PRIVACY-AWARE DESIGN IN THE SMART GRID: TECHNOLOGICAL AND ECONOMIC PERSPECTIVES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Dipayan Ghosh

August 2014

PRIVACY-AWARE DESIGN IN THE SMART GRID: TECHNOLOGICAL AND ECONOMIC PERSPECTIVES

Dipayan Ghosh, Ph.D.

Cornell University 2014

Technologists today aspire to apply data to solve a wide array of problems to advance quality of life and prosperity. However, actors across civil society have been wary of new technologies that implement analytical approaches supported by granular data, arguing this new age in which data is readily processed for immediate profit exploits the civil liberties – in particular, the rights to privacy – of the citizen. The power industry, with novel data mining technologies pervading smart grid systems throughout the United States and the world, is no stranger to this revolution. This dissertation demonstrates the consumer privacy concerns of smart grid technologies and addresses them, first by proposing a technical framework for their privacy-aware design, and then by examining the economic conditions required for privacy adoption by relevant stakeholders. We first illustrate the privacy hazards of collecting temporally precise, fine-granularity data in advanced metering applications by showing that residential consumption data can readily be modeled statistically. We then consider privacy-aware guidelines to design smart grid networks that protect consumer data. Recognizing that consumers are not the only stakeholders in the contest for their personal data, we then consider a competitive game between the individual consumer and the utility company, in which each attempts to maximize profit, with payoffs including gains appreciated from fine-granularity power consumption data. With the aid of this game theoretic framework, we then determine the economic conditions required to motivate stakeholders in the power industry to adopt privacy-aware smart metering at equilibrium. To determine the optimal regulatory framework for introducing smart metering technology to the power industry, we subsequently consider a

set of potential regulatory regimes and examine consumer choices under each as single-player decision processes. Finding that the average consumer's valuation of smart metering privacy is essential in determining the ultimate adoption rates of privacy-aware smart metering systems under each regulatory regime, we present the results of a national survey conducted to estimate this valuation level. Finally, considering these results, we present a series of policy recommendations to address privacy concerns as they relate to power consumption data and, more widely, collection of bulk data.

BIOGRAPHICAL SKETCH

Dipayan grew up in Storrs, Connecticut, and graduated from Edwin O. Smith High School. He attended the University of Connecticut, where he received a Bachelor of Science in Electrical Engineering, graduating summa cum laude as a University Scholar. He began his pursuit of a Doctor of Philosophy in electrical and computer engineering at Cornell University in 2010. There, he conducted interdisciplinary research oriented on privacy and involving engineering, economics, and policy as a National Defense Science and Engineering Graduate fellow with the Department of Defense. In 2012, Dipayan was a scholar at the University of California, Berkeley, studying the economics of privacy with Professor Shankar Sastry and the National Science Foundation's Team for Research in Ubiquitous Secure Technology. He spent summers in Armonk with IBM and in New York with Thomson Reuters. After graduation, he will work as an advisor in technology policy with the National Economic Council and the Office of Science and Technology Policy in the Executive Office of the President at the White House. There, he will focus on issues including the preservation of individual rights to privacy and civil liberties in the digital realm; the expansion of and increased competition in the national broadband network of operators; and the promotion of invention, innovation, and entrepreneurship through reform of the nation's intellectual property and patent system. Concurrently, he will continue his study of privacy economics as a fellow with the Center for Information Technology and Policy at Princeton University and the School of Information at the University of California, Berkeley.

To my parents, Chinmoy and Bedabati Ghosh, and my brother, Debraj Ghosh

ACKNOWLEDGMENTS

This thesis would not have been possible without the help of my friends and colleagues. Thank you to everyone who has supported me in Ithaca.

First and foremost, I would like to thank my doctoral committee including Professors Stephen Wicker, Timothy Mount, and William Schulze, as well as Professor Dawn Schrader, who have always provided their invaluable guidance. I have certainly encountered many obstacles in my research and would not have been able to overcome them without the help of my advisors.

It has been a distinct honor to work with Professor Wicker, my primary advisor, over the past three years. I greatly admire his insight, strive for knowledge, and passion for students, all profoundly rare qualities I hope to emulate. I appreciate his willingness to allow me the flexibility to pursue my studies in my own way. His confidence in me has allowed me to drive my doctoral work in the directions most important to me.

I would like to thank Professor Shankar Sastry and Larry Rohrbough in the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley, who provided the opportunity to study there in 2012, which proved to be one of my most productive periods as a graduate student.

I am grateful to my undergraduate advisors, Professors Peter Luh and Krishna Pattipati, who introduced me to research. Professor Luh helped launch my academic career in giving me the opportunity to work as an undergraduate researcher in his Manufacturing Computing Lab at the Booth Engineering Center for Advanced Technology.

My work has been supported by the Department of Defense (DOD) through the three-year National Defense Science and Engineering Graduate Fellowship (NDSEG), administered by the American Society for Engineering Education (ASEE), which has given me a great deal of freedom in my academic pursuits. I am extremely grateful to ASEE and DOD for their support with the fellowship, and profoundly honored by the federal government's recognition of my academic work. I am also grateful to the National Science Foundation Team for Research in Ubiquitous Secure Technology (TRUST) and the National Science Foundation Trustworthy Computing Program (TC) for funding I have received. Additionally, the first year of my doctoral work was funded by Joan and Irwin Jacobs of Qualcomm through the Jacobs Scholar Fellowship, administered by Cornell. I greatly appreciate the opportunity this fellowship provided to enjoy the early academic experience and determine a course of study.

Finally, I owe deep gratitude to my family for helping me realize my goals. They motivated me to strive for excellence from a young age and have always believed in me. My brother, who the set the bar high by finishing his Ph.D. the year before I started mine, has always challenged me to think critically and question the status quo. My parents have only wanted me to apply myself and put my heart and brain behind anything I endeavor to do, an attitude I hope to retain throughout my career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

1.1 OVERVIEW

Modern technologists have one overriding goal in mind: to maximize the functionality of their creations. From an economic perspective, there is no ill logic to this. Consumer satisfaction is driven by product experience, and if the engineering generates highly useful results, the designer is bound to succeed. This basic concept has driven the success of history's greatest consumer technology feats, whether Apple's innovation in early personal computers that brought them to average households, Microsoft's design of an operating system and software suite with global reach, or Google's development of a search algorithm that facilitates worldwide access to information. Modern startups are no different, with new-age chief executive officers and business development managers thinking continuously about user experience and increasing functionality of their products to push the technology needle further forward for the benefit of an increasingly global consumer market.

Indeed, technology is advancing at breakneck pace. In a few short decades, we have seen the Internet transform from an arcane research tool appreciated by computer scientists and information technology professionals to a commodity that has disrupted the world's economy to the point that there is sincere philosophical debate whether Internet access is a fundamental human right [1]; a new era of mobile devices ushered in as sleek laptops, cellphones, and tablet computers that have become common fare even in parts of the developing world; and novel

technology-powered industries established, from the various peer-to-peer sharing economy markets served by providers like Airbnb and Uber to the massive open online courses that deliver high-quality learning materials to hundreds of thousands of students around the world in real time. Unsurprisingly, in this content-crazed environment, consumers have come to expect frequent innovation from the technology industry. Steve Jobs encapsulated the sentiment technologists face in meeting incrementally progressing consumer needs when he noted aptly "you can't just ask customers what they want and then try to give that to them. By the time you get it built, they'll want something new" [2]. Technologists continuously think critically about their consumer bases to better reach them and provide more useful services. Where they come from, what their hobbies are, who their friends are, what shows they watch, and how their needs can best be met through sound product design are all of great interest to designers. Thus when opportunities arise to improve consumer experience, technologists pounce.

This frame of mind often leaves consumers in the lurch. When engineers, product managers and marketers think primarily about building functionalities into products to improve them without giving due consideration to the nuances of how design innovations can offset other factors of interest to consumers, consumers can inadvertently and indirectly be harmed. Often, this realization comes late to the technologist, who may consider redesign or, as is more often the case, supplementation of existing functionalities with technical workarounds that may add unwanted complexities and otherwise decrease efficiency.

Perhaps the most notable of such overlooked consumer interests is that of privacy. Technology today makes use of novel and highly sophisticated methods to collect, retain, and analyze data

in ways that can lead to the development of detailed insights into consumer behavior – a trend that has been led by industry and government over the past several years and is often referred to as "data mining" or "big data," which according to the White House's 2014 review of big data and privacy led by Counselor to the President John Podesta has "unprecedented computational power and sophistication [to] make possible unexpected discoveries, innovations, and advancements in our quality of life" [3]. The power of data is increasingly providing such opportunities to companies around the world. In recent years, the breadth, depth, and granularity of data associated with individual consumers available to commercial enterprises has vastly increased, a trend expected to continue long into the future. The International Data Corporation estimates less than one exabyte (10^9 gigabytes) of data was stored globally in 2009 [4]. This number grew to five exabytes in 2013 and will boom to forty by 2020. Meanwhile, IDC estimates that only half a percent of today's five exabytes has been analyzed to generate value, suggesting there is still vast potential for companies to exploit currently untapped databases.

This dormant data can be very powerful if businesses leverage it by applying new analytical methods. Analytics can be used to generate business intelligence and insights that in turn can be used to drive technological growth. Market research studies confirm that advanced analytical processes with data are becoming increasingly vital for businesses to entrench into their workflows. In a recent survey of business executives administered by the IBM Institute for Business Value and the MIT Sloan Management Review, organizations with significant experience using analytics across a range of functions were three times likelier to indicate they substantially outperform their industry peers than organizations that are far from achieving their analytical goals [5]. According to the aforementioned White House report, though, "these

capabilities, most of which are not visible or available to the average consumer, also create an asymmetry of power between those who hold the data and those who intentionally or inadvertently supply it." Indeed, there exist substantial privacy implications associated with big data; its use empowers one to collect large amounts of information about consumers. As Jay Stanley of the American Civil Liberties Union notes, "when you combine someone's personal information with vast external data sets, you can infer new facts about that person…And when it comes to such facts, a person a) might not want the data owner to know, b) might not want anyone to know, c) might not even know themselves." Big data technologies play their role, too: "The fact is, humans like to control what other people do and do not know about them – that's the core of what privacy is, and data mining threatens to violate that principle," Stanley writes [6].

Nonetheless, new technology platforms are continually being developed, which will only increase the breadth and depth of collected data associated with individual consumers. For instance, firms are quickly recognizing the ubiquity and impact of mobile technologies, and are fast ramping up their mobile operations to tackle opportunities in around the world, in both developed nations as well as emerging markets. A good example of the disruptive power of mobile data is the collection of information associated with usage of mobile applications. Such usage metrics internal to apps can readily be collected and used to drive further growth of mobile businesses, as knowledge of how individuals use apps can help businesses determine how to improve user experience. Businesses are increasingly using intelligence gathered from such data to improve products, strive for wider user adoption, and develop powerful new monetization strategies.

In this tech-obsessed environment, it is no wonder that enterprises are whole-heartedly tackling the opportunities lying in large databases to explore their potential value. "Data scientists," regarded by the Harvard Business Review as prime careers for those interested in both intellectual pursuits and high-paying salaries, are comprised of a breed of doctorate-wielding mathematicians who use their technical expertise to derive insights out of previously idling, untouched repositories of public and proprietary data [7]. A wide range of companies that conduct business through which they collect consumer data is employing them, including Internet and tech companies like Amazon, Facebook, Google, Palantir, Twitter, and Yahoo, to established corporations like American Express, Bank of America, ExxonMobil, Target, Thomson Reuters, and Wal-Mart.

Government is taking advantage of newly created opportunities, too; the Central Intelligence Agency, National Security Agency, various federal and state law enforcement agencies, and many others presently employ data science to derive insights from warehouses of public and proprietary data so as to improve government and protect our national security interests. Data-intensive initiatives abound outside the intelligence and law enforcement communities as well, with open government and open data initiatives commenced by President Obama and spearheaded by his Chief Technology Officer [8, 9].

Firms in the power industry, including electric utility companies, electricity market regulators, and various others, have extensively applied data analytics to improve demand and supply forecasting for many years, but new, innovative approaches inspired by data science and the availability of increasingly powerful databases are now piquing their interest, too. Data analytics companies are coming into existence left and right to satisfy the demand for data

analytics spurred by smart grid initiatives. Innovators in this space abound, like AutoGrid, who have developed platforms to analyze grid and metering data so that utility companies and consumers can control routing and consumption of power [10]; Viridity Energy, whose software analyzes energy profiles and links them to data from energy markets so that electricity customers can take advantage of new savings opportunities [11]; and Stem, who develop software to analyze electricity consumption in buildings so that administrators can better monitor their budgets and manage building operations [12]. Large corporations also have a significant presence in energy data analytics, including IBM, who have developed analytics suites that can integrate structured and unstructured data to equip utilities with more data to refine power routing and delivery [13]; Intel, whose processors pervade throughout the industry, including at the Pike Powers Laboratory & Center for Commercialization, which deploys Intel software and hardware to service companies that require high-powered computing facilities to analyze troves of disaggregated residential energy data [14]; and General Electric, who provide platforms for professionals to quickly access and visualize massive data sets to facilitate power system operations [15]. Indeed, data analytics is the future of the power industry.

These and many more companies are deeply invested in modernizing the electricity grid through the use of massive data sets on electricity consumption and cutting-edge analytical methods. However, these new technologies place consumer privacy at risk; as with any system involving the collection, use, retention, and sharing of personal data, they require an earnest assessment of the potential privacy risks. In 1998, Nicholas Negroponte, Chairman of MIT's Media Lab, wrote that "no one [entity] has a complete model of us, and it is hard for them to share the parts" [16]. Unfortunately, this may no longer be true. Particularly in the big data

ecosystem, enterprises must be attentive to the potential risk of a "mosaic effect" in which datasets (including those that may not appear to contain personally identifying information) may be combined to identify individuals or disclose unintended personally identifiable information [3]. Companies like Amazon and Google know many sensitive details about us; such companies automatically comb accounts to draw out information that can be used in targeted advertisements, which often forms the basis of their business models. The more detailed a profile of an individual they are able to compile, the more valuable that person's account becomes. With a little capital and a small team of data mining and machine learning experts, smaller-scale firms can also start gathering and monetizing personal information.

U.S. federal surveillance programs have also come under close scrutiny since the disclosures by Edward Snowden of the past year. It was argued by many that the federal government's PRISM may place personal freedom at risk. Many activists in the Middle East have, since the news first broke, claimed that widespread government surveillance – which could be supported by collection of fine-grained power consumption data – is the first step toward a police state and the repression of individuality. They further claim government may start with minimizing terrorism, but then gradually move toward fighting crime, and that such situations could quickly progress into the suppression of movements not aligned with the state. Comparisons of any such regime to Oscar Gandy's vision of the panopticon are not unfounded, and as such, the United States must ensure the right balance of security and privacy are achieved [17]. Cybersecurity and foreign affairs political experts will long debate whether programs like PRISM are right for the federal government, which was forced to increase national security efforts following the events of September 11, 2001.

The Obama Administration's Consumer Privacy Bill of Rights, threat to veto such bills as the Cybersecurity Information Sharing and Protection Act (CISPA), and release of the White House report on big data and privacy show the President is leading the country in the right direction by preserving the civil liberties of global citizens [18, 19]. Further, officials throughout government have taken note and acted upon risks associated with any potential use of combined data sets to derive insights about individual consumers that could be secondary to the original purpose of collection of individuals' data; the Open Data Policy published in 2013 advises that "before disclosing potential [personally identifiable information] or other potentially sensitive information, agencies must consider other publicly available data – in any medium and from any source – to determine whether some combination of existing data and the data intended to be publicly released could allow for the identification of an individual or pose another security concern" [20].

Nevertheless, the bulk collection of personally identifiable information – especially in the context of the power industry – and its use for government and corporate purposes is of great concern, especially as the advancement of data storage and analytical technologies continually outpaces the development of new policy to keep practices of data stewards in check. Analysis of power data can provide a view of private life not determinable through analysis of other types of readily available data, like browsing history or GPS-precise location data. For instance, it has been established that smart metering data can reveal when an individual is at home; when the microwave is used; when showers are taken; where in the home residents are; and whether they have guests visiting at certain times. However, granular data collection via smart metering systems is also necessary for the optimal routing of power, one of the key

features of the future smart grid. Is it possible to enable this degree of data collection, which is no doubt potentially beneficial to society, without placing consumer privacy at risk?

More often than not, achieving maximal functionality with the assurance of privacy is technically possible, but privacy-preserving solutions are not always desirable by all stakeholders in society. Therefore technical solutions alone are not enough to push the adoption of privacy in, for instance, power infrastructures. This thesis will tackle these difficult issues, starting first with a technical approach to preserving individual privacy in power infrastructures, and then discussing the economics of privacy adoption and determining policy mechanisms in the way forward that could support adoption of privacy technologies.

1.2 CONTRIBUTIONS

In this thesis, we will consider privacy in technology from a bottom-to-top perspective in which we will move from study to study, each time working closer toward understanding how the adoption of privacy-aware technologies can be achieved. First, we study raw, fine-grained high-resolution power consumption data associated with a typical household to understand the inferences that can be made about the residents of the household from their data. Second, we consider privacy-aware solutions to protect consumer data, and develop a privacy-aware framework for application to a novel power infrastructure, that of vehicle-to-grid technology. Third, we will investigate the stakeholder issues associated with the adoption of privacy in power infrastructures using game theory. Finally, we will seek a method to determine the

valuation of consumer privacy in the smart metering context to help shape policy around the adoption of privacy in smart metering.

To structure this thesis, chapter two begins by illustrating that an individual household's power consumption behaviors can be predicted using the application of jump diffusion processes. Chapter three seeks to address these issues of privacy by first outlining a methodology for the general design of privacy-aware design of networks, and then by applying this method to privacy-aware design of vehicle-to-grid systems. Chapter four applies a game theoretic approach to investigate the stakeholder issues associated with privacy adoption and endeavors to determine the societal and economic conditions required to assure its adoption is a Nash equilibrium outcome given the interests of all players, and chapter five outlines regulatory regimes that can be pursued to promote adoption policies. Chapters six and seven present results and analysis of a survey implemented to determine the value consumers place on protection of privacy in the smart metering and cellular data contexts, respectively. Finally, chapter eight provides concluding remarks.

1.3 REFERENCES

[1] E. Bucy, "Social Access to the Internet," *The Harvard International Journal of Press/Politics*, vol. 5, no. 1, January 2000.

[2] B. Burlingham and G. Gendron, "The Entrepreneur of the Decade: An interview with Steven Jobs, Inc.'s Entrepreneur of the Decade," *Inc. Magazine*, April 1989.

[3] Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values*, May 2014.

[4] J. Gantz and D. Reinsel, *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East*, International Data Corporation, December 2012.

[5] S. LaValle, R. Lesser, R. Shockley, M. Hopkins and N. Kruschwitz, "Big Data, Analytics and the Path From Insights to Value," *MIT Sloan Management Review*, vol. 52, no. 2, 2011.

[6] J. Stanley, "Eight Problems With "Big Data," *American Civil Liberties Union*, April 2012.

[7] T. Davenport and D.J. Patil, "Data Scientist: The Sexiest Job of the 21st Century," *Harvard Business Review*, October 2012.

[8] Exec. Order No. 13642, *Making Open and Machine Readable the New Default for Government Information*, 3 C.F.R., May 2013.

[9] G. Ferenstein, "Obama's chief tech officer: Let's unleash ingenuity of the public," *CNN*, June 2012.

[10] A. Lesser, *How energy data will impact the smart grid*, GigaOM Pro, March 2013.

[11] A. Freifeld, *Innovative Tools to Increase Efficiency and Create New Energy Revenues: The Diversity of New DG and DR Business Models*, Viridity Energy, 2012.

[12] J. St. John, "Stem Pulls Behind-the-Meter Storage Into an Aggregated Demand Response Resource," *Greentech Media*, June 2014.

[13] IBM Corporation, *Managing big data for smart grids and smart meters*, May 2012.

[14] K. Temple, "Chip Shot: Intel Powers Big Data Analytics at New Energy Lab," *Intel Newsroom*, June 2013.

[15] General Electric Company, *The Case for an Industrial Big Data Platform: Laying the Groundwork for the New Industrial Age*, 2013.

[16] N. Negroponte, "Being Anonymous," *Wired*, Iss. 6.10, October 1998.

[17] O. Gandy, *The Panoptic Sort: A Political Economy of Personal Information (Critical Studies in Communication and in the Cultural Industries)*, Westview Press, January 1993.

[18] The White House, *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*, February 2012.

[19] Executive Office of the President, *Statement of Administration Policy: H.R. 624 – Cyber Intelligence Sharing and Protection Act*, April 2013.

[20] Executive Office of the President, *Open Data Policy*, May 2013.

CHAPTER 2.

