A Case Study on the Trade-Offs between Security, Scalability, and Efficiency in Smart Home Sensor Networks

T. Reichherzer, A. Mishra, E. Kalaimannan, N. Wilde
Department of Computer Science
The University of West Florida
Pensacola, Florida, USA
e-mail: {treichherzer, amishra, ekalaimannan, nwilde}@uwf.edu

Abstract— **This paper presents the design of a case study that examines the trade-offs between security, scalability and efficiency for sensor networks in IoT environments. Among the trade-offs we will investigate throughput and real-time response for sensor data collected by wireless sensors and transmitted into a scalable Cloud environment for storage and processing. The case study will be performed on a Smart Home prototype system developed at the University of West Florida that functions as a testbed system of a comparable IoT environment. The Smart Home prototype system provides sensor network and sensor data collection and processing services designed to infer activities in a home but it lacks security and scalability of its services. We anticipate that results of the case study will provide generalizable insight into the design of secure and scalable IoT applications.**

Keywords—sensor networks; case study; scalability; security; Amazon Elastic Compute Cloud

I. Introduction

Smart devices are increasingly present in our daily lives as they monitor, analyze, and report data including energy usage, personal activity levels, and interactions with objects in an environment. Smart devices are projected to be progressively in demand over the next 20 years for applications in private homes and public areas [1]. However, the devices provide little to no protection from attackers that may compromise the devices, their availability, or the data that they store. A lack of standards and regulations has contributed to various proprietary communication protocols being used which may or may not provide adequate protection to the privacy of the data. However, even if protection is provided on the data stored and transmitted over the network, the overhead of the protection of the data is not well understood. In response to these challenges we propose a case study that will examine the trade-offs between security and efficiency for sensor networks. The study is based on a prototype system of a sensor network that includes multiple wireless sensors and services to process the data. This sensor network is currently used in another research project aimed at building a Smart Independent Living for Elders (SMILE) home [2].

Sensors in the SMILE home transport data over wireless and wired networks to a server that processes the collected environmental data using machine-learning algorithms to infer activities in the home and provide smart assistance. The sensor technology is based on off-the shelf sensor and chip components including Arduino boards, light, temperature, infrared, pressure sensors, and the Nordic nRF24L01+ wireless Radio Frequency (RF24) transceiver module. The electronic components are assembled into individual, battery-driven sensor boxes that measure environmental data and transport them wirelessly to a Raspberry Pi computing device, also connected to an RF24 transceiver module, which functions as a base station for data collection in the home. All wireless devices together form a mesh network to deliver data from sensors to the base station using other sensors as intermediate communication hubs to reach the nearest Pi in the home that connects directly to the home's Internet service provider. From this base station sensor data in the home are transported via a wired Internet connection to a Cloud computing environment.

In this system, the Arduino board executes the Sensor Data Collector (SDC) software for collecting sensor data from attached sensors. The Raspberry Pi executes the Sensor Data Distribution (SDD) software that receives data packets from the mesh network on a continuous basis and forwards them to sensor processing units called Sensor Agents. The Sensor Agents execute on a Pi platform and communicate internally over a UDP communication channel with the SDD. Designed to collect specific sensor data from the SDD, they can be shut-down, replaced, and restarted remotely to control the data collection in the home. A proprietary but unencrypted communication protocol is used for wireless communication between the SDC and the SDD. From the SDD data are transported to Sensor Agents. The role of the Agents is to filter redundant data and upload them to a Cloud-based Web Store via a Web Service that executes on an application server. Communication between the Agents and the Web Service is unencrypted. The Web Service connects to a Web Store for back-end data storage using JDBC. The Web Store is implemented by a MySQL database server. The Sensor Agents and the Web Service are implemented in Java and Java EE respectively. Figure 1 illustrates the currently implemented hardware and software components of the system along with the communication channels that links them. This system builds the experimental platform for our case study.

