SET YOUR TITLE USING \title

I. Results and Analysis

1.1 Overview

This section describes the results obtained using **CITIoT!** (**CITIoT!**) during the experimentation described in Chapter 4. Results are discussed in four configurations: (i) Bluetooth Low Energy (BLE), (ii) Wi-Fi with no mitigation, (iii) combined BLE and Wi-Fi with no mitigation, and (iv) Wi-Fi with mitigation. Configurations three and four are used when analyzing **CITIoT!**'s overall accuracy and performance. Section 1.2 discusses the accuracy of **CITIoT!** by examining the applicable metrics. The performance of **CITIoT!**, as defined by the **PT!** (**PT!**) and **HDU!** (**HDU!**) response variables, is reported in Section 1.3. Alibis are provided for each failure and results are discussed to provide insight into smart home leakage and its security ramifications.

1.2 CITIoT Accuracy

This section analyzes **CITIoT!**'s accuracy against **SHAA!** (**SHAA!**) (discussed in Section 4.7) using the response variables **DCSR!** (**DCSR!**), **EISR!** (**EISR!**), **EIFP!** (**EIFP!**), **EIFN!** (**EIFN!**), and **ULSR!** (**ULSR!**). Results are presented for all four configurations and are calculated using the script presented in Appendix **??**. The script compares a trial's event identification output from **CITIoT!** with the logs to provide a list of true positives, false positives, and false negatives. The script also calculates the values for each of the response variables per trial and configuration.

Tables **??**, **??**, **??**, and **??** provide results from experiment trials for each configuration respectively. The R script in Appendix **??** produced the plots provided in Appendix **??** and Appendix **??**. The first set of plots consist of four scatter plots per trial day in which **MIoTL!** (**MIoTL!**) was not operating: a plot for the entire day, the morning, the afternoon, and the evening with each showing logged and identified BLE and Wi-Fi events. The second set of plots include one scatter plot per trial day in which **MIoTL!** was operating. As depicted in the example plot provided in Figure **??**, a point with a star and circle depicts a true positive (the tool successfully identified an event), a point with only a circle represents a false negative (the tool failed to identify an event), and a point with only a star shows a false positive (the tool identified an event that did not occur). The R script also calculates the standard deviation and 95% confidence interval for each response variable. Table **??** summarizes the mean results of **CITIoT!**'s accuracy across all trials in each configuration. User tracking is only accomplished via Wi-Fi devices, therefore, the **ULSR!** of BLE trials is not considered. The rest of this section discusses and analyzes the significance of each of the response variables pertinent to the tool's accuracy.

1.2.1 DCSR!.

The **DCSR!** response variable, as presented in Section 4.3, measures **CITIoT!**'s ability to accurately classify devices within **SHAA!**. Table **??** provides the overall mean success rate for each configuration while Appendix **??** supplies the tool's device classification output. The **CITIoT!** tool classified BLE devices with a success rate of 75.0% per trial. The Bike Lock, Apple iPhone 6+, and Apple TV were not identifiable via BLE traffic from the smart home. The Bike Lock used a pseudonym as a device name "00000b67", while both Apple devices did not provide a device name and the BLE Media Access Control (MAC) addresses were randomized. Without mitigation

active, the tool successfully classified all 8 Wi-Fi devices providing a 100% success rate. Of the total 18 internet of things (IoT) devices employed within **SHAA!**, the tool was able to identify 17 using one of the wireless protocols for an overall success rate of 94.4%. With mitigation employed, however, the tool was only able to classify 6 out of 8 Wi-Fi devices for a success rate of 75%. The traffic spoofed by **MIoTL!** caused **CITIoT!** to categorize each device as an outlet. This feature of **MIoTL!** hides the existence of motion sensors and cameras within the smart home from an eavesdropper using **CITIoT!**. The success rates were consistent across all trials for each configuration, therefore the standard deviation of this response variable is not significant.

1.2.2 EISR!.

The **EISR!** response variable defined in Section 4.3, measures **CITIoT!**'s true positive rate, or the ability of the tool to successfully identify events that occur in **SHAA!**. An event is correctly identified if the time and device of the event recognized matches an event in the log. For sensor, camera, and BLE events, an event is considered successful if it the identified time is ± 1 minute from the log time. This success interval provides enough precision for the problem presented in this thesis. Table **??** provides the overall mean **EISR!** results for each configuration. Appendix **??** supplies the tool's event identification output listing each event successfully identified. Over all trials, **CITIoT!** identified 162 out of 170 BLE events for a mean **EISR!** of 94.9%. For trials without mitigation, **CITIoT!** identified 162 out of 173 Wi-Fi events for a mean **EISR!** of 93.7%. For BLE and Wi-Fi trials combined without mitigation, the tool identified a total 324 out of 343 events for a mean **EISR!** of 95.0%. For Wi-Fi trials with mitigation, the tool identified a total of X out of X events for a mean **EISR!** of X%. The standard deviations and 95% confidence intervals, listed

in Table **??**, show that the **EISR!** for each trial was close to the mean of all trials per configuration. This provides a high confidence that the tool can consistently identify events at a combined rate greater than 90% when **MIoTL!** is not operating. **CITIoT!** was still effective in identifying events during mitigation and still operated at a success rate greater than 90%.

