Real-Time Data Collection and Processing of Utility Customer's Power Usage for Improved Demand Response Control

Shawyun Sariri(✉), Volker Schwarzer, Dominik P.H. Kalisch, Michael Angelo, and Reza Ghorbani

2540 Dole Street Homes Hall #302, Honolulu, HI, USA
{shawyun,volkers,mangelo,rezag}@hawaii.edu,
dkalisch@trinity.edu

Abstract. A large growth in energy demand has increased renewable energy penetration into existing power grid infrastructures, as well as spurring increased research into demand response programs. But before implementing an efficient demand response program, it is first necessary to understand the power usage behaviors of a consumer. This paper presents a real-time data acquisition system for the collection and storage of power data that will allow the study of demand response in an urban area. Demand response programs are an ideal alternative to costly energy storage and spinning reserves. Detailed power consumption data is necessary to study proper demand response programs and implement efficient control decisions. A pilot system has been implemented on the island of Oahu in Hawai'i to prove the feasibility of a data collection system in a dense urban environment. The pilot program has implemented a smart metering device that is collecting power data at a high resolution and transmitting it to a server for load forecasting analysis. The architecture of the system will be discussed as well as preliminary results and scalability of the pilot system as it relates to the implementation of the system into a large urban center.

Keywords: Demand response · Load forecasting · Power profile signature · Urban center

1 Introduction

A 2013 report by the American Society of Civil Engineers (ASCE) gave the American electrical grid a "D+" rating, on an A to F scale [1]. Operation failures were mentioned as a main source of outages across the country because of congestion in transmission lines. Utility companies are relying on a current grid infrastructure that still has components from the 19th century. Expanding US energy capacity after 2020 will be a main concern for the utilities, and one way to alleviate some of the pressure of capacity expansion will be to increase consumer side power generation using renewable resources, but the addition of more generation comes with logistical issues, especially when transmitting power generated from stochastic sources. Rather than invest large amounts of money replacing the current grid infrastructure, the ASCE suggests research

A. Hameurlain et al. (Eds.): TLDKS XXVII, LNCS 9860, pp. 48–63, 2016.
DOI: 10.1007/978-3-662-53416-8_4

in smart grids and real-time forecasting as alternatives. Smart grids will be necessary in areas containing dense populations such as large urban centers.

With a majority of the world population living in urban areas by 2030 [2], cities themselves will need to become large power generators. This is because in times of peak power draw, factors such as lengthy transmission lines and lack of fuel supply can have crippling effects on a large urban populations, as was the case in 2014 during the US Polar Vortex [3, 4]. Cities have been turning to distributed generation (DG) as a way to become more self-sufficient in regards to power generation [5]. This is because many DG units now allow for more reliability, increased efficiency and cost effectiveness as well as an opportunity to use renewable generation sources [6]. Hybrid renewable systems being used as distributed generation (DG) provide a way for utility companies to move peak loads and deliver reliable power transmission [7].

Utilizing DG can allow a more reliable and cost-effective solution to consumers, and in cases where renewable generation sources are installed, a more maintainable and ecofriendly alternative to fossil fuels [6]. However, with more DG generation becoming interconnected into the current grid infrastructure, and DG sources potentially feeding power back into the current grid system, utilities will need to be able to better monitor different points within the grid to ensure grid stability. As renewable energy generation becomes more abundant and affordable, distributed generation use will only increase and become more interconnected with current grid infrastructure, necessitating a further need to collect large amounts of data to analyze and predict grid states in real-time. To do this, smart meter devices will be needed to collect large amounts of grid data to be analyzed. Thus contributing to the development and maintenance of demand response programs.

Utility companies today are needing to evolve from their historic position of producing energy, to managing energy production not only from the supply side, but from the consumer side as well. The topic of this paper revolves around the implementation of a pilot system that allows a power producer to collect large data and analyze it in order to create cost effective energy management strategies for urban centers.

1.1 A Smarter Grid

The transition to a "smarter" grid will grant utilities the ability to become more proactive in how they manage power supply in the transmission infrastructure. In the past, utility companies have needed to increase spinning reserves, and invest in generators with faster start up times to counter intermittent generation created by renewable energy sources [8–10]. Demand response is an option to alleviate the issues that come with renewable energy penetration, and are an alternative to costly large scale energy storage [10]. Even though there has been research into the feasibility of renewables into the current grid infrastructure, utilities and policymakers find themselves still requiring ways to understand the benefits and drawbacks of demand response programs [11, 12].

The North American Electric Reliability Corporation categorized demand response as a "subset" of Demand-Side Management (DSM), which looks to create efficient energy programs focused on the consumer end (node) of power consumption [13]. Many current grid infrastructures have a utility generating energy at a plant, and sending it

through a network to the consumer [14]. In a demand response program, the consumer has a direct connection to the utility, whether it be through Direct Control Load Management (DCLM), or and Interruptible Demand. DCLM involves the utility having the ability to remotely turn on/off, or cycle devices within a home, or business, thereby reducing demand on the consumer side. Interruptible demand is an agreement between the consumer and the utility where the utility can request that a consumer curtail their energy use during peak hours, or have the ability to remotely trip devices within the consumers property as long as notice is given beforehand. In exchange, a consumer will receive discounts and/or credits towards their energy bills.

