - i.e., oBIX history service. The retrieved list data, 300 KB, is included in the acknowledge message in the IA step (M_{ack}). Four messages, M_{IR}, M_{AR}, M_{AA}, and M_{IA}, are 5 KB, 24 KB, 5 KB, and 400 KB in size, respectively (see Section **??** for notations). As the server is assumed to have unique attribute in its private key, we can minimize the number of attributes in T_{bruin} (N_{Ab}) for IR and AA. We set N_{Ab} = 2 in the experiments. N_{Ai} for T_{invoke}, on the other hand, can change along with security policies, and our experiments vary the value from 5 to 20. Note that the processing time measured includes latency to generate and parse XML data.

Figure 4.4 illustrates the results showing that AR dominates the entire processing overhead, and its influence grows as N_{Ai} increases. When N_{Ai} = 20, it accounts for 73% of the

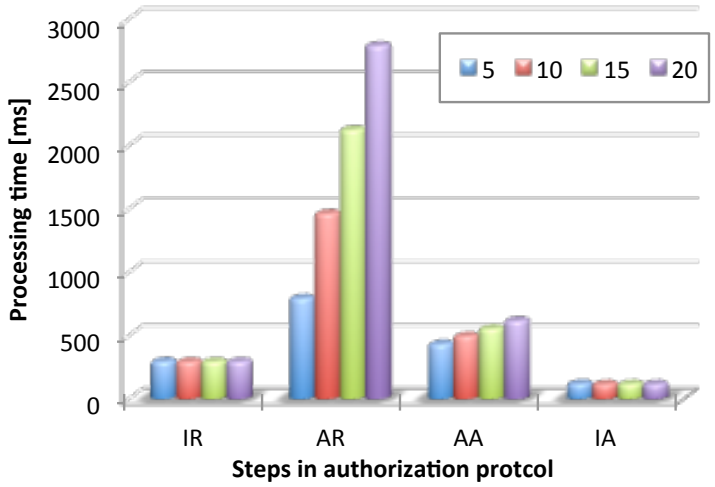


Figure 4.4: Processing time of individual step in authorization along with increasing numbers of attributes.

whole overhead. The overhead of AR mainly comes from the attribution base encryption, i.e., M_y. When looking at the results from previous experiments together, the overhead of M_y accounts for 94.3% in AR. The overhead of AA increases slowly along with N_{Ai} - 434.6 ms when $N_{Ai} = 5$ and 618.6 ms when $N_{Ai} = 20$. Such growth is mainly attributed to the decryption of M_y. Another major portion of overhead in AA comes from encryption using T_{bruin} that generates M_{AA}. Unlike AR and AA, the overheads of IR and IA barely change. IR encrypts data using T_{bruin} whose number of attributes is 2 all the time. In IA, the algorithm involves one attributed based decryption and one symmetric encryption, which enables the processing time to keep below 120 ms. The results together conclude that reducing N_{Ai} is very critical to run RCSec in an efficient way. Note that our running system currently uses 4∼8 attributes.

4.2 Weighted Privilege: Prioritized Authorization

A public report recently identifies that access control, including authentication, authorization, and verification, is the most frequent security action occurring in the electric sector [14]. One of the most promising mechanisms to realize access control is to exploit multiple factors. In Multi-Factor Authentication (MFA), for instance, a subject (or user) presents two or more authentication factors when accessing an object. Each factor is then verified against who/what it claims to be by other authentication parties. These factors include something the user knows (e.g., password); something the user has (e.g., smartphone or ATM card); and something the user is (e.g., fingerprint). MFA seeks to decrease the probability that the user presents false evidence of her identity. In this sense, the number of factors implies the strictness of MFA. The more factors an authentication process requires, the more accurately the user is verified.

Although the existing Multi-Factor (MF) technique enhances the assurance level in security mechanisms, it benefits smart grid communications in a very limited way. First, the MF technique has been applied only to authentication, which works well in conventional Internet applications that aim at sharing data. However, the smart grid involves a number of operations and actions to be taken, and this requires stringent authorization rules. Say, a compromised employee alone turns off several substations in power transmission, which causes an irrevocable disaster. The existing authorization practice of Role-Based Access Control (RBAC) alone cannot solve the problem. We believe that Multi-Factor Authorization could mitigate the risk. Second, the MF mechanisms mostly require active involvement of human beings. For example, fingerprint authentication requires the user to be present all the time. However, the smart grid communications require minimum human intervention and are automatic and real-time interoperations via Machine-To-Machine (M2M) communications amongst energy objects (resources). Last, the MF mechanisms rely on a conventional user centric identification (e.g., user ID), which works fine in small network applications under one authority. However, the smart grid is a giant-sized network in which multiple organizations collaborate together via inter-domain communications. An identity issued by

one authority might not be managed or used in another authority's domain in the same manner. The user centric identification is not scalable enough to fit to smart grid M2M communications.

To resolve the issues, we propose *Multi-Factor Authentication and Authorization (MFAA)* that performs the MF technique on both authentication and authorization without any human intervention. MFAA addresses the scalability concern by adopting the concept of attribute-based identification. Instead of a single user ID, a user is identified by a set of attributes, each of which represents the user's characteristic. In particular, MFAA exploits multi-authority attribute-based encryption. The user is granted a factor, a private key consisting of a set of number-assigned attributes, from an authority. She can obtain multiple factors from different multiple authorities. Such factors are linked together using the user's global identification, which prevents collusion attacks. The user initiates an access operation with the factors, which can be easily realized in M2M communications. An object develops own rules for authentication and authorization by combining arbitrary number of threshold attributes that involve multiple factors. It can determine the access control rule impromptu, which enables to limit accesses dynamically according to changing contexts. Upon receiving an access request from the user, the object challenges her to see if her attributes and factor keys can satisfy the rule. This requires the user to be privileged from multiple authorities in advance and thus helps thwart deceptive threats such as forgery and impersonation as well as reduce the possibility of verification error.

To demonstrate the feasibility and applicability of MFAA, we implement a `C` library and apply it to fine-grained access control in a smart building scenario. When a user accesses an energy resource (object) such as LED lightings and office appliances, she is required to present a single factor to read its energy usage data. However, when trying to control the resource, she is verified more strictly based on two factors. This distinction is reasonable since the resource control influences the residents' daily activities directly. We deploy a testbed in our laboratory to realize the access control scenario in which we also run MFAA on Android based Smartphones. Through conducting several experiments and analyzing their results, we evaluate performance of MFAA and illustrate an automated smart building

control.

4.2.1 Technological Background

4.2.1.1 Access Control in Smart Grid

Cyber security is one of the cross-cutting issues in the smart grid [12, 83]. Recently, the National Electric Sector Cybersecurity Organization Resource (NESCOR)[3] publishes an interesting report [14]. It identifies potential cyber security failure scenarios in six application domains, and analyzes their impacts on the electric sector.

- Advanced Metering Infrastructure (AMI)
- Distributed Energy Resources (DER)
- Wide Area Monitoring, Protection, and Control (WAMPAC)
- Electric Transportation (ET)
- Demand Response (DR)
- Distribution Grid Management (DGM)

The report also develops mitigation strategies for each failure scenario, and then ranks them to identify those mitigations with the greatest potential for the utilities. To this end, those mitigations are normalized to a common form and categorized into 22 mitigation action groups. Then, the report counts the occurrences of each action group across all failure scenarios. Figure 4.5 displays the top 10 action groups and the frequency of their occurrence across all application domains. An interesting observation is that the top four most frequent action groups relate to access control including authentication, authorization, and audit of authorized activities as below. This rank highlights the importance of access control in the smart grid.

[3]http://www.smartgrid.epri.com/nescor.aspx

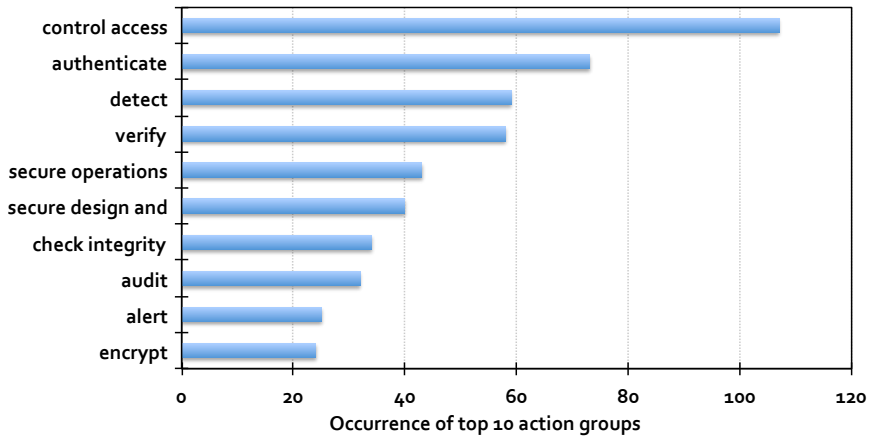


Figure 4.5: The NESCOR project ranks the mitigation action groups in the order of their occurrence across all the failure scenarios.

1. *Control access* - block unauthorized access
2. *Authenticate* authorized access
3. *Detect* abnormal or unauthorized activity
4. *Verify* that systems operate as they should

Another interesting observation is that the report recommends a multi-factor technique as one of the most promising mitigation mechanisms. For instance, the "require multi-factor authentication" is the most frequent security strategy in the group of *Authenticate*[4], and "require 2-person rule" is the one in the group of *Verify*. In a general term of Multi-Factor Authentication (MFA), a subject (or user) presents two or more authentication factors when accessing an object. Each factor is then verified against who/what it claims to be by other authentication parties. These factors include something the user knows (e.g., password); something the user has (e.g., smartphone or ATM card); and something the user is (e.g., fingerprint). MFA seeks to decrease the probability that the user presents false evidence of her identity. In this sense, the number of factors implies the strictness of MFA. The more factors an authentication process requires, the more accurately the user is verified.

[4]We do not count 'authenticate users' because it is a policy, not a mechanism.

4.2.1.2 Attribute-Based Identification

Today, a myriad of devices are connected to the information network, generating huge volumes of data every second that is replicated over the network. Such a deluge of data makes our management of individual communicating partners and information a non-trivial task. By using attribute-based identification, each user does not remember the ID of another user that may store data of her interest. Instead, she describes the data of interest using keywords and/or attributes and communicates with any users whose identifications match the description. Current research in networking, for instance, utilizes an Information Centric Networking (ICN) to search and transmit data contents using the attribute of content names instead of a user's ID or IP address [45].

The attribute-based identification has also been studied in the field of security research. In an attribute-based encryption [23, 40, 80], a user encrypts data using a set of descriptive attributes, not using a specific user's public key. Any users who can present credentials that correspond to the attributes can decrypt the ciphertext. Using attributes rather than the user's public key enables scalable key management in a large network. Authors in [79] apply the concept to develop a security framework for data sharing in the smart grid. An attribute-based signature is another active research theme [62, 65, 72]. A user signs a message using a signing key as other signature mechanisms do. The signature, however, ensures that the signer is granted a set of attributes from a trusted authority and possesses them legitimately.

4.2.2 Multi-Factor Authentication and Authorization

This section briefly introduces elementary theories and concepts on which we develop MFAA, and then describes the operations of the proposed MFAA. Due to space limitation, we omit mathematical details of each elementary technology. Instead, we show them while describing the operations.

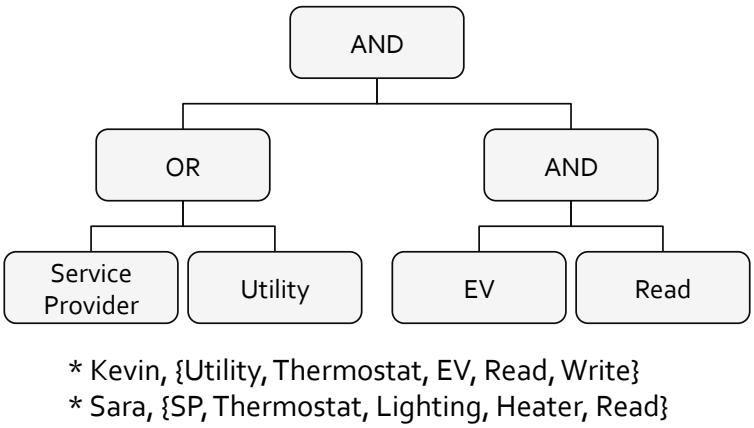


Figure 4.6: Revisit an access policy tree - Alice (data owner) creates an access policy tree when encrypting data.

