Smart Home Repos: Taking Data from Devices

PRANAV MAHAJAN, University of Cincinnati, United States
HOWARD HALL, University of Cincinnati, United States
SHANE HALSE, University of Cincinnati, United States
JESS KROPCZYNSKI, University of Cincinnati, United States
NATHAN ELROD, University of Cincinnati, United States

ABSTRACT

This is an ongoing research project focused on creating a framework for capturing various artifacts concerning Internet of Things devices. Research has shown a severe lack of frameworks focusing on collecting data from and about IoT devices. Mozilla's WebThings Gateway focuses on collecting this information from the devices. This project expects to find methods of IoT data collection through a proposed test-bed utilizing the Webthings Gateway.

Additional Key Words and Phrases: Internet of Things, Test-beds, Data Collection

ACM Reference Format:
Pranav Mahajan, Howard Hall, Shane Halse, Jess Kropczynski, and Nathan Elrod. 2020. Smart Home Repos: Taking Data from Devices . In *IT Research Symposium '20: School of Information Technology IT Research Symposium, April 14, 2020, Cincinnati, OH*.https://scholar.uc.edu/

1 INTRODUCTION

The Internet of Things (IoT) is an immensely popular suite of internet-based technologies with an extensive amount of contributions over the many years since its inception; as of August 2019 there are approximately 26.6 billion active IoT devices worldwide. However prevalent the IoT technology may be, because of its novelty it can often be ill-defined. Different organizations define the IoT in slightly different ways; for our definition we turn to the internationally recognized engineering consortium IEEE, who define IoT simply as "a network of items—each embedded with sensors—which are connected to the Internet" [4].

In spite of the prolific use of IoT devices, capacity for the unified testing of them has been lagging behind. Wireless sensor network (WSN) test-beds for IoT devices, such as MoteLab [8], CitySense [3], and Kansei [2] have been developed and innovated upon through research efforts working towards hosting such devices. The test-beds are also developed for specific use cases depending on the scale of the network. One such case is smart home networks compared to a smart city, which uses high throughput nodes and long range radios to cover a much wider area [3]. However, developing the means of safe communication between these mediums is a difficult task due to many aspects of these connections such as device and software compatibility and connection security.

Furthermore, security measures must be put in place to protect the data analyzed by the IoT devices. IoT devices have low computational power which creates a lack of strength in encryption and an inability to host antivirus software. On top of this, IoT devices are always trying to secure connections, stretching bandwidth thin and creating holes in the defenses of the device.

IT Research Symposium '20, April 14, 2020, Cincinnati, OH

As applications for these devices are developed, a medium is needed for testing functionality of the device and its application, which creates certain issues. As mentioned previously there are billions of devices originating from numerous manufacturers, none of whom have subscribed to a standard connection protocol. Therefore one of the requirements of the environment for testing, is the ability to host a wide range of devices and connection types so that the application tested can be universally connected to various devices. In doing this, the platform serves as an adapter for a device and interface that possess differing requirements from one another.

Our approach begins with research into the commonalities between the test-beds that have either been proposed or implemented in previous research. After learning of other test-beds [5], studying the architecture, security, compatibility, and other various pros & cons, we compile a list of common features and limitations. Thus, allowing us to propose improvements using the WebThings Gateway, then to address our research question; *What methods of IoT data collection exist, and how can requirements be ascertained from them for development of a test-bed to record these data artifacts?*

2 LITERATURE REVIEW

As work began on data collection through the WebThings Gateway, other test-bed documentations and papers were reviewed, as well as other literature on the challenges posed by testing environments. Through this search, several commonalities have emerged:

2.1 Test-bed security vulnerability

The nature of smart devices leaves them open to attack; they are limited in memory, CPU power, and bandwidth [9]. This means few encryption or authentication methods can operate on these devices. Public encryption poses a challenge to smart devices as well due to high computational requirements, with little remedy in the way of reducing the computational overhead [9]. Back doors are also easy to put in place on a victim's device, as Intrusion Detection Systems (IDS) or Antivirus software simply can't be handled by the device, once again because of the computational power required [9].

On top of this, smart devices typically pair through long distance connections like cellular networks, or through short/medium distance range protocols such as Wi-Fi, Zigbee, or Bluetooth. Constant attempts at securing these connections leads to a great use of bandwidth and creates larger openings of infiltration through an application with little in the way of implemented security [7].

These IoT devices are widely exposed through weaker security and their purpose to communicate through a gateway and application. The integrity of this software is entirely dependent on it's developer, and these concerns can cause great risk to devices if there are no auxiliary protections from the test-bed.

2.2 Heterogenic device design impeding universality

The heterogeneous nature of IoT devices creates another issue when developing a test-bed. Heterogeneity is described as a difference in compatible hardware, software or configurations between a test-bed and an individual smart device [7]. This incompatibility can come at any level of the TCP/IP model [5] (the layered design of networking used for Internet). A few examples would include two devices being unable to connect because one device uses ZigBee connection while another device uses Bluetooth, or a smart device being unable to physically hook into a computer because the hardware is incompatible.