Because demand response is relatively new solution to controlling peak loads, large data collection with high sampling rates will be necessary to provide as much detailed data as possible. The necessity for large amounts of data comes from the fact that there is still a lack of experience with long term demand response programs [15].

Demand response for a large urban area is hard to model as it is complex and multilayered, so data is needed to properly simulate demand response in a densely populated area [15]. To better understand the factors that affect demand response programs, data relating to consumer behaviors, as well as external factors such as weather, price sensitivity, and the changing of seasons must be obtained, and researched. An outline of the demand response logic as it pertains to the pilot system is displayed in Fig. 1.

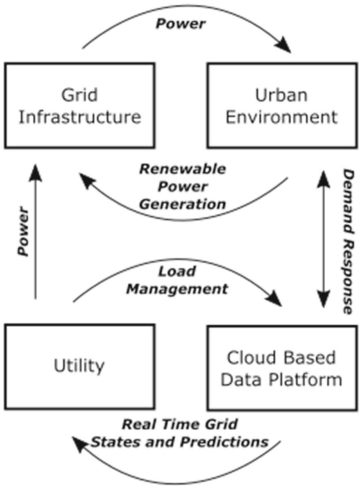

Fig. 1. The system demand structure for a data collection system is presented. A cloud based platform will store and analyze data collected from a home, or business, in real-time, allowing for quick control decisions in demand response programs.

Devices that measure power consumption have been used in research, however, most studies do not offer high frequency data with the resolution to detect small transient changes. Current research on the pilot system collects and analyzes data at higher resolutions. A 1 Hz resolution, or better, will provide a good sampling rate for large data

collection and the ability to see transient patterns in power usage, such as the warming of a stove, or the brightness of a television. Results from the pilot system have shown that different devices such as a stove top, or a water heater, create a specific power profile signature when their power draw is monitored. This signature can be thought of as a "power fingerprint." Having the ability to determine device usage from power data allows cost efficiency in power monitoring because rather than installing a power monitoring meter on each device within a building, software can instead analyze and determine which devices on a property are in use based on the power signatures found within an aggregate power data set for an entire home, or business.

Power producers will be able to monitor a home, or business, and understand which devices can be cycled during peak loads to relieve grid pressure, especially in high energy consumption areas like urban centers where large percentages of a population tend to live. In order to accomplish this, a device is needed to record a consumer's power usage. A pilot program has been created at the University of Hawai'i that currently involves monitoring aggregate power usage from 20 homes on the island of Oahu using a smart power meter (SPM). The components, challenges and scalability of the pilot system will be discussed, as well as future work pertaining to demand response programs, which will be discussed in the following sections.

1.2 Related Research

The study and feasibility of demand response as it relates to power grids is ongoing, and the pilot program looks to contribute to that research in the areas of large data collection, storage and analysis [12, 13].

Demand response programs allow for increased peak load reduction as well as the ability to balance supply and demand of energy in power grids [12]. Stability and load shifting are two factors that are important in maintaining grid stability, which can be accomplished through demand response programs. Cost efficiency is another benefit of demand response because there is no need to maintain spinning reserves and large power storage infrastructure [8].

Similar research is being done on smart meters to collect and analyze data. A group from the University of Bath investigated the use of smart metering devices in combination with voltage control techniques. Their research focused on analyzing the consumer side of demand response as a way to create cost efficiency for a consumer as well as a tool to restore grid system faults and maintain transmission stability. The Lon Local Operating System (LonWorks) and ZigBee Wireless Network Standard were two suggestions for creating a system of communication between smart meters and controllers to handle real-time data [31].

A research group in Europe proposed the use of local area networks (LAN) and wireless local area networks (WLAN) in combination with KNX communication standards as an option to set up communication between smart metering devices. The use of ZigBee and KNX components were deemed feasible to monitor load consumption of devices in order to create a timetable of *shiftable loads*. The load shifts refer to the rescheduling of device usage from peak hours to times that do not provide large strains on the grid. Real-time analysis and visualization would allow consumers to make the

proper choices in energy consumption that are related to cost efficiency. An algorithm based on tariffs was the basis for the load timetables [32].

Researchers in Canada proposed a smart metering system based on load disaggregation where a power signal is analyzed into the various device components that produce it. Their research focused on the factors that affect load disaggregation such as noisy signals, simultaneous loading, computational costs and privacy issues. They noticed that devices produced different power signals when cycled, for example, constant vs. periodic loads. To train algorithms in detecting a device, the research group suggested algorithm training based on probabilities and the clustering of individual devices. The research group deemed the definition of *deferrable actions* as necessary in their proposed system. *Deferrable actions* are those relating to devices whose utilization is not a priority and cycling can instead be scheduled at an alternative time, which would allow for load shedding. These devices include washer/dryers, ovens and dishwashers [33].