THE PRIVACY RISKS OF GRANULAR DATA COLLECTION

The widespread use of data mining techniques throughout the power industry increases the reliability and security of U.S. power systems, but also can place consumer privacy at risk. Smart meters in particular have recently been highlighted in the literature and media as a technology that collects highly sensitive information associated with individual consumers. To arrive at a more precise understanding of the privacy risks of smart metering, this chapter examines temporally precise, fine-granularity power consumption data collected from a typical household. We first note that basic statistical analysis can be conducted to show that basic behavioral patterns can retroactively be determined from historical power consumption data. Furthermore, we motivate the use of a jump-diffusion model to encapsulate historical observations as a stochastic process. Maximum likelihood estimation is then used to compute optimal parameters for the jump-diffusion model. The resulting model is found to be meaningful and can be used to model and predict residential power consumption, thereby furthering the privacy concerns of collecting fine-grained residential power consumption data.

2.1 INTRODUCTION

Smart metering technology has the capability to revolutionize the energy industry. The installation of advanced metering infrastructures at distributed consumer locations will enable two-way communications between consumers and their regional utility companies. This, in turn, will allow for a demand response network that will potentially levelize the aggregate power demand curve and thereby greatly lower the production cost of electricity.

But the future of smart metering isn't as bright as it should be. Several public backlashes against the technology have recently made the news and have led to the suspension or early termination of smart metering initiatives throughout the United States and Canada. In January 2013, BC Hydro and Power Authority sent out 85,000 letters to their electricity customers stating they would replace older analog residential models with smart meters. Outcry ensued, forcing the company to quickly retract the letters and postpone installations indefinitely [1]. Meanwhile, in Maine, a state in which most commercial and residential buildings have already been fitted with advanced metering systems, homeowners have taken their grievances to the Maine Supreme Court to oppose Central Maine Power's smart metering initiative [2]. In Lubbock, Texas, local residents have raised concerns about the use of smart metering in residential locations, calling on the city's leaders to reconsider moving forward in negotiations with Lubbock Light & Power over planned installations of smart metering [3].

While various reasons for such backlashes have been cited, one lies at the heart of them all – privacy. Numerous studies have recently been conducted to show that there are significant privacy concerns associated with smart metering. This has not gone unnoticed by consumers, watchdog groups, and industry. The issue has been a highly frustrating one for many electric

utility companies like BC Hydro, who have invested hundreds of millions of dollars into smart metering programs. Installation of smart meters will afford utilities great advantages in the optimization of power flow through the system, but it seems unlikely that smart metering will gain ground in many regions until these privacy concerns are addressed by regional public utility companies. Until smart metering becomes a reliable source of information for utility companies to better optimize the transmission of power, the smart grid will remain far from achievable.

In this chapter, we study the type of data collected by utility companies via smart metering modules to enable demand response programs. This data comes in the form of temporally precise power consumption readings for a household. Readings of watts of power consumption are recorded in fifteen-second intervals for the household. Most notably, spikes in the dataset occur for time intervals that are associated with special events within the household, including the use of a microwave or the start of a cooling cycle of the refrigerator, suggesting that analysis of the power consumption curve can lead to meaningful inferences about behaviors and routines within the household – information that is sensitive and should remain private to the residents. This leads us to analyze the power consumption curve as a Poisson jump process, an application that can allow us to predict resident behavior in the household. By showing that behavior can be predicted, we illustrate that the collection of residential power consumption data presents a unique privacy risk that has not yet been discussed in the literature – the ability to predict a homeowner's presence, location, and activity within the home.

2.2 LITERATURE REVIEW

The collection of a household's fine-grained power consumption data presents a consumer privacy risk. It has been shown by Lisovich et. al. and others that, given access to a household's power consumption data, various inferences can be made about the consumer's activities within the home [4]. This is because of the unique nature of different current-drawing household appliances. For instance, refrigerators, microwaves, and televisions all have unique power consumption signatures, and using fine-grained consumption data, it can easily be determined when a consumer is using one of these appliances. In many cases, the power signatures can yield very detailed information, including the model or age of the appliance. Furthermore, analysis of fine-grained power consumption data can help determine behavior within the household. Information including a resident's presence at home, the timing of power-consuming events (such as turning on the television set or using the microwave), the types of appliances used, the resident's time-specific location in your home, and the presence of visitors in the home can all be detected through analysis of data collected by smart meters. Using this information, a profile chronicling the resident's unique lifestyle, behavior, preferences, interests, and beliefs can be composed.

Such information can readily be monetized by entities like utility companies. Potential uses of this information about consumers abound. Some privacy-infringing applications include the following.

- *Targeted advertising and marketing*. Different individuals respond to different types of advertising, in terms of both content and method. Advertising agencies continually try to glean information about individual consumers so that they

identify consumers who are most likely to respond positively to particular types of advertising. Information about a consumer's activity within the household is very revealing in this respect and therefore highly valuable in the advertising industry [5].

- *Law enforcement*. Police organizations presently use power consumption data recorded by simple analog metering devices to identify potential growers of in-house marijuana. As discussed earlier, smart metering devices can be used to infer a far greater amount of information about consumers than analog meters, and it is likely that police organizations would be able to use smart metering data to identify criminals in new ways. While it is important for law enforcement agencies to be able to apprehend criminals efficiently, it is critical that the police are not given access to too much information. Police require search warrants to search a person or private location because the state protects the right to privacy of the citizen. This right may be infringed upon with the expanded amount of information that police would have access to via data from smart meters. This issue mirrors a recent case in which police attached a GPS-tracking device on a suspected drug dealer's vehicle without his knowledge, and then used data collected using the device in court proceedings. The judge duly threw the evidence out as it was deemed a search warrant was required for this data to be collected, and the conviction of the drug dealer was reversed due to the breach of his right to privacy [6].

Clearly, fine-grained power consumption data holds value to various entities. Processing historical data allows collectors of the consumption data to make inferences about the potentially sensitive lifestyle of electricity consumers.

Going one step further, it has not yet been shown that consumer behavior within the household can also be predicted using fine-grained power consumption data. Prediction of a consumer's behavior can be of great interest to certain entities. Consider a residence fitted with a smart meter, and imagine that power consumption data is collected in real-time at a central data warehouse, perhaps at the site of the utility company. If that data were processed to learn what the consumer's next action will be, then the consumer may be posed with significant new privacy and security risks.

For instance, imagine a criminal group that either has access to or has hacked into the smart metering modules in a neighborhood. (Such cyber-attacks on data have been shown to be possible either wirelessly or physically [7]). This criminal group can then process the fine-grained data to identify households that own certain items of interest, such as expensive power-consuming appliances in the household. Households that have such items can be designated by criminals as potential targets for burglary. Criminals would be further empowered to break and enter a residence if they knew whether the homeowners are at home or not. This necessitates the criminal's ability to predict the presence of consumers at home. If this were possible using fine-grained data, then criminals could potentially execute sophisticated burglary attacks on households identified as prime targets.

The power to predict a consumer's actions can be put to use in various other ways that would in the end be harmful to the consumer. Our goal in this chapter is thus to show that electricity

consumers are vulnerable to such attacks by illustrating that prediction of household behavior is possible given fine-grained consumption data.

To that end, we will employ a Poisson stochastic process to encapsulate sample consumption data in an ARCH-based predictive model. This basis for the model is predicated by the time-varying volatility illustrated in the data. Further examination of the data suggests that certain events within the household can induce spikes, or jumps, in the fine-grained power consumption curve (Fig. 1). This in turn leads us to introduce an additional term to the model involving a jump process.

Stochastic processes are used in an array of financial applications to model the price movement of various instruments. The Black-Scholes model, for example, prices options using a traditional diffusion process, by which changes in the price are modeled as relatively small and random movements [8]. Later, Cox and Ross proposed that options do not adhere to a pure diffusion process, but instead continually exhibit "jumps," or large price movements over short time intervals. This view is formalized in the Cox-Ross-Rubinstein binomial options pricing model [9]. Merton subsequently combined the two approaches to options pricing and argued that routine releases of information correspond to traditional random diffusion behavior, while sudden bursts of high-impact information correspond to jumps in prices [10]. Further, Kou provided a novel double exponential jump-diffusion model to specifically incorporate the effects of leptokurtosis and a common feature known as the volatility smile, by which options in the high- and low-moneyness regimes exhibit higher volatility than at-the-money options [11]. Meanwhile, various others have used models adapted from Merton's and provided empirical results on the impact of information release on market prices. For example, Balduzzi

et. al. illustrated that eight different types of economic announcements have a significant impact on prices in the United States treasury market [12]. Similarly, Das showed that jumps in the federal funds interest rate are induced by intervention from the Federal Reserve, and Jorion showed comparable effects in forex markets. Both studies thus employed a jump-diffusion model to mitigate the non-linearity effects in price drifts [13, 14].

In this chapter, we will follow on these studies and borrow the jump-diffusion model from pricing theory to model residential power consumption. This is motivated largely by the fact that residential power consumption typically exhibits random diffusion behavior, but sudden jumps occur because of occasional increased power draw from appliances in the home. Using this methodology, we the will predict the path of the power consumption curve to show that behavior in the household can also be predicted with confidence.

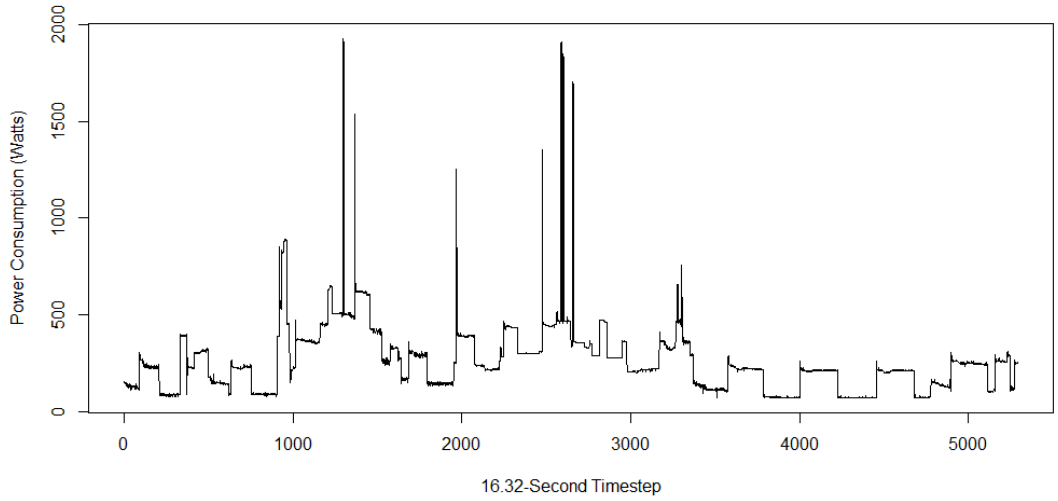

Figure 1: A plot of raw fine-grained power consumption data.

2.3 THE DATA

Advanced metering systems are used to collect temporally precise, fine-grained power consumption data. Lisovich et. al. conducted an experiment in which such temporally precise data was collected for statistical analysis, and we use their dataset in this chapter. To replicate the data collected by smart meters, Lisovich et. al. fitted a Brultech EML energy usage monitor to the breaker panel of a graduate student's residence in Ithaca, New York. Total power consumption level was recorded in 16.32-second time intervals with a 1-watt resolution and sent wirelessly to a data server. The experiment was conducted over a two-week period. In this chapter, a single day's data (amounting to 5294 observations) is henceforth used for simplicity in empirical analysis. The data corresponds to a weekday in December in Ithaca during which heating, refrigeration, television, microwave, other and daily appliances were used.

An initial survey of the data (depicted in Figs. 2 and 3) reveals the appropriateness of employing a jump-diffusion model (Table 1). Three key features lead to this assertion.

1) *Higher moment behavior*. This sample of power consumption data exhibits strong kurtosis and skewness characteristics. In particular, the data is highly leptokurtic, suggesting that a jump model may be appropriate to capture this behavior.

2) *Volatility*. Some volatility is exhibited in the short term. We will apply ARCH models to accommodate this feature of the data.

3) *Mean reversion and autocorrelation*. The data exhibits both effects to some degree, likely due partly to the physical nature of electricity.

Feature (1) shall be treated separately from (2) and (3) in our model since jumps are caused by the sudden draw of electricity from a device.

Table 1: Basic statistics on raw power consumption data.

	Power consumption (watts)	Power consumption (differenced)
Mean	262.6219	0.01833
Median	229	0
Variance	32332.77	5005.742
Standard deviation	179.8132	70.75127
Skewness	3.863219	0.664583
Kurtosis (excess)	29.00798	304.972
Minimum	72	-1438
Maximum	1929	1418

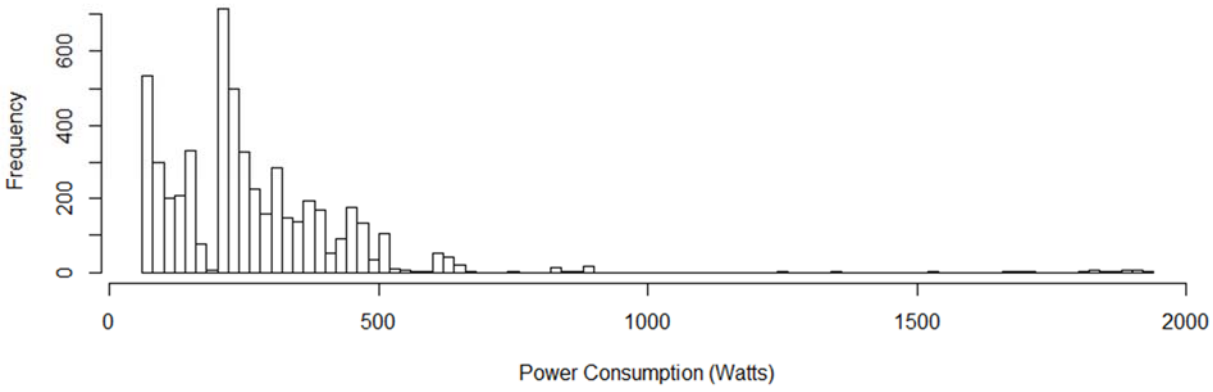

Figure 2: Histogram of power consumption data.

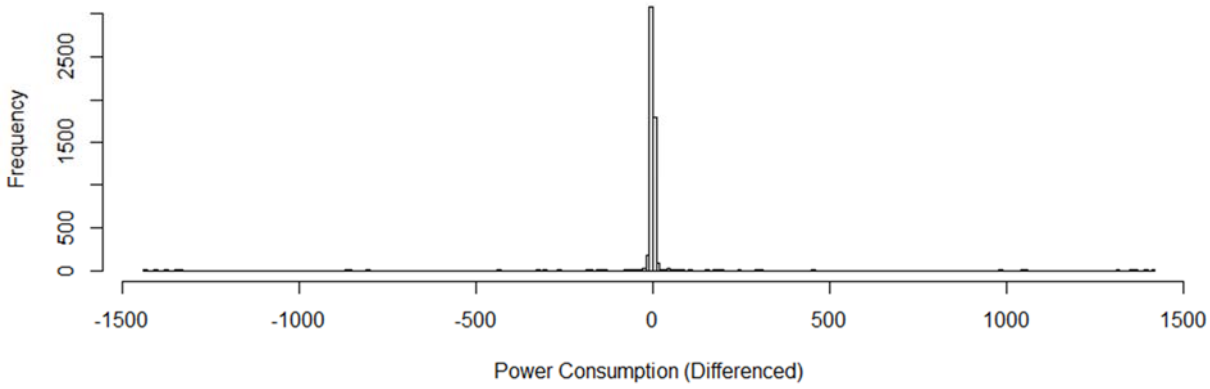

Figure 3: Histogram of differenced power consumption data.

2.4 JUMP DIFFUSION MODEL

The econometric specification of our model encompasses the three sample features discussed in section 2.3. Note the presence of mean reversion requires that the probability function for the stochastic function is dependent on both timing and magnitude of the jumps, as opposed to only the magnitude.

The basic stochastic model we employ for temporally precise power consumption is the following:

$$dy = k\,(\theta - y\,) + v\,dz + J\,d\pi\,(h)\,.$$

Here, y is the level of power consumption in watts. The first term in the expression captures the effect of the mean reversion drift, where theta is the long-run equilibrium level of power consumption and k is the rate at which reversion occurs. The last two terms of the expression encapsulate the randomness of the power consumption process. The first of these is a standard

diffusion process for which the variance coefficient is v^2. The second of them models the jump feature, with jumps arriving with magnitude J as per a Poisson process. The frequency of jumps is specified by h, and the magnitude of the jumps J can either be constant or a random variable with an independent probability distribution. Note that the diffusion and Poisson jump processes are independent. The discretized form of this function will be used for estimation and subsequent analysis.

2.4.1 The characteristic function of the stochastic model

Furthermore, we shall make use of the characteristic function of the basic stochastic model given above as well as its moments later in the chapter as a means to check the diagnostics of our estimated jump-diffusion model, specifically the variance, skewness, and kurtosis. The expression for the characteristic function is given below. Here, the characteristic function $F(y, T; s)$ for our stochastic model has characteristic function parameter s. To determine the expression for $F(y, T; s)$ we are required to solve the Kolmogorov backward equation subject to the condition that $F(y, t = 0; s) = \exp(iys)$. This computation is not repeated here but comprehensive details can be found in Duffie et. al [15].

$$F(y, T; s) = \exp\left[A(T; s) + y\,B(T; s)\right]$$

$$A(T; s) = \int \left[k\theta\, B(T; s) + \frac{1}{2} v^2\, B(T; s)^2 + h\, E\left[\exp\left(J\, B(T; s)\right) - 1\right] \right] dT$$

$$B(T; s) = is \exp(-kT)$$

2.4.2 The moments of the stochastic model

The moments can be determined by successively computing the derivative of the characteristic function F (y, T; s) with respect to the function parameter s and evaluating each result at $s = 0$. Here, μ_n denotes the n^{th} moment. The expressions for the first three moments are provided here but comprehensive derivations for similar computations can be found in Singleton [16] and Das [13].

$$\mu_1 = \left(\theta + \frac{h\,E[J]}{k}\right)(1 - \exp(-kT)) + r \exp(-kT)$$

$$\mu_2 = \frac{v^2 + h\,E[J^2]}{2k}(1 - \exp(-2kT)) + {\mu_1}^2$$

$$\mu_3 = h\,E[J^3]\left(\frac{1 - \exp(-3kT)}{3k}\right) + 3\mu_1\left(v^2 + h\,E[J^2]\right)\left(\frac{1 - \exp(-2kT)}{2k}\right) + {\mu_1}^3$$

2.5 ESTIMATION

Perhaps the first example of estimation applied to jump-diffusion processes was developed by Naik and Lee [17]. In their work, the specific application is to continuous changes in interest rates. They show that fluctuations in the business cycle of the aggregate economy in

combination with subsequent monetary policy actions impact not only the short term interest rate, but also the underlying term structure itself. Various others expanded on this work in later years. Notably, Chacko and Viceira [18] and Singleton [16] have developed characteristic functions and methods of estimation for jump-diffusion models in the context of interest rates. We will employ similar discrete-time methods for estimation given the data on fine-grained power consumption data described in the previous section. Descriptive characteristics for the sample, given in Table 1, indicate a high degree of kurtosis, validating the use of a jump-diffusion model.

We employ a discrete-time approach developed by Ball and Torous [19] for estimation of the model with normally distributed jumps. In that work, the authors make the assumption that in each time interval, there may be either one or zero jumps. This assumption holds for the data used in this study, as it has been collected with high frequency, though if the frequency were lower, this approach may be disputable. Note also many have noted that estimation bias arises for models in which the continuous time stochastic differential equation developed in the previous section for estimation is discretized. However, it is established that this is negligible for high-frequency data. Further techniques on the minimization of estimation bias are discussed in many works including that of Phillips and Yu [20].

The discretized form of the jump-diffusion process can be expressed as:

$$\Delta y = k\,(\theta - y)\,\Delta t + v\,\Delta z + J\,(\mu, \gamma^2)\,\Delta\pi\,(q)\,.$$

Here, Δz is a standard normal shock term. Corresponding to the continuous-time expression, v^2 is the variance of the Gaussian shock. Meanwhile, in the last term, $J\,(\mu, \gamma^2)$ is the jump

shock. The jump shock follows a normal distribution, and carries mean μ and variance γ^2. Further, $\Delta\pi(q)$ is the discrete-time Poisson increment and is approximated using a Bernoulli distribution with parameter $h\Delta t + O(\Delta t)$. The transition probabilities for power consumption level can then be expressed in the transition density function:

$$f\left[y(s) \mid y(t)\right] = q \ \exp\left[\frac{-\left(y(s) - y(t) - k\left(\theta - y(t)\right)\Delta t - \mu\right)^2}{2(v_t^2\,\Delta t + \gamma^2)}\right]\left(\frac{1}{\sqrt{2\pi\,(v_t^2\,\Delta t + \gamma^2)}}\right)$$

$$+ (1 - q) \ \exp\left[\frac{-\left(y(s) - y(t) - k\left(\theta - y(t)\right)\Delta t\right)^2}{2v_t^2\,\Delta t}\right]\left(\frac{1}{\sqrt{2\pi\,v_t^2\,\Delta t}}\right).$$

For estimation, we use maximum likelihood to maximize L given the sample consisting of T observations, where

$$L = \prod_{t=1}^{T} f\ \left[y\ (t + \Delta t) \mid y(t)\right].$$

Taking the logarithmic form, we reformulate the objective function as:

$$\max_{\Omega = [\,k,\,\theta,\,v,\,\mu,\,\gamma^2,\,q]} \sum_{t=1}^{T} \left(\log\ \left(f\left(r\ (t + \Delta t) \mid r\ (t)\right]\right)\right).$$

Constraints for this maximization problem as noted in Cramer's technical conditions include 1) that the weights for the jump and non-jump regimes sum to one, and 2) that $0 \leq q \leq 1$, as q

is the probability that a jump occurs in a particular time interval. The first of these is already satisfied in the formulations above.