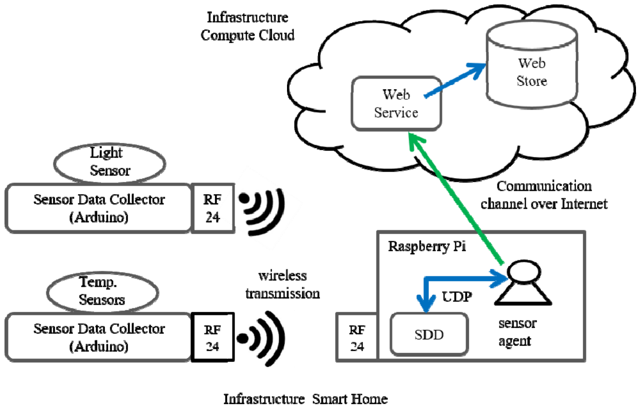

Figure 1: The Experimental Model for the Case Study.

II. Related Work

Smart IoT devices are poised to become increasingly mainstream, shaping technology innovation to application areas ranging from healthcare to retail and transportation. Gartner, Inc. forecasts that 6.4 billion connected things will be in use worldwide in 2016, up 30% from 2015, and will reach 20.8 billion by 2020. In 2016, 5.5 million new things will get connected every day [3]. Gartner estimates that the Internet of Things (IoT) will support a total spending of $235 billion on services in 2016, up 22 percent from 2015. Currently, IoT is an ecosystem composed of specialized hardware, network connectivity, and cloud counterparts all designed to facilitate data collection and processing.

Many challenging issues still need to be addressed before the IoT idea attains wide acceptance [4] including full interoperability of interconnected devices, providing them with a higher degree of smartness while guaranteeing trust, privacy, and security. The growing awareness regarding loss of privacy makes the security of personal data a fundamental consideration for IoT [5]. Presently, there is no agreement among IoT product developers on how to implement security in IoT devices [6]. IoT is included by the US National Intelligence Council in the list of six ‘‘Disruptive Civil Technologies” with potential impacts on US national power [7]. Landmark IoT security breaches such as the leakage of millions of healthcare records [8], CSX Train Signaling System attack [9], hacking of a car leading to recall [10], smart energy meter tampering [11] are some that indicate that IoT security needs to be taken seriously.

Jing et. al [12] presented a comprehensive survey on the perspectives and challenges about the security of IoT. The authors examined and addressed the cross-layer heterogeneous integration and security issues in detail.

III. A System Model for Security & Scalability

The Internet of Things, in general and Smart Home Network systems, in particular face two main challenges: security and scalability. The stringency of the requirements in both categories depends primarily on the kind of data being collected and the specific applications that utilize them. The Smart Home sensor network model described in this paper spans broadly over the wired and wireless network security issues. Our model addresses several of the main tenets of information security which include *authentication, availability, confidentiality, integrity and non-repudiation.*

We focus primarily on those security issues that arise from the distributed and networked nature of IoT systems, leaving aside the physical security of the Raspberry Pi and the infrastructure security of operating systems, database servers, and application servers. In this context, confidentiality will be an important concern in almost all Smart Home model contexts, since even the simplest sensors for measuring temperature or illumination can reveal occupants living habits and thus infringe upon their privacy. Authentication, availability, integrity and non-repudiation will become more important depending on the application; if a fall detector in a geriatric household is relied upon to summon an ambulance service then defense against spoofing or denial of service attacks becomes essential. Also, data integrity is a serious concern particularly to the sensors, when they transmit, record and store data at various nodes inside the Smart Home system architecture. Any kind of modification or tampering of the data while in transmission, will create a significant threat to the system and the inhabitants.

Scalability requirements also depend greatly on the application and its design. If the Smart Homes are localized to a single gated community, then the corresponding data volumes may be anticipated with relative ease. On the other hand, applications that span a city or a country need to be able to handle large and variable volumes of data without failure. Cloud infrastructures would serve as a potential solution to address this kind of challenges in both security and scalability. Thus, the proposed case study deploys and evaluates central parts of the system into the Cloud.