A UK-based power utility, National Grid, looked into the affect the power usage of certain devices had on the grid. They found that millions of kettles are cycled around 5 pm, knowledge such as this allows a utility to know when to cycle specific loads within home. National Grid uses the aforementioned knowledge to maintain grid frequency. Aggregating these cycling patterns with the loads of other houses in a neighborhood, or region, allow for the ability to maintain grid stability throughout sections of a power grid [30].

2 SPM Pilot System

Because of the island's geography and dense population, Oahu provides an ideal location to understand renewable energy penetration into an existing power grid, and how it relates to demand response programs. Several factors allow for Oahu to be the location to implement the pilot system, these factors include high solar radiation on the island, access to a dense urban populations, and Oahu being an isolated power grid. In 2015, the Hawaii state legislature voted to have 100 % energy generation from renewable sources by 2045 [16, 17]. Hawaii's commitment to alternative energy sources allows for a continued study of an urban area with high renewable energy generation, and the effects of this generation on demand response. Because most buildings have circuit breaker boxes, a common interface is already in place to install the SPMs. The device collects data at one-second intervals and sends it through a local WiFi network to a remote cloud server using a SSH tunnel. Data storage, analysis, forecasting and control can all occur within the cloud. The server will have the ability to send control signals based on analysis of the power data to the consumer, where an installed client can cycle devices in accordance with demand response programs to reduce peak loads. Figure 2 illustrates the overall pilot system.

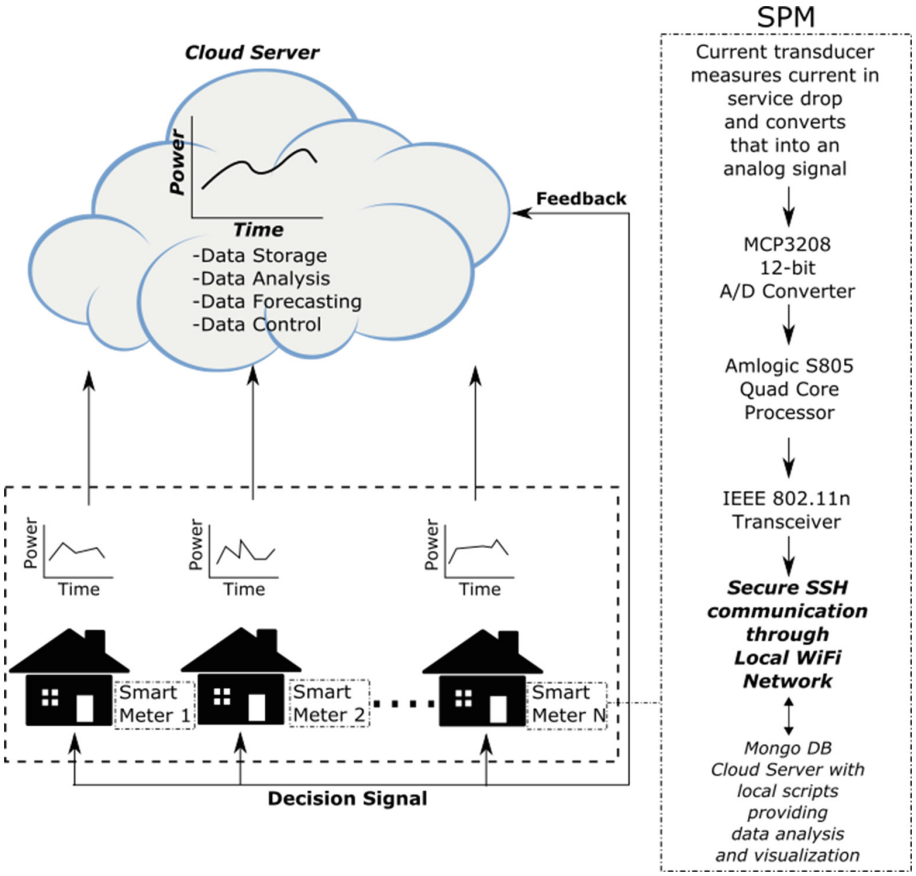

Fig. 2. The setup of the proposed system implements a SPM to monitor and transmit data from a circuit box. Data is then transmitted to a server for analysis. The current server can be scaled to cloud storage, so that more nodes can participate in the pilot program and provide more data for load forecasting analysis.

2.1 Data Acquisition

The data acquisition is performed by a power metering device at the local consumer level. The device can fit within a circuit breaker box, is non-invasive, and allows for easy installation, setup and maintenance while delivering accurate power measurement, data preprocessing and server communication. The SPM is powered through the circuit breaker box. Two current transducers, one connected to each service drop wire within the circuit breaker box, measure current signals, which are transformed into analog voltage signals, and sent to a MCP3208 12 bit analog digital converter (ADC), which collects data at 80kSps. Images of an installed device are shown in Fig. 3.