Estimation was computed using the gradient optimization method in R for statistical programming. Additionally, results were confirmed using the OPL modeling language and the ILOG CPLEX optimization suite. Code for both of these applications is provided in the chapter appendix. Optimization results are given in Table 2.

Table 2: Results of the maximum likelihood estimation.

Parameter	Estimated value
k	1.0254E-03
θ	1.1973E+02
v	3.3094E+00
μ	-2.8597E+02
γ	4.1441E+02
q	1.0117E-01
Log-likelihood	16393.78

The estimation procedure was carried out in much the same way as in studies by Jorion [14] and Das [13] on foreign exchange markets and interest rates respectively.

2.6 APPLYING THE MODEL

Notably, the model estimates the value of q at 0.1012. Under the Bernoulli model, this is simply the probability of a jump occurring in a particular time step. This indicates that a jump in power consumption occurs in one in every ten 16-second time intervals. That is, in this

particular graduate housing unit, students used some power-consuming appliance that caused a draw on power not captured by the stochastic volatility component of our model under this estimation method.

To determine how well the model fits the data, we proceed to calculate the conditional variance, skewness, and kurtosis. Here we use the expressions for the moments calculated earlier. Denoting the time interval between observations as T, we calculate conditional variance as

$$\mu_2 - \mu_1^2 = \frac{v^2 + hE\left[J^2\right]}{2k}\left(1 - \exp\left(-2kT\right)\right).$$

Skewness is computed as:

$$\frac{E\left(J - \mu_1\right)^3}{\left(\mu_2 - \mu_1^2\right)^{3/2}} = \frac{2\sqrt{2k}\exp(-kT)\left(1+\exp(kT)+\exp(2kT)\right)hE(J^3)}{3\left(1+\exp\left(kT\right)\left(v^2+hE\left(J^2\right)\right)\sqrt{\left(1-\exp\left(-2kT\right)\right)\left(v^2+hE\left(J^2\right)\right)}}$$

and the kurtosis is:

$$\frac{E\left(J - \mu_1\right)^4}{\left(\mu_2 - \mu_1^2\right)^2} =$$

$$\frac{\left(\exp\left(2kT\right)-1\right)\left(3h^2\,E\left(J^2\right)^2 + 6hv^2\,E\left(J^2\right) + 3v^4\right) + khE\left(J^4\right)\left(\exp\left(2k\,T+1\right)\right.}{\left(\exp\left(2kT\right)-1\right)\left(v^2+hE\left(J^2\right)\right)^2}.$$

Note that if $h = 0$, signifying the absence of jumps, the kurtosis is 3 as expected. As there are 5294 observations recorded for a single day, we use $T = 1/5294$ as the horizon. Further, given the jump probability distribution parameters μ and γ^2, we can compute the values:

$$E(J) = \mu,$$

$$E(J^2) = \mu + \gamma^2,$$

$$E(J^3) = \mu + 3\mu\gamma^2,$$

$$E(J^4) = \mu^4 + 6\mu^2\gamma^2 + 3\gamma^4.$$

The values of the three diagnostic statistics, given in Table 3, show that the model corresponds reasonably well to the data.

Table 3: Jump-diffusion model statistics.

	Model	Raw data
Conditional variance	58701.34	32332.77
Skewness	2.571	3.863
Kurtosis	14.911	29.008

As a further test of the model, a time series plot was generated using R (Fig. 4). The code used to generate this plot appears in the chapter appendix.

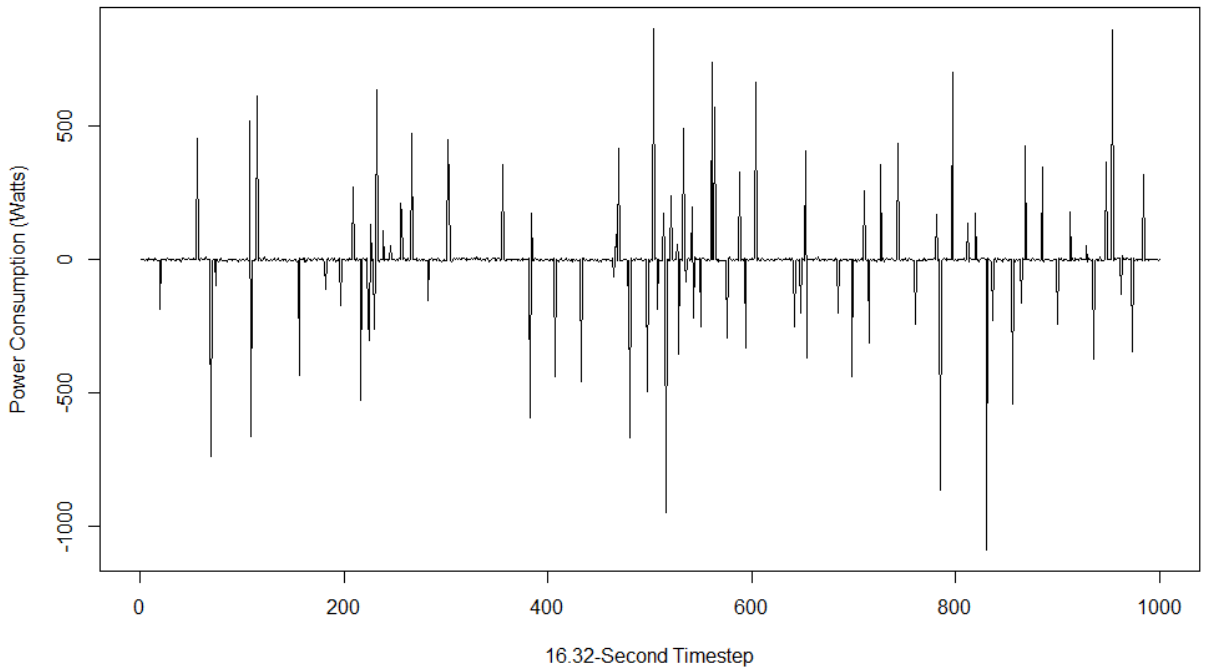

Figure 4: Time series plot of the jump-diffusion process over 1000 time steps.

Clearly, the plot of the model has the characteristics of a typical jump-diffusion model. There are two noteworthy items about this plot.

Firstly, the nature of the jumps in the model seems to fit the data quite well. In particular, the frequency of jumps is predicted well, thereby validating our use of the Poisson process to model the arrival of jumps. Given the accuracy of the jump arrival process as well as the meaningful diagnostic statistics found earlier in this section, this seems to suggest that a version of our model can effectively be used to model and predict consumer behavior within the household based on their power consumption patterns. Furthermore, the spread of the magnitude of jumps occurs in the approximate range from 100 watts to 1000 watts. The raw data appears to replicate this to some extent, although several jumps of approximately 1300 watts also occur

in the raw data. This suggests that a normal distribution may not be the most accurate for modeling jumps in application to the power consumption data associated with our example residence. It should also be noted, though, that the distribution of the magnitude of jumps in power consumption data may vary by residence. This would add a more random element on jump magnitude and predicate the use of a normal distribution more positively.

A second item of note is the mean reversion behavior of the process. The raw data shows a cyclical tendency toward the mean, no doubt caused by refrigeration cooling cycles. This cyclical behavior in the mean has not been overlaid here on the model, but as it is independent of jump behavior, it would not significantly affect our analysis of consumer behavior within the residence. Meanwhile, the model clearly exhibits strong mean reverting behavior.

2.7 DISCUSSION

We have shown that power consumption data can be used to develop a reasonable model of consumer behavior, which can potentially be predicted by a sophisticated attacker. We also described various ways in which this information can be exploited by third parties to the harm of the consumer and profit of the public utility company. Power consumption data can readily be monetized.

Various improvements can potentially be made to our model to better simulate the use of residential. As mentioned earlier, the jump arrival process seems to cover the random arrival of jumps in the raw data quite well. That said, jump magnitude may be better represented by something other than the normal distribution we employed. These factors may also vary by

household. Furthermore, the use of a cyclical overlay on the model may be predicated given the presence of a cooling pattern in residential refrigeration units.

In the future, more research needs to be conducted in the way of analyzing residential power consumption data so that we can fully understand the inferences that can be made about consumer behavior in the household and thereby comprehensively understand the privacy risks associated with smart metering.

2.8 CONCLUSIONS

We will conclude by noting that consumer privacy is critical for smart metering applications. A lack of privacy in this space facilitates monetization of consumer information and places individual customers at risk of attack. The technical design of smart meters and their related smart grid architectures must therefore be developed with consumer privacy in mind. In this light, Rial and Danezis have developed a design for a privacy-preserving smart meter – one that can provide all the technical functionality of a standard smart meter, but also protect the privacy of the consumer [21]. It makes use of such tools as cryptography and trusted computing; by scrambling data using a cryptographic process known as public key infrastructure, sending it via a wireless area network, and then decrypting and aggregating it at a secure data processing point, data can be used for the benefit of the smart grid while protecting consumer privacy. However, it is still only a proof of concept and requires testing. Further, the seemingly impossible barrier of industry consensus on private smart metering is far from

overcome. This calls for more work studying the regulatory policy surrounding the electricity industry, particularly with respect to smart metering.

2.9 REFERENCES

[1] "BC Hydro backs down on smart meter installation," *CBC News*, January 2013.

[2] A. Regalado, "Rage Against the Smart Meter," *MIT Technology Review*, April 2013.

[3] A. Young, "Committee to explore smart meter option for Lubbock," *Lubbock Avalanche-Journal*, February 2013.

[4] P. McDaniel and S. McLaughlin, "Security and Privacy Challenges in the Smart Grid," *IEEE Security & Privacy*, Vol. 7, Iss. 3, May 2009.

[5] K. Chauvin and M. Hirschey, "Advertising, R&D Expenditures and the Market Value of the Firm," *Financial Management*, Vol. 22, No. 4, 1993.

[6] A. de Vogue, "GPS Tracking Requires Warrant, Supreme Court Rules," *ABC News*, January 2012.

[7] J. Lerner and D. Mulligan, "Taking the 'Long View' on the Fourth Amendment: Stored Records and the Sanctity of the Home," *Stanford Technology Law Review*, Vol. 3, February 2008.

[8] F. Black and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, Vol. 81, No. 3, June 1973.

[9] J. Cox, S. Ross, and M. Rubinstein, "Option pricing: A simplified approach," *Journal of Financial Economics*, Vol. 7, Iss. 3, September 1979.

[10] R. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, Vol. 3, Iss. 1-2, January-March 1976.

[11] S. Kou, "A Jump-Diffusion Model for Option Pricing," *Management Science*, Vol. 48, No. 8, August 2002.

[12] P. Balduzzi, E. Elton and T. Green, "Economic News and Bond Prices: Evidence from the U.S. Treasury Market," *Journal of Financial and Quantitative Analysis*, Vol. 36, Iss. 4, December 2001.

[13] S. Das, "The surprise element: jumps in interest rates," *Journal of Econometrics*, Vol. 106, Iss. 1, January 2002.

[14] P. Jorion, "On Jump Processes in the Foreign Exchange and Stock Markets," *The Review of Financial Studies*, Vol. 1, No. 4, 1988.

[15] D. Duffie and K. Singleton, "Modeling term structures of defaultable bonds," *Review of Financial Studies*, 12 (4), pp. 687-720, 1990.

[16] K. Singleton, "Estimation of affine asset pricing models using the empirical characteristic function," *Journal of Econometrics*, Vol. 102, Iss. 1, May 2001.

[17] V. Naik and M. Lee, "The Yield Curve and Bond Option Prices with Discrete Shifts in Economic Regimes," *Social Science Research Network*, September 1994.

[18] G. Chacko and L. Viceira, "Spectral GMM estimation of continuous-time processes," *Journal of Econometrics*, Vol. 116, Iss. 1-2, September-October 2003.

[19] C. Ball and W. Torous, "A Simplified Jump Process for Common Stock Returns," *Journal of Financial and Quantitative Analysis*, Vol. 18, Iss. 1, March 1983.

[20] P. Phillips and J. Yu, "Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance," *Handbook of Financial Time Series*, pp. 497-530, 2009.

[21] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proceedings of the 10th ACM Workshop on Privacy in the Electronic Society*, pp. 49-60, October 2011.

2.10 APPENDIX

The code used to produce the maximum likelihood estimation for the stochastic model as well as its corresponding time series plot is provided here.

Estimation

Estimation was executed in R using the gradient method and subsequently confirmed using the ILOG CPLEX OPL optimization package.

In R:

```
dat = read.csv("/project-data.csv")
y = dat$Total.W
yd = diff(dat$Total.W)

llikjump = function(omega, y)

{

  k = omega[1]
  theta = omega[2]
  v = omega[3]
  mu = omega[4]
  gamma = omega[5]
  q = omega[6]
  n = length(y)

  llik = 1

  for(t in 2:n){

    llik = llik + -log(q*exp(-(y[t] - y[t-1] - k*(theta -
    y[t-1]) - mu)^2 / (2*v^2 + 2*gamma^2)) /
    sqrt(2*pi*(v^2+gamma^2)) + (1-q)*exp(-(y[t] - y[t-1] -
    k*(theta - y[t-1]))^2 / (2*v^2)) / sqrt(2*pi*(v^2)))

    }

  - llik

}

out = optim(par =c(1,250,110,100,179,.2), fn = llikjump,
    y=y,control = list(trace = 0, fnscale = -1))
out
```

In CPLEX:

```
int nTechs = 5294;
range techs = 2..nTechs;

int ys [1..nTechs] =  ...;

dvar float q in 0..1;
dvar float k;
dvar float theta;
dvar float v;
dvar float gamma;
dvar float mu;

maximize
  sum (t in techs) (log(q*exp(-(ys[t] - ys[t-1] - k*(theta -
    ys[t-1])   -   mu)^2   /   (2*v^2   +   2*gamma^2))   /
    sqrt(2*3.1415*(v^2+gamma^2)) + (1-q)*exp(-(ys[t] - ys[t-
    1]   -   k*(theta   -   ys[t-1]))^2   /   (2*v^2))   /
    sqrt(2*3.1415*(v^2))));
```

CHAPTER 3.

PRIVACY-AWARE NETWORKS

In Chapter 2, we found that fine-granularity power consumption data can be used to both make inferences about as well as predict consumer behavior in the home. Given that household activity is sensitive information, smart meters can clearly leave consumer privacy at risk. The next question that arises is, how can we provide a framework for the design of the privacy-aware technology – that is, technology that protects consumer privacy by default – to factor out such privacy risks in emerging power system technologies?

While smart metering is a prime example of an emerging technology in the power industry that will make use of high-resolution customer data, it is not the only one that has raised concerns for privacy advocates. In this chapter, we will explore the issue of privacy-aware design of vehicle-to-grid (V2G) technology which, like smart metering systems, requires the collection of fine-grained power data to provide technical functionality. We open the chapter by introducing a set of privacy-aware design methods developed and based upon the Fair Information Practices. Subsequently, we apply these design guidelines to V2G technology to illustrate that privacy-aware solutions can technically be achieved through relatively simple means for upcoming smart grid applications.

3.1 PRIVACY-AWARE DESIGN METHODS

In designing privacy-aware systems, we require model guidance to ensure proper consideration for any privacy risks at a technological level. Wicker and Schrader have developed a framework for the design of information networks that protects the privacy of consumer data [1]. In this section, we provide an overview of their framework and apply it to the specific example of V2G technology, outlining the actions that should be taken by the V2G service provider to ensure consumer privacy.

3.1.1 Provide full disclosure of data collection

- Description requirement: Provision of an adequate explanation of the type of data that will be collected must be required; this will help close the information gap between V2G service provider and consumer, and allow consumers a view into the purpose for collection of their data and the privacy and security practices the service provider will pursue to protect its consumer base.

- Enforceability requirement: There must exist an enforced threat of punishment for inadequate descriptions of the data to be collected. This helps to keep service providers in check. Any punishment assigned to service providers for misuses or deficient security protocols on user data should be adequately significant so as to induce service providers to adopt practices that better protect the interests of V2G consumers.

- Irrevocability requirement: Service providers must communicate to the customer how collected data will be treated while retained. V2G data may be collected for various reasons, whether billing, planning, or otherwise. If any of the data is retained, the service provider must be required to justify its retention to consumers, outlining the requirement for the data to assure technical functionality.

- Intelligibility requirement: Service providers must be required to assure the customer has a clear understanding of the terms of data collection. Electricity customers are often not well-informed that their data could be collected and processed, and others may not even know about any relevant privacy concerns. Electricity companies and V2G service providers should be required to explain to consumers via a public platform and through individual communications in simple terms.

3.1.2 Require consent to data collection

- Opt-in requirement: Consumer information should not be collected unless customers opt in to the data collection program. This differs from opt out programs in which customer information is collected by default. Many have noted the stark difference between implementation of opt-in versus opt-out consumer choice mechanisms; generally they yield very different results in adoption of new technology because consumers lack information or do not wish

to make efforts to opt out of new programs they are assigned to by default. Consumers should not have this additional hassle, though; to protect privacy to the utmost, consumers should be required to opt in to new data programs. To increase opt in rates, service providers can engage consumers in informative messaging strategies.

- Acknowledgement requirement and the opt-in requirement: The customer should be required to acknowledge their data is being collected or processed before it is accessed by V2G service providers. This is in addition to opt-in programs, so that whenever service providers wish to collect new types of information or process data in new ways under their data collection programs, consumers are informed. This is akin to enforcing a requirement that service providers inform consumers about changes to the terms of service.

3.1.3 Minimize collection of personal data

- Functional requirement for collection: It must be established that data collection or processing is a technical necessity for functionality of the service. This could be enforced through a certification program led by technical staff at a government or regulatory agency. This requirement would ensure consumer data is only collected for reasons related to the technology, and that no data is collected strictly for the purpose of advancing misuses of data on the part of the service provider, in for instance marketing or targeted advertising campaigns.

- Distributed processing requirement: When fine-grained consumer data requires processing for the technical functionality of the V2G system, it should be done locally (i.e. at the site of the customer's electric vehicle) as opposed to centrally at the site of the service provider to whatever extent possible. This prevents unnecessary bulk collection of V2G data. As we will describe later, this can be accomplished through use of a set of cryptographic processes on an electronic chip installed in the charging modules of electric vehicles.

3.1.4 Minimize identification of data with individual consumers

- Non-attribution requirement: Service providers should attempt to anonymize potentially identifying data wherever possible. In some cases, identifying data is required; for instance, service providers will need to know the periodic bills attributed to individual consumers to provide a sound and accurate billing mechanism. However, in many cases, identifying data is not required to tag bulk collection of V2G data. Planning capabilities enabled through collection of fine-grained consumer data do not necessarily require identification of the data with individual consumers, as we will discuss later in this chapter. Note, in some cases, fine-grained data itself can be used to identify individuals. For example, GPS data, though not tagged with an identity, can clearly be associated with individuals through inference of home and work addresses. Similar examples would exist with fine-grained V2G data, which can be used to infer detailed

transportation profiles. It is critical engineers and policymakers are wary of this nuance and anonymize such fine-grained data to the extent possible.

- Separate storage requirement: Financial and functional data should be stored in separate places such that customer identification details cannot be associated with use of V2G service. This serves as a de-aggregation step that can help dissociate identifying data with fine-grained data that can be used to infer specific interests, behaviors and preferences of individual consumers.

3.1.5 Minimize and secure retained data

- Functional requirement for retention: Just as service providers should be required to establish a necessity for data collection and processing in the way of providing greater system functionality, they should be required to establish the necessity for any period of retention of the data. Retention should raise immediate flags for consumer and policymakers, as it increases exposure for attack, and increases likelihood that service providers could share the data with third parties in the future. Service providers should document the technical reasoning if any data needs to be retained for any period of time.

- Security requirement: Service providers should ensure adequate security of any retained data. Any security protocols implemented by the service provider should receive a full technical assessment and be regularly audited by a regulatory entity.

- Non-reusability requirement: Any data stored by V2G service providers should be retained in such a way that its reuse by the service provider in an undisclosed manner is not possible. This will help ensure consumer data is not misused by the service provider.

In our development of emerging power system technologies, as detailed in the remainder of this chapter, we attempt to follow these guideline to assure privacy-aware design.

3.2 PRIVACY-AWARE DESIGN FOR VEHICLE-TO-GRID SYSTEMS

Vehicle-to-grid (V2G) technology promises the capability to employ plug-in hybrid electric vehicle batteries for storage capacity and ancillary services, but the collection of fine-grained data on the charge and discharge of individual batteries to support vehicle-to-grid functionality presents a unique privacy risk to owners of electric vehicles. Fortunately, widespread adoption of this technology need not be at the expense of consumer privacy. The remainder of this chapter presents the design for a high-level privacy-aware framework informed by the theoretical concept of contextual integrity that can be implemented through a novel automated auction mechanism for batteries and an application of secure wallets backed by trusted platform modules.