IV. A Case Study of Security and Scalability

Technologies exist to handle many security and scalability concerns. However, these almost always introduce some loss of efficiency. In designing sensor networks trade-offs must be made to balance conflicting goals. The objective of our case study is to provide data that will aid developers in making these choices.

The experimental model for the case study as shown in Figure 1 considers a Smart Home, which contains sensors of different types that communicate through a wireless channel to Sensor Agents located in the home. The Sensor Agents communicate over a wired connection to a Web Service provided by an application server that in turn connects to a Web Store such as a database server for storing the sensor data. For simplicity of measurements, we consider unidirectional data flow from the sensor, via the Sensor Agents, to a Web Service located in a public infrastructure Cloud. Applications would then use the stored information to perform whatever tasks were necessary. Since some of these tasks may have stringent real time constraints, efficient and reliable end-to-end flow is highly desirable.

The case study we plan to conduct considers four variations of the experimental model that progressively improve the security and scalability of the system. Rather than developing large amounts of software ourselves, we will use proven industrial services for scalability since these are likely to be well engineered and similar to products that would be used in a commercial Smart Home development. Thus, for the Compute Cloud infrastructure we will use Amazon's Elastic Compute Cloud EC2™ to run Web Services and the Web Store. For the Smart Home infrastructure we will use the already implemented software and hardware built for the Smart Home including a Raspberry Pi for running Sensor Agents and various physical sensor boxes. The four different configurations of our case study include:

1. Base System with Minimal Security and no Scalability Technologies: This configuration forms the base case of our studies to compare all other configurations. The Web Service and the Web Data Store will both be hosted on a single AWS instance of type t2.micro running Ubuntu server, LTS 14.04 with 8 GB of memory. Communication between the Sensor Agent and the Web Service will be by unsecured HTTP messages. The Web Service will be implemented in Java and will run on a GlassFish application server, version 4.1. The Web Data Store will be a MySQL database server, version 5.5, that maintains the sensor database. The Web Data Store is not directly accessible by a Sensor Agent. It must be accessed through a Web Service implemented by the Glassfish application server that uses an unsecured Java Database Connection (JDBC) to authenticate and communicate with the MySQL database server.

2. Basic Security, No Scalability: This configuration extends the base case by adding secure communication and authentication between Sensor Agents on the Raspberry Pi and the Web Service as a communication point. This is achieved by modifying the Java code that implements the Sensor Agents and the Web Service to use a Secure Socket Layer (SSL) for communication and authentication. The modification of the Web Service is based on the Apache Shiro security framework that adds a security manager to the Glassfish server platform. Each Sensor Agent deployed in a Smart Home is given a unique username and password that it uses when making the Web Service call to authenticate with the Web Service. The Web Service uses the provided account information in the call to validate the authenticity of the sensor data by using the Web Store that records sensor data besides the username and password of all Sensor Agents. If the provided username and password matches the records in the Web Store, the Web Store accepts the sensor data and stores them in the Web Store's environmental data table.

3. Basic Security, Scalable Authentication: This configuration replaces the Web Store for authentication with the Amazon Identity and Access Management (IAM) [13] system to provide scalable services for validating sensor data.

4. Scalable Web Service, Scalable Authentication: In this configuration the Web Service will be implemented using the AWS Lambda platform. Lambda executes code in response to events, such as the arrival of a message. The platform automatically scales computing capacity by creating copies of the code as needed [14]. The Web Data Store will be implemented using Amazon DynamoDB, a NoSQL database that promises "single-digit millisecond latency at any scale" [15]. As in the previous configuration, authentication will be provided by Amazon's Identity and Access Management (IAM) service.