Fig. 3. A SPM meter is installed in the circuit breaker box of a home taking part in the pilot project.

An Amlogic Quad Core processor computes the power consumption for each phase. Power is calculated assuming a constant voltage. The median power pertaining to one second of collected data is obtained for each phase, and sent to the cloud server for storage and analyzing. Figure 4 describes data collection and transmission on the consumer level.

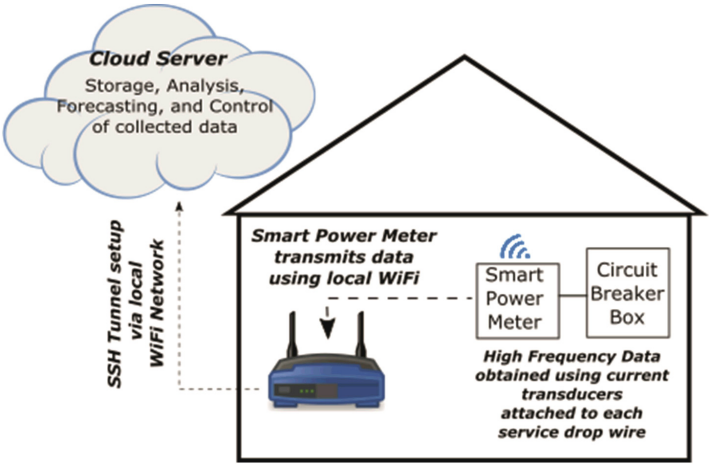

Fig. 4. Utilizing preexisting WiFi connections within a home allow for a cost effective solution for data transmission. Circuit breaker boxes are usually located in a remote area of a building, so it is necessary to utilize a wireless connection to allow for a robust system to monitor and transmit data from a node. A secure SSH connection allows for safe and reliable transmission of data to a server in real-time.

2.2 Communication

After the power data is collected and preprocessed by the SPM, the data is then transmitted to a remote server using a secure SSH tunnel via a local WiFi network. The advantage of this communication setup is that the SSH tunnel provides an added layer of security for what is confidential information. While the utilization of a preexisting local WiFi connection takes advantage of an already existing network, thus eliminating the added cost of building a new communication infrastructure. Data is stored directly into a MongoDB database hosted on a cloud server. Because data is being sent from multiple locations, each data set needs to be identified by the node it originated from, this is accomplished when the SPM assigns a node identifier to each outgoing data set. When there is a disturbance in the WiFi connection, or a communication delay, the SPM will buffer until a connection is reestablished to minimize data-loss. Despite the 1 Hz transmission rate of the SPM, bandwidth and storage requirements are kept minimal. Each database query consists of just three integers, which total 24 bytes of data per second on a 64 bit system. Households are currently transmitting approximately 2 MB/d. The island of Oahu has a population of approximately 950,000, assuming 200,000 households, 400 GB of power data would be sent to the servers each day at a rate of 4.63 MB/s.

2.3 Data Storage/Analysis

The MongoDB database on the cloud server, is a document based open-source database. It is utilized as a multiuse agent that acts as a central node where large amounts of power data is collected, streamed and queried for data analysis of real-time system states and forecasting.

Document based databases yield high scalability and data storage flexibility, which is quintessential for power analysis of large complex urban centers. Streams of real-time and recent data, as well as data queries for historical data must be performed as efficiently as possible to create predictions that will analyze data in real-time, thus allowing for fast and efficient conclusions and decisions. These conclusions will be utilized in future work to create control decisions to be sent back to the consumer where devices within a property can be controlled using a client. Thus granting the ability to create forecasts that enable efficient demand response programs to be implemented, which will reduce peak loads and ensure reliable power transmission within the grid infrastructure.

2.4 Control

Future work revolves around enabling the cloud server to analyze real-time and historic data in order to determine, and send control decisions for demand response programs. Smart control decisions enable the ability to better ensure grid stability and power transmission reliability. These commands include, but are not limited to, ON/OFF commands, as well as time constraint commands. The control clients executing the commands will have the ability to send feedback data to the cloud.

The server itself can be utilized by the consumer as an interface to monitor power consumption, or override control decisions.

3 Data Analysis

Data collection is currently in progress using a total of 20 nodes and has been ongoing since August 2015. Participants volunteered (not compensated) to participate in the study and the household sizes range from two to six members. The backgrounds of the various participants are varied, however, specific details are kept confidential for privacy reasons. There was no criteria for selecting participants, the only requirement was that they had an accessible circuit breaker box within their home.

Each phase in the circuit breaker box is measured, and the power for each phase is plotted. Figure 5 gives an example of data from a node for one day. Phase one and two are plotted in red and black, respectively.