V2G systems efficiently manage rechargeable electric vehicle batteries as distributed storage modules that can reliably be called upon to provide power during periods of high electricity

demand. This reduces the need to dispatch expensive peak-shaving generators in the real-time market to meet unexpected spikes in demand. Meanwhile, car batteries can be used to absorb energy from the grid when storage is required, which also recharges the battery in the process. The storage capacity provided by batteries can be a much-needed facilitator for the commercialization of the renewable energy industry, as wind and solar power tend to be intermittent and unpredictable sources of energy [2, 3].

In the absence of a V2G system, expensive generators providing regulation service are required to satisfy the charging demands of electric vehicle owners. Regulation service is typically used to ensure system stability by adjusting parameters in the grid in real-time – a very expensive process. However, the dispatch of expensive generators providing regulation service could be avoided through the introduction of V2G. Intelligent control and aggregation of batteries using V2G systems allow the V2G manager to intelligently charge each constituent battery and levelize aggregate demand levels, helping circumvent the need to resort to expensive regulation service providers [4].

V2G has the potential to provide wider benefits to the economy and environment [5]. A V2G manager has access to a large amount of storage capacity in the form of the electric vehicle batteries. These batteries bring with them a depth of information, including their historical fine-grained state of charge (or the current percentage level of battery charge) and consumer charging preferences. Using this high-level control architecture, electric vehicle batteries collectively act as suppliers to the electricity market via the V2G manager. While it is unlikely for an aggregation of batteries to be profitable in the base load power or peak power markets, it has been argued that it may be competitive in ancillary services markets [6].

Present V2G concepts call for a V2G manager, which may be independent of the utility company, to control the charging and discharging of all regionally distributed electric vehicle batteries. While a single battery does not have enough power to have a significant impact on the system, thousands of batteries that are intelligently controlled and collectively represented by an aggregator can participate competitively in ancillary service markets. In this scheme, the V2G manager serves as the interface between the wholesale electricity market and the electric vehicle owners providing the service of high-level battery management and commercial representation. This arrangement allows electric vehicle owners to earn profits by availing their car batteries to the grid for electricity storage, thereby further incentivizing electric vehicle ownership.

However, V2G managers pose a unique privacy risk to owners of electric vehicles. In order to ensure V2G functionality, managers would require fine-grained, temporally precise data associated with individual car batteries to inform their decisions on individual battery management. In the wrong hands, such data can be manipulated to reveal private information related to the original owners of electric vehicles.

In the remainder of this section, we will identify the privacy risks associated with V2G managers and discuss the information they can glean about individual vehicle owners from battery data. Finally, using the set of privacy-aware guidelines for the design of a vehicle-to-grid network presented in the section 3.1, we will present a framework for a privacy-aware vehicle-to-grid system informed by the theoretical concept of contextual integrity. In this framework, we make use of an automated, agent-based auction mechanism conducted among individual batteries, by which we obviate the need for batteries to reveal fine-grained data to

the V2G managers. This framework also makes use of various cryptographic processes enabled by trusted platform modules to allow for a secure, digital wallet for individual V2G customers.

3.3 V2G'S RISKS TO PRIVACY

Current V2G frameworks call for a specific set of data to be sent from individual car batteries to the V2G manager in real time. While these data sets vary from design to design, they all call for temporally precise, identifying data which may be harmful to individual privacy. More specifically, it has been argued that V2G managers require five points of data for each grid-connected battery to ensure the functionality of the system, including:

1) an identification number unique to each battery to enable a reliable billing mechanism;
2) the binary connection status of each battery;
3) the charging preferences of each grid-connected electric vehicle's owner;
4) the real-time state of charge for each battery; and
5) the measure of power flow from each battery to the grid, or vice versa.

These five points of data provide a comprehensive summary of the real-time status of each battery, and enable key V2G functionalities including intelligent charge management, system planning, and accurate billing [4].

3.3.1 Potential inferences from battery charging data

Temporally precise, fine-grained charging data sent to the V2G manager to support efficient system operation procedures can reveal a great deal of information about the electric vehicle owner. This type of information can be used to determine, among other things, the consumer's behavior and preferences. The potential threat to privacy is compounded both by a trend of incorporating advanced telemetry and sensors in electric vehicles as well as a general computing environment in which a wide variety of personal data is produced and stored. Consider the following inferences that can be made about consumers using a historical profile of battery charging information.

Ownership of an electric vehicle

Ownership can be used as an indicator of personal wealth, level of sophistication, or awareness of the environment. While some owners may not feel overly sensitive about this information, others will. A recent study found that a very particular subset of American car owners drive electric vehicles: "well-educated, upper-class white men in their early fifties with ideal living situations for [electric vehicle] charging" – presumably, garages or sheltered locations where their cars could be charged overnight [7]. Researchers at the University of California, Davis have further classified owners into groups, including environmentalists, early adopters and technology enthusiasts, individuals who wish to reduce their fuel bills, and individuals who wish to ease national dependence on oil [8]. Whether or not electric vehicle owners wish to be associated with such groups is not the issue at hand; instead, the underlying V2G technology should seek to dissociate owners with any particular background, as such information could

potentially be used in targeted advertising or marketing campaigns at the vehicle owners' expense.

Your presence at home

When a vehicle is not plugged in to the grid at home, it can be assumed that the owner is away. Numerous inferences can be made from this simple information. The owner's work pattern is one, as battery charging information can be used to determine when the car is plugged in to the grid. Although charging information for a single day may not reveal life routines, aggregated historical information over several days can do so, given V2G managers can readily apply econometric models to data to make reliable inferences from it. Similarly, the times the owner leaves home on leisure trips can easily be derived. In the wrong hands, such information can be used in crimes including robbery and others.

Distance to your places of travel

Distance traveled can easily be derived from the battery's state of charge, as the battery will discharge while the car is traveling. The battery's change in state of charge between two points of travel can therefore be translated to miles driven. If the car battery's charging physics are known, then the percentage change in state of charge between two points driven can be used to determine the miles driven, which in turn can be used to infer whether the owner traveled to

work or elsewhere. Such monitoring of an individual's whereabouts would place the vehicle owner's privacy at risk.

3.3.2 Potential inferences from charging location

Further information can be revealed if the battery's location of charging is known by the V2G manager [9, 10]. If significant penetration of electric vehicles is eventually achieved, as has been shown possible, it is expected that commercial charging stations will become common and widespread [11]. Even among current electric vehicle owners, 43 percent charge their vehicles away from home [7]. It is expected this number will increase as electric vehicles become more popular and commercial charging stations become more common. To enable the billing mechanism taking into account non-residential charging, the charging station will need to communicate the battery's identification code to the V2G manager, revealing the address-precise location of the vehicle. Consider the following information that can be inferred from the location of charging.

The location of your residence and workplace

Given your address, financial and socioeconomic details about your life can be inferred. Correlating information about the neighborhood's average age and affluence can all be used to derive otherwise private information about its owner. Further, if your vehicle is charged at the

workplace, it is possible to derive occupational specifics such as a position in organizational hierarchy, period of employment, salary details, and frequency of changing employment.

Identities of friends

Visitors to your house are likely to use it for charging, especially if they have traveled a long distance to arrive there. A network of friends and acquaintances, including information about strength of ties, can be reconstructed based on inferences drawn from frequency of visits, demographics of visitors based on their residential and occupational information, as well as time, duration and reciprocity of visits. Vehicles owners wishing for their social life to remain private would be at serious risk if inferences about their network of friends were to be derived through locational analysis of battery charging.

Medical information

If you use your electric car to visit the doctor, the doctor's identity could be determined by the V2G manager. Various types of sensitive medical information can be derived from your visits to the doctor or the hospital. For instance, you may visit specialists in addition to your general physician. Correlating information about frequency of visits to different physicians can provide clues about both general health and specific ailments. This information is invaluable to insurance companies and various others interested in your medical records and may result in detrimental changes to premiums or discrimination based on certain medical conditions.

General interests, preferences, and beliefs

Given the combination of all of the places you visit, a profile of your likes and dislikes can be inferred. In addition to your residence and workplace, you may charge your car at various other commercial charging locations. This address-specific data can be used to reveal your interests, preferences, and beliefs, as the businesses you frequent most often, be they shopping malls, doctor's offices, or liquor shops, can easily be inferred. This information can in turn be used to infer individual preferences, political or religious affiliations, and many other specifics, which in turn can feed targeted advertising campaigns that do not necessarily benefit the vehicle owner.

3.3.3 Entities that may be interested in V2G data

It is clear that data collected through V2G systems can be correlated with other available information (such as power consumption data collected by electric utility companies) and used to construct a more detailed profile of each electric vehicle owner and V2G participant in a distribution network. Whether the V2G manager is the utility company, a third party aggregator, or another organization, information may be sold to entities that have an interest in acquiring information about individual consumers. These entities will likely have no interest in acquiring this information for any reason other than to use it for selfish gains.

Some have suggested that automotive manufacturers may wish to serve as V2G managers because they already have vehicular telemetry systems that are used for repair services [6]. V2G management would be cost-effective for manufacturers like General Motors, who

presently use the Onstar link in the Chevy Volt to gather extensive data on the operation of Volts on the road. It has also been argued that cellular network providers may have a natural incentive to serve as V2G managers as they already have depth in knowledge on the deployment of wireless communication technologies. However, without sound regulation, cellular and automotive companies have no inherent incentive to protect consumer privacy.

Those individuals whose travel history is of particular interest are also at great risk. V2G implementations that do not consider these issues leave the transportation profiles of such individuals at risk of public exposure. Consider the 2011 United States Supreme Court case involving an alleged drug dealer whose vehicle was tracked using a GPS device for four weeks. This GPS-aided vehicle tracking provided authorities with incriminating information that they would not have had otherwise, but recent controversy about whether information acquired in this way can be upheld in court has attracted considerable media coverage. While V2G technology does not require a time-stamped location, it would provide location-based information about electric vehicles when they are connected to the grid, from which a great deal of sensitive information can still be acquired.

More generally, there is a great potential threat to privacy at a systemic level beyond bounded relationships between specific entities that seek information and individuals that are their subjects. Oscar Gandy has described how rampant collection of personal information can lead to subtle and invidious forms of discrimination that ultimately impinge on democratic processes and individual freedom [12]. While the aim here is not argue the merits and boundaries of privacy in general, it is important that a technology designed to improve the automobile, a

technology representative of personal freedom and choice, does not inadvertently become a backdoor to impinge personal freedoms.

3.4 PRIVACY-AWARE GUIDELINES FOR V2G

Wicker and Schrader have developed a framework based upon *Records, Computers, and the Rights of Citizen*, published in 1973 by the US Department of Health, Education, and Welfare, for the design of information networks that protects the privacy information associated with individual consumers [1]. This framework, outlined at the beginning of this chapter, is also generally representative of privacy concerns and protections incorporated in policies such as the Organisation for Economic Co-operation and Development's Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, the European Union Data Protection Directive, and the Asia-Pacific Economic Cooperation's Privacy Framework, all of which are applicable in other parts of the world that are currently target markets for electric vehicle technology. In this section, we provide a few key insights from their framework and apply them directly to the V2G concept. Further, we describe the actions that should be taken by the V2G manager and regulation agencies to ensure consumers of data privacy.

3.4.1 Consumer knowledge of and consent to data collection

It is critical for prospective buyers of electric vehicles to have a comprehensive understanding of the type of data that is to be collected about them should they partake in a V2G program.

To accomplish this, V2G operators must be required to provide a statement serving as a contract as to the type of data they will collect. This statement should be very precise, noting specifics like the level of granularity at which data will be collected. Such a statement allows the consumer to understand the privacy risks associated with V2G participation. Further, in certain areas, consumers may be presented with various choices for V2G service. In such cases, this statement will act as a point of comparison between all competitors. This level of competitive transparency can help to leverage market mechanisms to foster the protection of consumer privacy.

Additionally, V2G service providers must be required to receive the consent of consumers before consumers are subjected to any collection of their data. Specifically, before consumers decide to participate in V2G, they should be required to fully acknowledge that their data is being collected and processed. Furthermore, the inferences about their personal details that can be derived from that data must be made clear to consumers. Opt-in requirements can be used to achieve this, as they have been shown to affect participation rates [14]. In an opt-in scheme, consumers would be default not be registered for V2G service. To participate, they would be required to first acknowledge their understanding of the data that is to be collected, and then opt into participation in V2G.

3.4.2 Minimization of data collection

The more data on an individual that is collected, the more it can be used to infer sensitive information about the individual. This requirement is thus critical to restrict data collection,

but it is typically difficult to determine what amount of data collection is fair. To approach this issue, we argue that only the amount of data that is required to ensure the technical functionality of the V2G system should be collected. Any collection of data above this level must be deemed unnecessary and therefore a breach of privacy. The argument for this stance stems from the fact that because the excess data is not necessary for V2G functionality, its collection does not benefit the consumer in any way.

Though data may be required for technical functionality, there may be cases in which it is not required at the physical site of the V2G manager. For instance, in the case of wireless sensor networks, autonomous sensors are spread across an area and each one monitors the status of its locality using a built-in processor. Such distributed processing systems can afford various benefits, including privacy of information. When data is processed locally as opposed to centrally, it is not immediately accessible to the service provider. An emphasis on distributed processing can therefore enhance the privacy of individual consumers.

3.4.3 Minimization of identification of data with individual consumers

Data collected about a consumer's battery charging patterns describes that consumer's individual use of the service. Critically, though, even if the charging status of each battery were to be required by the V2G manager, there may be no need to associate it with an individual consumer. For example, the V2G manager may wish to know how much storage capacity is available in a particular neighborhood in real time. This can be determined if the state of charge of each individual battery in the neighborhood is known, but there is no need for this granularity

of information. The V2G manager can provide the same technical functionality simply by knowing the aggregate storage capacity in the neighborhood. V2G systems should be designed with this in mind.

We use these privacy-preserving design concepts in the development of a privacy-aware V2G infrastructure later in this chapter.

3.5 CONTEXTUAL INTEGRITY

A recent theoretical framework for understanding privacy dubbed "contextual integrity" has been proposed by political philosopher Helen Nissenbaum. A decision heuristic based on contextual integrity posits that "we locate contexts, explicate entrenched informational norms, identify disruptive flows, and evaluate these flows against norms based on general ethical and political principles as well as context specific purposes and values" [15]. While the Fair Information Practices discussed above provide guidance for ethical and political principles, they do not in and of themselves present an appropriate contextual framework for the application of V2G.

Although it is often a non-trivial task to determine the appropriate contextual parameters for a novel technical application, in this case we feel that the common information flows associated with refueling a vehicle powered by internal combustion provide a viable baseline for evaluating a V2G system. We choose this context because these are the two refueling options

that a consumer has to choose from – either to recharge an electric car via a charging station, or to refill a gasoline-powered car at a gas pump. Given this context, a consumer choosing to opt into a V2G program should be able to do so without concern that they are under any greater privacy threat than if they were to purchase gasoline or diesel at a local station. This transaction generally requires a consumer to divulge minimal personal information. If paying with cash, information flows are limited to common interactions between station attendants and other patrons. There may be greater privacy implications if paying with credit card or another form of payment, but this is done purely by the customer's volition. The presence of surveillance cameras may present an additional privacy threat that falls outside of the scope of this chapter, but even then, the information from such devices is commonly understood to be maintained locally for a limited amount of time and is not stored in databases that allow for correlation of data.

3.6 PRIVACY-AWARE INFRASTRUCTURE FOR V2G

We propose a V2G system that preserves the full benefits of V2G while ensuring consumer privacy. To achieve this, we stress the importance of distributed processing, by which the processing of data takes place at the battery. We also call for use of Trusted Platform Modules to encrypt data and store keys, as well as an auction mechanism to provide high-level control of battery charging and storage.

We refer to three processes that are crucial to V2G functionality: charging, planning, and billing, all of which are required for intelligent battery management. We first discuss the

technical requirements for each of these processes by describing the functionalities required of a V2G system. Then, using a combination of distributed processing, the Trusted Platform Module, and the auction mechanism for charging and discharging management, we present novel privacy-aware methods for charging, planning and billing.

Here we refer to the V2G manager as an as an independent agent separate from the utility company or power system operator. The V2G manager may, however, be a subsidiary of either of these organizations. It is assumed that the V2G manager also reports on the aggregate status of its constituent electric vehicles to the utility company and power system operator.

3.6.1 Charging

The most prevalent motivation for V2G implementation is that it would enable the manager to monitor constituent batteries when connected to the grid and schedule charging and discharging intelligently. In current V2G frameworks, temporally precise charging information is reported in real-time to the V2G manager, which schedules the charging of batteries optimally. We have illustrated that this can compromise consumer privacy.

We propose that each battery acts as an intelligent agent that can autonomously make the decision to charge, idle, or supply power to the grid. The battery's decision to charge at a particular time will depend on information that is provided by the utility company via a public broadcast of real-time and projected prices. The battery will charge according to consumer inputs including desired finish time of charging and state of charge. Because the battery is able

to manage itself autonomously using broadcasted information, direct communication from battery to V2G manager for charging purposes is avoided.

Two significant logistical issues arise from the use of batteries as intelligent agents. The first is to determine how all batteries in a region can collectively be managed so that system-wide electricity demands are met when batteries make decisions independently. The V2G manager would have prior agreements with the electricity markets on the necessary provision of various amounts of power at different times of the day [4]. If these prior agreements on power flows are not satisfied, the V2G manager would face financial penalty.

This problem arises from the implementation of batteries as independent agents as many batteries may decide to take identical actions at particular times in the day. For instance, if real-time prices suddenly become very low, many batteries may instantaneously decide to charge. Similarly, if prices are suddenly high, then many batteries may decide to discharge and sell electricity back to the grid. What must be devised is a method for the V2G manager to grant requests in a way that ensures that their arrangements with the electricity market are also satisfied.

To address this issue, we propose implementing an auction mechanism in which the V2G manager receives requests from individual electric vehicle batteries to charge or sell electricity to the grid. This auction can be conducted hourly, in an hour-ahead market. A request may be either a bid to buy or an offer to sell electricity, and requests are submitted each hour for the hour-ahead auction. Both bids and offers of electricity consist of three basic parameters, including (1) a buying or selling price, (2) an upper bound and (3) a lower bound on power levels that correspond to the amount that the battery prefers to charge or discharge itself. After

receiving these bids and offers, the V2G manager selects a cost-minimizing dispatch for regional battery charging.

To illustrate this, consider a time interval in the near future in which the forecasted electricity price is very high and the V2G manager has received a large number of requests from batteries to supply electricity to the grid. Assume that the total supply of power from the requests to sell is greater than the total power demanded of the V2G manager by the market in that future interval. Given this problem, the V2G manager conducts the auction mechanism to select only certain batteries. Based on predetermined consumer preferences, some batteries submit sell offers for high prices, while others submit offers for low prices. Finally, the V2G manager selects the subset of offers that satisfies the market demand and minimizes the total cost of selected offers. The batteries may be compensated using either the market-clearing price or pay-as-bid settlement schemes. This auction mechanism can be conducted very quickly. The V2G manager can broadcast pricing information and batteries can prepare bids or offers that correspond to consumer preferences in an automated fashion very quickly. Once the V2G manager receives all requests, determining the dispatch is trivial.

The second issue that arises from the use of batteries as independent agents is that communications from the battery to the V2G manager, such as charging requests, must not reveal the identity of the battery. If the battery's identity were revealed to the V2G manager, the battery's charging profile would be compromised. This is an issue that can be addressed by a direct anonymous attestation protocol implemented by cryptographic hardware.

3.6.2 Planning

In current V2G concepts, the manager uses charging information to aid in planning the operation of the power system. Data collected by the V2G manager provides power system operators with a better understanding of electricity demand levels, but if batteries were developed as intelligent agents that are not directly monitored at a high level then the V2G manager would not have access to their temporally precise charging information. Aggregate battery demand and supply would still need to be communicated to the V2G manager to facilitate system planning.

To solve this issue, it is necessary to understand the structure of the power grid. The transmission system is a mesh network of power lines that transport energy over long distances. Transmission lines are connected by nodes called substations, which step high voltages down to distribution-level voltages. The distribution network is connected to the transmission network through substations. However, while transmission lines are arranged in a mesh network, distribution lines are arranged as radially outward laterals from the substation to service local consumers (Fig. 5). For the purposes of system operation and planning, the utility company requires aggregated demand levels for each distribution lateral [16]. Therefore, collecting individual consumer data is unnecessary.

In this light, we propose use of a lateral aggregator to aggregate the total electricity demand, storage capacity, and electricity supply for each distribution lateral in the V2G-enabled grid (Fig. 5). Transmission of this data occurs between each car battery and the lateral aggregator via encrypted communication. The required level of encryption can be achieved using an

electrical chip called the Trusted Platform Module (TPM), which is installed as a component of the battery electronics specifically for the purpose of privacy protection.