4.2.2.1 Preliminary

Decentralized Attribute Based Encryption (DeABE). DeABE is a decentralized extension of Attribute-Based Encryption (ABE) that encrypts data using user attributes [61]. In ABE, each user is assigned a private key having a set of attributes from a Certificate Authority (CA), and all users share one public key [23]. A data sender encrypts data using the public key and an arbitrary set of attributes. The encryptor creates an access policy tree, representing a Boolean formula defining the combination of attributes in the ciphertext. If a user presents a proper credential, i.e., any combination of her attributes satisfies the tree, she recovers the secret and is authorized to read the data. Figure 4.6 illustrates an example of a policy tree that consists of two types of Boolean logic gates and four attributes at the leaf positions. Two users, Kevin and Sara, have own sets of attributes. Two users, Kevin and Sara, have 5 attributes in their private keys. A decryption process begins from the leaves by matching their attributes, and each gate returns true to its parent if children satisfy the logic. If the root returns true, then the user recovers data successfully. In this way, Kevin accesses Alice's data, but Sara cannot. ABE leverages a bilinear map based Pairing-Based Cryptography (PBC) for the secrecy of the attributes and leverages a polynomial interpolation technique to guarantee information theoretic security in the tree.

In DeABE, there are more than one CAs in the network, and each CA functions entirely

independently. A user maintains multiple sets of attributes, each of which is granted by different CAs. A data sender also creates an policy tree using such heterogeneous attributes when encryption data. For instance, Alice in Figure 4.6 may use two attributes of Service Provider and Utility issued by the state of California and another two attributes of Electric Vehicle (EV) and Read issued by the city of Los Angeles. Unlike ABE, DeABE converts the Boolean formula of the tree into an equivalent Linear Secret-Sharing Schemes (LSSS) matrix [20] to represent the policy in a mathematical form. The accessing user must present a proper set of attributes to recover data, which requires to obtain authorized attributes from two CAs independently.

Bell LaPadula (BLP) model. BLP is a well-known, simple computer security model for enforcing access control policies that corresponds to military classifications consisting of Clearances and Classifications. A subject (e.g., user or program) has a clearance that describes how trusted the subject is. Examples include unclassified, confidential, secret, and top secret. Each object (e.g., file or database entry) has a classification that describes how sensitive the object is. It uses the same categories as clearances. Informally, only subjects with the same (or higher) clearance should be able to access objects of a particular classification. BLP specifies two rules for a subject S to read and write an object O as followings. l_X represents the level of an entity X.

$$S \text{ can read } O \text{ if and only if } l_S \geq l_O.$$
$$S \text{ can write } O \text{ if and only if } l_S \leq l_O.$$

Decentralized Access Control Entry (DeACE). DeACE is a decentralized extension of a filesystem Access Control List (ACL) [51]. In ACL, each file (object) maintains an ACE that predefines three subject classes of *user*, *group*, and *others* and determines which privileges (*read*, *write*, and *execute*) are assigned to each subject class. The privilege for each class is represented by a 3-digit code so that a subject is authorized to read from, write to, or execute the object. DeACE decentralizes the ACL by exploiting the concept of Capability-based security and ABE. That is, a subject is granted a capability `object_1=101`

by a CA. When the subject accesses the object_1, it is authorized to read and execute, but not to write to. The capability is realized by number-assigned attributes using ABE.

4.2.2.2 MFAA Operations

In a MFAA network, there exist three types of entities - authorities $CA_j \in \mathcal{CA}$, users $u \in \mathcal{U}$, and objects $\tilde{u} \in \tilde{\mathcal{U}}$. Each authority CA_j maintains its own set of attributes, L_j, that can be in various forms. A primitive attribute form may describe a type of energy resources in a smart building such as lighting, air conditioner, fan, solar panel, and EV. Or, it may directly represent a specific employee in a company, e.g., employee ID and name. In MFAA, we add a new form of attribute, clearance. The category for clearance include *manager*, *administrator*, and *member*, to which a number is assigned. The number indicates the clearance level within the attribute, e.g., `manager=300`. Each authority can determine the categorization and the levels for its own control.

A user u contacts the authority CA_j that grants a factor $\mathsf{F_{j,u}}$ after authenticating and authorizing the user. The factor is comprised of a set of attributes, $L_{j,u} = \{a_1, \ldots, a_m \mid a_i \in L_j\}$. Each attribute a_i is represented by a factor key, $K_{i,u}$.

$$\mathsf{F_{j,u}} = \langle L_{j,u} = \{a_1, \ldots, a_m\},\ \{K_{1,u}, \ldots, K_{m,u}\}\rangle$$

In this way, u can communicate with multiple CAs to be granted multiple factors and form her factor set $\mathsf{F_u} = \langle \mathsf{F_{1,u}}, \ldots \mathsf{F_{n,u}}\rangle$. We note that any subset of attributes can be used for the user's identification regardless of the number of factors involved.

Upon accepting an access request from the user u, an object $\tilde{u}$ requires u to present one or more factors. To authenticate and authorize the request, it maintains a Challenge, an access policy tree as shown in Figure 4.6. Attributes in the tree may belong to one authority, or each attribute may belong to different authorities. As mentioned, the number of factors used determines the strictness of the access authorization. For the number-assigned attributes, $\tilde{u}$ sets condition for the clearance level by using an inequality like `manager`$\geq$`200`. Once u presents her factors and attributes that satisfy the Challenge, the access request is verified

and accepted. One advantage of this approach is that the object can dynamically adjust the Challenge by changing the number of factors and the combination of attributes according to applications' contexts, which makes the access control mechanism more flexible. The followings present mathematical equations for the operations of MFAA.

Setup. MFAA selects a prime N, groups of G and G_T of order N and a bilinear map $e : G \times G \rightarrow G_T$. Then, it publishes a generator g of G along with a hash function $H : \{0,1\}^* \rightarrow G$ that maps the identities of users to G.

An authority $CA_j \in \mathcal{CA}$ maintains a set of attributes L_j. For each attribute $i \in L_j$, the authority chooses two random exponents $\alpha_i,\ y_i \in \mathbb{Z}_N$. It keeps the secret key SK_j and publishes the public key PK_j.

$$SK_j = \{\alpha_i, y_i \mid \forall i\},\ PK_j = \{e(g,g)^{\alpha_i}, g^{y_i} \mid \forall i\}$$

Factor generation and distribution. The authority CA_j grants a factor $\mathsf{F_{j,u}}$ consisting of a set of attributes $L_{j,u}$ to a user u. To this end, it generates the factor key

$$K_{i,u} = g^{\alpha_i} H(u)^{y_i} \text{ for attribute } i \in L_{j,u} \subset L_j$$

Challenge generation. The user u chooses a random $\hat{a} \in \mathbb{Z}$ and transmits $g^{\hat{a}}$ to an object $\tilde{u}$. Then, $\tilde{u}$ generates a challenge used for authenticating and authorizing u. To this end, $\tilde{u}$ chooses two random $\hat{b}$ and $\hat{c} \in \mathbb{Z}$, takes own managing sequence number z_s, computes $z_1 = g^{\hat{b}}$ and $z_2 = enc_{z_s}(\hat{c}, (g^{\hat{a}})^{\hat{b}\hat{c}})$, where $enc_x(y)$ is a a symmetric key algorithm that encrypts a plaintext y with a secret x. It generates a challenge message $M_c = \langle z_1, z_2 \rangle$.

Then, $\tilde{u}$ chooses the factors and a list of attributes, and takes corresponding factor keys from authorities' public keys. It first creates an access policy tree and converts it to an $n \times l$ access matrix $\mathcal{A}$[5] and selects and computes the following variables.

- Choose a random $s \in \mathbb{Z}_N$

[5] l is number of leaves in the access tree, and n is determined by the tree shape [61].

- Choose two random vectors, $v \in \mathbb{Z}_N^l$ with s as the first entry and $w \in \mathbb{Z}_N^l$ with 0 as the first entry
- Compute, $\lambda_x = \mathcal{A}_x \cdot v$ and $w_x = \mathcal{A}_x \cdot w$, where $\mathcal{A}_x$ is row x of $\mathcal{A}$
- Choose a random $r_x \in \mathbb{Z}_N$ for each row $\mathcal{A}_x$ of $\mathcal{A}$
- Obtain $\rho(x)$, a mapping from $\mathcal{A}_x$ to the attribute i that is located at the corresponding leaf of the access tree.

$\tilde{u}$ generates the challenge:

$$C_{\tilde{u}} = \langle \mathcal{A},\ \rho,\ C_0 = M_c e(g,g)^s,$$
$$C_{1,x} = e(g,g)^{\lambda_x} e(g,g)^{\alpha_{\rho(x)} r_x}\ \forall x,$$
$$C_{2,x} = g^{r_x}\ \forall x,\ C_{3,x} = g^{y_{\rho(x)} r_x} g^{w_x}\ \forall x \rangle.$$

Challenge response. The user u responds to the challenge $C_{\tilde{u}}$ and sends a response $R_u(C_{\tilde{u}})$ back to $\tilde{u}$ that, then, verifies the response. To respond, u takes the matrix $\mathcal{A}$ and ρ. If it has the factor keys $K_{\rho(x),u}$ for the set of attributes $\{\rho(x) : x \in \mathcal{A}_x\}$, then it proceeds. Otherwise, u stops responding since it is not qualified. It chooses constants $c_x \in \mathbb{Z}_N$ such that $\sum_x c_x \mathcal{A}_x = (1, 0, ..., 0)$. It computes

$$val(x) = \frac{C_{1,x} \cdot e(H(u), C_{3,x})}{e(K_{\rho(x),u}, C_{2,x})} \text{ for each } x,$$
$$M_c' = \frac{C_0}{\prod_x (val(x))^{c_x}}.$$

After obtaining z_1' and z_2' from M_c', u generates a response message $M_r = \langle (z_1')^{\hat{a}'}, z_2' \rangle$.

Verification. $\tilde{u}$ obtains $dec_{z_s}(z_2') = \langle \hat{c},\ (g^{\hat{a}})^{\hat{b}\hat{c}} \rangle$ from M_r. It verifies the access request from u by validating if

$$(g^{\hat{a}})^{\hat{b}\hat{c}} = ((z_1')^{\hat{a}'})^{\hat{c}}$$

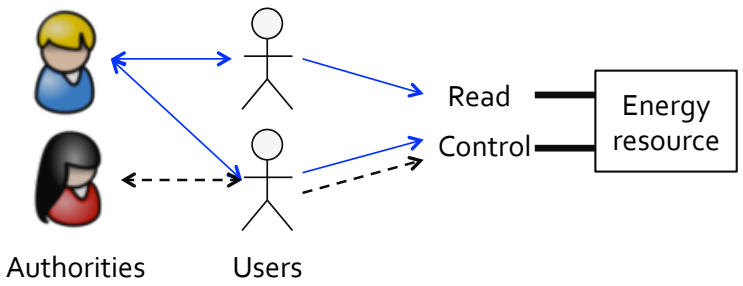


Figure 4.7: Fine-grained access control to energy resources using MFAA.

4.2.3 MFAA Application in Smart Grid

4.2.3.1 Fine-grained Access Control

The concept of the proposed MFAA can be applied to various situations in many applications. This paper considers a fine-grained access control as a sample scenario [51]. In the smart grid, a user (subject) accesses an energy resource (object) to read data and/or to control the resource. Such different access actions induce different operation consequences and indicate different levels of privacy violation. So, the authors in [51] distinguish permission to data reading from permission to resource controlling. In general, data reading affects the owner's daily life less immediately than resource controlling. Say, you might be okay that a contracted service provider reads the energy usage of your air conditioning system, but you do not allow the provider to turn it off on a hot summer day.

To such differently weighted access actions, we apply MFAA in the following way. For the access of data reading, a user is required to present one type of credential factor, whereas the user must present two types of factors to control energy resources. This implies, as shown in Figure 4.7, that the user must be authenticated and authorized in advance from two independent authorities in order to gain privilege to control the resource. As mentioned earlier, requiring multiple factors for the access control process can reduce the probability of operation failure and thus security violation. In a similar manner, we can prioritize access actions in the smart grid and protect them differently according to their priorities[6]. Changing

[6]One of the most critical security failure scenarios in the electric sector is that a utility server broadcasts a control command to a number of smart meters, which simultaneously turns all of them off [14]. We believe that MFAA can mitigate the failure risk.