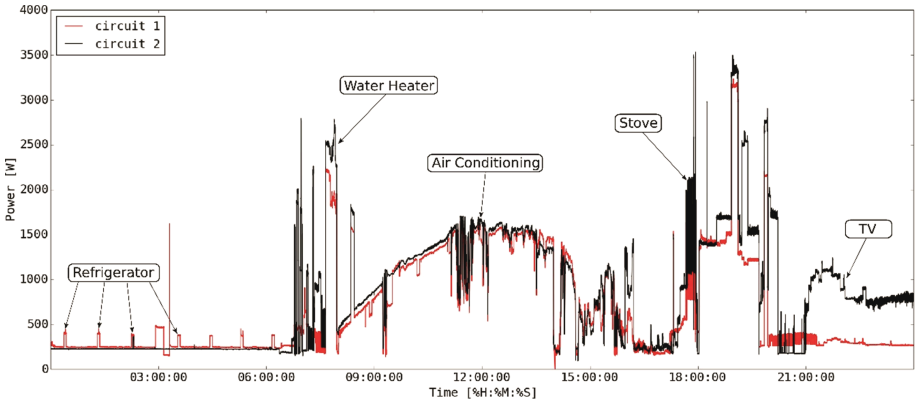

Fig. 5. Devices produce specific power signatures when in use. It can be observed when certain devices are cycled. The cycling of loads within a node displays the behavior and patterns of a consumer that can be used to predict and schedule power generation. (Color figure online)

It can be seen that there are unique device signatures throughout the day, which correspond to a combination of specific devices within the node. In the displayed example, from midnight to 7 am, the only signal that stands out is the refrigerator cycling, which is due to the fact no other major loads are present at the respective time interval. During the day air conditioning is the dominant load, which correlates to the heat in Oahu at midday. Evening loads are dominated by consumer electronics such as TV. Detailed power profiles over extended time periods grant an observer the ability to understand the energy needs of a consumer and predict when to schedule loads. Such is the case in Fig. 6 where a week of data has been plotted.

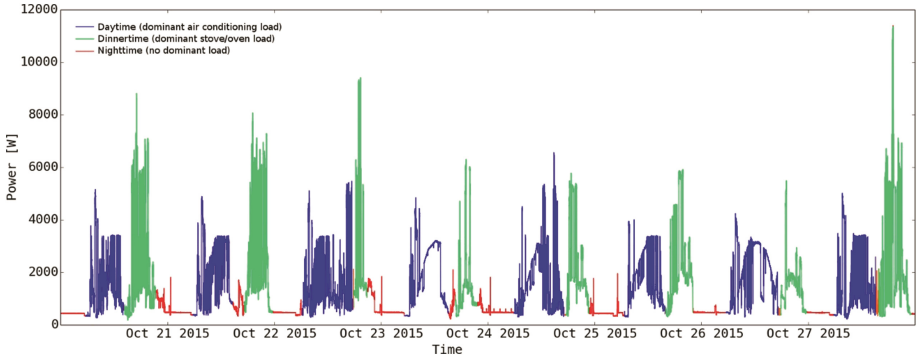

Fig. 6. One week of total power consumption is plotted for one family home. Consumer pattern behavior is evident from the increases in power consumption. (Color figure online)

The node displays a clear pattern of power consumption throughout a week. Dominant loads throughout the day are shown in blue and green, correlating to air conditioning and dinner-related activities, respectively. The family exhibits a fixed pattern of power consumption throughout the week that can be used for load prediction. Air conditioning loads dominate the day while cooking-related activities dominate evening loads. The two main load patterns stemming for air conditioning and cooking are repeated daily throughout the week. Nighttime loads are reduced to a bare minimum because of inactivity at night.

Aiding in the study of demand response it the fact that each device produces a specific power signature, or fingerprint, when spectral analysis is performed on the plotted power signal obtained by the SPM, as shown in Fig. 7.

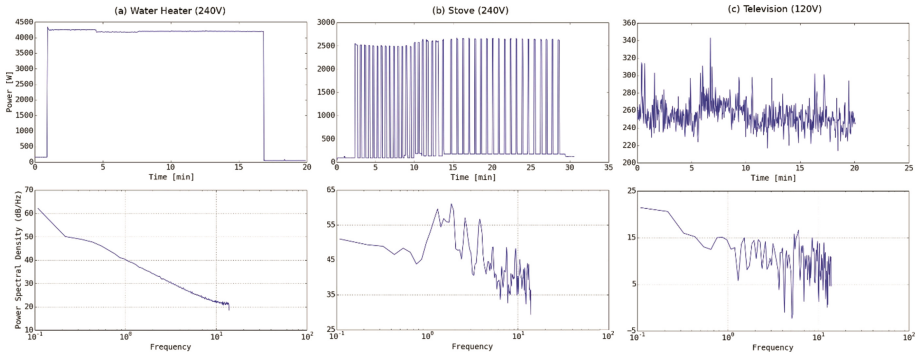

Fig. 7. The first row displays the power measured and transmitted by the SPM to the cloud server for a water heater, stove and television. Spectral analysis of the power signals correlating to each device are shown in the second row. A time based signal is converted to a frequency spectra that allows the ability to locate unique frequency signatures related to the time series signal.