V. Wired and Wireless Experiments

For each of the four configurations described above, we will run controlled experiments involving physical sensor boxes from the SMILE home and simulation of sensors to increase the number of sensors in a home and evaluate the scalability of the system. Sensor data will be time-stamped at three points in the system to study the impact of improved security. The first timestamp will be at the Sensor Agent where sensor data are collected in the home, then at message arrival on the Glassfish server where messages from all Sensor Agents across many homes are received, and finally at the Glassfish server again after the Web Store confirms that sensor data have been stored. Thus, as our indicator of efficiency we will have overall elapsed time as well as its breakdown between the main system components. We use the Network Time Protocol (NTP), to ensure that clocks between the different components of the system are synchronized.

The security and scalability issues are different in the wireless and wired parts of the network. The case study above identifies the security and scalability concerns from the point of the Sensor Agents to the Web Store. It does not address the security concerns that arise between sensors and the Sensor Data Distributor (SDD) on the Raspberry Pi. Hence, we will conduct a complementary study independent from the case study described above to examine methods of securing traffic between sensors and the SDD and their respective overhead.

In these experiments, sensor data will be generated either through the physical sensor boxes or through simulation by a number of Sensor Agents. The data will be transported from the source to the Web Store and elapsed time will be computed to study efficiency for messages processed and transported through the network.

A. *Wireless Network Experiment*

In the wireless network for a single home, scalability is not a major concern since the number of sensors is fairly small and predictable. The main security concerns are the integrity and confidentiality of sensor data. It should not be possible to "spoof" a sensor to introduce spurious data, or to "sniff" wireless data to gather information about activities in the home. Both concerns may be addressed by encrypting message data using a key shared between the sensor and the SDD.

Currently, the sensor boxes in a Smart Home sense physical quantities and transmit their unencrypted values to Sensor Agents over wireless links. This unencrypted data on air can be tapped into by a tuned receiver, making eavesdropping and man-in-the-middle attacks possible. We propose to encrypt this data before transmission and thus to reduce the chances of success of attacks through the air interface. The sensor data is high in volume, and encryption attempts could result in delays due to local processing at the sensor boxes. We propose to measure the delay in communication due to encryption before transmission, and compare the unencrypted and encrypted throughputs. We intend to focus just on the security aspect of the sensor data, and not on scalability.

Management of energy consumption of the wireless part of the network for resourceful operation is a major challenge faced by Smart Home designers. The sensor boxes are battery-powered, and the use of traditional batteries as a power source involves keeping an eye on the battery life and frequent battery replacements. The sensor boxes consume energy in their transducer unit, communications unit, and in the computing unit for the processing of data. The encryption of data would also mean additional processing at the boxes, which would require energy. We propose to study this additional energy consumption due to encryption.

B. Wired Network Experiment

For the wired network between the Sensor Agents and the Cloud services, both security and scalability are concerns. Thus, the experiment will include sensors generated by the physical environment to evaluate real-world data, as well as simulation of sensor data to scale-up the environment to large number of sensors and to many homes. The physical sensor boxes generate sensor values at the rate of 2 seconds per sensor. However, the software in the Sensor Data Collector (SDC) only transports sensor data to the Sensor Data Distributor (SDD) on the Raspberry Pi when the observed sensor value change. However, for the experiment, we will configure the physical sensor boxes to force data to be transmitted at all time. The experiment will be run for a one hour period of time for each of the four cases described above.

For the simulation of the sensor data, the Sensor Agents will be modified to generate random data and submit them to the Web Store via the Web Service. We will increase the number of Sensor Agents executed from multiple servers on the UWF network up to the point where we observe significant time delays between the time the data are generated and the time they are recorded in the Web Store. Finally, we will conduct the physical and simulation experiments at four different times during the day to account for fluctuations in Internet traffic and usage of Cloud services.