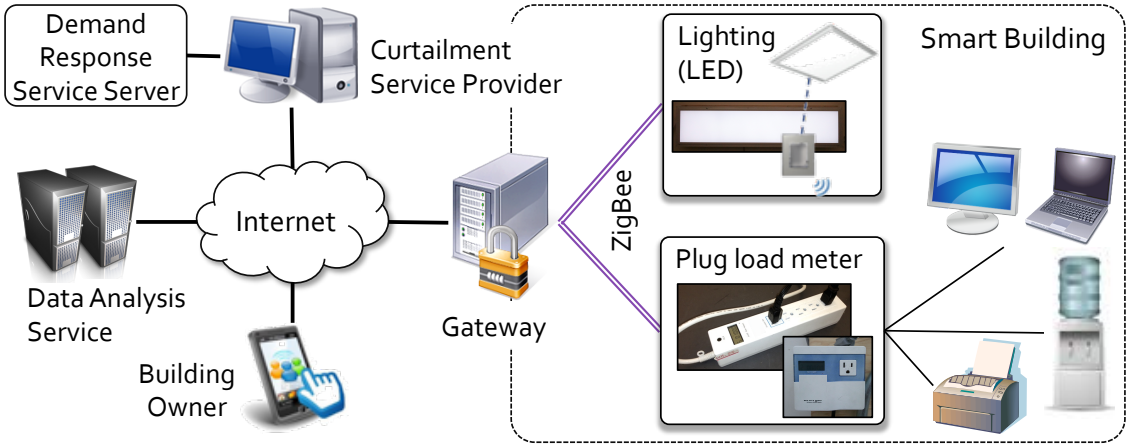


Figure 4.8: Testbed and weighted fine-grained access control.

the number of credential factors via MFAA realizes various protection levels.

4.2.3.2 Testbed Implementation

We develop a `C` library of the proposed MFAA on top of the PBC library[7] and apply it to our smart building testbed [52]. To demonstrate the application scenario, the weighted fine-grained access control, we reconfigure our testbed as shown in Figure 4.8. The testbed consists of several types of energy resources and external users, and they communicate with each other via a gateway that places at the boundary of the smart building.

We deploy office appliances such as monitor, printer, and water dispenser that are plugged into plug load meters. The meters are capable of measuring their energy usage every one minute and turning on/off the input power with embedded relay. We also deploy LED lights that can adjust own operations beyond a simple on/off status. Each LED operates with 8 steps of brightness and temperature that affect its power consumption directly. All the resources communicate with the gateway via IEEE 802.15.4 ZigBee communication. The testbed includes three external users, each of which represents different levels of access privilege to the energy resources. The building owner uses her own smartphone to control all the resources as well as to read data. To this end, we implant our library to an Android platform. The data analysis service is only able to read energy usage from all the resources.

[7] http://www.crypto.stanford.edu/pbc

It collects data, analyzes the building's energy usage pattern, and makes a recommendation for building operations to reduce energy bill. The Curtailment Service Provider (CSP) makes an Automated Demand Response (ADR) service contract with the building owner, in which a group of energy resources are registered to the service. It also contracts with a local utility company running the service server. Upon receiving a DR signal from the server, the CSP directly controls the registered resources to reduce the building's energy usage down to a pre-contracted level. There are three authorities that issue private factor keys to the users in the testbed (not shown in the figure). They are the utility company, the building owner, and a city council to which the service providers are registered.

4.2.4 Experiments and Results

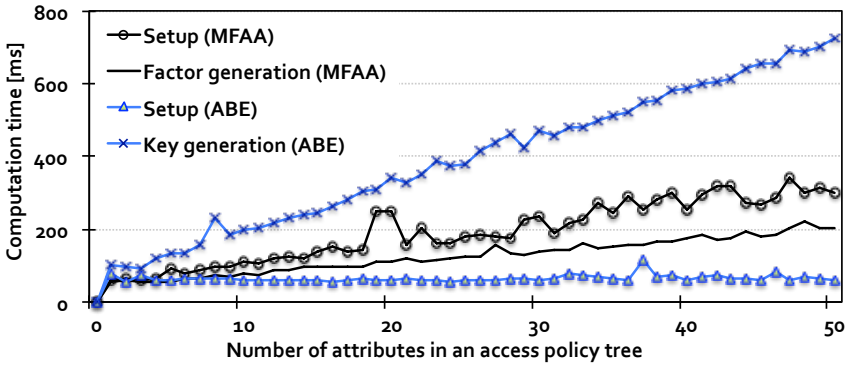
MFAA provides advanced authentication and authorization by exploiting the PBC library that is computationally expensive. Thus, this section evaluates its operational performance by measuring computation cost. We also illustrate experimental results from our smart building testbed.

4.2.4.1 Computation Cost

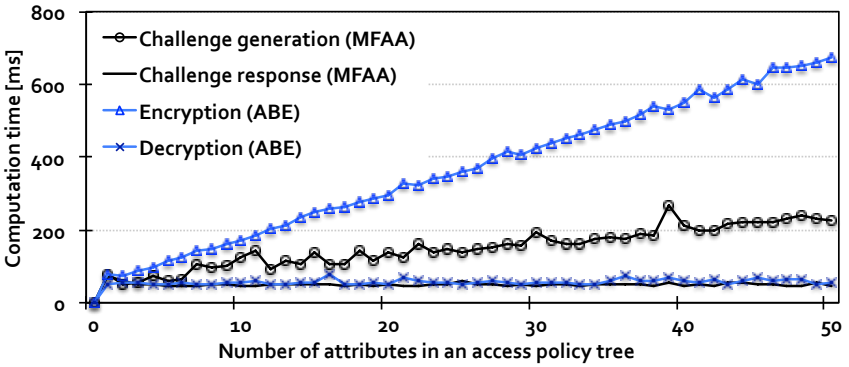
We implement MFAA on laptops that use Intel Core 2 Duo running with 2.26 GHz of clock speed and 8 GB of memory. To measure the computation time in millisecond, this experiment varies the number of attributes involved in the access policy tree. For comparison, we run an ABE implementation[8] on the same machines.

Figure 4.9a compares the computation time for the operations of setup, factor generation (MFAA) and key generation (ABE). When comparing two setup times, the difference is mainly attributed to the different scopes of setup procedures. In ABE, an authority generates one public and secret key, which does not consume any attributes. This makes the curve even over varying numbers of attributes. On the other hand, in MFAA, each authority generates a public-secret key pair for each attribute belonging to itself. The computation cost of a

[8]http://hms.isi.jhu.edu/acsc/cpabe/



(a) Setup, factor generation, and key generation.



(b) Challenge generation/response and data encryption/decryption.

Figure 4.9: Measurement on the computation cost.

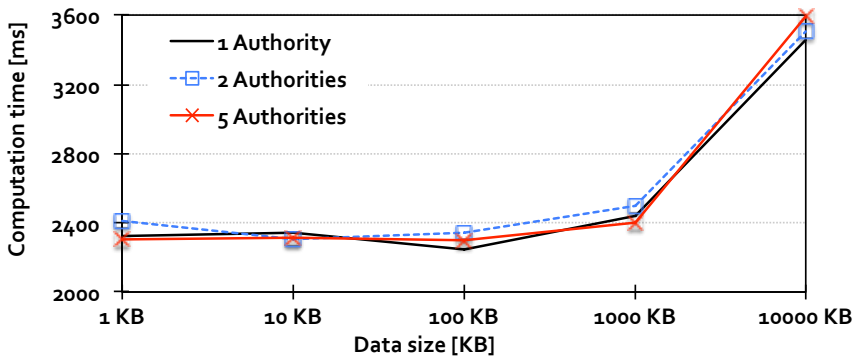


Figure 4.10: Experiments on a Smartphone. We measure computation cost on challenge generation.

bilinear pairing, one of the most expensive portion, influences the performance. However, the overhead in MFAA due to pre-computation on attributes is rewarded in the next operations, factor generation in MFAA vs. key generation in ABE. The pre-computation simplifies the cost for the factor generation. ABE takes 3.27 times and 3.55 times longer than MFAA with 25 and 50 attributes, respectively.

The cost of the main operations in MFAA and ABE are illustrated in Figure 4.9b. The cost of challenge generation in MFAA can be compared with that of encryption in ABE, since it includes encryption process. The challenge generation takes around 200 ms with 50 attributes, which is quite reasonable for conventional applications. On the other hand, the running time of encryption in ABE reaches up to 671 ms with 50 attributes. This is mainly attributed to the exponentiations required for each leaf in the access tree. Challenge response in MFAA and decryption in ABE show quite similar performance. Challenge response demonstrates 49.7 ms on average with 4.63 of standard deviation, whereas decryption shows 56.8 ms on average with 6.52 of standard deviation. Such fast running time is mainly attributed to computational optimization as authors in [23] also mentioned. One of such optimization is merging internal computations or using caches intermediary values, which can reduce the number of highly expensive computation of exponentiations in the operations.

4.2.4.2 MFAA on Smartphone

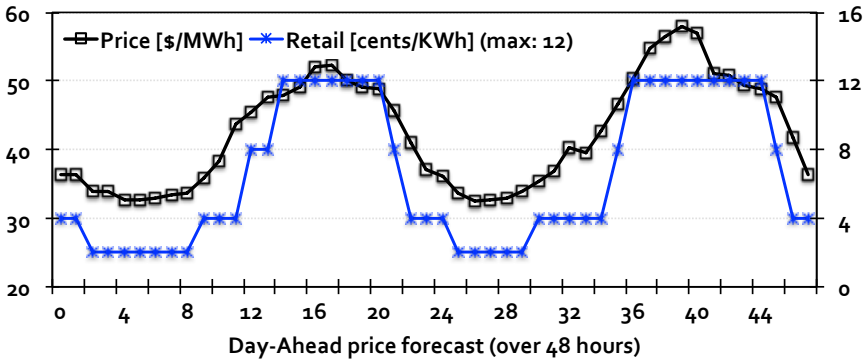
As also introduced in our testbed in Figure 4.8, an increasing volume of users uses their Smartphones to access energy objects. To accommodate such a trend, we deploy the MFAA library into an Android Smartphone and run experiments to measure its performance. We assume that a Smartphone user hardly becomes an authority, and thus this experiment focuses on the performance in challenge generation and response.

This experiment varies the number of authorities in the access tree from 1 to 5 and then measures the running time of challenge generation. We fix the number of attributes to 5. As shown in Figure 4.10, three curves draw almost same performance, concluding that the number of authorities, i.e., factors in access control, does not affect the computation cost in MFAA. This is mainly because all the attributes are processed in the same way even though different authorities issue them. We note that the average running time is 2,347 ms with 5 attributes. This can be acceptable in some applications, but there remain rooms to be further enhanced. We report that the computation cost of challenge response is also not affected by the authorities and shows 237 ms on average.

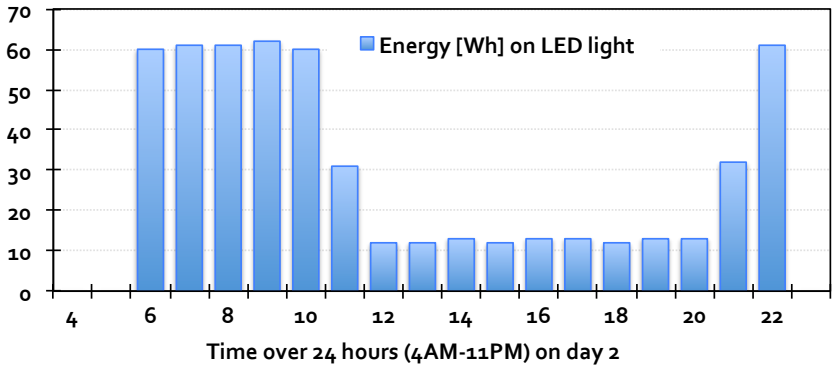
With Smartphones, we extend the scope of our application scenario to include Device-to-Device communications, where a direct 'data reading' occurs frequently whereas a 'control' action barely occurs. If this is the case, MFAA can be optimized by replacing the challenge message M_c in challenge generation with real data and by eliminating the challenge response operation. To accommodate the scenario, we vary the data size in our experiments, and the results are shown in the figure. MFAA shows constant performance until 1000 KB, but starts jumping at 10 MB. We think that this is a limitation due to data size, but not to the algorithm's computation cost.

4.2.4.3 Resource Control over Real Time Price

This experiment runs our smart building scenario described earlier. In particular, we conducts experiments for the ADR service, where energy resources in the smart building, the



(a) The DR server obtains price forecast from a California wholesale market and determines real-time power prices for a retail market.



(b) The CPS responds to the changes of the power price by controlling the brightness on an LED light directly.

Figure 4.11: Experimental results with the RTP-based ADR scenario.

Curtailment Service Provider, and the utility company are involved[9].

To this end, we deploy an OpenADR 1.0 server by exploiting the open source [11]. Instead of manual event generation, our DR server implements an automated Real Time Pricing (RTP) program. The DR server acquires a power price forecast from a wholesale market in California [1]. Figure 4.11a draws a curve of Day-Ahead Prices (DAM) over 48 hours. The wholesale market price is 42.4 [$/MWh] on average with max of 58 and min of 32.5 during the period. The server, then, determines a retail market price based on the price and other factors. Say, a unit price is 4 [cents/KWh]. The star-marked line in the figure draws the changes of power price in the retail market. The server, then, generates a DR event of the RTP program. Taking values on day 2, the event starts at 11am and lasts until 9pm. This event information is generated one hour before the event - so it is an hour-ahead RTP DR program.