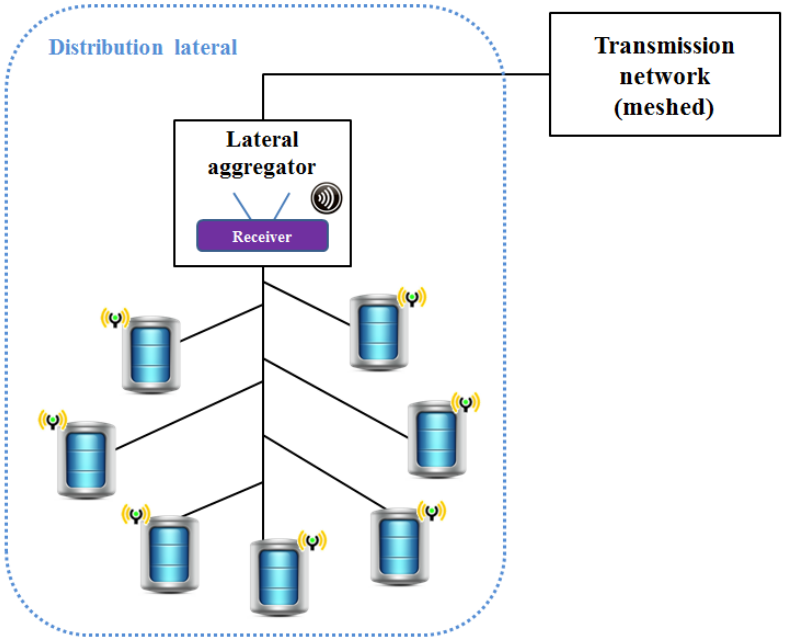

Figure 5: An aggregator is assigned to each distribution lateral to communicate with individual batteries, enabling the anonymized data collection process.

The TPM was developed by the Trusted Computing Group (TCG), who established a set of design principles for trustworthy computing [17]. Use of a TPM on each battery would allow for encryption using public key infrastructure (PKI) through a cryptographic process known as binding, by which a nonmigratable key unique to the TPM is used to encrypt messages. Thus, any message encrypted using the nonmigratable key is "bound" to the TPM, thereby enabling secure communication of charging status to the lateral aggregator. Recent TPM specifications have been updated to include a direct anonymous attestation protocol (DAA) [18]. The DAA allows for a decoupling of identity and transmitted data through the use of separate DAA issuers

and verifiers. The issuer is able to provide a secure pseudonym to a TPM client that can then be verified by a verifier. Using DAA, the battery can upon request generate a software authentification certificate for the V2G manager, enabling the manager to attest the identity of a battery without knowing the actual identity of the battery. This allows for the lateral aggregator to receive the charging status of each battery on its lateral, but leaves it unable able to associate any data it receives to an individual consumer. Additionally, all communication the lateral aggregator receives is aggregated. Despite some degree of anonymity being provided by the TPM, aggregation can provide protection against more sophisticated privacy threats such as timing and intersection attacks based on correlation of pseudonymous charging data. Thus, this process adds two levels of anonymization to protect the consumer: TPM-based anonymity and data aggregation.

3.6.3 Billing

To realize the full potential of the V2G system, electric vehicles must be developed as controllable load and storage modules. This suggests that real-time pricing needs to be implemented to serve as a driving force for a shift toward higher consumption during off-peak periods and lower consumption during peak-load periods. In current V2G concepts, the real-time state of charge of each battery is communicated directly to the manager. This allows the manager to apply real-time prices to real-time consumption levels. As we have argued, this compromises consumer privacy.

Here we again rely on functionality of the TPM to provide a privacy-aware payment system. As informed by our baseline contextual analogy to current fossil fuel market practices, we propose that payment associated with charging and discharge is settled per session when a battery is disconnected from a charging station. When a battery is connected to the grid, the lateral aggregator can query the TPM on the battery to ensure its authenticity and to attest its functionality to prevent rogue batteries from connecting to the grid. The authenticated batteries then monitor the results of bidding and record charging and discharging along with the respective prices that are broadcast by the utility. When the battery is disconnected from the charging station, it calculates a bill and supplies it to the aggregator. The total amount of the bill is then credited or debited to a payment mechanism associated with a TPM-backed wallet [19]. The utility company need not be concerned with the individual batteries in this case and can charge the aggregator using whatever periodicity and settlement mechanisms it normally uses to bill customers.

TPM wallets allow consumer flexibility by accommodating a variety of payment mechanisms, including e-cash or credit cards. There are different privacy implications to the customer depending on the specific payment method they choose to use. This choice is left to the consumer in this case in the same way that they are free to choose how they pay for petrol or diesel at a pump. TPM-based attestation allows the aggregator to verify that the software included with the battery is designed to verify sufficient funds before making bids ensuring that the consumers are not able to "charge and run" without settling their account at disconnect.

3.6.4 A privacy-aware V2G design

To protect consumer privacy, critical technologies need to be implemented at the high and low levels. Specifically, (1) a local computing device is required at each charging station to manage the battery; (2) a lateral aggregator is required for each distribution lateral to aggregate regional demand and supply from electric vehicle batteries; and (3) a TPM is required to apply cryptographic keys sent from the battery and for remote attestation. This adapted V2G system design is illustrated in Figure 6.

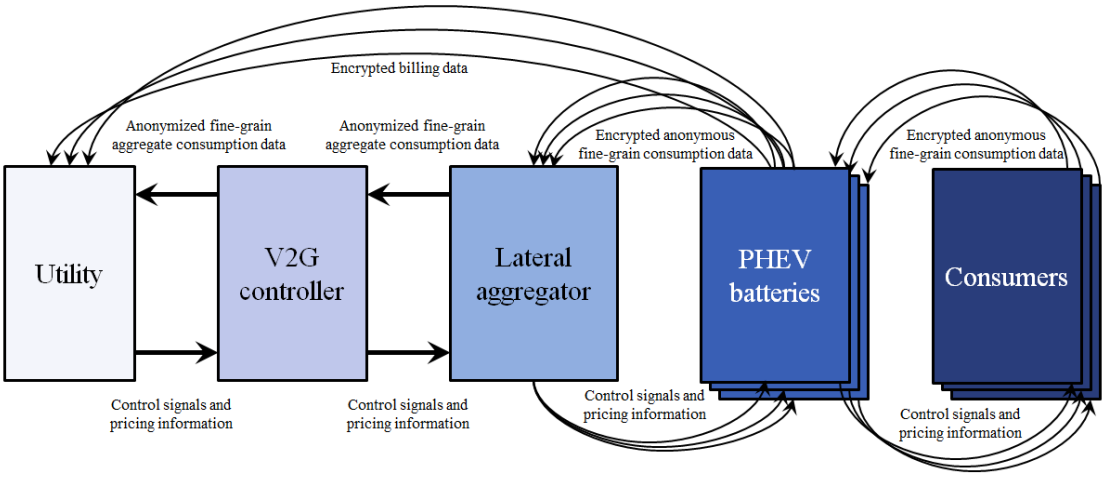

Figure 6: The privacy-aware V2G infrastructure

3.7 CONCLUSIONS

V2G technology can reduce stress on the power grid by providing a ready supply of storage capacity in the form of electric vehicle batteries. It will allow us to develop a market in which aggregations of batteries can sell power to the grid at a profit, which in turn will heighten the incentive for consumers to buy electric vehicles. It is critical, though, that the privacy risks of

V2G are comprehensively understood. Only by understanding them can V2G systems be designed to minimize privacy risks. To conclude, we wish to emphasize the importance of privacy not only with respect to V2G but with regard to all upcoming technologies, and that privacy-aware design of upcoming technologies must be used wherever possible to mitigate privacy risks.

3.8 REFERENCES

[1] S. Wicker and D. Schrader, "Privacy-Aware Design Principles for Information Networks," in *Proceedings of the IEEE*, Vol. 99, Iss. 2. pp. 330-350, 2011.

[2] S. Hadley and A. Tsvetkova, "Potential Impacts of Plug-In Hybrid Electric Vehicles on regional Power Generation," *The Electricity Journal*, Vol. 22, Iss. 10, pp. 56-68, 2009.

[3] P. Kelly-Detwiler, "Electric Cars and the Power Grid: How Are They Coming Together?," *Forbes*, 2013.

[4] C. Guille and G. Gross, "A conceptual framework for the vehicle-to-grid (V2G) implementation," *Energy Policy*, Vol. 37, Iss. 11, pp. 4379-4390, 2009.

[5] R. Sioshansi and P. Denholm, "Emissions Impacts and Benefits of Plug-In Hybrid Electric Vehicles and Vehicle-to-Grid Services," *Environmental Science & Technology*, Vol. 43, pp. 1199-1204, 2009.

[6] W. Kempton and J. Tomic, "Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy," *Journal of Power Sources*, Vol. 144, Iss. 1, pp. 280-294, 2005.

[7] C. Woodyard, "Smart, Rich People Buy Electric Cars," *USA Today*, November 2012.

[8] R. Heffner, K. Kurani, and T. Turrentine, "Symbolism in California's Early Market for Electric Vehicles," *Transportation Research*, Vol. 12, Iss. 6, pp. 396-413, 2007.

[9] S. Wicker, "Digital Telephony and the Question of Privacy," *Communications of the ACM*, Vol. 54, No. 7, pp. 88-98, 2011.

[10] S. Wicker, "The Loss of Location Privacy in the Cellular Age," *Communications of the ACM,* Vol. 55, No. 8, pp. 60-68, 2012.

[11] C. Farmer, P. Hines, J. Dowds, and S. Blumsack, "Modeling the Impact of Increasing PHEV Loads on the Distribution Infrastructure," *IEEE Hawaii International Conference on System Sciences 43*, 2010.

[12] O. Gandy, *The Panoptic Sort: A Political Economy of Personal Information (Critical Studies in Communication and in the Cultural Industries)*, Westview Press, January 1993.

[13] Department of Health, Education, and Welfare, *Records, Computers, and the Rights of Citizens*, 1973.

[14] S. Bellman, E. Johnson, and G. Lohse, "To Opt-In or Opt-Out: It Depends on the Question," *Communications of the ACM*, Vol. 44, Iss. 2, pp. 25-27, 2001.

[15] H. Nissenbaum, "A Contextual Approach to Privacy Online," *Dædalus*, 2011.

[16] H. Willis, *Distributed Power Generation: Planning and Evaluation*, New York, Marcel Dekker, Inc., pp. 1-34, 2000.

[17] Trusted Computing Group, *TPM Main, Part 1 Design Principles, Specification Version 1.2*, Level 2 Revision 103, 2007.

[18] E. Brickell, J. Camenisch, and L. Chen, "Direct Anonymous Attestation," in *Proceedings of the 11th ACM conference on Computer and Communications Security*, pp. 132-145, October 2004.

[19] S. Gajek, H. Löhr, A. Sadeghi, and M. Winandy, "TruWallet: trustworthy and migratable wallet-based web authentication," *ACM workshop on Scalable Trusted Computing*, pp. 19-28, 2009.

CHAPTER 4.

GAME-THEORETIC ANALYSIS OF SMART GRID PRIVACY

Thus far, we have seen that collection of fine-grained, temporally precise smart grid data can pose serious privacy risks to electricity consumers if the data is accessible to utility companies or third parties. Specifically, data collected via smart metering and vehicle-to-grid technologies among others can reveal a great deal about behavior, beliefs, and preferences of consumers. In the previous chapter, we found that systems can be designed so as to factor out such privacy risks for many applications in the smart grid, including smart metering systems as well as vehicle-to-grid technology. Technical solutions for privacy can often be developed quite inexpensively while still adhering to the privacy-aware design guidelines presented in section 3.1.

Nonetheless, privacy-aware solutions are not always in the interests of all parties. In the smart metering example, various entities exist that would potentially be affected by the use of cryptography in smart metering. The two stakeholder groups that would be most affected by such a change would like be the individual electricity consumers and the electric utility companies. On one hand, consumers want to protect their privacy, and therefore will opt for privacy-aware smart meters. On the other, electric utility companies wish to collect as much fine-grained consumer data as they can to streamline transmission and power routing and to cut electricity production costs. It is also conceivable utility companies could sell such data to interested third parties.

Therefore, the development of privacy-aware solutions is not enough to ensure their adoption. To understand the legal standards and regulations necessary to promote the adoption of privacy in the smart metering market, we must understand the economics of the stakeholder dynamic behind smart metering, with specific focus on the aforementioned struggle for data between individual consumers and electric utility companies. In this chapter, we will analyze this stakeholder problem from a game-theoretic perspective in order to determine the economic conditions and regulatory standards necessary to promote the adoption of privacy in the market.

Adopting a game-theoretic model, we consider utilities that offer both privacy-aware and non-privacy-aware AMI. A non-cooperative game is developed in which a representative consumer strategizes against the utility. The regulatory measures required for the desired privacy-facilitating Nash equilibrium of the game are discussed, and recommendations for policymakers are presented. In particular, it is found that a combination of regulation and consumer awareness must overcome the financial benefit arising from the sale of consumption information to third parties.

4.1 INTRODUCTION

Power meters presently used in most electricity markets around the world employ old technology. In general, these meters collect power consumption data, and the total usage over a period of time is aggregated. The meter then reports this total amount to the billing entity, usually an electric utility company that provides electricity transmission services. The company typically bills consumers at a predetermined flat rate [1].

In recent times, though, the industry has shown strong interest in adopting a demand response (DR) system by which consumers would be billed at the real-time energy price as opposed to the flat rate [2]. The real-time price, which varies greatly over the course of a day, would be reported to the consumer, and if the consumer deems that price is too high, he or she may switch a power-intensive application off, and vice versa. It is expected such changes in individual consumer activity would significantly impact the daily pattern of aggregate demand such that during the on-peak hours, aggregate demand would be lower than it would be without DR, and similarly, higher during the off-peak hours, thereby resulting in a flatter demand curve. This in turn would reduce the required system-wide generation capacity and lower the need of frequently turning marginal generators on and off. Production costs would thus decrease, leading to lower final consumer payments for electricity [1, 3, 4].

DR requires a method of reporting temporally precise power consumption data to the utility for billing purposes, which calls for new technology in metering. Advanced Metering Infrastructure (AMI), an intelligent metering technology, would enable DR, but also carries with it potential harm for consumers [5]. It has been established that if a third party has access to a consumer's temporally precise usage data, then the nature and timing of his or her power-

consuming activities can easily be determined [6]. This information may be valuable to third parties who, for example, wish to target a particular demographic in a campaign. A third party may obtain this data by either hacking an unsecured AMI module, or by purchasing it from the utility. Clearly, then, temporally precise usage data is valuable to both the representative consumer as well as the utility.

Thus, in order for AMI to enter the market, the technology must be acceptable to consumers, utilities, and policymakers. Specifically, consumers must be certain of their privacy, while utilities must be compensated in some way for the costs associated with switching to the newer technology. To ensure their privacy, consumers would require privacy-aware AMI that encodes and secures usage data such that the meters cannot be hacked for information. Furthermore, public key infrastructure would be used to hide usage data at the individual consumer level from the utility.

Meanwhile, use of non-privacy-aware AMI would be preferable to the utility, as fine-grained consumption data could then be sold on to third parties. This competition for benefit between the consumer and utility can be modeled as a two-player game in which consumers have the option of either continuing to use the standard aggregating power meter or choosing to upgrade to an advanced meter.

In this chapter, the pros and cons of each meter type are discussed. Based on these, a matrix game between the representative consumer and the utility is developed. The policy requirements for the desired privacy-friendly Nash equilibrium strategy tuple are then analyzed under various combinations of regulatory policies. Based on these results, and in view of the desired Nash equilibrium, recommendations for the unbiased regulatory regime are provided.

4.2 LITERATURE REVIEW

Many studies of the potential impact of DR exist in the literature. They generally conclude that with an appropriate intelligent metering technology installed, DR is enabled, resulting in lower peak demand levels by about 10% and vast reduction of production costs at the aggregate level [3, 4].

Critically, the collection of fine-grained power consumption data creates a serious privacy problem for the consumer. Lisovich et. al. illustrated that, given a household's temporally precise power consumption data, patterns in the household member's private activities can easily be determined [6]. In this light, Wicker et. al. proposed a privacy-aware AMI infrastructure that protects consumer information using public key infrastructure. In the proposed infrastructure, fine-grained consumption data is not visible to utilities, but utilities are still able to bill consumers properly [7]. Wicker et. al. have conducted further study on proper design principles for information networks from a privacy-aware perspective [8]. Thus, in this chapter, it is assumed that usage data is perfectly invisible to utilities for the PA-AMI, while it is visible for the NPA-AMI. We also assume here that privacy-aware AMI, which we denote "PA-AMI," is available to the utility for deployment. The alternative, "NPA-AMI," represents the currently available, non-privacy AMI.

To date, no studies have been conducted to analyze the viability of a privacy-aware AMI infrastructure from an economic and regulatory perspective. This chapter presents a novel method of analyzing this problem using a game theoretic framework.

4.3 THE BENEFITS AND COSTS OF PRIVACY-AWARE AMI

In the present study, the economic effects of implementation of three different types of meters, the standard aggregating meter (SM), the NPA-AMI, and the PA-AMI, will be investigated. To accomplish this, the benefits and costs associated with each type of meter to both consumers and utilities must be identified.

We assign the amount v as the value of privacy of usage information to the representative consumer, who gains the payoff component v only when the meter used is the SM or the PA-AMI. For the unsecure NPA-AMI, this value is conceded to the utility, which can profit from the sale of the consumer's usage data to third parties.

In the case that the utility deploys NPA-AMI but chooses not to sell consumption data to third parties, consumers would still have a natural privacy-related aversion to the NPA-AMI, which can be hacked by a third party. This uncertainty suggests that consumers would potentially back out of the DR program for reasons of privacy, thereby preventing the utility from receiving the benefits associated with AMI deployment. Thus, with NPA-AMI, the utility incurs an expected cost associated with the risk of having its DR program being shut down due to the possibility of a public outcry against the NPA-AMI. We introduce the variable n to denote this cost.

An advanced meter is expected to save the consumer a significant amount of money, denoted by s. This payoff is received only when the consumer has chosen to install either type of AMI. Meanwhile, there are two effects experienced by the utility with AMI installation. It gains the amount l due to the cost-savings it achieves from the superior planning ability offered by installation of AMI. However, it incurs the cost c of installation of an AMI module. Because there is no physical difference between the PA- and NPA-AMI, it is assumed that the cost of

installation is the same for both types. Furthermore, the utility must pay this cost of installation, although in reality the utility may be subsidized for this effort.

With a successful sale of the representative consumer's data, the utility can expect to earn amount *g*. It is assumed that the utility can potentially earn this revenue only if NPA-AMI is in use, as this is the only type of meter for which the utility would have access to individual consumers' fine-grained usage data.

Two key regulatory choices will determine the economic efficacy of each meter type. The first of these choices regards the legality of selling information to third parties. If the sale of information is made illegal, then selling data would be very risky for utilities due to the imposed criminal punishment, which is modeled here as a large financial penalty. Meanwhile, there is a certain probability that the utility will not be discovered while completing a sale. Thus, for each sale under the condition that selling data is illegal, the utility incurs the cost *r* associated with the risk of discovery. If the sale of data is deemed legal, then *r* is zero.

The second of the key regulatory choices concerns the nature of recruitment of new participants to the DR program. The regime may decide that, by default, all consumers will participate in the program. In such an "opt-out" scheme, consumers must actively decide to leave the program by notifying the utility; this action requires the consumer to incur the cost e for the effort required to leave the program. Alternatively, the regime may decide to employ an "opt-in" scheme, by which consumers would be required to notify the utility if they wish to participate in the DR program. In this case, consumers would incur the cost *e* in joining the program [9]. Finally, the utility may decide to enforce the policy that the switch to AMI is mandatory for all consumers. It has been shown that the expected participation rates differ

greatly between the opt-in and opt-out cases, implying that e is indeed a significant cost for consumers to bear.

The benefits and costs to consumers of each meter type are summarized in Table 4.

Table 4: The benefits and costs of each metering type.

	Meter type		
	SM	**NPA-AMI**	**PA-AMI**
For consumer	$+ v, + e$	$+ s$	$+ v, + s$
For utility	0	$+ l, - c, - n$	$+ g, + l, - c, - r$

4.4 GAME THEORETIC FORMULATION

Given the benefits and costs associated with each metering infrastructure developed in section 4.3, it is possible to consider the two-player non-cooperative game between the representative consumer and the utility in which each player makes a decision regarding the type of metering infrastructure to be employed in the market.

It is assumed that the consumer initially uses a standard aggregating meter and must decide whether to continue using the standard meter or switch to an advanced meter. If the consumer settles on former, the utility has no choice but to continue employing the standard meter. If the consumer decides to use an advanced meter, then the utility has the option of installing either an NPA- or PA-AMI module.

This formulation can thus be described by the player set P and corresponding strategy set S, with

$$P = \{P_{consumer}, P_{utility}\} \text{ and}$$

$$S = (\{SM, AM\}, \{SM, NPA\text{-}AMI, PA\text{-}AMI\}) .$$

Figure 7 illustrates this game in the normal form. Consistent with the model mentioned earlier, strategy tuples $\{AM, SM\}$, $\{SM, NPA\text{-}AMI\}$, $\{SM, PA\text{-}AMI\}$ are invalid states. Payoffs are computed according to the costs and benefits as described in subsection 4.2.3.