VI. DISCUSSION & CONCLUSIONS

The proposed case study and experiments measure processing overhead as security in the prototype network is improved and as more scalable services are added to the system. The experiments rely on existing frameworks, software tools, and services, as well as our own prototype software and hardware that implements the Smart Home's sensor network. From this perspective, we believe that our results depend heavily on the efficiency of the various tools involved as well as the AWS services. Alternative implementations may give different performance results for transporting sensor data to the Web Store. However, beyond the efficiency limitations of our own code and hardware, the existing tools and frameworks used in the proposed experiment are standardized and in wide-use by the industry. This will allow us to assess the generalizability of the results in terms of the cost of security and the scalability of services with respect to other Smart Home or IoT systems.

Smart Home security implications create hurdles for their wider adoption. Attacks on Smart Home applications vary from eavesdropping on transmissions including traffic analysis or disclosure of message contents, to modification, fabrication, and interruption of transmissions through node capturing, routing attacks, or flooding. Research within UWF and outside has shown that fitness trackers and home automation IoT devices have serious security issues because sensitive data is transmitted over wireless links with little or no security. We propose to study improvement on the air interface security through encryption techniques and the tradeoff involved.

REFERENCES

[1] D. L. Hoffman and T. P. Novak, Thomas, "How to Market the Smart Home: Focus on Emergent Experience, Not Use Cases, Social Science Research Network, January, 2016. Available at SSRN: http://ssrn.com/abstract=2840976

[2] T. Reichherzer, S. Satterfield, J. Belitsos, J. Chudzynski, L. Watson, "An Agent-Based Architecture for Sensor Data Collection and Reasoning in Smart Home Environments for Independent Living," In Khoury, R., Drummond, C. (Eds) Lecture Notes in Computer Science, Vol. 9673, Springer-Verlag, pp. 15 – 20, 2016.

[3] G. Press, "Internet of Things (IoT) Predictions, From Forrester, Machina Research, WEF, Gartner, IDC", Forbes Magazine, January 2016. Available at http://www.forbes.com/sites/gilpress/2016/01/27/internet-of-things-iot-predictions-from-forrester-machina-research-wef-gartner-idc/.

[4] M. Botterman, "Internet of Things: An Early Reality of the Future Internet", Report of the Internet of Things Workshop, for the European Commission, Information Society and Media Directorate General, Networked Enterprise & RFID Unit – D4, Prague, Czech Republic, May 2009.

[5] R. H. Weber, "Internet of Things – New security and privacy challenges", Elsevier computer law & security review 26, pp. 23 - 30, 2010.

[6] D. Giusto, A. Iera, G. Morabito, and L. Atzori (Eds.), "The Internet of Things", Springer, 2010. ISBN: 978-1-4419-1673-0.

[7] National Intelligence Council, "Disruptive Civil Technologies – Six Technologies with Potential Impacts on US Interests Out to 2025", Conference Report CR 2008-07, April 2008.

[8] David Harries and Peter M. Yellowlees. Telemedicine and e-Health. January 2013, 19(1): 61-66. doi:10.1089/tmj.2012.0022.

[9] M. Niland, "Virus Disrupts Train Signals", The Associated Press, 2003, Available here: http://www.cbsnews.com/stories/2003/08/21/tech/main569418.shtml

[10] A. Greenberg, "Hackers Remotely Kill a Jeep On the Highway - With Me In It", Wired Magazine, July, 2015, Available here: http://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/.

[11] B. Krebs, "FBI: Smart Meter Hacks Likely to Spread", Krebs on Security, 2012, Available here: http://krebsonsecurity.com/2012/04/fbi-smart-meter-hacks-likely-to-spread/.

[12] Q. Jing, A.V. Vasilakos, J. Wan, J. Lu and D. Qiu, "Security of the Internet of Things: perspectives and challenges", Wireless Networks, Vol. 20, Issue 1, pp. 2481-2501, 2014.

[13] Amazon, "Identity and Access Management", https://aws.amazon.com/iam/.

[14] Amazon, "AWS Lambda FAQs", https://aws.amazon.com/lambda/faqs/

[15] Amazon, "Amazon DynamoDB - NoSQL Cloud Database Service", https://aws.amazon.com/dynamodb/