Once the CSP receives the DR signal and notices that the price goes up, it performs a predefined DR strategy. In this experiment, we register one LED light to our strategy, and the CSP directly controls its brightness. As the price goes up, the LED gets dimmed proportionally. Since the power draw of the LED is proportional to the brightness level, we can easily observe the change of energy usage during the DR event, which is depicted in Figure 4.11b. We note that we pre-scheduled the LED light to turn on/off at 6AM and 11PM, respectively.

[9]Due to space limitation, we show portions of our experimental results.

4.3 Physical Layer Security in Wireless Smart Grid

The smart grid consists of heterogeneous systems and devices over various domains. For instance, the customer domain includes smart meters, appliances, and thermostats, and the operations domain includes Supervisory Control And Data Acquisition (SCADA) systems. These devices are interconnected through public and private networks and corresponding protocols. An enterprise network is built on the Ethernet using the Internet Protocol (IP) whereas the SCADA system uses specialized protocols. The largest public network, the Internet, may inter-connect them. In this way, the information network forms a hierarchical structure amongst small networks having different ownerships and various underlying communication technologies. Therefore, there is no doubt that multiple communication technologies will be deployed in the smart grid environment. Some potential candidates include the Ethernet, Power Line Carrier (PLC), and RF mesh over unlicensed spectrum.

Using wireless communication in the smart grid offers significant benefits over wired, such as untethered access to information, cost and complexity reduction for installation and maintenance, and mobility and remote application support. Figure 4.12 illustrates a brief comparison of communication technologies for the smart grid. However, wireless technologies induce additional vulnerability. In addition to cyber security issues that have been investigated extensively [83], we must consider *physical layer security*. This concern is attributed to the feature of shared wireless medium, in which a node with an off-the-shelf radio module can send/receive packets as well as freely listen to other communications within the radio range. For instance, an adversary can inject bogus signals into the wireless medium, which prevents legitimate users in the radio range from receiving wireless signals correctly. This deteriorates network performance and eventually threatens desirable operations of the smart grid. Moreover, it can steal private and confidential information. Therefore, principal research on potential risks and proper protection solution must be carried out before deploying wireless technologies in the smart grid. In this chapter, we propose a novel mechanism exploiting a random spread-spectrum technique that ensures wireless communications against physical layer attacks. Thus, the proposed solution will protect customers' private information and

Networking Technology	Advantage	Disadvantage
Wireless	▪ Untethered access to information ▪ Reduced cost and complexity for installation and maintenance ▪ Support for mobility	▪ Vulnerable to environmental effects ▪ Physical layer attack
Optical Fiber	▪ Reliable and robust communication ▪ Easily scalable	▪ Expensive cost ▪ Complex installation and maintenance
PLC	▪ Using existing electrical power network	▪ PLC signals cannot pass through transformers ▪ Limited support of peer-to-peer communication ▪ Not scalable

Figure 4.12: Three candidates of communication technologies provide own advantages and disadvantages.

system data that facilitates desirable smart grid operations.

4.3.1 Opportunity of Wireless Communication in Smart Grid

The National Institute of Standards and Technology (NIST) has investigated existing and emerging physical media for wireless communications and published a set of guidelines for smart grid system designers [71]. It lists a set of parameters and metrics (e.g., link availability) that are used to characterize wireless technologies. The list is followed by description of factors to consider when a designer makes a decision to apply wireless technologies to any set of applications. The NIST also develops a method for assessing their appropriateness for various smart grid applications [84]. The authors identify requirements of smart grid applications, convert them into link traffic characteristics, perform a coverage analysis that estimates the maximum radio range given environmental constraints, and model a MAC/PHY layer to evaluate network performance with respect to reliability, delay, and throughput. Parikh *et al.* list a set of potential wireless technologies[10] and then identify smart grid applications where each can be applied [73]. Patel *et al.* specify communication requirements of five

[10]These include IEEE 802.11 based wireless LAN, IEEE 802.16 based WiMAX, 3G/4G cellular, IEEE 802.15 based ZigBee, IEEE 802.20 based MobileFi, etc.

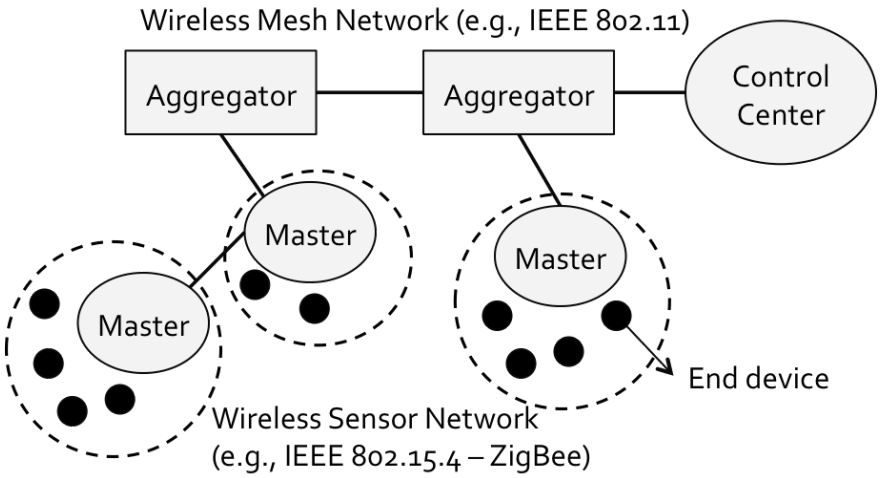


Figure 4.13: A three-layer wireless network architecture in smart grid.

representative applications[11] in the distribution and customer domains [74].

4.3.1.1 Wireless Network Architecture

Among many smart grid functionalities, communication networks for Advanced Metering Infrastructure (AMI) and Distribution Automation (DA) are considered as good candidates for wireless communication due to their scope and scale [71]. To support them, a three-layer wireless network architecture has been proposed, where each layer uses different wireless technologies as shown in Figure 4.13. The lowest layer consists of a set of end devices and a master node collecting data from the devices. They form Wireless Sensor Network (WSN) using IEEE 802.15 family and/or ZigBee profiles. When the AMI interfaces with Home Area Network (HAN), a set of devices compliant with ZigBee Smart Energy Profile (SEP)[12] can constitute the last-mile communication network. Neighboring masters (including an aggregator) exchange any information of abnormal events, which defines layer-2 communication. At the top layer, IEEE 802.11-based Wireless Mesh Network (WMN) enables aggregators to communicate with a control center. Current active research topics in WMN include development of inter-domain routing protocols, which are expected to improve network latency.

[11] AMI, automated demand response, feeder automation, electric vehicle charging, and mobile workflow management

[12] http://www.zigbee.org/Standards/ZigBeeSmartEnergy/Overview.aspx

However, as noted in [71], many issues still remain unaddressed.

4.3.1.2 Understanding Wireless Communication

In order to identify physical layer vulnerability and devise optimal solutions, it is fundamental to understand the principle of wireless communication. This section gives a brief overview of two lowest networking layers, data link layer and physical layer, which are directly associated with physical layer attacks. In particular, we review a medium access control protocol used in IEEE 802.11 and 802.15 families and discuss implication of power strength in signal propagation. Note that any networking functionalities on the upper layers are beyond the scope of this paper. This implies that an application prepares data to send, and a network layer determines the location of the communication partner in the network. Once receiving the "packet" from the upper layer, the medium access control protocol determines when to start transmitting it and then the physical layer actually emits analog signals representing the packet over the wireless.

Medium Access Control. Because all nodes in the network share wireless medium (or channel), Medium Access Control (MAC) protocols have scheduling algorithm that controls nodes' resource access. The IEEE defines specifications of Carrier Sense Multiple Access/-Collision Avoidance (CSMA/CA) protocols for 802.11 and 802.15 families. In CSMA/CA, when a node is ready to transmit a packet, it listens to wireless channel - *carrier sensing* to find whether another node is transmitting data on the wireless. If the channel is "idle", the node is allowed to transmit the packet. Otherwise, the node defers its transmission until the existing transmission stops and further random period of time - *random backoff*. After the backoff, it retries to transmit the packet again by resuming the carrier sensing. After several trials more than a predefined threshold, the node gives up transmission and notifies communication failure to the upper layer. In any way, postponed transmission increases network latency, which can lead to violation of application requirements.

Physical layer. Once the node is allowed to transmit data, the physical layer (PHY) modulates the digital representation of the packet to (analog) signals, each of which consists of a sequence of discrete complex symbols in the form of sinusoid (or sine wave). Each symbol can represent or convey one or several bits of data according to the modulation scheme used. Suppose that n^{th} complex symbol transmitted from a sender is denoted as x[n]. When a receiver receives the signal, say y[n] and its interpretation is same to x[n], the transmission is said to be successful.

However, due to unexpected wireless environment, the transmitted signal is often blocked or distorted over the air. In communications theory, the receiver sensitivity and the standard Signal-to-Interference-plus-Noise-Ratio (SINR) model are used to determine the ability of successful data recovery. That is, a signal is transmitted with standardized power level at the sender[13], and its strength attenuates as traveling. When the signal arrives at the receiver and its remaining power level is greater than a threshold, *receiver sensitivity*[14], the receiver is able to recover the signal. If the signal is too weak, then the receiver cannot derive correct bit information from it.

Noise and interferences also affect the capability of signal reception. When a signal propagates, it is affected by unexpected radios. They alter the original form of the signal. The degree of distortion is computed as the ratio of the received signal power to the combined power of noise and interference at the receiver, i.e., SINR. The SINR value is used to compute the bit error rate (BER): a large SINR implies a stronger signal and thus few bit errors. Then, the BER is used to calculate the packet error rate (PER). The receiver can decode the transmitted packet if the SINR value is above a given threshold τ[15]. Otherwise, it handles the signal as an error.

Wireless channel. As well known, 802.11b uses the 2.4 GHz frequency band in U.S. The band is divided into 13 channels in a similar way how the radio and TV broadcast bands are

[13]It is 10-20 dBm and 0 dBm in 802.11 and 802.15.4, respectively.

[14]It is around -80 dB in 802.11 devices.

[15]To achieve a PER of 1% in 1Mbps rate of 22MHz 802.11b network, τ of at least 10 dB above the noise threshold is required without considering a processing gain.

sub-divided. A channel has own center frequency (e.g., channel 6 is centered at 2.437 GHz) and has a bandwidth of 22 MHz. Transmitting a wireless signal implies that the sender uses the center frequency of a pre-defined channel, and such signal does not interfere with any non-overlapping channels within the band. Therefore, the receiver must sit on the same channel in order to receive the signal correctly. Note that we use the terms "channel" and "frequency" interchangeably from this point.

4.3.2 Physical Layer Security

A physical layer attack is defined as malicious behavior disturbing legitimate communication on a wireless network. Unlike conventional security threats, an adversary disrupts wireless medium by simply injecting false messages into the network. The shared nature of the wireless medium even empowers the adversary to disable all data transmissions within the radio range. In this sense, the PHY attack is regarded as a wireless version of Denial-of-Service (DoS). We classify the PHY attacks into four groups according to objectives and behaviors: eavesdropping, jamming, restricting access, and injecting.

4.3.2.1 PHY attacks

Eavesdropping. Since a wireless signal propagates over open space, any network node within the radio range is able to capture the signal. Moreover, as legitimate nodes follow communication standards, the neighboring nodes can obtain meaningful information from the captured signal. These openness and standards are misused. That is, an unauthorized node eavesdrops data transmission and accesses credential information. By doing so, the eavesdropping attack violates confidentiality requirement. The attack shows two properties. First, it can be quickly launched. With technological advancement, an off-the-shelf radio module and a short line of script can make an eavesdropping device. Second, the attack is not easily detected because the adversary does not expose its activity. The eavesdropping can be mitigated by advanced cryptography. Encrypting packets hinders unauthorized nodes from reading data easily.

Jamming. As the most typical form of the PHY attack, a jamming attack fulfills the wireless medium with noise signals. This affects a legitimate node in two ways. First, when the node performs the carrier sense before transmitting a packet, the channel is always sensed "busy". This defers its transmission, and the node eventually gives up communication. Second, the node may fail to receive packets. Suppose it is receiving packets from a legitimate communication partner. The noise signal can distort the data signal, which aggravates the SINR value, and the node cannot recover messages out of the damaged packets. The goal of the jamming is to deteriorate availability requirement.

A jammer can emit noise signals continuously to completely block wireless channel, i.e., *proactive jamming*. As transmitting radio signals wastes energy, however, it is not efficient. Moreover, the proactive jammer is easily detected due to its suspicious behavior. Instead, in *reactive jamming*, the jammer listens to the radio channel first and launches a jamming attack only when sensing signals on the channel. In this situation, the legitimate node cannot clearly distinguish whether its packet error results from attacks or normal collision.