$P_{consumer}$ \ $P_{utility}$	*SM*	*NPA-AMI*	*PA-AMI*
SM	$v + e, 0$	--	--
AM	--	$s,$ $g + l - c - r - n$	$v + s, l - c$

Figure 7: Game theoretic analysis of AMI game between a representative individual consumer and the utility. The desired Nash equilibrium for implementation of privacy-aware AMI is {*AM, PA-AMI*}.

Recall the cost r associated with risk of discovery is variable depending on the legality of selling usage data. Also, note that the cost of effort to the consumer e is a positive term in the $\{SM, SM\}$ strategy tuple (Fig. 7). In the case of an opt-in market in which the consumer incurs

this cost if he upgrades from *SM* to the *AMI*, *e* is positive. For the opt-out market, *e* is negative. If consumers are required by the regulatory regime to use AMI, *e* is zero.

Both the privacy-aware and the non-privacy-aware strategy tuples will now be examined further to understand the conditions necessary for the market to reach equilibrium at each of these states.

4.5 GAME THEORETIC RESULTS

4.5.1 Privacy-aware solution

The strategy tuple that facilitates privacy awareness is {*AM*, *NPA-AMI*}. Using the Nash equilibrium condition that no individual player can benefit from unilaterally deviating from the equilibrium strategy tuple, it is clear that the privacy-aware equilibrium requires the conditions

$$s > e \tag{1}$$

$$r > g - n \tag{2}$$

to be satisfied. In other words, for the utility to choose to use PA-AMI, its expected financial risk of selling information (r) must be greater than difference between the financial gain it would receive from selling information (g) and the expected cost associated with the risk of the DR program being shut down due to public outcry against NPA-AMI (n).

Note that the effort *e* required of the consumer takes on different values depending on the opt-in/opt-out regulatory regime. In general, it is required that the expected savings (s) from using

the AMI module are greater than the effort associated with the switch to advanced metering. This value is zero under the mandatory regime and so the required inequality is satisfied. The inequality is also satisfied in the opt-out case, where e takes on a negative value, but is not necessarily satisfied in the opt-in case.

4.5.2 Non privacy-aware solution

If NPA implementation were the equilibrium strategy for the utility, and {*AM*, *NPA-AMI*} were the equilibrium tuple, then the following conditions would be required:

$$s > v + e \quad (1)$$

$$g - n > r \quad (2)$$

Thus, if $v + e > s$ (as it could be reasonable to say that the value of privacy would be greater than the expected savings or consumers from using AMI), then the utility would have to compensate the consumer by at least the amount δ, where $v + e = s + \delta$, in order to ensure that {*AM*, *NPA-AMI*} would be the equilibrium strategy tuple.

4.6 POLICY RECOMMENDATIONS

Given the conditions necessary for the desired Nash equilibrium {*AM*, *PA-AMI*} described in section 4.4, specific recommendations on the most appropriate market structure can be made to the regulatory regime.

First, it is necessary to ensure that the consumer's expected savings with advanced metering are greater that the effort required for the consumer to choose to install an advanced meter. This condition is automatically satisfied for the opt-out and mandatory regulations, but it may not be under the opt-in scheme. Previous literature indicates that the opt-in scheme satisfies the Fair Information Practices best [10]. Thus, the regulatory regime must maximize $s + e$ with $e < 0$, subject to economic constraints.

Second, the regulatory regime must set the penalty for illegally selling information sufficiently high to entice utilities to choose to install privacy-aware meters. The expected penalty r is the product of the probability of discovery of data sale and the monetary fine imposed on the utility. Thus, in order to ensure that utilities are deterred from choosing NPA-AMI, this fine must be set sufficiently high given some probability of discovery.

4.7 CONCLUDING REMARKS

A preliminary study of the economic impact of AMI implementation was conducted from a privacy-aware perspective. A game theoretic formulation with a representative consumer versus the utility company was used for economic analysis. It was found that two particular regulatory conditions need to be satisfied in order to push the market toward privacy-aware AMI introduction at equilibrium. Recommendations were made accordingly. Future studies will focus on practical development of the technology and regulations required for privacy-aware AMI implementation.

4.8 REFERENCES

[1] S. Borenstein, "The Long-Run Efficiency of Real-Time Electricity Pricing," *The Energy Journal,* vol. 26, iss. 3, pp. 93-116, February 2005.

[2] A. Vojdani, "Smart Integration," *IEEE Power and Energy Magazine*, Vol. 6, Issue 6, pp. 71–79, November-December 2008.

[3] Federal Energy Regulatory Commission, *A National Assessment of Demand Response Potential*, June 2009.

[4] Federal Energy Regulatory Commission, *2009 Assessment of Demand Response and Advanced Metering*, September 2009.

[5] Electronic Power Research Institute, *Advanced Metering Infrastructure (AMI)*, February 2007.

[6] M. Lisovich, D. Mulligan, and S. Wicker, "Inferring Personal Information from Demand-Response Systems," *IEEE Security & Privacy*, vol. 8, no. 1, pp.11-20, January-February 2010.

[7] S. Wicker and R. Thomas, "A Privacy-Aware Architecture for Demand Response Systems" in *Proceedings of the 44th Hawaiian Conference on System Science (HICSS-44)*, Kauai, Hawaii, January 2011.

[8] S. Wicker and D. Schrader, "Privacy-Aware Design Principles for Information Networks," in *Proceedings of the IEEE*, vol. 99, no. 2, pp. 330, January 2011.

[9] J. Sovern, "Opting In, Opting Out, or No Options at All: The Fight for Control of Personal Information," *Washington Law Review*, 74, 1033, 1999.

[10] Department of Health, Education, and Welfare, *Records, Computers, and the Rights of Citizens*, 1973.

CHAPTER 5.

REGIMES FOR SMART METERING INTRODUCTION

Privacy-aware smart metering modules can be used to avoid the privacy risks of fine-granularity data collection by protecting an individual consumer's data using public key infrastructure. But while privacy-aware AMI would be preferred by consumers, utilities would naturally prefer non-privacy-aware modules, as they could profit from the sale of consumer usage data. As we have seen in chapter 4, it is possible to push the market toward privacy-aware smart metering by treating privacy adoption as the Nash equilibrium in a competitive game between the individual consumer and the electric utility company, but it is still not clear what regulatory structure should be implemented in introducing privacy-aware smart metering. In this chapter, we examine two possible regulatory regimes using consumer decision theory and determine the economic conditions required for privacy-aware AMI adoption at equilibrium under both regimes.

5.1 INTRODUCTION

Wicker et. al. have proposed the use of privacy-aware AMI, which protects consumer data by employing public key infrastructure but also enables real-time pricing and proper billing [1, 2]. While their technical development of privacy-aware AMI is sound, there remains a question as to how it should be introduced to consumers. No prior studies to determine the method of introduction of AMI that should be used by the regulatory regime have been conducted. In this chapter, we introduce two possible regulatory structures and analyze each using a consumer decision process. Subsequently, we determine the economic requirements for PA-AMI adoption by solving for the optimal strategy for the consumer under each regulatory structure.

We approach this issue by considering two regulatory regimes that differ in how privacy-aware (PA) and non-privacy-aware (NPA) AMI are made available to consumers. We begin by developing the regulatory framework for each regime into a decision process for a representative consumer. Subsequently, we determine the economic conditions required to push the representative consumer toward opting for PA-AMI adoption under each case.

5.2 REGULATORY REGIME STRUCTURES

When initiating the DR program, the regulatory regime must decide how AMI is to be provided to consumers. Discussion of two possible regimes follows. Variables that detail payoffs to either the utility company or consumer that are referenced in the remainder of this chapter correspond with the variables defined in section 4.3 and illustrated in Table 4.

5.2.1 Regulatory Regime 1: Consumer can use AMI or retain SM

Under Regulatory Regime 1, consumers using standard metering (SM) are presented with an offer to switch to AMI by the utility. The utility makes two offers – the NPA-AMI with an additional financial incentive *d* for consumers, and the PA-AMI. It is assumed that the utility must inform consumers up front that their privacy is at risk if NPA-AMI is installed. The consumer may refuse AMI altogether and choose to retain SM. The utility offers the incentive with NPA-AMI because otherwise, the consumer would immediately choose the PA-AMI, as it provides both energy savings and privacy from the utility. This process is illustrated in Figure 8. Payoffs are expressed in terms of the values introduced in the previous section, with the payoff to the consumer listed first, followed by that of the utility. Probabilities associated with each of the possible consumer decisions are expressed in terms of x and y, with $0 \leq x \leq 1$, $0 \leq y \leq 1$.

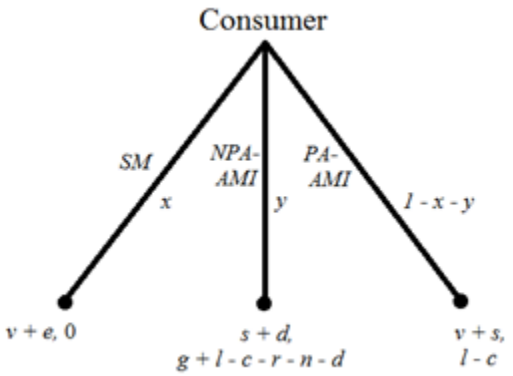

Figure 8: AMI Regulatory Regime 1.

5.2.2 Regulatory Regime 2: Consumer must switch to AMI

In the case of Regulatory Regime 2, the utility offers NPA-AMI plus the financial incentive d, and PA-AMI. In this regime, the consumer must switch to AMI and so does not have the option to retain SM. The utility offers the financial incentive d with NPA-AMI for the same reason as in the previous regime. The consumer decision process for this regime is shown in Figure 9. Again, payoffs are shown in terms of the variables of the previous section, with consumer payoffs listed first. The probability the consumer makes a decision is expressed in terms of z, with $0 \leq z \leq 1$.

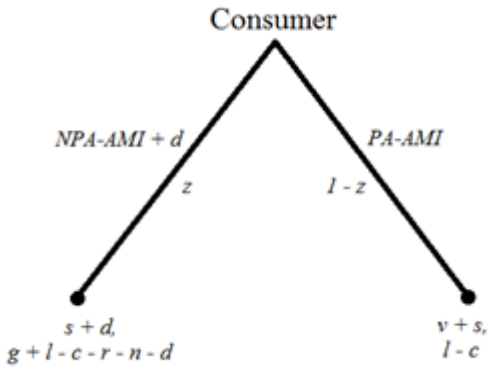

Figure 9: AMI Regulatory Regime 2.

5.3 THEORETICAL RESULTS

Given the decision processes discussed in the previous section, it is possible to determine the requirements under each regime for PA-AMI to be the representative consumer's optimal strategy. As described earlier, for each process, a distinct probability is assigned to each branch

on the decision tree. This value represents the probability that the consumer will decide to follow that branch of the decision tree. Subsequently, backward induction is applied to each process given that the optimal consumer choice is for PA-AMI. Applying this method yields the following requirements on the benefits and costs associated with the various metering technologies.

5.3.1 Regulatory Regime 1 requirements for PA-AMI adoption

$$v + s - 2xv - xs - xe - yv - ys > 0$$

$$v + s - xv - xs - yv - 2ys - yd > 0$$

5.3.2 Regulatory Regime 2 requirements for PA-AMI adoption

$$v + s - zv - 2zs - zd > 0$$

5.4 DISCUSSION

Results indicate that if it is possible to assign numerical values to *x, y, z, s, d*, and *v*, then it is possible to solve for the privacy adoption rate under each of the regulatory regimes. The first three of these can simply be found by conducting a survey to determine the percentage of individuals who would take various actions when presented with options on AMI. Furthermore, variable *s* can be determined through analysis of historical electricity market data,

and variable *d* can be developed using the analysis discussed in [3] or by comparing this game to incentive offerings in the real world, such as those made by the Pacific Gas and Electric Company. However, the last of these variables, *v*, measures the consumer valuation of privacy. This variable is critical in order to solve for the adoption rate of privacy in either regulatory regime, but is difficult to determine, as there are limited options in arriving at a numerical value for this variable. In the next chapter, we will provide a novel method for determining the value of *v* through use of a conditional choice regression model implemented through a survey. Using these results, policy recommendations will be made on the regulatory regime that would be the best option for the market.

5.5 REFERENCES

[1] S. Wicker and R. Thomas, "A Privacy-Aware Architecture for Demand Response Systems" in *Proceedings of the 44th Hawaiian Conference on System Science (HICSS-44)*, Kauai, Hawaii, January 2011.

[2] S. Wicker and D. Schrader, "Privacy-Aware Design Principles for Information Networks," in *Proceedings of the IEEE*, vol. 99, no. 2, pp. 330, January 2011.

[3] O. Gandy, *The Panoptic Sort: A Political Economy of Personal Information (Critical Studies in Communication and in the Cultural Industries)*, Westview Press, January 1993.

CHAPTER 6.

CONSUMER PRIVACY VALUATION

Smart metering initiatives have caused division among electricity consumers and utility companies since their inception. Advocates of smart metering expound the technological functionalities of smart meters that can help lower electricity production costs and reduce dependency on fossil fuels. But smart meters can also collect sensitive information about the power consumption habits of consumers, which has led critics to argue that consumer privacy is left exposed. How much do consumers value their privacy in this context? Do the benefits of smart metering outweigh the loss of privacy?

In the previous chapter, we saw that this valuation is specifically the most critical term to determine in deciding a regulatory regime to follow for introduction of a smart metering option for consumers. In this chapter, consumer willingness to adopt smart metering is investigated. A set of research questions around consumer decision-making related to smart metering and privacy is first developed. To address these questions, a national survey was implemented to examine consumer willingness to adopt standard smart metering and privacy-aware smart metering. Results indicate that the average consumer is willing to pay $11 each month to ensure privacy protection in smart metering. Several other key insights are generated from the survey results, which are used to suggest policy recommendations for smart metering.

6.1 INTRODUCTION

Smart metering programs have recently come under criticism from consumers, policymakers, and academics, partly due to privacy risks. Privacy is becoming a major concern for utility companies as well, many of which have invested hundreds of millions of dollars over several years in their smart metering programs. Backlashes against the technology are extremely expensive to handle for utility companies, and recent incidents indicate that some may even face the possibility that their programs will be ousted due to continual public outcry [1-3]. The reality, though, is that the technical functionality, efficiencies, and cost savings offered by smart meters are too much for the industry to continue to resist.

Researchers have suggested the idea of the privacy-aware smart meter [4-6]. This type of meter would be able to provide all of the functionality of the standard smart meter – including two-way wireless communication, demand response enablement, and reliable billing – while preserving the privacy of the consumer. Meanwhile, relatively little is known about how consumers feel about their smart metering privacy. What leads some consumers to accept smart meters and the privacy risks that come with them, while others continue to resist the technology? Is the minority simply more aware of how much their smart metering data can reveal about them? Or, do the consumers who accept installations of smart meters at their residences prefer the convenience and cost-savings offered by the program? What other issues, including demographic factors, exposure to positive or negative media, or level of trust in the utility company might factor into consumer decisions on smart metering and privacy?

To address these issues, a national survey of American homeowners and electricity consumers was designed to answer the following set of research questions:

1. *How aware are electricity consumers of the privacy risks of smart metering?*
2. *How much is smart metering privacy worth to the average electricity consumer?*
3. *Does owning a smart meter change the consumer's relationship with the utility company?*
4. *How does exposure to media change a consumer's decisions on smart metering?*

In the survey, a conditional choice model was used to determine the average participant's willingness to pay for the smart metering privacy. Also included in the survey were questions asked to determine the participant's feelings and beliefs related to privacy and trust in the utility company. In section 6.2, an overview of the design of the survey study is provided. Results are then presented in the next four sections, with section 6.3 focusing on consumer awareness of privacy risks, section 6.4 on consumer willingness to pay for privacy, section 6.5 on the effect that owning a smart meter has on the consumer smart metering choices, and section 6.6 on the effect of media on consumer smart metering choices. In section 6.7, analysis of the results from the survey is used to provide policy recommendations on smart metering adoption and privacy. Section 6.8 provides concluding remarks on this chapter.

6.2 EXPERIMENTAL DESIGN

In this study, a survey was administered to a national sample of household owners who pay regular electricity bills to better understand electricity customers' views on the privacy of their power consumption data. The survey was aimed at determining their valuation of the privacy of their fine-grained power consumption data, which is the type of data that is typically

collected by a smart meter. A conditional choice model was employed to determine privacy valuation. In other parts of the survey, participants are asked about their preferences related to electricity bills, environment issues, and utility companies. Approval from the Institutional Review Board was obtained prior to the study, which qualified for the exempt protocol.

6.2.1 Survey design

In the design of the survey, various methods in line with current survey methods were used to engage the participants and maximize response rate [7]. In the first part of the survey, questions were designed to be interesting and fun to answer, increasing the likelihood for participants to remain interested throughout the survey. Further, the survey was written in a conversational tone, and the use of advanced terminology was avoided as much as possible. Answer formats were restricted to multiple choice single answer, multiple choice multiple answer, and open-ended.

6.2.2 Recruitment

The survey participation group consisted of 233 individuals from across the United States. Participants were required to be paying regular (i.e., monthly) electricity bills to their local electric utility company at the time of the survey, but could either rent an apartment or own a house. To obtain a random sample of participants fitting these requirements, StudyResponse

Project, a non-profit social science research service that facilitates survey research, was solicited.

Initially, a prescreening survey was sent to 5,200 individuals. In addition to age, gender, location of residence, and StudyResponse ID number, the prescreening survey asked the following four questions:

1. *Do you or your family own your residence?*
2. *If not, do you live in a rented apartment?*
3. *What kind of electricity meter do you have at your residence?*
4. *Has your electricity company replaced or serviced your electricity meter in the past five years?*

The answers were in multiple choice format. To aid participants in answering question 3, each answer choice (including "standard meter" and "smart meter") was accompanied by a photo. The last two questions were used to determine whether participants had an analog meter or smart meter at home. To avoid false answers, no information was provided on the type of metering device required for participation in the final study.

After receiving 1365 responses, the prescreening survey was closed. Using the criteria above, 150 individuals with smart meters and 150 with analog meters were invited to participate in the study via email. The content of the emails included a consent statement in which the participants were told (1) what the research purpose of the survey was, (2) what they would be required to do as participants, (3) that they would receive $12 to participate in the survey, and (4) that their participation was voluntary and that their information would remain confidential.

The survey was launched in July 2012 and responses were recorded over three weeks. Of the 300 individuals invited to partake, 233 completed the survey, resulting in a response rate of 77.3%.

6.2.3 Participant demographics

Ages of the 233 individuals who completed the survey ranged from 24 to 78 years, with mean 41 years. Females comprised 48.9% of the sample, and males 51.1%. Average household income was about $150,000, and the average household size was 3.34 people. 110 people reported that they lived in urban areas, 93 in suburban areas, and 30 in rural areas. Participation in the survey was limited to people living in the United States. Of the 300 study invitees, half had indicated in the prescreening survey that their residential metering device was a smart meter, and the other half had indicated that it was an analog meter. In the final subset of 233 participants, 126 indicated that they have smart meters at their residences, and 107 indicated that they have analog meters.

The original survey can be accessed using the link in the chapter appendix. To analyze the survey results, 20 variables were defined, each of which represents a key demographic or behavioral indicator for consumer decisions. Each of these variables is based on the answers to one or more questions in the survey. When any of these variables is referred to in this chapter, monospace type is used. For example, trust, which measures consumer trust in the utility company, is based on participant answers to question 39 in the survey.

To illustrate a significant relationship between two variables, *p*-values resulting from Pearson chi-squared (χ^2) contingency tests are reported, which were performed of the null hypothesis that the two tested variables are independent. Typically, a *p*-value less than 0.05 rejects the null hypothesis, suggesting that the value of one variable affects the value of the other.

6.3 AWARENESS OF PRIVACY RISKS

Parts A and B of the survey focused on collecting information about participants' attitudes and preferences related to their electricity bills and environmental issues, respectively. Prior to Part C, participants were not educated on the privacy risks of smart metering. In Part C, participants were asked to consider a hypothetical situation as follows:

> *Your local electric utility company is interested in giving you a new meter, with free installation. You receive a phone call at your home and are given the following two options:*
>
> *Option 1: Retain your standard meter. (If you choose this option, your bill stays the same as before. You will be charged for electricity at a predetermined fixed rate.)*
>
> *Option 2: Choose the Smart Meter. (You will be charged at the current market price, which is usually high during the day and low during the night. Choosing a Smart Meter would allow you to reduce your CO2 emissions and lower your electricity bills.)*

The participants were then asked a series of four questions, all using the following wording:

If you could save X% on your electricity bill if you chose the Smart Meter, what would you decide?

The percentage *X* refers to savings on the participant's electricity bills. The value of *X* was progressively altered over the four questions. In the first question of the series, *X* was set to 0; in the second, 5; in the third, 10; and in the fourth, 20. Results are shown in Table 5.

Table 5: Smart metering choices without knowledge of privacy risks of smart metering.

Savings on electricity bill (*X*)	Percentage that chose standard meter	Percentage that chose smart meter
0%	39.5	60.5
5%	30.9	69.1
10%	21.9	78.1
20%	13.7	86.3

At the start of Part D, participants were asked to read a paragraph that outlined the privacy risks of smart metering. This paragraph mentioned that smart metering data can be used to determine when someone sleeps, takes a shower, or uses a microwave. Participants were then placed in same hypothetical situation and asked the same series of four questions as in Part C. Results are shown in Table 6.