Self-learning algorithms, such as ANNs, can be taught to detect power fingerprints in large data sets such as those shown in Figs. 5 and 6. Knowing which devices are in

use, and when, will allow for scripts installed on a server to calculate optimal load schedules to cycle devices, such as water heaters and HVAC units within a node. Being able to distinguish when, and how often, a consumer uses a device will enable a power provider the ability to shed peak loads while not creating an interruption to a consumer's power usage. The capability to cycle a load can be automated, so that a client within a home can obtain decision signals from a cloud based server and implement the signals in real-time.

The results show that it is possible to determine which devices are consuming power at a given time. It is also clear that large quantities of data from a node permit the observation of consumer patterns as they relate to power usage. Combining the historical and real-time power data from multiple nodes within a section of the grid, allows a power producer to understand the needs of the consumer while providing efficient load management. However, it should be noted that the observed data can be sensitive as it displays patterns and behaviors of consumers, which must remain confidential to protect privacy.

4 Scalability

A large and flexible database is necessary for bulk amounts of data being collected from an urban center. MongoDB is a "NoSQL" cloud database where large data collection will be stored and analyzed when the pilot system is scaled.

A "NoSQL", or "non SQL" database is an alternative to the relational databases that use the Structured Query Language (SQL). There are alternative "NoSQL" databases such as Apache Cassandra and Couchbase, but recent studies have shown MongoDB to be more efficient in terms of reduced latencies when it came to read and update workloads [18, 34–36]. MongoDB contains a document database architecture, which provides the flexibility needed for scalability as the pilot system grows to include more nodes.

The use of a single server would lead to scalability issues as more data is collected and processed, MongoDB overcomes these issues with the potential to add more servers to accommodate large data as well as the utilization of automatic sharding, meaning that data is spread throughout multiple servers. Automatic sharding permits data to be accessed easier, and managed faster [19]. MongoDB utilizes a flexible data model, which allows the opportunity for easier development and scalability. This is because MongoDB does not use a rigid database schema, which determines how data is logically grouped. Rather, documents within MongoDB are assigned a primary key (id), which allows for a flexible schema where data can be easily queried. This is advantageous because factors that are not originally in prediction algorithms, but are later proven to be vital (as more data is collected) in load forecasting, can easily be added to the existing database (using key value pairs) and be used in prediction algorithms [20]. The "NoSQL" database MongoDB also takes advantage of bucket streaming URI (Uniform Resource Identifier), which is based on chunked transfer encoding. The streaming transfers permit data to be sent directly to the cloud server whenever data is available for transfer. This is because data is not buffered and saved to an isolated file, thus allowing for faster data transmission to the cloud servers [21].

There are drawbacks to using MongoDB, one being that the database performs poorly when it comes to aggregate functions, such as medians, modes, and sums. However, current research has not deemed this to be a problem when implementing algorithms into the cloud server. MongoDB also struggles with non-key values, but this too has not been deemed an issue in current research related to the pilot system. Because MongoDB is a "NoSQL" database, its implementation will require more effort than a SQL database due to the fact the schema in a "NoSQL" database is not as rigid. And because "NoSQL" databases have only recently gained in the popularity they have today, there is less support and literature as compared to a SQL database, which in many industries is considered a standard [20].

4.1 Data Security

Analyzing data will grant the ability to understand the behavior of a consumer, and as the pilot system is scaled up to include thousands of users within an urban environment, it will be necessary to protect sensitive information. The information is sensitive because it can reveal what a person, or persons, are doing at a specific time in the day. Many activities can be monitored, such as a person cooking, taking a shower, or working on the computer. It can also be determined when a person is home based on their air conditioning and heating usage. The monitoring of data can even analyze the power spectrum of a television, allowing for the TV power signal to be compared to the TV signatures of known channels, and from there determine what TV programs a person is watching. Unauthorized disclosure of this potentially sensitive information could allow an unauthorized agent to study the habits and routines of an end-user, thus creating potential threats to the privacy of the consumer.

Currently, the pilot system utilizes a single server, however, when scaling up the system to include consumers from a dense urban population, a cloud server will be used. Once the computational and storage limits of the single server are reached, the pilot system will be scaled to cloud computational storage. The use of cloud services has been increasing due to a number of factors, some of these factors include; the potential for scalability, geographic reach, cost savings and higher availability [22]. With the growth of cloud service and usage comes the need to address potential for security risks.

4.2 Addressing Threats

On the local WiFi network level, it will be important for the owner of the network to make sure their WiFi network has a strong network password as well as the most current firmware available for their router. In some cases, a user may disable the Wi-Fi Protected Setup (WPS) that is vulnerable to brute force attacks on the WPS PIN. It will also be important to apply "secure by design" principles when creating levels of security within a local WiFi network. One example will be determining who will have "root," or "admin," privileges in regards to the network, and whether these privileges will apply to the entire network, or in an "isolated environment" where only certain functions are available [23].

4.3 Unauthorized Internal Users and External Hackers

Two threats to the cloud server include unauthorized internal users, as well as external hackers. An unauthorized user, whether intentionally, or accidently, can access and manipulate data within the server. To prevent this, proper security protocols must be implemented that insure only authorized persons can access sensitive data within the servers. These protocols may include, but are not limited to, physically protecting servers, encryption tools, and anomalous behavior pattern detection [24].