Restricting access. Another type of attack tries to disrupt the MAC protocol by simply preventing nodes from initiating legitimate MAC operations or by causing packet collisions. This is conceptually similar to the reactive jamming - i.e., a jammer starts an attack only when necessary to block wireless channel. However, the access attack targets at multi-user access procedure. The attacker sets own backoff timer very short so that it occupies wireless channel first all the time. Other nodes will sense the channel busy and postpone their transmission - this portion is similar to the proactive jamming. In this way, the attack disturbs legitimate communication.

Injecting. An injecting attack inserts *formatted messages* into the wireless network, whereas two previous attackers can use bogus signal. Impersonation and replay attack fall into this category.

The adversary impersonates either a legitimate sender or a receiver to obtain unauthorized access to a wireless network. A typical impersonation is a device cloning. In terms of physical layer, the cloning is done via MAC address spoofing. The unauthorized access can

cause secondary vulnerability such as de-association or de-authorization attack. A rogue aggregator may imitate a legitimate aggregator in Figure 4.13. The rogue aggregator confuses a set of subscribers (masters) trying to get service through what they believe to be a legitimate entity. It may result in long disruptions of service. A valid data transmission is maliciously or fraudulently repeated or delayed by an adversary - replay attack. A duplicated notice of error event to the control center can trigger inappropriate emergency response or cause an inadvertent error.

In addition, the injecting attack can be used like a SYN flooding attack. That is, when receiving too many fake messages, a victim can be burdened with processing them. Then, the overheaded system resource cannot respond legitimate requests any more. This violates availability requirement. In the injecting attack, the messages remain readable by the receiver so that it is not easy to prevent the attack. Only, pertinent authentication schemes at higher layers can alleviate its impact.

Internal attack. All the described attacks can be also initiated by an internal node that is compromised or suffers inadvertent malfunctioning. Because the node is generally considered as a legitimate user, a typical solution may not protect precious resources appropriately. Up to date, few solutions can resolve the issue of internal attacks.

4.3.2.2 Spread Spectrum based Communications against PHY Attacks

Over the years, spread spectrum techniques such as FHSS (Frequency Hopping Spread Spectrum) have been actively investigated as countermeasures against PHY attacks [78]. FHSS is a method of transmitting wireless signals by hopping a carrier among many frequencies, using a common sequence that is pre-shared between a sender and a receiver. When detecting attacks on the current frequency, they jump to next available frequencies. Or, they hop over multiple frequencies together regardless of the presence of attackers. However, resiliency of the FHSS scheme relies on a pre-shared secure key (e.g., a common hopping sequence), and establishing secure key pairing under PHY attacks requires another secure communication; this creates a circular dependency.

One promising solution to break the dependency between data communications and key establishment is to leverage Pseudo-random Frequency Hopping (PFH) that is also used in connection establishment of Bluetooth [43]. Before the start of data communications, a sender and a receiver randomly switch over multiple frequencies. Upon meeting on the same frequency by chance, they exchange the key. Performance of the connection set up phase (in terms of latency) relies on the probability of encounters. In PFH, the sender hops 10 times faster than the receiver, which may increase the rendezvous probability. Nevertheless, PFH takes considerable time. It takes up to $10s$ by default and up to several tens of seconds under moderate PHY attack environment. The slow operation is primarily attributed to the fact that a successful data transmission depends on accidental rendezvous amongst nodes.

By definition, random frequency hopping systems fail to provide an upper bound for rendezvous latency; a key establishment phase using PFH could be too slow for the application at hand. Furthermore, time latency is aggravated under severe PHY attacks. An adversary in the worst may block the rendezvous slot. This observation motivates us to devise a new frequency hopping method built on the concept of guaranteed rendezvous. To solve this issue, we propose a novel key establishment scheme harnessing a quorum system.

4.3.3 Fast and Robust Communication

This section represents the proposed scheme, *Frequency Quorum Rendezvous (FQR)*. The novelty of FQR is that it coordinates two random hopping sequences using a quorum system so that it guarantees them to rendezvous within a bounded time. Moreover, FQR is a distributed algorithm and does not require any pre-shared knowledge among network nodes.

4.3.3.1 Quorum System

FQR exploits a quorum system that is a tool for increasing the availability and efficiency of replicated services in distributed computing.

Definition and property. We provide a brief definition of a quorum system and describes its two fundamental properties: *intersection property* and *rotation closure property*. For definitions, we borrow terminologies from [46, 64, 69].

DEFINITION 1. Given a finite universal set $U = \mathbb{Z}_N = \{0, 1, ..., N-1\}$ of N elements, a *quorum system* Q under U is a collection of non-empty subsets of U, which satisfies the *intersection property*:

$$\forall G, H \in Q;\ G \cap H \neq \emptyset.$$

Each G or $H \in Q$ is called a *quorum*, and $\mathbb{Z}_N$ represents a set of non-negative integers less than N.

DEFINITION 2. Given a non-negative integer i and a quorum H in a quorum system Q under $U = \{0, ..., N-1\}$, we define:

$$rotate(H, i) = \{(h + i) \bmod N \mid h \in H\}.$$

DEFINITION 3. A quorum system Q under $U = \{0, ..., N-1\}$ has the *rotation closure property* if the following holds:

$$\forall G, H \in Q \text{ and } i\ \in \{0, ..., N-1\};\ G \cap rotate(H, i) \neq \emptyset.$$

For example, a quorum system $Q = \{\{0,1\}, \{0,2\}, \{1,2\}\}$ under $U = \mathbb{Z}_3 = \{0,1,2\}$ has the rotation closure property. On the other hand, $Q' = \{\{0,1\}, \{0,2\}, \{0,3\}\}$ under $U' = \mathbb{Z}_4 = \{0,1,2,3\}$ does not hold the rotation closure property.

Cyclic quorum system. All quorum systems hold the intersection property. Yet, some of them satisfy the rotation closure property which is essential to support asynchronous operations. In particular, the proposed FQR leverages a *cyclic quorum system* [64] to construct a set of hopping sequences.

DEFINITION 4. A subset $D = \{a_1, ..., a_\kappa\} \subset \mathbb{Z}_N,\ a_i \in \{0, ..., N-1\}$ and $\kappa \leq N$, is called a *cyclic* (N, κ) *difference set* if for every $d \not\equiv 0 \pmod N$ there exist at least one pair of elements (a_i, a_j) such that $a_i - a_j \equiv d \pmod N$.

1 period = 7 time slots

Time slot	0	1	2	3	4	5	6
Sender A	c	c	$\hat{c}$	c	$\hat{c}$	$\hat{c}$	$\hat{c}$
Receiver B	c	$\hat{c}$	c	$\hat{c}$	$\hat{c}$	$\hat{c}$	c

Figure 4.14: A *quorum-time mapping* strategy in a frequency hopping system.

Given any N, Jiang *et al.* [46] proved that $\sqrt{N} \leq \kappa \leq N$. When selecting the minimum κ ($=\lceil\sqrt{N}\rceil$), it is called a *minimal (N,κ) difference set.*

DEFINITION 5. Given a (N, κ) difference set $D = \{a_1, ..., a_\kappa\} \subset \mathbb{Z}_N$, a *cyclic quorum system constructed by* D is $Q = \{G_0, ..., G_{N-1}\}$, where $G_i = \{a_1 + i, a_2 + i, ..., a_\kappa + i\} \pmod N$ and $i = 0, ..., N-1$.

For a (7,3) difference set $D = \{0, 1, 3\} \subset \mathbb{Z}_7$, for instance, the set $\{0, 1, 3\}$ modulo 7 yields $d = \{1, ..., 6\}$. Then, a cyclic quorum system $Q = \{G_0 = \{0, 1, 3\}, ..., G_6 = \{6, 0, 2\}\}$ is constructed from D.

Quorum-based frequency hopping system. The quorum system has been used for rendezvous in wireless ad hoc networks. In a single frequency environment [46], each element in a quorum is mapped to a time slot index. In Figure 4.14, node A and B choose the quorum $G_0 = \{0, 1, 3\}$ and $G_6 = \{6, 0, 2\}$, respectively. The quorum determines active time slots at which each node wakes up and listens to the frequency, denoted by c in the figure; thus, they meet at time slot *0*. The quorum system, however, cannot be directly applied to frequency hopping. A quorum is a set of numbers, whereas frequency hopping needs two independent variables, *time slot index* and *frequency index*. In other words, *two nodes should be on the same frequency at the same time to rendezvous*. For instance, if A and B in Figure 4.14 visit frequency *5* and *2* at time slot *0*, then they miss each other.

To overcome this challenge, Quorum-based Channel Hopping (QCH) [24] proposed a *quorum-time mapping* algorithm for a control channel establishment among secondary nodes in a Cognitive Radio (CR) system. QCH allows nodes to rendezvous as many times as

possible, but it is not appropriate to a hostile, jamming environment. Moreover, QCH cannot guarantee rendezvous among nodes unless synchronized strictly. The hopping pattern is so simple that a sophisticated jammer can easily predict the sender's hopping sequence and achieve efficient jamming attacks. This paper leverages a quorum system in a different manner to generate a set of frequency hopping sequences.

4.3.3.2 Frequency Quorum Rendezvous

Problem definition. To discuss frequency hopping for key establishment, we assume that time is divided into slots and a frequency hopping period consists of t time slots. We also suppose that a set of N sub frequencies (i.e., sub-channels) is known. A frequency hopping system is constructed by assigning frequencies to t time slots and thus it determines frequency hopping sequence in one period, X, which is denoted:

$$X = \{x_0, ..., x_{t-1}\} = \{(0, f_0), ..., (t-1, f_{t-1})\},$$

where $x_i \in X$ contains a tuple of (*time slot index*, *frequency index*) and $f_i \in \{0, ..., N-1\}$ represents the frequency index at time slot i in a period. Given two frequency hopping sequences X and Y, they are said to *rendezvous* if they have at least one common element, that is, $x_i = y_i$ $(0 \leq i \leq t-1)$. If a pair of nodes selects the sequences of X and Y respectively, then they are guaranteed to be on the same frequency at the same time slot at least once within a period.

Quorum rendezvous frequency hopping. Algorithm 2 is the construction algorithm for the FQR system. It generates two different sequences, i.e., sending and receiving, by assigning frequencies to time slots. We introduce the algorithm using a simple example that has $N = 7$, $\kappa = 3$, and a *minimal (N,κ) difference set*. The procedure is following:

1. Construct a universal set $U = \mathbb{Z}_7 = \{0, ..., 6\}$ and determine a (7,3) difference set D, $(\sqrt{7} \leq 3 \leq 7)$.

2. Construct a cyclic quorum system $Q = \{G_0, ..., G_6\}$ from D.

Algorithm 2 FQR System Construction Algorithm

Require: $N, \kappa, U = \mathbb{Z}_N$, and a cyclic quorum system Q

Ensure: Sending sequence X and receiving sequence Y

1: Select i randomly, where $i \in U = \{0, ...N-1\}$
2: Obtain a quorum $G_i = \{g_0, ..., g_{\kappa-1}\}$, where $G_i \in Q = \{G_0, ..., G_{N-1}\}$
3: $X = \emptyset$ and $Y = \emptyset$
4: **for** $j = 0$ to $\kappa^2 - 1$ **do**
5: $\quad m \leftarrow j \bmod \kappa$
6: $\quad n \leftarrow (j - (j \bmod \kappa))/\kappa$
7: $\quad x_j = (j, g_m)$, where $g_m \in G_i$
8: $\quad y_j = (j, g_n)$, where $g_n \in G_i$
9: $\quad X \leftarrow X \cup x_j$
10: $\quad Y \leftarrow Y \cup y_j$
11: **end for**
12: Permute frequency indexes in each κ frame
13: **return** $X = \{x_0, ..., x_{\kappa^2-1}\}$ and
14: $Y = \{y_0, ..., y_{\kappa^2-1}\}$

3. A node A selects a random number *i=1* from U, and then obtains a quorum $G_1 = \{1, 2, 4\}$ from Q.

4. The following equation assigns a frequency to the time slot j using the quorum $G_1 = \{1, 2, 4\}$.

$$x_j = (j, g_m) \text{ and } y_j = (j, g_n)$$

where $m = j \bmod \kappa$ and $n = (j - (j \bmod \kappa))/\kappa$.

5. Repeat step (4) for all 9 ($= \kappa^2$) time slots. This constructs a sending sequence $X = \{(0,1), (1,2), (2,4), (3,1), (4,2), (5,4), (6,1), (7,2), (8,4)\}$ and a receiving sequence $Y = \{(0,1), (1,1), (2,1), (3,2), (4,2), (5,2), (6,4), (7,4), (8,4)\}$. Then, permute frequency indexes within every frame of X.