Table 6: Smart metering choices with knowledge of privacy risks of smart metering.

Savings on electricity bill (*X*)	Percentage that chose standard meter	Percentage that chose smart meter
0%	53.2	46.8
5%	39.5	60.5
10%	29.6	70.4
20%	21.9	78.1

Comparison of Tables 5 and 6 reveals three key insights. Firstly, a significant subset of participants was not aware that there are privacy risks associated with smart metering, because many people changed their decisions after learning about what smart metering data can reveal about their lifestyles. Secondly, the results indicate that many people would place significant value on their smart metering privacy, because they chose to retain the standard meter, which in the hypothetical situation would keep their bills high, in order to protect their privacy. Thirdly, although participants chose the smart meter less often after they had knowledge of the privacy risks, they were still willing to use the smart meter provided the savings on their bill were high enough.

These results indicate that participants placed a price on their privacy, and that given sufficient compensation, they would be willing to give up their personal information. What is this threshold price? That is, what is the average electricity consumer's willingness to accept money for smart metering data? This and other questions are tackled in section 6.4.

6.4 THE PRICE OF PRIVACY

While power consumption data holds value to utility companies and advertising agencies among other corporations, it also holds significant value to consumers, especially if the data reveals sensitive information about in-home activity. But how much do individuals value this information? Can a monetary value be assigned to this valuation for the average electricity consumer? And what factors, demographic or otherwise, are indicators of this valuation across the population of consumers?

Part E of the survey used conditional modeling to answer these questions. Participants were asked a series of questions, all using the following wording:

> *If you could save X% on your electricity bill if you chose the Smart Meter, or you could save Y% if you chose the Privacy-Aware Smart Meter, what would you decide?*

The percentages *X* and *Y* refer to savings on the participant's electricity bills. The monetary values shown here as *X* and *Y* were progressively altered. In the first question of the series, *X* and *Y* were set to 0 and -1 (that is, a 1% increase), respectively; in the second, 5 and 4; in the third, 5 and 2; in the fourth, 10 and 9; in the fifth, 20 and 19; and in the sixth, 20 and 17. Results showed that participants responded positively to the protection of their privacy, with higher proportions of participants opting for the privacy-aware smart meter as its differential price was decreased over the four questions, as shown in Table 7.

Table 7: Smart metering choices with knowledge of privacy risks of smart metering and privacy-aware smart metering option.

X and *Y*	Percent that chose analog metering	Percent that chose standard smart meter	Percent that chose privacy-aware smart meter
$X = 0$; $Y = -1$	34.9	31.1	34.0
$X = 5$; $Y = 4$	23.8	29.8	46.4
$X = 5$; $Y = 2$	19.6	33.2	47.2
$X = 10$; $Y = 9$	14.5	32.3	53.2
$X = 20$; $Y = 19$	9.4	31.5	59.1
$X = 20$; $Y = 17$	11.9	31.5	56.6

The questions were developed in this way such that the random utility conditional logit model could be applied using the results from Table 7 [8]. In the survey, participants were asked how

much money they would be willing to pay to preserve their privacy by opting for the privacy-aware smart meter. Therefore, each participant revealed his or her willingness to pay for the privacy of personal smart metering data.

To derive an economic model for the average consumer's willingness to pay for privacy, a utility function is used in separable form to represent consumer utility as follows:

$$U_{njt} = -\alpha_{njt} p_{njt} + \beta_n' x_{njt} + \varepsilon_{njt} ,$$

where p is the value of the monthly electricity bill and x is a set of dummy variables indicating whether or not the metering choice protects (1) protects privacy and (2) is a smart meter. These variables have coefficients α and β, which determine how impactful p and x are to consumer utility. Indices n, j, and t are used to specify a consumer number, survey question, and answer choices for question j, respectively. Furthermore, because there are unobservable factors that may affect a participant's decisions, the variance of ϵ_{njt} is allowed vary across different subjects. This model is based on the results of the panel data from the conditional choice model described earlier. Therefore, it cannot be assumed that ϵ_{njt} is independent and identically distributed. To overcome these issues, the model is transformed to:

$$U_{njt} = - (1_n / k_n) p_{njt} + (\beta_n^1 / k_n)' x_{njt} + \varepsilon_{njt} .$$

Because scale of utility does not change behavior, this model can be assumed to estimate the choices of the participants effectively. In the context of this survey, the model can be written as:

$$U_{njt} = -\alpha_n \, Monthly\ Electricity\ Bill_{njt} + \beta_n^1 \, Private_{njt} + \beta_n^2 \, Smart_{njt} + \varepsilon_{njt} .$$

The negative sign on the coefficient for *Monthly Electricity Bill* was included to restrict the coefficient to be positive in the estimation calculations. *Private* and *Smart* are dummy variables and denote whether or not the participant chose the privacy-aware smart meter and whether the participant had a residential smart meter, respectively.

A mixed logit model was used to estimate coefficients to determine the correlation structure in the error term. Estimation results provided distributions of coefficients in the population. In the following estimate, it is assumed that the coefficient for *Monthly Electricity Bill* is log normal and the other two coefficients are normal:

$$U_{njt} = -e^{a}{}_{n} \, (Monthly\ Electricity\ Bill)_{njt} + b_n{}^{1} \, Private_{njt} + b_n{}^{2} \, Smart_{njt} + \varepsilon_{njt} \, .$$

Table 8: Estimation results for logit model coefficients.

Random Coefficients			*Coefficient Distribution*		
Coefficient	**Mean**	**Standard Deviation**	**Coefficient**	**Mean**	**Standard Deviation**
a_n	-3.380 (0.275)	2.461 (0.186)	α_n	0.0340	14.50
$b_n{}^{1}$	2.134 (0.377)	4.458 (0.481)	$\beta_n{}^{1}$	2.134	4.458
$b_n{}^{2}$	1.882 (0.386)	-3.836 (0.389)	$\beta_n{}^{2}$	1.882	3.836

This model estimates that consumers are willing to pay approximately $11 per month to ensure the privacy of their electricity usage data. In other words, the average consumer is willing to pay $11 to use privacy-aware smart metering instead of standard smart metering.

This result is critical, as it illustrates that the market for privacy exists, and that it is up to utility companies and energy regulators to provide it to consumers. Recent events involving public

outcries against smart metering have shown that offering privacy up front may ultimately be in the interest of utility companies.

6.5 OWNING A SMART METER

Survey participants who had residential smart meters responded to questions in the survey having to do with their relationship to the utility company very differently from others. Notably, participants with smart meters believed the utility company uses smart meters to monitor residential power usage (encapsulated as the variable *eu_monitors* in Table 9) more often than the rest of participants. Not surprisingly, participants with smart meters also valued privacy less (*priv_value*) and trusted their utility company (*trust*) more than the remainder of the sample. Similarly, they cared about the utility monitoring their smart metering data (*care*) less than others did and on average felt that it is convenient if the utility company possesses their data (*convenience*) more than the general population. These results are summarized in Table 9.

Table 9: *p*-values and correlation values for *has_sm* variable.

	eu_monitors	***priv_value***	***trust***	***care***	***convenience***
χ^2 *p*-value	0.00285	0.00208	2.71E-06	0.000565	2.15E-10
Correlation	0.205	-0.273	0.335	0.236	0.425

One interpretation of these results is that most consumers who already have smart meters become used to them and appreciate the benefits of the utility company having their data, since this potentially lowers their bills and benefits the environment. If there are no significant

privacy-related incidents during the time they have smart meters, then these consumers also don't have reason to doubt the utility company's intentions in collecting smart metering data. While the argument can still be made that public outcries against smart metering are started by a minority of consumers with residential smart meters, no significant evidence was found to support the existence of such a group.

6.6 THE ROLE OF MEDIA

The effectiveness of videos in affecting public opinion on smart metering was tested in this study. Two videos with opposite agendas were developed, and the survey participants were made to watch them prior to taking the survey. In the first video, a middle-aged female executive at a public utility company describes the benefits that a household can derive from having a residential smart meter. The video, which uses footage produced by a major American utility company, features a mother and daughter monitoring their power consumption on a computer and identifying ways in which they can reduce their electricity bills. This first video (*sm_vid*) can be thought of as media that a public utility might produce to promote smart metering. In the second video, an experienced male electricity consumer highlights the privacy risks of smart metering. In this video (*priv_vid*), the speaker uses a very powerful tone and incites emotions reminiscent of moral panic. The video represents media that might be proliferated by privacy watchdogs in an effort to provoke a public outcry against smart metering and public utility companies. Links to the videos are provided in the chapter appendix.

The 300 survey invitees were randomly divided into four treatment groups. Prior to answering the survey questions, each group was made to watch both videos, only the video on smart metering benefits, only the video on privacy risks, or neither video. It was found that none of these treatment groups answered questions significantly differently from the other groups, and neither video had a statistically significant effect on how people felt about smart metering or its related privacy issues. These results are shown in Table 10.

Table 10: *p*-values and correlation values for *sm_vid* and *priv_vid* with privacy-related variables.

		priv_value	***trust***	***care***	***retain***
sm_vid	χ^2 *p*-value	0.294	0.882	0.346	0.272
	Correlation	0.0406	0.0730	-0.0716	0.0875
priv_vid	χ^2 *p*-value	0.865	0.837	0.522	0.980
	Correlation	0.0165	-0.0586	0.0519	-0.00403

This result suggests that partisan rhetoric has limited effects on consumer psychology. Though further evidence is needed to support this hypothesis, it seems it is important for stakeholders in smart metering and privacy to assess the media they use to reach consumers and determine whether it really answers questions that consumers have about the technology, or whether it will simply be perceived as overtly biased material.

6.7 IMPLICATIONS

The results presented in previous sections reveal a wealth of information about consumer decision-making on smart metering and privacy. This section highlights implications of the survey results.

6.7.1 Provide a market for smart metering privacy

The fact that the economic model estimates that electricity consumers value their monthly privacy at $11 indicates that there is a significant market for privacy. The result shows that consumers are very willing to pay a significant premium in order to ensure the protection of their information. Over a large network of electricity consumers, the premium for privacy would amount to a significant share of the revenue for a utility company.

Utility companies may be hesitant to provide privacy-aware smart metering options because of the initial development costs of the meter and other physical infrastructures. They may also wish to collect fine-grained data on consumers so that they can create a market for advertising directed at target demographics. However, the results suggest that utility companies may still be able to make money out of providing privacy-aware metering, depending on development costs.

Furthermore, by implementing privacy-aware smart meters, utility companies can gain the trust of their customers. As shown in Figure 10, only 53% of the survey participants indicated that they trust their utility companies somewhat or very much. Ensuring privacy can show customers that the utility companies care about protecting their customers while providing them with the technical functionalities of a smart meter. This can raise customer trust in their utility

companies, which in turn can lead to decreased risk of public outcries and privacy lawsuits against smart metering.

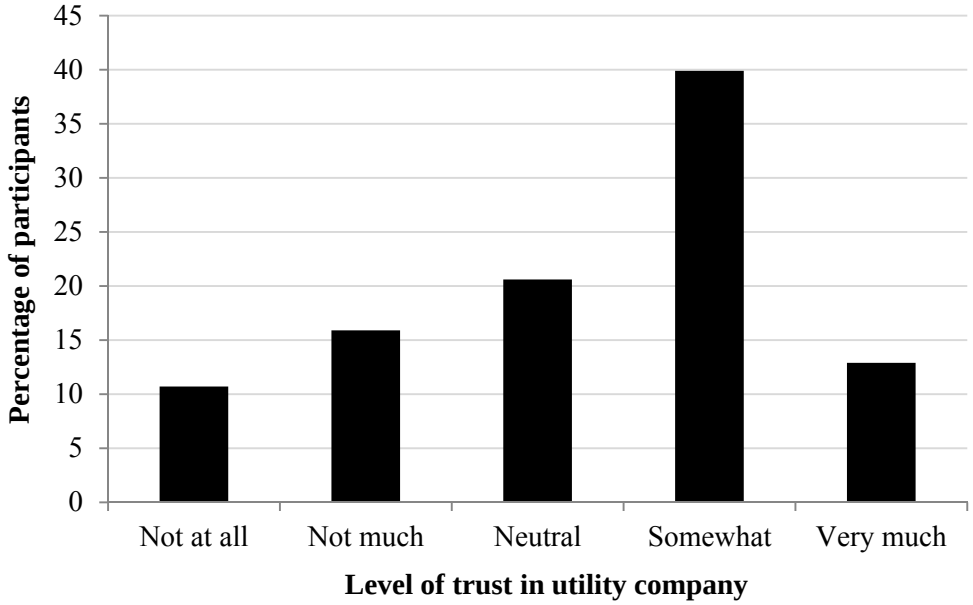

Figure 10: Distribution of level of trust in utility companies among survey participants.

6.7.2 Educate consumers on privacy

As has been shown, watching either the video on the benefits of smart metering or the video on the risks of smart metering prior to taking the survey had no significant impact on how participants responded to the survey questions. The content of both videos was overtly partisan, with one reflecting the opinions that an electric utility company might have and the other reflecting those of a consumer privacy lobbyist. The survey results suggest that, when making

decisions on power metering privacy choices, consumers will not be affected by exposure to biased rhetoric.

Notice, though, that it is not the case that survey participants did not understand that there are privacy risks associated with smart metering. Instead, participants placed a great deal of value on their personal privacy. Presumably, there were other information sources from which participants learned about the value of their smart metering data, including information presented to the participants within the survey itself, and information gathered prior to administration of the survey. The results presented in sections 6.3 and 6.4 suggest that, if informed, consumers do place value on their privacy, which will accordingly affect their power metering choices.

In order to empower consumers to make the choices that are best for themselves, it is critical to inform them objectively of all the benefits and risks associated with smart metering while avoiding biased rhetoric. Utility companies or their regulation agencies that wish to implement smart metering can do this through direct mailings to consumers detailing smart metering benefits and risks. Additionally, the impact that smart metering has on privacy and health should clearly be communicated to consumers. This information should include details on the granularity of data that is to be collected and to what extent collected data can be used to make inferences about the lifestyle of the consumer.

6.8 CONCLUSIONS

It is critical to emphasize the importance of considering consumer privacy during the development stages of new technologies, and the role of motivation in the adoption of those technologies. With widespread dispersal of private information, corporations seeking wide adoption of a technology can utilize that information in organizing movements and designing media that would specifically and unwittingly influence consumers to adopt their product. This imbalance of corporate over individual power for the profit of the corporation threatens the very process of consumer decision making. Implementation of privacy-aware technology must be pursued whenever possible to obviate this risk.

6.9 REFERENCES

[1] "BC Hydro backs down on smart meter installation," *CBC News*, January 2013.

[2] A. Regaledo, "Rage Against the Smart Meter," *MIT Technology Review*, April 2012.

[3] A. Young, "Committee to explore smart meter option for Lubbock," *Lubbock Avalanche-Journal*, February 2013.

[4] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proceedings of the 10th ACM Workshop on Privacy in the Electronic Society*, pp. 49-60, October 2011.

[5] S. Wicker and R. Thomas, "A Privacy-Aware Architecture for Demand Response Systems" in *Proceedings of the 44th Hawaiian Conference on System Science (HICSS-44)*, Kauai, Hawaii, January 2011.

[6] M. Lisovich, D. Mulligan, and S. Wicker, "Inferring Personal Information from Demand-Response Systems," *IEEE Security & Privacy*, vol. 8, no. 1, pp. 11-20, January–February 2010.

[7] D. Dillman, *Mail and Internet surveys: The tailored design method*, New York, Wiley, 2000.

[8] D. McFadden, "Econometric Models for Probabilistic Choice Among Products," *The Journal of Business*, Vol. 53, No. 3, pp. S13-S29, July 1980.

6.10 APPENDIX

1. A live version of the Smart Metering Privacy survey administered in this study can be found at:

   https://cornell.qualtrics.com/SE/?SID=SV_8Cj1Vlj2cnz0HZO&Preview=Survey&BrandID=cornell.

2. The video on the benefits of smart metering can be found at:

   http://www.youtube.com/watch?v=BEUaVikl1vQ.

3. The video on the privacy risks of smart metering can be found at:

   http://www.youtube.com/watch?v=OUoFDJ5pO3s.

CHAPTER 7.

LOCATIONAL TECHNOLOGY PRIVACY SURVEY

The results of the last chapter showed us that on average, consumers value their smart metering privacy at approximately $11 per month. But in order to provide additional basis for the privacy valuation method discussed in Chapter 6, as well as to extend our results to other cases in which consumers and corporates compete over the value of individual consumer data, it proves useful to conduct a similar study for the consumer-corporate stakeholder problem in at least one more context to better advise our final recommendations on the introduction of privacy-aware technologies to consumer markets. As we saw in chapter 3 on assurance of privacy in vehicle-to-grid technology, locational privacy issues are of great significance in the protection of consumer privacy, as location can be used to infer an individual's likes and dislikes.

In determining valuation for locational privacy through a consumer survey, it is best to refer to a technology that makes use of fine-grained location data that survey participants use every day and can understand easily. But while locational privacy is a critical matter that smart grid engineers will need to consider in the way forward, there currently are no popular or widespread technologies in power grids to which laypeople can relate.

Smartphone technology provides a suitable proxy for upcoming smart grid technologies that may use location data, such as those for vehicular applications, and is particularly interesting due to its wide consumer base and the richness of data associated with its use. Further, it is a sufficiently different context from the smart metering application studied earlier, but still

encompasses the broader issues associated with of the consumer-corporate situation. Thus in this chapter, we will discuss the design and results of a national survey that was administered to 300 smartphone users in the United States.

Privacy risks abound for users of smartphones, but they are often unaware of these risks or undervalue their impact. This chapter presents the results of a national survey on smartphone privacy that investigates user awareness of privacy risks related to smartphones, including location privacy, cellular data privacy, and network usage privacy issues. Additionally, we shall employ a conditional choice model to determine how much location privacy is worth to smartphone users. Our findings suggest that many users are either unaware of the privacy risks associated with smartphone use or do not know how to effectively protect their privacy. Further, we find that on average, participants are willing to pay about $12 each month to protect their location privacy in the smartphone context. We use these results to design policy recommendations for stakeholders in the business of collection of fine-grained locational data.

7.1 INTRODUCTION

Mobile phones have become ubiquitous. In the United States, the penetration rate for mobile phones in 2012 was 102.4% [1]. There are currently about 130 million smartphones in use across the country for everything from making phone calls to engaging in social media through third-party applications. Smartphone users are increasingly centralizing their communications, work and play activities onto one device that goes everywhere with them, a concept known as cellular convergence [2]. Data is constantly being collected about users' activities through their mobile phones, some of which is potentially invasive to individual privacy. The nature of cellular technology requires handsets to communicate with cell towers in order to route calls and enable online activities when Wi-Fi is unavailable. To facilitate incoming calls and handoffs, this requires the network to know the geographic location of a given handset.

In some cases, additional location-based tracking capabilities are added to mobile phones and activated without the knowledge of their users. This was the case when researchers discovered in 2011 that a file stored on Apple's iPhone contained time-stamped user-location data under the auspices of improving location-based services [3]. A similar case surfaced later that year when a developer discovered that Carrier IQ software, which could monitor keystrokes, came preinstalled on Android devices. The company contended that its software was designed to help service providers diagnose network and device issues and that no personal communications were collected, but this is a position that has been contested by privacy advocates.

Given the level of precision and granularity at which data can be collected about users, there are significant privacy risks associated with using mobile phones. The focus for this study is to assess mobile users' awareness of privacy risks and how this relates directly to 1) users' value

of mobile phone privacy and 2) how cell phones are routinely used. To explore these questions, we conducted an online survey issued to 300 participants. The survey included questions about respondents' use of mobile apps, knowledge of mobile location tracking capabilities and general views on privacy law.

Results of the survey reveal that individuals who are aware of data collection relating to mobile phone usage have greater privacy concerns than those who do not. Perceived privacy risks and trust in cellular service providers play a role in determining the types of smartphone activities that users engage in. To better understand this relationship between awareness, perceived risks and trust, we contextualize the survey results using the Internet Users' Information Privacy Concerns (IUIPC) [4]. The IUIPC is a theoretical framework developed by Malhotra et. al. to systematically identify information privacy concerns among Internet users. IUIPC divides users' information privacy concerns into different dimensions. These dimensions can then be used to explain how privacy concerns affect users' trust of cellular service providers and users' perceived privacy risk in using mobile phones. Trust and risk ultimately impact the decisions that users make about how to use their mobile phones and the information they choose to share.

7.2 BACKGROUND

7.2.1 Related work

There is a vast literature on cellular privacy. While previous research has explored individuals' privacy concerns around mobile phone use, no study has sought to examine the relationships between awareness of privacy risks, privacy concerns and user trust of related third parties

(cellular service providers, phone manufacturers, app providers) and how these factors influence mobile phone use.