Due to the absence of a set of standardized criteria which should be met by smart IoT devices, there exists a plethora of devices from different vendors (for example, Google Home vs. Amazon

Alexa), each utilizing proprietary methods for processing data communication, as illustrated by figure 1. A truly universal test-bed would encompass all communication technologies utilized by IoT devices such as Wi-Fi, BLE (Bluetooth Low Energy), Cellular, Zigbee/Z-wave or any other methods of connection.

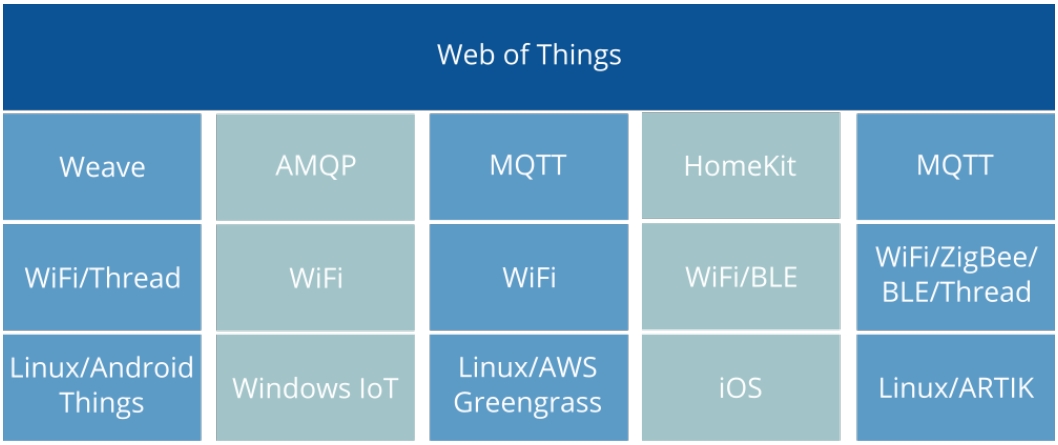

Fig. 1. Illustrating technologies used in IoT devices, via Mozilla (https://iot.mozilla.org/about).

In addition to being able to support all the connection technologies, the test-bed should be modular and dynamic allowing it to adapt to new IoT technologies; that is, until a new standard is developed and adopted, which would encourage uniformity and allow for the creation of better, more comprehensive test-beds.

3 METHODOLOGY

3.1 The Mozilla WebThings Gateway

Due to the lack of standards in architecture for test-beds and smart devices, the problems discussed prior - especially those of heterogeneity - have hindered the creation of a universal IoT test-bed. The open source WebThings Gateway from Mozilla [6] is a work-in-progress that utilizes a set of standards proposed by the World Wide Web Consortium (W3C) [1]. The WebThings Gateway is an open source framework supported by Mozilla which enables consumers to monitor and control their smart devices. It is expected the standards proposed will be taken in and integrated by smart device manufacturers, so that uniformity issues with connections, data collection, and others can be solved to create a comprehensive test-bed.

This model (Figure 2), developed by the authors, illustrates the proposed architecture of the system we are developing in order to create a framework for making an attempt towards creating a comprehensive IoT data repository. The system is currently compatible with Wi-Fi enabled IoT devices supported by the WebThings Gateway. The system captures both aspects of the data associated with an IoT device:

The network data, captured by an open source packet analyzer. The system employs the Wireshark application in order to capture and store the network traffic. The IoT data from the device itself is stored and logged by the gateway.

This creates a more comprehensive database providing another dimension to the data. As previously mentioned, this project will be expanded to cover other connection protocols and other smart devices as the number of devices supported by the gateway increases.