		One time period								
		Frame 1			Frame 2			Frame 3		
	Time slot	0	1	2	3	4	5	6	7	8
A takes X	Quorum G_1	4	1	2	2	1	***4***	1	4	2
B takes Y′	Quorum G_3	3	3	3	4	4	***4***	6	6	6

Figure 4.15: FQR with (7,3) difference set under $\mathbb{Z}_7$. The node A, as a sender, uses the sending sequence X, and the node B uses the receiving sequence Y'. They rendezvous on *frequency 4* at *time slot 5*.

6. A node B repeats step (4-5) with a selected quorum $G_3 = \{3, 4, 6\}$, and then, construct two hopping sequences X' and Y'.

Figure 4.15 illustrates rendezvous of the FQR system when the nodes A and B choose the sequence X and Y', respectively. As shown, they rendezvous on *frequency 4* at *time slot 5*.

FQR has three distinctive characteristics. First, FQR exploits a *quorum-frequency mapping* strategy. The elements in a selected quorum are mapped into frequency indexes. For example, when a quorum $G_1 = \{1, 2, 4\}$ is selected, a node assigns frequencies 1, 2, and 4 from $\mathbb{Z}_7$ to consecutive time slots. Second, after selecting a quorum randomly, a node generates two different frequency hopping sequences: a *sending sequence* and a *receiving sequence*. If a node has data to transmit, it hops frequencies according to the *sending sequence*. Otherwise, the node follows the *receiving sequence*. Last, the quorum size determines the length of one time period. Say, given quorum size κ, one period consists of $|\kappa|$ frames each of which contains $|\kappa|$ time slots. In short, the length of one time period = κ^2 time slots. The length of the period indicates upper bound since FQR guarantees at least one rendezvous within one period. To build a hopping sequence, FQR makes use of the *minimal* (N, κ) *difference set* where κ approximates its lower bound $\lceil\sqrt{N}\rceil$. Thus, the upper bound of time cost for rendezvous in FQR, κ^2, approximates N given N frequencies.

4.3.3.3 Key Establishment under Jamming Attack

A key pairing (key establishment) phase enables communicating nodes to establish an initial common secure key (hopping sequence in FHSS) which will be used subsequently to mitigate the effects of jamming attacks during data transmission. Because of the presence of jammers, the key pairing should be established securely. Moreover, time is of essence in tactical and emergency applications.

For *secure establishment* set up, a node verifies communication partners via an authentication process. A popular key agreement scheme, the Elliptic Curve Diffie–Hellman (DH) protocol is used for FQR. In the DH authentication, two nodes exchange two messages, say M_A of A and M_B of B. Each message may contain credentials such as a public key signature, and an encrypted common key. The message is split into l fragments each of which is transmitted as a separate packet: $m_1, m_2, ..., m_l$. It is possible to transmit l packets in a single time slot. Yet, a long data transmission is vulnerable because it is easily detected by jammers. We assume that the time slot duration is short enough to prevent detection by jammers with significant probability. Thus, a node must use multiple time slots to deliver l packets. The node is assumed to have two independent radio interfaces, so that it sends and receives packets simultaneously [77]. A repeatedly sends the chain of l packets and, in parallel, receives incoming packets. Upon receiving the l packets from M_A, B validates A's signature. If verified, B transmits its signature to A, M_B. A terminates sending the packet chain when a timeout expires or A verifies B. Successful transmission of $2l$ packets is necessary for the DH authentication. After the authentication, A and B extract the shared key from which they construct a common hopping sequence. Then, they transmit data hopping along the shared frequency.

The key pairing needs *fast establishment* to reduce time latency. Having observed that nodes are required to exchange $2l$ packets for authentication, the next issue is how quickly A and B are able to transmit the packets in the presence of jammers. Since nodes can transmit packets only when they rendezvous, the rendezvous probability P_{rdvs} of the frequency hopping scheme clearly affects performance. In PFH, different hopping speeds amongst nodes affects

P_{rdvs}. Let η the ratio of the hopping speed of A to that of B[16]. Then, rendezvous probability that A rendezvous B in η time slots becomes $\frac{\eta}{N}$. Yet, this cannot guarantee rendezvous since they still selects the hopping sequences arbitrarily. In contrast, FQR guarantees at least 1 rendezvous within κ^2 time slots. This proves that FQR assures data transmission of the $2l$ packets within a bounded time if no adversary jams the channel.

4.3.4 Performance Evaluation

4.3.4.1 Jamming Attack Models

We assume that a jammer is capable of transmitting noise signals on a subset of different frequencies (F_b) simultaneously, so it can block all those frequencies. At the same time, it can listen to multiple frequencies (F_s) to sense ongoing transmissions. The jammer is assumed to be so smart that it does not listen and block the same frequency at the same time. The transmit and receive operations on frequencies in F_b and F_s are independent; thus the jammer can switch transmissions in F_b regardless of the listening operation. One restriction is switching latency of τ, in both. Given the total number of available frequencies N in a network, we have $N = n_b + n_s + n_c$, where n_b and n_s are the numbers of blocked and sensed frequencies (i.e., $|F_b|$ and $|F_s|$) and n_c is the number of clear frequencies. Now, we define n_j as the number of jammed frequencies, the total number of frequencies that are being blocked or are possibly blocked after being sensed. It is computed: $n_j = \alpha n_b + \beta n_s$ $(0 \leq n_j < N)$, where α and β are estimated differently according to various jamming attack models. We note that there is at least one subset of non-jammed frequencies through which data can be successfully transmitted: $N - n_j \geq 1$. The probability of jamming attack on an arbitrary slot P_j is given by: $P_j = \frac{n_j}{N}$.

A legitimate node is assumed to hop over multiple frequencies during each of which it stays for δ (dwell time) with a switching latency of τ. We assume that packets are transmitted over the entire dwell time. The jammer requires ε seconds to detect data transmission on a listening frequency. If the jammer jams the packet transmission for at least $\rho\delta$ duration,

[16] A is a sender, and B is a receiver.

where $\rho > 0$, then a receiver cannot decode the packet correctly, signifying the success of the jamming attack.

External jammer. An external jammer is unaware of frequency hopping behaviors of A and B. According to its sensing capability, it is classified into two models: an active jammer and a responsive jammer.

An *Active External (AE) jammer* does not sense frequencies ($n_s = 0$). Instead, it arbitrarily selects and blocks target frequencies (F_b) for $\rho\delta$ duration. Then, the jammer switches to other target frequencies on which it jams again. One packet transmission (δ) can be jammed $\frac{\delta}{\tau+\rho\delta}$ times. Thus, we compute $n_j = (\frac{\delta}{\tau+\rho\delta})n_b$. To be precisely, $n_j = min((\frac{\delta}{\tau+\rho\delta})n_b, N-1)$.

A *Responsive External (RE) jammer* is able to listen to frequencies (F_s) as well as block other frequencies (F_b). Note that $F_s \cap F_b = \emptyset$. With respect to F_s, the RE jammer initially monitors ongoing transmissions on F_s and launches an attack only when a signal has been detected. Thus, the jammer takes $\varepsilon + \tau + \rho\delta$ seconds to jam a packet transmission of δ duration. The impact of F_b on n_j is same to that in the AE jammer. The similar reasoning for n_j applies, and we compute $n_j = (\frac{\delta}{\tau+\rho\delta})n_b + (\frac{\delta}{\varepsilon+\tau+\rho\delta})n_s$. Then, $n_j = min(n_j, N-1)$.

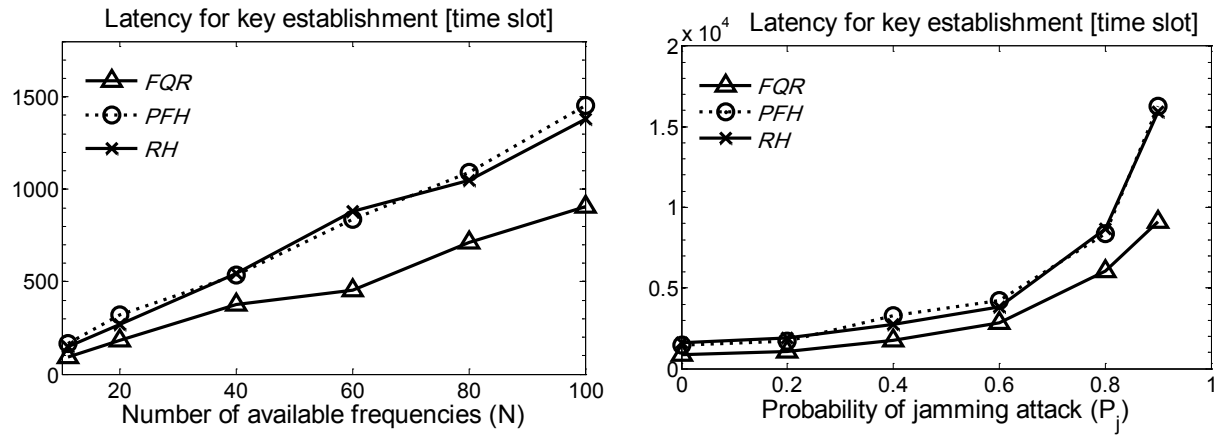
Internal jammer. The hopping frequencies in FQR are determined by the quorum system, whereas PFH selects frequency randomly. Therefore, an internal jammer, who knows the (N, κ) difference set of the quorum system used for key pairing, should be considered to investigate effects of its knowledge on key pairing performance. We do not consider the jammer having valid credentials or any forged credentials of a legitimate node, however. Two types of internal jammers are introduced below. We assume that there is one jammer in the network for easy representation.

An *Impersonating Internal (IP) jammer* constructs a hopping sequence using the quorum system and pretends to be a legitimate node. Here, we assume $n_b = n_s = 1$. On rendezvous, the jammer tries to establish a fake key pairing with a communication partner. Fortunately, the authentication process eventually detects the IP jammer. Nonetheless, the attack wastes time slots and increases time and computation overhead. In particular, impersonating a

receiver consumes transmission of $2l$ packets before being detected, which is worse than mimicking the sender. Given P_j, time overhead of packet transmission relies on P_{rdvs} of the hopping scheme.

Time overhead can be alleviated by reducing the number of packets, $2l$, for authentication. We consider two alternate authentication approaches. The one is that a sender A temporarily generates a pseudo random hopping sequence and transmits it as a plain text. After that, A listens to the begin of the sequence to check if a legitimate receiver B has rendezvous. Upon receiving the message, B replies with M_B by hopping along the sequence. Then, the authentication process completes after A transmits M_A to B along another temporary hopping sequence. Since credentials are not included in the transmitted message, this hugely reduces the size of the initial message to transmit. One packet over one time slot duration might deliver the sequence data over P_{rdvs}. Because both M_A and M_B are delivered via the shared temporary sequence, their transmission is not affected by P_{rdvs}. If A rendezvous with the jammer before B, however, A receives the jammer's fake signature, say M_J, and wastes time slots for l packets which also do not suffer from P_{rdvs}. More jammers probably waste more slots due to increased fake messages. The other alternative assumes a group key used only for data encryption. The temporary sequence data is encrypted with the group key which also contains M_A. B, as a group member, only can decrypt data and send M_B along the sequence, but the jammer cannot. Note that encryption increases the size of the packets suffering from P_{rdvs} whereas M_B are not affected by P_{rdvs}.

An *Intelligent Internal (IT) jammer* is able to listen on multiple frequencies in F_s simultaneously. It predicts the next hopping frequency that A is likely to jump based on the listening records and its knowledge of the quorum. The jammer, having knowledge of N quorums, initially tries to find which quorum A is using for its hopping. Each listening record tells a frequency index that belongs to a specific quorum. As the jammer observes more records, therefore, it may predict A's quorum with higher probability. It, then, estimates A's next frequency within the quorum. Thereafter, the jammer attacks the estimated frequency with n_b=1. Due to limited space, detailed analysis and evaluation of the IT jammer are not included. We refer [59] to interested readers.



(a) Latency of frequency hopping schemes with varying number of available frequencies (N).

(b) Latency of frequency hopping schemes with varying probability of jamming attack (P_j). N=100.

Figure 4.16: Latency performance of frequency hopping systems.

4.3.4.2 Experiments and Results

FQR is implemented on MATLAB. It exploits the cyclic quorum system using minimal (N, κ) difference sets [64]. For performance comparison, we implement PFH with $\eta = 10$ (also known as UFH) and Random Hopping (RH) which is PFH with $\eta = 1$. Like Bluetooth, the hopping speed of a sender is $1.6kHz$; $\delta = 625\mu s$. We also set $\tau = 80\mu s$ and $\varepsilon = 200\mu s$ [19]. For reliable communications, an error-correcting code such as Turbo code or Reed-Solomon code determines the ρ value which is set to 0.2 in this work. Given the configuration, we compute $P_j = \frac{3n_b}{N}$ for the AE jammer and $P_j = \frac{3n_b+1.5n_s}{N}$ for the RE jammer. Network bandwidth is $1Mbps$ and 1024 bits M_A is split into 6 packets ($l = 6$).