A comprehensive survey of over 3,000 smartphone users' privacy concerns conducted by Felt et. al. resulted in the ranking of 99 different privacy risks. Their analysis discovered that location-related risks are in the bottom half of all risks perceived by Android users. They were also able to understand how users would react to various privacy incidents if they were to actually occur [5]. With regards to privacy concerns of location-based services using mobile phones, Barkhuus and Dey found that individuals had more privacy concerns about location-tracking services (which involve the tracking of user activity by a third party such as a service provider) compared to position aware services (which are based on the device's own knowledge of its position) [6]. Muslukhov et al. explored the different types of data that mobile phone users are concerned about protecting and examined their current practices for securing this information [7]. When Knijnenburg et al. examined the privacy concerns involved in personalizing the mobile experience through an application recommender system, they found that when individuals were more adequately provided with information disclosure notices, they were less likely to disclose this information, less trusting of the app provider, and less satisfied with the recommender system itself [8].

7.2.2 Internet users' information privacy concerns (IUIPC)

The framework describes three different dimensions for privacy concerns: (1) collection, (2) control, and (3) awareness of privacy practices. Collection is defined as the extent to which

individuals are concerned about the amount of personal data aggregated by others in relation to the value of the benefits of this tradeoff. Control is the ability of individuals to make decisions about who has access to their personal information and how it is used. Awareness is the degree to which individuals are informed and understand how organizations use the data that they collect about them and what types of privacy-related risks this might entail.

IUIPC is based on social contract theory, which articulates that activities and transactions between two parties in society are based on informed consent and framed by opportunities for expression and rights by either party to exit such a relationship. Thus, "a firm's collection of personally identifiable data is perceived to be fair only when the consumer is granted control over the information and the consumer is informed about the firm's intended use of the information".

The survey results we present in the following section reveal that cellular technology is a black box to most mobile phone users. This framework illustrates that without users being aware of what type of data is collected, the way in which it is collected and what it is used for afterwards, users may not even realize that they have a reason to be concerned about their privacy, or that control of their personal data (information about their communications and mobile phone usage) has been delegated to a third party.

Malhotra et. al. also created a causal model to describe how IUIPC affects a consumer's decision to release or not release personally identifiable data. For the purposes of this study, the causal model can be used to understand how IUIPC influences general smartphone usage, given that many activities require disclosure of such data. The causal model affirms that within

a specific context, individuals who trust a third party receiving their cellular data will have a lower perceived risk of privacy issues, making them more likely to share personal information.

7.3 EXPERIMENTAL DESIGN

We administered a survey to a national sample of smartphone users who pay regular cellular usage bills in order to determine user valuation of the privacy of their cellular data. A conditional choice model was employed to determine privacy valuation. Additionally, participants are asked about their preferences related to cellular usage bills, environment issues, and utility companies. IRB approval was obtained prior to the study, which qualified for the exempt protocol.

7.3.1 Survey design

In the design of the survey, various techniques in line with current survey methods were used to engage the participants and maximize response rate [9]. The survey opened with a motivational screen, followed by simple instructions on how participants should proceed. In the first part of the survey, questions were designed to be interesting and simple to answer, increasing the likelihood for participants to remain interested throughout the survey. Further, the use of advanced terminology was avoided as much as possible. Answer formats were restricted to multiple choice single answer, multiple choice multiple answer, and open-ended. A link to the full survey is provided in the chapter appendix.

7.3.2 Recruitment

The survey participant group consisted of 300 individuals from across the United States. Each participant owned a smartphone and paid regular (i.e., monthly) cellular usage bills at the time of the survey. StudyResponse Project, a non-profit social science research service that facilitates survey research, was used to obtain a random, national sample of participants fitting these requirements.

Initially, a prescreening survey was completed by 520 individuals. In addition to age, work status, location of residence, and StudyResponse ID number, the prescreening survey included the following questions:

- *Do you own a cell phone?*
- *If you do own a cell phone, is it a smartphone?*
- *Can you browse the web using your phone?*
- *Can you check your emails on your phone?*
- *Who is the manufacturer of your phone?*
- *What is the model name of your phone?*

If participants answered that their phone is a smartphone, then they would see additional questions. The last several questions were used to determine whether the participant had a smartphone or not. If participants' answers to the last several questions suggested that they owned a smartphone, then it was assumed that they did.

After receiving 520 responses, the prescreening survey was closed. Using the criteria above, 300 smartphone owners were randomly selected and invited to participate in the study via email. The contents of the emails included a consent statement in which the participants were told (1) what the research purpose of the survey was, (2) what they would be required to do as participants, (3) that they would receive $12 to participate in the survey, and (4) that their participation was voluntary and their information would remain confidential.

The survey was launched in November 2012 and responses were recorded over three weeks. Of the 300 individuals invited to participate, 232 completed the survey, a response rate of 77.3%.

7.3.3 Participant demographics

The ages of the 232 individuals who completed the survey ranged from 20 to 77 years, with mean 39 years. Average household income was about $130,000. 118 people reported that they lived in urban areas, 91 in suburban areas, and 23 in rural areas. Participation in the survey was limited to people living in the United States. Participants took an average of 28 minutes to finish the survey.

The results of data analysis presented in the following sections refer to specific survey questions. To illustrate a significant relationship between two variables, we report *p*-values resulting from Pearson chi-squared (χ^2) contingency tests, which were performed on the null hypothesis that the two tested variables were independent. In this case, variables are the collective set of participants' answers to a question in the survey. Typically, a *p*-value less than

0.05 rejects the null hypothesis, suggesting that the value of one variable affects the value of the other. The chi-squared results presented in the following sections therefore indicate that participants' responses to one question affected their responses to another question.

7.4 FINDINGS AND IMPLICATIONS

7.4.1 The value of locational privacy

Prior to and during the survey, participants were educated on the risks of privacy associated with using a smartphone. Participants were then asked a series of four questions aimed specifically at determining how much utility they would theoretically derive from a private smartphone – one that uses a set of cryptographic keys and trusted computing to enable location-based services but specifically protect location privacy – as opposed to a standard non-privacy-aware smartphone. These questions all took the following format:

> *If the monthly service contract price were X for a standard smartphone or Y for a private smartphone, which would you choose?*

The monetary values shown here as *X* and *Y* were progressively altered. In the first question of the series, *X* and *Y* were both set to \$50; in the second, *X* was \$50 and *Y* was \$55; in the third, \$50 and \$60; and in the fourth, \$50 and \$70. Results showed that participants responded positively to the protection of their privacy, but with lower numbers of participants opting for privacy as its differential price increased over the four questions. Results are shown in the table below.

Table 11: Responsiveness to smartphone location privacy.

X; Y	Percentage that chose standard smartphone	Percentage that chose private smartphone
$50; $50	29.7	70.3
$50; $55	31.5	68.5
$50; $60	48.3	51.7
$50; $70	64.7	35.3

The questions were developed in this way such that the random utility conditional logit model developed by econometrician Daniel McFadden could be applied using the results from Table 11 [10]. The explanatory variables used in this regression included (1) the dollar cost savings, and (2) whether or not the choice retains privacy. Using this approach, it was estimated that the average participant was willing to pay $12.20 per month to ensure the privacy of location data collected by their smartphones.

This result is critical. It illustrates that the market for privacy exists, and that it may be beneficial to cellular service providers and phone manufacturers to provide it to consumers. In fact, recent events involving public outcries against smartphone privacy have shown that offering privacy up front may ultimately be in the interest of market stakeholders.

In terms of the IUIPC model, the results demonstrate that the ability for consumers to trust that cellular service providers will preserve their privacy is integral to ensuring that consumers will utilize the various capabilities of their smartphones without limiting such activities due to privacy concerns. This will work toward achieving the sense of fairness that the IUIPC model affirms must be present in order for social contracts or transactions between consumer and a third party provider involving personal information (such as location) to be successful [3].

7.4.2 Privacy awareness issues

The survey asked participants various questions about their awareness of the collection of data by cellular service providers. In this section, we highlight three key trends discovered from the analysis of the survey results.

Lack of knowledge of location technology

Participants were asked to rate their level of awareness that GPS location data is collected by service providers and third parties in question 24 of the survey. 53% answered that they were "not at all unaware" to "moderately aware" (Fig. 11).

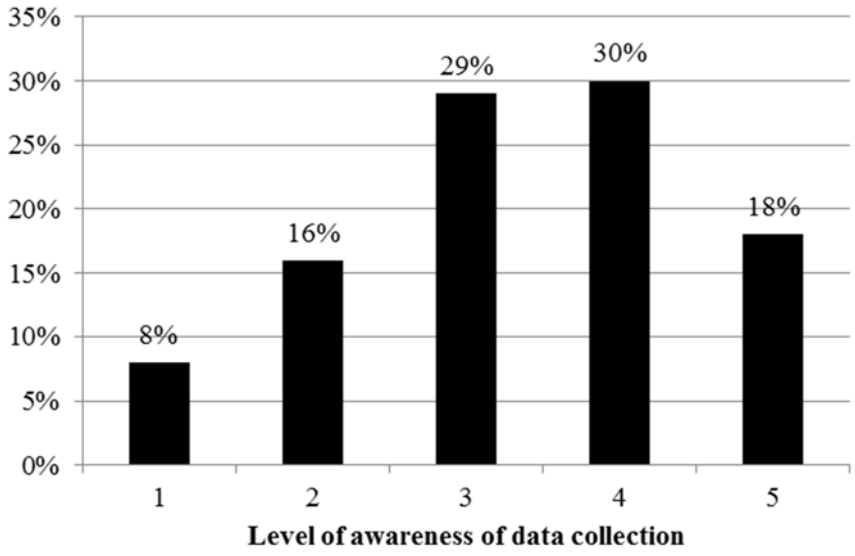

Figure 11: Distribution of data collection awareness among survey participants.

Meanwhile, participants were asked whether they think that cell phones have location determination capabilities even if GPS features are turned off. 30% of participants indicated that there are no other location-determining capabilities in a cell phone besides GPS. Although 48% of respondents claimed they were either "mostly aware" or "completely aware" of data collection activities, the data revealed that many people in this group did not know that location can be determined via cellular towers without the help of GPS technology (χ^2 p-value = 0.0111; corr. = -0.203).

This finding raises two issues. First, a large subset of smartphone users is to some extent unaware that location data is collected by service providers and permitted applications. Second, there are many people who claim to know that GPS data is collected who are unaware that location can also be determined via cellular network technology as well. This indicates a gap in understanding of how the technology works and suggests that many people may have a false sense of confidence that they know how their cellular usage data is handled.

The IUIPC model articulates that in order to achieve awareness of how data is collected, who has access to it and what is done with it after the aggregation process, a certain level of transparency and granularity must exist in order for the consumer to be well informed. The survey results suggest that this type of clarity, with regards to the technical capabilities of mobile phones for data collection and communication is greatly lacking.

Lack of awareness of currently available privacy-preserving smartphone settings

In order to determine whether smartphone users are aware of the privacy-preserving choices available to them, participants were asked a series of questions related to a set of privacy-related settings and choices available on most smartphones. Participants were presented with each setting and asked to indicate whether they currently apply the setting, do not apply it, or were unaware or unsure of how to activate it. Results are shown in Table 12 below.

Table 12: Use of privacy-preserving phone settings.

	Uses setting to protect privacy (%)	Aware but doesn't set setting (%)	Unaware or unsure of setting (%)
Turning "mobile data" off	36.0	42.6	21.5
Private browsing in mobile web browser	40.1	28.1	31.8
Enabling passcode lock on your phone	46.3	38.4	15.3
Turning location services or GPS off	35.5	45.9	18.6
Disabling sending of diagnostic data	33.9	35.5	30.6

Table 12 indicates that many smartphone users are unaware of certain privacy options or do not know how to activate them. This suggests that not enough is done to inform consumers of available privacy-preserving options on smartphones, and that many people may be giving up cellular data to service providers, phone manufacturers, and third parties unknowingly.

These privacy-preserving phone settings are mechanisms of control for the consumer, according to the IUIPC model. Because control over personal information is one of the key dimensions of privacy concerns, it is disconcerting that such a substantial proportion of those

surveyed are unaware of existing tools that they can use to actively protect their privacy. Perhaps a portion of privacy concerns expressed by individuals would be assuaged if they were better informed about these features.

Data collection awareness and use of smartphone applications

The survey results indicate there is a strong correlation between the use of applications and awareness that data is collected. Specifically, analysis revealed that the frequency with which an individual uses applications used is positively correlated with the individual's awareness of data collection (χ^2 p-value = 0.0237; corr. = 0.219).

One interpretation of this result is that there is a causal relationship between the use of advanced features of smartphones and the awareness of data collection practices. Advanced users of smartphones, including individuals who download and actively use applications, may somehow become aware of data collection practices through their use of smartphones. This may happen because they see more privacy policies, are more exposed to related news stories and incidents, or are subjected to collection of data more often than others.

Data collection awareness and trust in cellular service provider

Results from the survey also showed that data collection awareness and trust in the cellular service provider are positively correlated with each other (χ^2 p-value = 1.605×10^{-6}; corr. = 0.138). This may suggest that consumers appreciate being well-informed about data collection

practices. Accordingly, IUIPC finds that trust and risk beliefs mediate the impact of a consumer's privacy concerns upon behavioral intentions. Thus, the trust established between a consumer and a service provider as a result of greater data collection awareness can affect the decisions made by the consumer about mobile phone usage.

7.5 RECOMMENDATIONS

In this section, we outline recommendations for cellular stakeholders based on our findings from the survey.

7.5.1 Inform consumers about privacy risks

Results from the survey show that there are many smartphone users who are unaware of the privacy risks associated with data collection. There is also a large group of consumers who are unaware or unsure of how to enable privacy-preserving processes on their smartphones. Many social experts have noted technology has metaphorically been described as a "black box," an artifact that has a certain social value, but the inner workings of which are a mystery to most individuals [11]. Further, others have found that users typically prefer to secure their smartphone data as much as they can, but find it inconvenient to do this due to a present lack of viable solutions [7]. To close this gap and enable consumers to open the black box, service providers and phone manufacturers should aim to improve consumer understanding of cellular technology through various methods, including advertising campaigns detailing smartphones

and the benefits and risks of cellular data collection to consumers. The known impact that smartphone use has on privacy should clearly be communicated to consumers. This information should include details on the type and granularity of data that is to be collected and to what extent collected data can be used to make inferences about the lifestyle of the consumer.

7.5.2 Establish a market for privacy

The fact that the model of subsection 7.4.1 estimates that smartphone users value their location privacy at about $12 per month indicates that there is a significant market for privacy. The result shows that consumers are very willing to pay a significant premium in order to ensure the protection of their information. Over a large network of users, the premium for privacy would amount to a significant share of the revenue for a cellular service provider.

Nonetheless, consumers will be deterred by fees associated with privacy assurance. As such, the premium charged to assure privacy must be minimized to facilitate privacy adoption.

7.6 CONCLUSIONS

These results suggest consumers are relatively unaware of the many privacy risks associated with mobile phone usage. Because of this lack of awareness, consumers are unable to make educated decisions about the types of activities in which to engage on their smartphones. The result of such uncertainties is that consumers may hesitate to use certain applications or services unnecessarily or disclose an inordinate amount of personal information due to a false sense of

security, whether in the mobile context or others in the power industry and beyond. Consumers' relatively high privacy valuation for locational data suggests that a valuable opportunity exists for stewards of locational data to build trust with their consumers by being more transparent about their data collection and processing techniques and educating consumers about the real privacy risks associated with use of technologies that make use of fine-grained location data for functionality. Consumers' perceived risk is not always the same as actual risk; creating policy infrastructures toward aligning these two categories of risk should be an objective for service providers.

7.7 REFERENCES

[1] CTIA – The Wireless Association, *Year-End Semi-Annual Wireless Industry Survey*, May 2012.

[2] S. Wicker, *Cellular Convergence and the Death of Privacy*, Oxford University Press, August 2013.

[3] A. Alasdair and P. Warden, "Got an iPhone or 3G iPad? Apple is recording your moves," *O'Reilly Media*, April 2011.

[4] N. Malhotra, S. Kim, J. Agarwal, "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale and a Causal Model," *Information Systems Research*, 15(4), 336-355, December 2004.

[5] A. Felt, S. Egelman, and D. Wagner, "I've Got 99 Problems, But Vibration Ain't One: A Survey of Smartphone Users' Concerns," Technical Report, University of California, Berkeley, May 2012.

[6] L. Barkhuus and A. Dey, "Location-based Services for Mobile Telephony: a study of users' privacy concerns," in *Proceedings of the INTERACT 2003, 9th IFIP TC13 International Conference on Human-Computer Interaction*, Zurich, Switzerland, September 2003.

[7] I. Muslukhov, Y. Boshmaf, C. Kuo, J. Lester, and K. Beznosov, "Understanding Users' Requirements for Data Protection in Smartphones," in *ICDE Workshop on Secure Data Management on Smartphones and Mobiles*, December 2012.

[8] B. Knijnenburg, A. Kobsa, "Privacy in Mobile Personalized Systems: The Effect of Disclosure Justifications," *Workshop on Usable Privacy & Security for Mobile Devices (U-PriSM)*, July 2012.

[9] D. Dillman, *Mail and Internet surveys: The tailored design method*, New York, Wiley, 2000.

[10] D. McFadden, "Econometric Models for Probabilistic Choice Among Products," *The Journal of Business*, Vol. 53, No. 3, pp. S13-S29, July 1980.

[11] W. Bijker, T. Hughes, and T. Pinch, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, Cambridge: MIT Press, 2012.

7.8 APPENDIX

A live version of the locational data privacy survey can be found at:
http://bit.ly/19WPLyC.

CHAPTER 8.

CONCLUDING REMARKS

Ensuring the protection of individual privacy for consumers is a difficult challenge for technologists, policymakers, and government alike in today's digital economy. Modern technology makes use of sophisticated methods to collect, retain, and analyze data to reveal detailed insights on consumer behavior, often resulting in an asymmetry of power between powerful entities possessing vast amounts of information and average consumers. Engineers have developed solutions to counter this worrisome trend with "privacy-aware" technology – comprising technologies that provide functionalities while protecting the privacy interests of the end consumer. But many have noted that even though privacy-aware solutions can be developed for most applications, various conflicting economic factors are at play, and it will be difficult to find solutions that are acceptable to all stakeholders.

This dissertation has studied and brought together several issues associated with the provision and adoption of privacy in electricity consumer markets. Starting with a philosophical grounding, we moved on to establish that bulk collection of power consumption data using smart meters presents problems for individual privacy, referencing results that show with fine-grained consumption data, it is possible to determine when consumers are using water, using the microwave, or switching on lights. Further, we illustrated through a stochastic model that it is possible to predict such behaviors. Drawing on the specific example of vehicle-to-grid networks, we then presented a set of guidelines for the design of privacy-aware systems that make intensive use of data in the power industry such that privacy risks can in certain cases be

factored out of these emerging technologies while still preserving their key functionalities. Subsequently, we considered the issues associated with the consumer versus corporate stakeholder contest over access to power consumption data. After discovering that consumer privacy valuation is the key variable that needs to be determined in order to advise an optimal regulatory regime for the introduction of privacy-aware technology, we conducted a survey on consumers in which a conditional choice model was used to estimate valuation of privacy in the context of smart metering. The results showed that consumers valued the privacy of their fine-grained power consumption data at approximately $11 per month, or about ten percent of the participants' average monthly utilities bill. Further results from a separate survey indicated that consumers value the privacy of their locational data at about $12 per month. These two values should in conjunction advise any regulatory policy on the control of individuals' data in the power industry.

The most meaningful line of work related to this research is yet to come. Given the results of this work, the critical question to answer will be how policy should be shaped so that the potential economic opportunities presented by smart metering and other emerging technologies in the smart grid can be realized while still protecting the privacy of individual consumers. Various regulatory structures can be employed. For instance, policymakers in this space will need to determine whether or not opt-out regimes or opt-in regimes represent the interests of consumers; whether options to revert to analog systems are available to consumers; which entity will oversee changes to residential metering systems; and which parties will be responsible for administering newly-implemented intelligent technologies in the electrical system, among many others. Though these issues are to an extent addressed in this work, further analysis of these policy frameworks will be critical in the way forward to determine

how the electricity industry and others like it can best be managed by government and key stakeholders.

While pushing the boundaries of consumer technology is gradually making the world a better place by creating value for consumers, it is essential that engineers and policy administrators do not lose sight of the fact that consumer privacy must be given due consideration. We are living in an age of unbelievable growth in data storage capacity and processing power, and the cost to corporations of obtaining and storing information about individual consumers is dropping. Without adequate safeguards for consumers in their use of emerging technologies, we will continue to witness public outcry and backlash against such technologies that make undue use of data at the expense of individual consumers, which will in the long term hinder our pace of technological advancement and commercialization.

In an increasingly interconnected world, we must ensure we achieve the right balance of privacy and security to protect every user of technology. The Internet is spreading far and wide, and connections throughout the United States and the world are getting faster every day. Much of this is down to the sound work of U.S. and foreign policymakers, who have fought for and thus far largely secured an open Internet through which the world can access and disseminate information in democratic fashion. As the Internet and the vast sea of content it brings continue to spread, we must create and enforce a global understanding of trust whereby anyone in the world can feel safe to use technology with confidence.

Privacy-aware design is one way to accomplish this. Ultimately, engineers must learn to protect individual civil liberties by considering privacy not as an afterthought, but from the outset of design.