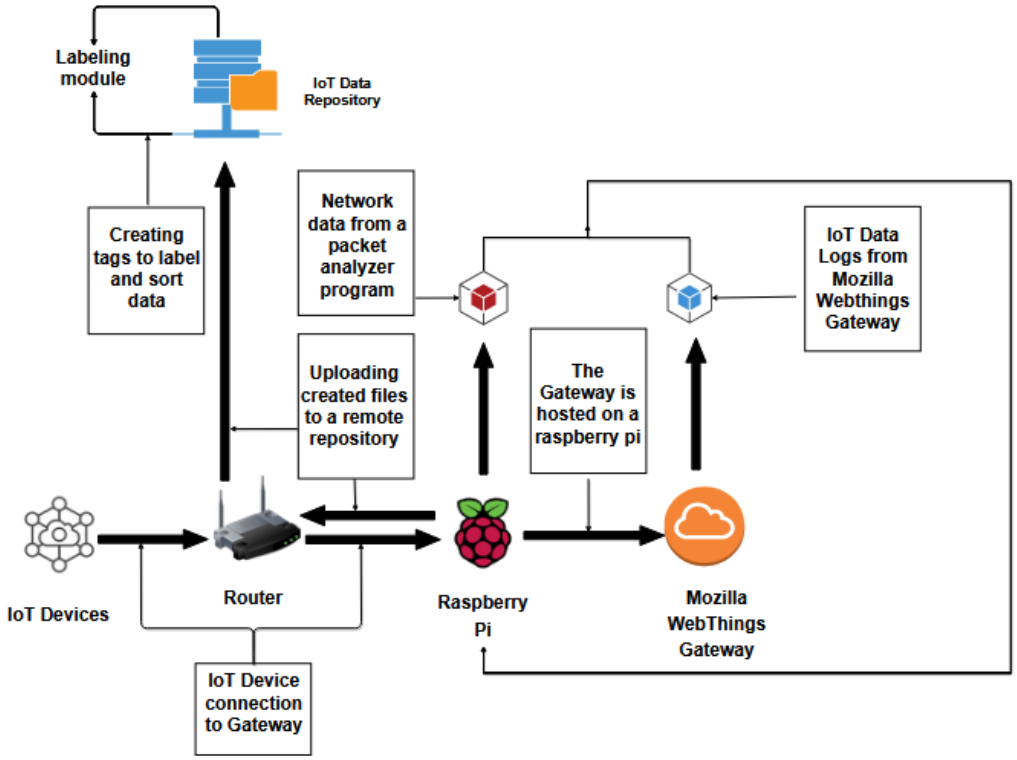

Fig. 2. Current framework in development

3.2 Gateway Setup

Following the illustration in Figure 1, the WebThings Gateway will be hosted on a raspberry pi which will be accessed by ssh enabling control over the operating system. The gateway is built every time it is run as a node module, downloaded from the github page for the project.

We then proceeded to configure the IoT devices, enabled access for the smart devices to the network and then installed the required add-ons for the gateway allowing us to monitor the devices.

3.3 Data Collection Procedure

The WebThings gateway has a feature that allows the gateway to log device data from the smart devices connected to it. However, we needed three things from the gateway:

- The ability to save these logs and upload them to an online repository.
- A method allowing us to capture network data and upload it to the repository as well.
- A method allowing us to upload the aggregated data.

The gateway itself supports logging of device data and storing it for a user defined time period. As the gateway does not have an official method to support sharing of logs we will focus on accessing the logs in the near future. The method to capture network data would be using Wireshark in order to capture the incoming network traffic and filtering it by the IP addresses of the smart devices in order to avoid capturing arbitrary data. We will use this in tandem with a script to appropriately

sort/tag the data leveraging the timestamps and then uploading the packet capture files along the log files time stamped by the day and labelled by the device to an online repository.

4 CONTRIBUTIONS/SIGNIFICANCE

The work sets up a proposal for a novel method using a standard set by the World Wide Web Consortium. The live repository that will be created as an end result of this project would contribute to data science as well as artificial intelligence communities as the applications of such a data set have no boundaries. This work is important as there is no existing solution focusing on data collection, rather the papers which have been reviewed as part of the research conducted inference that the projects focus on pointing out the security flaws in IoT devices.

The primary focus of the framework is to collect data in a comprehensive manner. The data will then be dedicated to be an online repository. If this follows expectation, we will have a framework to capture IoT device data along with other closely related artifacts about the data creating a comprehensive data set, which would be continually expanding.

REFERENCES

[1] Web of things (wot) architecture.

[2] A. Arora, E. Ertin, R. Ramnath, M. Nesterenko, and W. Leal. Kansei: a high-fidelity sensing testbed. *IEEE Internet Computing*, 10(2):35–47, March 2006.

[3] Josh Bers, Abhimanyu Gosain, Ian Rose, and Matt Welsh. Citysense: The design and performance of an urban wireless sensor network testbed. 01 2008.

[4] Abiy Biru Chebudie, Roberto Minerva, and Domenico Rotondi. *Towards a definition of the Internet of Things (IoT).* PhD thesis, 08 2014.

[5] A. Gluhak, S. Krco, M. Nati, D. Pfisterer, N. Mitton, and T. Razafindralambo. A survey on facilities for experimental internet of things research. *IEEE Communications Magazine*, 49(11):58–67, November 2011.

[6] Mozilla. mozilla-iot/gateway, Mar 2020.

[7] Shachar Siboni, Vinay Sachidananda, Yair Meidan, Michael Bohadana, Yael Mathov, Suhas Bhairav, Asaf Shabtai, and Yuval Elovici. Security testbed for internet-of-things devices. *IEEE Transactions on Reliability*, 68(1):23–44, 2019.

[8] G. Werner-Allen, P. Swieskowski, and M. Welsh. Motelab: a wireless sensor network testbed. In *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.*, pages 483–488, April 2005.

[9] Z. Zhang, M. C. Y. Cho, C. Wang, C. Hsu, C. Chen, and S. Shieh. Iot security: Ongoing challenges and research opportunities. In *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, pages 230–234, Nov 2014.