Comparison of frequency hopping schemes. In the first experiment, we vary the number of available frequencies (N) from 11 to 100 and measure the required number of time slots to transmit $2l$ packets with $P_j = 0$. This represents time latency of frequency hopping schemes. Figure 4.16a shows that FQR outperforms PFH and RH. The separation of lines becomes clear as N increases because nodes in PFH and RH randomly select frequencies among the increasing number of available frequencies. In particular, FQR reduces the time overhead of the key pairing by 35∼38% when N=100. Figure 4.16b depicts time latency

results when we vary probability of jamming attack (P_j) with N=100. PFH shows the same performance to RH although its η value is 10 times larger than RH's. As analyzed earlier, P_{rdvs} of PFH is influenced by N and is computed as 0.1 within 10 time slots in this experiment. This fails to make any difference from RH. Moreover, time latency grows rapidly as P_j increases. On the contrary, FQR demonstrates better performance than PFH and RH over all the ranges of P_j. We expect that higher rendezvous probability of FQR contributes to the reduction of latency which is measured by more than 7,000 time slots (up to 47%). It should also be noted that we are just reporting average latencies. On top of that, RH and PFH suffer from a potentially large variance that prevents the definition of tight delay bounds. In contrast, FQR has zero of very small variance.

Impact of authentication scheme. In addition to DH, an earlier section introduces two alternate authentications to reduce the time overhead: *plain text* and *group key*. For implementations, we assume that one packet is capable of delivering the temporary sequence without encryption, i.e., plain text sequence. We also take into account the case where a sender A rendezvous with the IP jammer before meeting a legitimate receiver B. Say, with P_j=0.7, A may meet B after 7 rendezvous with the jammer in the worst case. Then, the overhead includes time wasted by 7 packets containing the temporary sequence and 7 authentication messages of the jammer's. As for the group key, the encrypted message sent by the sender is assumed to be 1.5 times larger than M_A in DH.

The performance of FQR with three authentication schemes are compared with varying P_j in Figure 4.17. The group key reduces the time overhead by 34% compared to DH because M_B is transmitted along the temporarily shared hopping sequence. Since P_j only influences this transmission, the group key can achieve faster key pairing phase. However, having the encrypted message whose size is 2 times larger than M_A cancels the benefit (not shown in the plot). The plain text remarkably alleviates the time overhead; 10 times of reduction. This result implicitly shows the impact of P_{rdvs} of the frequency hopping scheme on the overhead. Recall that the plain text transmits the temporary sequence data in a packet, which is only influenced by P_{rdvs}, and M_A and M_B are delivered along the shared hopping sequence. Since

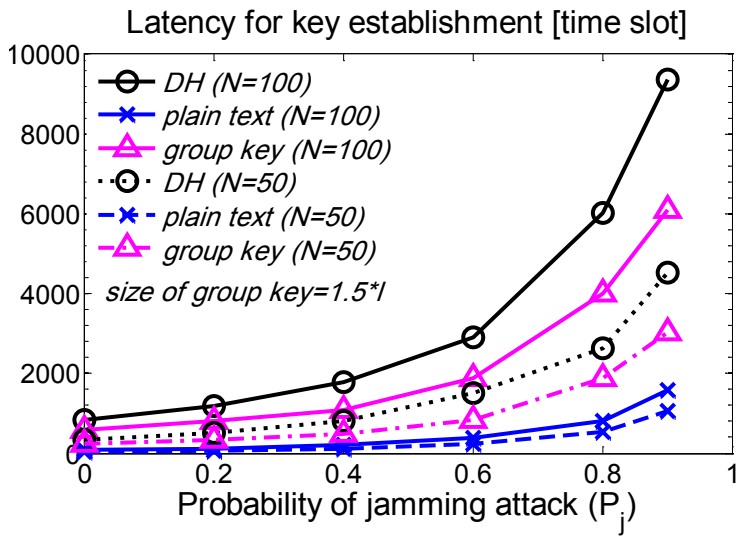


Figure 4.17: Authentication methods affect latency performance in FQR. The message size in the group key is 1.5 time larger than that in DH.

FQR assures rendezvous, it could minimize latency caused by P_{rdvs}. We also evaluate the benefit (see the three dotted curves in Figure 4.17) that each scheme enjoys with smaller N, after having learned that the smaller number of available frequencies improves performance from Figure 4.16a. DH and the group key show a performance enhancement of around 54% with N=50 while the plain text reduces the time overhead by 44%. This difference is also attributed to the number of packets suffering from P_{rdvs}.

CHAPTER 5

Conclusion

Summary. The smart grid, as the most widely accepted instantiation of the IoT, aims to maintain energy balance reliably over the entire power grid by facilitating interoperations among ubiquitous energy objects. While foundational object technologies have been studied for a while and matured, few researches studied their inter-networking problems. Thus, this dissertation investigated an interoperable and secure interaction system enabling the interoperations through three chapters.

Chapter 2 examined an interoperable interaction model for a standalone energy object system. Especially, it designed an Energy Service Interface, a middleware system for interoperable communication. We looked into four categories of design issues to support interoperable customer energy services. Two of them address functional requirements that the ESI supports as a service prosumer. The other two examine quality requirements for better energy service - security and ESI system architecture. To verify the issues, we built and deployed two ESI testbeds that also showed how the design issues are implemented in a real world. Through experiments with a couple of energy service scenarios, we have demonstrated the service interoperation and evaluated the performance.

Chapter 3 examined interoperable energy services to realize the IoT in a smart building context. It enhanced the existing energy management system in a building facility and proposed a Microgrid Platform. We discussed the autonomous operation of a microgrid with emphasis on energy forecasting and examined three types of energy services that realized smart grid interoperation. To demonstrate the feasibility of MP, we implemented a building-level testbed. The testbed deployed various types of energy resources and developed a few energy services. We conducted experiments and showed the capability of building facility

as a microgrid. We also deployed the testbed in a small customer facility and helped the building owner participate in an automated demand response service that a local utility company offered. By running the testbed in the real-world energy market, we demonstrated that the facility could contribute to the energy balance as a smart griddable microgrid.

Chapter 4 examined three security mechanisms to protect the interactions between the energy objects. The first section presented Resource Centric Security that provides fine-grained, scalable access control and encryption. To support fine granularity, it leverages the concept of a filesystem access control list so that individual energy resource maintains three privileges of read, write, and execute. Instead of having three predefined classes of accessing users, RCSec dynamically assigns a set of attributes to each privilege. And, an external user can only obtain permission to each privilege by showing that his own attribute set matches the resources set. To provide confidentiality, we implement RCSec on top of attribute-based encryption that encrypts data using the assigned set of attributes. RCSec scales well and fits to distributed smart grid environment because the ESI works without any prior knowledge of user information. The experimental results and following analysis discover that RCSec can provide a proper level of abstracted data protection with reasonable overhead.

Next section presented Multi-Factor Authentication and Authorization that employs a multi-factor technique for enhanced access control using attribute-based identification. A user is granted more than two factors consisting of attributes from independent authorities, and an object develops its own access control policy by combining arbitrary number of factors. When accessing the object (to read data or to control the resource), the user is challenged with the policy. If her attributes and factors satisfy the policy, she is qualified. We applied MFAA to a fine-grained access control scenario in our smart building testbed. The experiments measured its computation cost and showed reasonable performance although there remain rooms for further enhancement.

The last section presented a novel anti-jamming key establishment scheme, Frequency Quorum Rendezvous. It allows nodes to select hopping sequences in a pseudorandom fashion. This does not require prior knowledge, which guarantees protected communications. In addition, the hopping sequence is constructed from a quorum system, so nodes are guaran-

teed to rendezvous within a bounded time. Applying this concept to the key pairing phase in the presence of jammers reduces time latency and achieves fast communications. These benefits make wireless communication more robust against jamming attacks. The experimental results show that the proposed scheme outperforms existing methods under various jamming attack models.

Discussion. Based on the lessons learned from this dissertation research in the smart grid context, the followings discuss the insights about the Internet of Things.

Smart city. The ultimate goal of the IoT is to construct an intelligent environment that interacts with the existing information network so as to benefit people. In this sense, a tangible goal is to build a smart city accommodating the Internet of Energy, connected vehicles, etc.

Loosely-coupled connection. An object (or thing) freely interacts with any other systems, instead of being tightly coupled with a centralized server. This property requires each object system to operate in a more independent and autonomous manner. Moreover, the systems make stateless interactions due to hardware limitations.

Big data. Each object generates data every second, and there is no doubt that the total volume overwhelms that flowing on the Internet today. Analyzing the big data, creating value added information, and understanding our environment will be one of the integral parts in the future IoT research.

Intelligent network. Unlike the Internet, a network plays more intelligent roles in the IoT. Significant portions of data are generated and consumed within a local environment, and thus it is often inefficient to process huge amount of IoT data solely in a backend cloud. In this sense, the importance of in-network computing will increase. Moreover, the network will be more aware of surroundings' contexts and data contents and flexibly adjust its networking behaviors based on the dynamics in our environment. Recent emergence of information-centric networking, software-defined networking, and machine-to-machine communication addresses the network intelligence.

Security. Analyzing a huge volume of IoT data can disclose what we hear, look at, feel, and even think in very detail. This is a significant privacy violation that must be inspected in IoT research. Another unique concern in the IoT is physical domain security. It is attributed to the fact that the objects are not protected well physically unlike conventional servers and personal mobile devices. Since everything is inter-connected, a small threat can subvert the entire IoT infrastructure.

References

[1] Open Access Same-Time Information System (OASIS), California Independent System Operator (CAISO). http://oasis.caiso.com/.

[2] REQ.21-Energy Services Provider Interface (ESPI), North American Energy Standards Board (NAESB). http://www.naesb.org/ESPI_Standards.asp.

[3] Semantic Model Working Party, Smart Grid Architecture Committee. http://collaborate.nist.gov/twiki-sggrid/bin/view/SmartGrid/SGIPSemanticModelSGAC.

[4] UCA International user group (UCAIug) Home Area Network (HAN) System Requirements Specification, v2.0. http://osgug.ucaiug.org/sgsystems/openhan/default.aspx.

[5] Energy Independence and Security Act of 2007 - [Public Law No: 110-140], 2007.

[6] EIS Alliance, Customer Domain Energy Services Interface (ESI) Requirements, V3.01, 2010.

[7] EIS Alliance, Customer Domain Use Cases, v3.01, 2010.

[8] WXXM 1.1 Primer, Federal Aviation Administration / European Organization for the Safety of Air Navigation, 2010.

[9] EC Communication on Smart Grids - Smart grids: from innovation to deployment - COM(2011) 202, 2011.

[10] Machine-to-Machine communications (M2M), Functional architecture, ETSI TS 102 690 v.1.1.1, 2011.

[11] OpenADR Open Source Toolkit: Developing Open Source Software for the Smart Grid, LBNL-5064E. Technical report, Lawrence Berkeley National Laboratory, 2011.

[12] NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 2.0, Feb. 2012.

[13] The U.S. Energy Information Administration, Annual Energy Outlook, 2012.

[14] Electric Sector Failure Scenarios and Impact Analyses. *National Electric Sector Cybersecurity Organization Resource (NESCOR)*, 2013.

[15] Y. Agarwal, B. Balaji, S. Dutta, R. Gupta, and T. Weng. Managing Plug-Loads for Demand Response within Buildings. In *ACM BuildSys*, 2011.

[16] Y. Agarwal, R. Gupta, D. Komaki, and T. Weng. BuildingDepot: An Extensible and Distributed Architecture for Building Data Storage, Access and Sharing. In *ACM BuildSys*, 2012.

[17] Y. Agarwal, T. Weng, and R. K. Gupta. The Energy Dashboard: Improving the Visibility of Energy Consumption at a Campus-Wide Scale. In *ACM BuildSys*, 2009.

[18] P. Anderson and I. Geckil. Northeast Blackout Likely to Reduce US Earnings by $6.4 Billion. *AEG Working Paper*, (2003-2), 2003.

[19] P. Bahl, R. Chandra, and J. Dunagan. Ssch: slotted seeded channel hopping for capacity improvement in ieee 802.11 ad-hoc wireless networks. In *ACM MobiCom*, 2004.

[20] A. Beimel. *Secure Schemes for Secret Sharing and Key Distribution*. PhD thesis, Israel Institute of Technology, 1996.

[21] G. Bellala, M. Marwah, M. Arlitt, G. Lyon, and C. Bash. Towards an understanding of campus-scale power consumption. In *ACM BuildSys*, 2011.