External hackers may use techniques to prevent the proper function of a cloud server, or to obtain sensitive information. A common practice is the abuse of resources technique, which includes sending thousands of requests per second to disrupt and inhibit network computational resources. This technique is also known as a denial of service attack (DoS). The motivation for these types of attack are not to obtain data, but rather to overwhelm networks, and consumer computational resources. DoS attacks can be prevented by blocking an IP address where an attack (large number of requests within a certain amount of time) is found to originate, as well as implementing a security tool that can recognize when an attack takes place [25]. In addition to general hacking techniques, where someone will look to exploit perceived weaknesses in a system, external hackers will also employ password sniffing and man in the middle attacks (MITM). Password sniffing involves monitoring messages in a network with the goal of obtaining a password [26]. MITM attacks will involve a hacker impersonating two parties in hopes of obtaining confidential information, this technique can be applied to a server and user, which can lead to a compromised network [27]. Both password sniffing and MITM attacks can be prevented using strong passwords, ensuring a direct and authenticated connection to the server in use, as well as strong encryption techniques.

4.4 Vulnerabilities in Cloud Security

There are many vulnerabilities that are associated with cloud server use, a few will be mentioned to provide a foundation for future security protocols.

Data Interception. Data interception is a key concern due to the fact a large number of consumers will be sending sensitive data to a cloud server in the range of seconds. To remedy this, a secure shell (SSH) will implemented in the transfer of data from the consumer to the cloud. A SSH provides data encryption and the ability to implement a proxy for added security [28].

Data Leakage. There is potential for data within MongoDB to be leaked to unauthorized users, however, the developers of MongoDB look to actively recognize and address any issues relating to data leakage, which are usually related to versions of MongoDB that are outdated and unpatched. Other ways to prevent data leakage is using proper encryption methods, and recognizing when and where data is sent, so that it can be properly monitored. Physical protection of servers and personnel screening provide added security benefits [28].

Insecure or Ineffective Deletion of Data. When deleting data from a cloud server, there is always potential that data deletion may be incomplete, or insufficient. To counteract any potential issues from data deletion, it will be necessary to follow proper deletion protocols related to the cloud server platform, and in worst case scenarios, insure that a disk containing sensitive data is destroyed. Once again, proper encryption of data will decrease the risk related to ineffective data deletion [28].

Loss of Encryption Keys. The loss of encryption keys may be due to the accidental publication of a secret key, such as a secure socket layer (SSL), or a network password, which would create a vulnerability for potential threats. Several methods to mitigate encryption key loss are listed below [28]:

1. Storing encryption keys and the data in separate locations
2. Implementing audit trails to track who accesses data, and when the data is accessed
3. Backing up encryption keys onto a secured device
4. Encrypting the encryption keys themselves
5. Periodic changing of encryption keys [29]

Malicious Probes. In the case of a malicious probe, an unauthorized user may look to introduce a virus into the system. This can be prevented by creating database logs that record who and when someone attempts to access the database, so that any unauthorized user attempts may be blocked. Continually updating security patches will also insure that cloud security architecture is up to date. If a malicious probe does enter the system, historical data can be protected though periodical backup into a secured location [28].

5 Conclusion and Future Work

In order to provide the proper demand response program to a power grid, it is first necessary to collect large amounts of data in order to understand consumer behaviors and patterns. A pilot system was created with a smart metering device that can collect and transmit data at high frequencies (1 Hz or less) through a SSH tunnel to a server. A robust collection of data allows for the patterns and behaviors of a dense urban population to be analyzed. The small scale pilot system has proven the feasibility of data collection related to large-scale demand response. However, challenges will be present when scaling the pilot system to include more nodes. Topics to be addressed will include protecting sensitive consumer information, server infrastructure, security, and the management of big data.

Current research related to the pilot program is in the early stages of understanding consumer behavior. Human behavior is complex and is a study within itself, however we look to just understand device usage as it relates to demand response. Initial results are encouraging as patterns related to node power consumption can be detected. But because of the complex nature of human behavior, more data will need to be taken to see how external factors such as weather, holidays, and season affect consumer power consumption. However, the pilot system provides an initial foundation into the study of factors affecting consumer power usage. As more nodes are added to the current system

and data collection continues a better understanding of consumer behavior as it relates to demand response programs will be achieved.