[22] G. Bellala, M. Marwah, M. Arlitt, G. Lyon, and C. E. Bash. Towards an Understanding of Campus-scale Power Consumption. In *ACM BuildSys*, 2011.

[23] J. Bethencourt, A. Sahai, and B. Waters. Ciphertext-policy attribute-based encryption. In *IEEE Symposium on Security and Privacy*, Oakland, USA, May 2007.

[24] K. Bian, J.-M. Park, and R. Chen. A quorum-based framework for establishing control channels in dynamic spectrum access networks. In *ACM MobiCom*, 2009.

[25] M. Botts and A. Robin. OpenGIS Sensor Model Language (SensorML) Implementation Specification, OpenGIS Implementation Specication OGC 07-000 v1.0.0, July 2007.

[26] P. Cappers, C. Goldman, and D. Kathan. Demand Response in US Electricity Markets: Empirical Evidence. *Elsevier Energy*, 35(4), 2010.

[27] H.-P. Chao. Demand Response in Wholesale Electricity Markets: The Choice of Customer Baseline. *Journal of Regulatory Economics*, 39(1), Feb. 2011.

[28] C.-Y. Chung, P. Chu, and R. Gadh. Design of Smart Charging Infrastructure Hardware And Firmware Design of The Various Current Multiplexing Charging System. In *Global Conference on Power Control and Optimization*, Aug. 2013.

[29] C.-Y. Chung, J. Chynoweth, C. Qiu, C.-C. Chu, and R. Gadh. Design of Fair Charging Algorithm for Smart Electrical Vehicle Charging Infrastructure. In *Int'l Conference on ICT Convergence*, Oct. 2013.

[30] C.-Y. Chung, J. Chynoweth, C. Qiu, C.-C. Chu, and R. Gadh. Design of Fast Response Smart Electric Vehicle Charging Infrastructure. In *IEEE Green Energy and Systems Conference*, Nov. 2013.

[31] C.-Y. Chung, A. Shepelev, C. Qiu, C.-C. Chu, and R. Gadh. Design of RFID Mesh Network for Electric Vehicle Smart Charging Infrastructure. In *IEEE Int'l Conference on RFID Technologies and Applications*, Sep. 2013.

[32] C.-Y. Chung, E. Youn, J. Chynoweth, C. Qiu, C.-C. Chu, and R. Gadh. Safety Design for Smart Electric Vehicle Charging with Current and Multiplexing Control. In *IEEE Int'l Conference on Smart Grid Communications*, Oct. 2013.

[33] T. Considine and et al. oBIX 1.0 Committee Specification 01 - obix-1.0-cs-01, 2006.

[34] K. Coughlin, M. A. Piette, C. A. Goldman, and S. Kiliccote. Statistical Analysis of Baseline Load Models for Non-Residential Buildings. *Elsevier Energy and Buildings*, 41(4), 2009.

[35] S. Dawson-Haggerty, X. Jiang, G. Tolle, J. Ortiz, and D. Culler. sMAP-a Simple Measurement and Actuation Profile for Physical Information. In *ACM SenSys*, 2010.

[36] D. V. Dollen. Electric Power Research Institute (EPRI) Report to NIST on the Smart Grid Interoperability Standards Roadmap. 2009.

[37] A. Faruqui, R. Hledik, and J. Tsoukalis. The Power of Dynamic Pricing. *Elsevier The Electricity Journal*, 22(3), April 2009.

[38] R. Fielding and R. Taylor. Principled Design of the Modern Web Architecture. *ACM Transactions on Internet Technology*, 2(2):115–150, 2002.

[39] M. Gerla, E.-K. Lee, G. Pau, and U. Lee. Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds. In *IEEE World Forum on Internet of Things*, March 2014.

[40] V. Goyal, O. Pandey, A. Sahai, and B. Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *ACM CCS*, Oct. 2006.

[41] D. Hardin. Customer energy services interface white paper. In *Grid-interop Forum*, Dec. 2011.

[42] R. Huang, T. Huang, R. Gadh, and N. Li. Solar Generation Prediction using the ARMA Model in a Laboratory-level Micro-grid. In *IEEE Smart Grid Communications*, Taiwan, Nov. 2012.

[43] IEEE 802.15.1-2005 Standard for Information Technology. *Wireless MAC and PHY Specifications for Wireless Personal Area Networks (WPANs)*, 2005.

[44] International Electrotechnical Commission. Communication Networks and Systems in Substations, IEC 61850-161850-10. 2003.

[45] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and R. Braynard. Networking named content. In *ACM CoNEXT*, Rome, Italy, Dec. 2009.

[46] J.-R. Jiang, Y.-C. Tseng, C.-S. Hsu, and T.-H. Lai. Quorum-based asynchronous power-saving protocols for ieee 802.11 ad hoc networks. In *ACM Mobile Networks and Applications*, 2005.

[47] Y.-J. Kim, J. Lee, G. Atkinson, H. Kim, and M. Thottan. SeDAX: A Scalable, Resilient, and Secure Platform for Smart Grid Communications. *IEEE Journal on Selected Areas in Communications*, 30(6):1119 – 1136, Aug. 2012.

[48] A. Krioukov, G. Fierro, N. Kitaev, and D. Culler. Building Application Stack (BAS). In *ACM BuildSys*, 2012.

[49] R. Lasseter and P. Piagi. Microgrid: A Conceptual Solution. In *IEEE Power Electronics Specialists Conf. (PESC)*, June 2004.

[50] E.-K. Lee, P. Chu, and R. Gadh. Fine-Grained Access to Smart Building Energy Resources. *IEEE Internet Computing*, 17(6), Nov.-Dec. 2013.

[51] E.-K. Lee, R. Gadh, and M. Gerla. Resource Centric Security to Protect Customer Energy Information in the Smart Grid. In *IEEE Smart Grid Communications*, Nov. 2012.

[52] E.-K. Lee, R. Gadh, and M. Gerla. Energy Service Interface: Accessing to Customer Energy Resources for Smart Grid Interoperation. *IEEE Journal on Selected Areas in Communications*, 31(7):1195–1204, July 2013.

[53] E.-K. Lee, R. Huang, P. Chu, and R. Gadh. Microgrid Platform: Enhancing Customer Building Facility to Support Microgrid Operations. *IEEE Journal on Select Areas in Communications*, submitted, 2014.

[54] E.-K. Lee, R. Huang, P. Chu, R. Gadh, and M. Gerla. Enhancing Customer Building Facility to Support Microgrid Operations. Technical Report 140001, UCLA, 2014.

[55] E.-K. Lee, J. K. Lee, J. C. Bae, S. C. Choi, and J. I. Choi. Washington Square: Constructing A Building Energy Management System Using Smart Submeters. In *ACM MobiCom, poster*, April 2011.

[56] E.-K. Lee, J.-H. Lim, J. Joy, M. Gerla, and R. Gadh. Multi-Factor Authentication and Authorization using Attribute Based Identification. Technical Report 140003, UCLA, 2014.

[57] E.-K. Lee, S. Oh, and M. Gerla. Randomized Channel Hopping Scheme for Anti-Jamming Communication. In *IFIP Wireless Days*, Oct. 2010.

[58] E.-K. Lee, S. Oh, and M. Gerla. Frequency quorum rendezvous for fast and resilient key establishment under jamming attack. *ACM SIGMOBILE Mobile Computing and Communications Review*, 14(4), 2011.

[59] E.-K. Lee, S. Y. Oh, and M. Gerla. Fast and resilient key establishment using quorum rendezvous under jamming attack. Technical Report 110005, UCLA, 2011.

[60] E.-K. Lee, S. Y. Oh, and M. Gerla. Physical Layer Security in Wireless Smart Grid. *IEEE Communications Magazine*, 50(8):46–52, Aug. 2012.

[61] A. Lewko and B. Waters. Decentralizing Attribute-Based Encryption. In *EUROCRYPT*, May 2011.

[62] J. Li, M. H. Au, W. Susilo, D. Xie, and K. Ren. Attribute-based Signature and its Applications. In *ACM ASIACCS*, April 2010.

[63] A. Littman, G. Lyon, A. Shah, and J. Vogler. Exploring Advanced Metering Infrastructure Deployments for Commercial and Industrial Sites. In *ASME Conf. on Energy Sustainability*, July 2012.

[64] W.-S. Luk and T.-T. Wong. Two new quorum based algorithms for distributed mutual exclusion. In *IEEE International Conference on Distributed Computing Systems (ICDCS)*, 1997.

[65] H. Maji, M. Prabhakaran, and M. Rosulek. Attribute-Based Signatures. *CT-RSA, LNCS*, 6558:376–392, 2011.

[66] S. Mal, A. Chattopadhyay, A. Yang, and R. Gadh. Electric Vehicle Smart Charging and Vehicle-to-Grid Operation. *Int'l Journal of Parallel, Emergent and Distributed Systems*, 27(3):1–17, 2012.

[67] S. Mal and R. Gadh. Real-Time Push Middleware and Mobile Application for Electric Vehicle Smart Charging and Aggregation. *Int'l Journal of Communication Networks and Distributed Systems*, 10(4):351–378, 2011.

[68] E. Mills and P. Mathew. Monitoring-Based Commissioning: Benchmarking Analysis of 24 UC/CSU/IOU Projects. Technical Report LBNL-1972E, Lawrence Berkeley National Laboratory, June 2009.

[69] M. Naor and A. Wool. The load, capacity, and availability of quorum systems. In *SIAM Journal on Computing*, 1998.

[70] M. Neugschwandtner, G. Neugschwandtner, and W. Kastner. Web Services in Building Automation: Mapping KNX to oBIX. In *IEEE Conference on Industrial Informatics*, June 2007.

[71] NIST Priority Action Plan 2. Guidelines for Assessing Wireless Standards for Smart Grid Applications. Feb. 2011.

[72] T. Okamoto and K. Takashima. Decentralized Attribute-Based Signatures. *PKC, LNCS*, 7778:125–142, 2013.

[73] P. Parikh, M. Kanabar, and T. Sidhu. Opportunities and Challenges of Wireless Communication Technologies for Smart Grid Applications. In *IEEE Power and Energy Society General Meeting*, July 2010.

[74] A. Patel, J. Aparicio, N. Tas, M. Loiacono, and J. Rosca. Assessing Communications Technology Options for Smart Grid Applications. In *IEEE International Conference on Smart Grid Communications*, Oct. 2011.

[75] M. A. Piette, G. Ghatikar, S. Kiliccote, E. Koch, D. Hennage, P. Palensky, and C. McParland. Open Automated Demand Response Communications Specification v1.0. *California Energy Commission - PIER Program*, CEC-500-2009-063, 2009.

[76] M. Pirretti, P. Traynor, P. Mcdaniel, and B. Waters. Secure Attribute-Based Systems. *Journal of Computer Security*, 18(5):799–837, 2010.

[77] C. Popper, M. Strasser, and S. Capkun. Anti-jamming broadcast communication using uncoordinated spread spectrum techniques. *IEEE Journal on Selected Areas in Communications: Special Issue on Mission Critical Networking*, 28(5), June 2010.

[78] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2nd edition, 2001.

[79] S. Ruj and A. Nayak. A Decentralized Security Framework for Data Aggregation and Access Control in Smart Grids. *IEEE Transactions on Smart Grid*, 4(1):196–205, March 2013.

[80] A. Sahai and B. Waters. Fuzzy identity-based encryption. In *IACR Eurocrypt*, Aarhus, Denmark, May 2005.

[81] A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[82] A. Shepelev, C.-Y. Chung, C.-C. Chu, and R. Gadh. Mesh Network Design for Smart Charging Infrastructure and Electric Vehicle Remote Monitoring. In *Int'l Conference on ICT Convergence*, Oct. 2013.

[83] Smart Grid Interoperability Panel - Cyber Security Working Group. Introduction to NISTIR 7628 Guidelines for Smart Grid Cyber Security. Sep. 2010.

[84] M. Souryal, C. Gentile, D. Grifth, D. Cypher, and N. Golmie. A Methodology to Evaluate Wireless Technologies for the Smart Grid. In *IEEE International Conference on Smart Grid Communications*, Oct. 2010.

[85] J. Torres, A. Garcia, M. D. Blas, and A. D. Francisco. Forecast of hourly average wind speed with ARMA models in Navarre. *Solar Energy*, 9:65–77, 2005.

[86] A. Wright, P. Kalv, and R. Sibery. Interoperability and security for converged smart grid networks. In *Grid-Interop Forum*, Dec. 2010.

[87] T. Zhu, A. K. Mishra, D. Irwin, N. Sharma, P. Shenoy, and D. Towsley. The Case for Efficient Renewable Energy Management in Smart Homes. In *ACM BuildSys*, 2011.