References

1. 2013 Report Card for America's Infrastructure. ASCE, Reston, VA (2013)
2. Global Trends 2030: Alternate Worlds. National Intelligence Council, Washington, DC, vol. 5 (2012)
3. Tweed, K.: Polar Vortex Cripples Power Generation, But Grid Survives, 9 January 2014. http://spectrum.ieee.org. Accessed 1 Nov 2015
4. Duke Energy: Causes of Power Outages (2015). https://www.duke-energy.com. Accessed 5 Nov 2015
5. Calvillo, C.F., et al.: Distributed energy generation in smart cities. In: Paper presented at the International Conference on Renewable Energy Research and Applications, Madrid, Spain, 20–23 October 2013
6. The Potential Benefits of Distributed Generation and Rate Related Issues that may impede their Expansion. US Department of Energy, Washington, DC (2007)
7. Salameh, Z.M., Davis, A.J.: Case study of a residential-scale hybrid renewable energy power system in an urban setting. In: Paper presented at the Power Engineering Society General Meeting, Toronto, Canada, 13–17 July 2003
8. Impacts of Solar Power on Operating Reserve Requirements. NREL, Golden, CO (2012)
9. Wesoff, E.: What are the impacts of high wind and solar penetration on the grid? 25 September 2013. http://www.greentechmedia.com. Accessed 2 Nov 2015
10. Smart Grids and Renewables: A Guide for Effective Deployment. IRENA, Abu Dhabi, UAE (2013)
11. Lew, D., et al.: The Western Wind and Solar Integration Study Phase 2. NREL, Golden (2013)
12. Demand Response and Advanced Metering. FERC, Washington, DC (2008)
13. Demand Response Discussion for the 2007 Long-Term Reliability Assessment. NAERC, Atlanta, GA (2007)
14. Smart Grid (2015). http://energy.gov. Accessed 20 Oct 2015
15. O'Connell, N., et al.: Benefits and challenges of electrical demand response: a critical review. Renew. Sustain. Energ. Rev. **39**, 686–699 (2014). Elsevier
16. Press Release: Governor Ige Signs Bill Setting 100 Percent Renewable Energy Goal in Power Sector. http://governor.hawaii.gov. Accessed 2 Nov 2015
17. Namata, B.: New Law requires 100-percent renewable energy in Hawaii by 2045, 8 June 2015. http://khon2.com. Accessed 21 Oct 2015
18. Scalability Benchmarking: MongoDB and NoSQL Systems. USA, Pleasanton, CA (2015)
19. Cattell, R.: Scalable SQL and NoSQL Data Stores. ACM Sigmod Rec. **39**(4), 12–27 (2010)
20. Parker, Z., et al.: Comparing NoSQL MongoDB to an SQL DB. In: Paper Presented at the ACMSE 2013 Proceedings of the 51st ACM Southeast Conference, Article no. 5, Savannah, GA, 4–6 April 2013
21. Google: Concepts and Techniques (2015). https://cloud.google.com/storage/docs/concepts-techniques?hl=en#streaming. Accessed 30 Oct 2015
22. Cloud Computing Trends: 2014 State of the Cloud Survey. RightScale, Santa Barbara (2014)
23. Wikipedia: Secure by design (2015). https://en.wikipedia.org. Accessed 25 Oct 2015
24. Chou, T.S.: Security threats on cloud computing vulnerabilities. IJCSIT **5**(3), 79–88 (2013)
25. Weiss, A.: How to Prevent DoS Attacks, 2 July 2012. http://www.esecurityplanet.com. Accessed 1 Nov 2015

26. Armstrong, D.: Password Sniffing, 25 October 1996. http://cng.seas.rochester.edu. Accessed 27 Oct 2015
27. Karapanos, N., Capkun, S.: On the effective prevention of TLS man-in-the-middle attacks in web applications. In: Paper Presented at the Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, 20–22 August 2014
28. Computing, Cloud: Benefits, Risks and Recommendations for Information Security. ENISA, Crete (2009)
29. Protect your most critical data and your access to it by following these tips for securing encryption keys. http://aspg.com. Accessed 25 Oct 2015
30. National Grid: Frequency Response Services (2015). http://www2.nationalgrid.com/uk/services/balancing-services/frequency-response. Accessed 22 Nov 2015
31. Gao, C., Redfern, M.A.: A review of voltage control in smart grid and smart metering technologies on distribution networks. In: Paper Presented at the 46th International Universities Power Engineering Conference, Soest, Germany, 5–8 September 2011
32. Kunold, I., et al.: A system concept of an energy information system in flats using wireless technologies and smart metering devices. In: Paper presented at the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Prague, Czech Republic, 15–17 September 2011
33. Makonin, S.: The cognitive power meter: looking beyond the smart meter. In: Paper Presented at the 26th IEEE Canadian Conference Of Electrical and Computer Engineering (CCECE), Regina, Canada, 5–8 May 2013
34. Olavsrud, T.: 9 MongoDB success stories, 24 November 2015. http://www.cio.com/article/3008114/open-source-tools/9-mongodb-success-stories.html. Accessed 10 Dec 2015
35. Bhattacharjee, A.: NoSQL vs SQL – Which is a Better Option? 8 May 2014. https://blog.udemy.com/nosql-vs-sql-2. Accessed 11 Dec 2015
36. McNulty, E.: SQL VS. NOSQL- WHAT YOU NEED TO KNOW, 1 July 2014. http://dataconomy.com/sql-vs-nosql-need-know/. Accessed 10 Dec 2015