1.2.3 EIFP!.

The **EIFP!** response variable described in Section 4.3, measures **CITIoT!**'s false positive rate, or the rate at which the tool identifies events that did not occur in **SHAA!**. An event is a false positive if there is no corresponding entry in the logs. Table **??** provides the overall mean **EIFP!** results for each configuration. Appendix **??** supplies the tool's event identification false positives for each trial.

Of the 176 BLE events identified by **CITIoT!**, 8 did not occur providing a mean **EIFP!** rate of 4.1%. BLE false positives occur because **CITIoT!** identifies events based off of connection requests sent from a master to a slave which may occur outside of smart home events. For example, a phone and a temperature sensor might create a connection to pass battery information that **CITIoT!** would mistakingly identify as a smart home event. Of the false postives, all 8 were from one of the Eve devices and the wireless traffic is encrypted.

For trials without mitigation, 5 Wi-Fi events did not occur out of the 166 identified for a mean **EIFP!** of 2.9%. All 5 false positives were either motion or camera events. The **CITIoT!** tool identifies these type of events by summing the total number of packets sent within a minute. If the sum of these packets reaches a threshold, then an event is identified. After traffic analysis, it was observed that these false positives occur when the two devices send enought packets in a minute to trigger **CITIoT!** to falsely identify an event. The wireless traffic is encrypted so it is unclear if these false

positives were caused by actual events that failed to report to the Homebridge log or other status traffic. The mean **EIFP!** of 2.9%, however, is well within the desired results.

A combined 13 of the 342 events identified did not occur resulting in a mean **EIFP!** rate of 3.2%. With mitigation, however, X out of X total events identified by **CITIoT!** were false positives providing a mean **EIFP!** rate of X%. The standard deviations and 95% confidence intervals, provided in Table **??**, show that the **EIFP!** for each trial was consistent to the mean of all trials per configuration. Therefore, without mitigation **CITIoT!** consistently produces false positives with a combined rate less than 7%. The BLE **EIFP!** rate was higher than Wi-Fi and indicates that portion of the identifier may need to be altered. Mitigation increased the **EIFP!** rate for Wi-Fi devices to X% and made **CITIoT!** ineffective at differentiating real events from spoofed events.

1.2.4 EIFN!.

The **EIFN!** response variable, as provided in Section 4.3, measures **CITIoT!**'s false negative rate, or the rate at which the tool fails to identify events that occur in **SHAA!**. Table **??** provides the **EIFN!** results for each configuration. Appendix **??** supplies the tool's event identification output listing each false negative.

CITIoT! failed to identify 8 BLE events from the 170 total events resulting in a 5.2% mean **EIFN!** rate. There are three primary reasons the tool can fail to identify a BLE event: first, the connection request packet sent by the master may not be collected by the sniffers due to wireless interference; second, the connection request packet may have been collected by the sniffers, but corrupted; and third, the sniffer set to listen on the advertisement channel that a connect request packet was sent may already be following a different connection and, therefore, would not collect the

connect request packet. The cause of a false negative can be determined via traffic analysis. The lack of connection events around the time of an event can indicate the first cause. A malformed connect request packet at the time of an event can point towards the second cause. If a sniffer is following a connection at the time of an unidentified event can suggest the third reason. Traffic analysis of the 8 false negatives indicates that four of the events were probably caused by interference, while the other four could have been a result of the Ubertooth One sniffers following other connections.

Without mitigation, a total of 11 out of 173 Wi-Fi events were not identified for a mean **EIFN!** of 6.3%. The main reason the tool may fail to identify event is if the Alfa Card sniffer fails to capture the packet sent from the Rapberry Pi to a device due to congestion. There is no way to guarantee the cause, but the traffic analysis indicates this was the cause for all 11 false negatives.

Combined with no mitigation, the tool failed to identify 19 of 342 events for a mean **EIFN!** rate of 5.1% across all trials. With mitigation, X out of X events were not identified by **CITIoT!** resulting in a mean **EIFN!** rate of X%. The standard deviations and 95% confidence interval, provided in Table **??**, suggest that the **EIFN!** for each trial was consistent to the mean of all trials per configuration. Therefore, without mitigation **CITIoT!** consistently fails to identify events at a rate less than 10%.

1.2.5 ULSR!.

The **ULSR!** response variable measures the rate at which **CITIoT!** accurately determines if a user is in the smart house or not. Tracking is accomplished via Wi-Fi wireless traffic, so there are no results from the BLE configuration. Without mitigation, the tool successfully tracked the user's location for all but 9 minutes out

of 51 hours and 49 minutes. This resulted in a mean **ULSR!** of 99.7% across trials. Inaccuracies in timing can be explained by the Apple iPhone 6+ connecting to the access point (AP) as the user is walking up to or away from the smart home and prior to or after the times reflected in the logs.

1.3 CITIoT Performance

This section analyzes **CITIoT!**'s performance while operating against **SHAA!** using the response variables **PT!** and **HDU!** discussed in Section 4.3. The performance parameters are presented for the BLE and WI-Fi components separately and **CITIoT!** as a whole.

1.3.1 PT!.

The **PT!** response variable measures the average processing time for each component and the **CITIoT!** tool as a whole. Each component and trial processes a different number of packets so **PT!** is normalized by calculating the average **PT!** per 25,000 packets. The individual times are calculated while that component is operating alone. In operation, the BLE and Wi-Fi components operate simultaneously in two different processes. Therefore, the **CITIoT! NPT!** (**NPT!**) is calculated when both components are operating concurrently. Table **??** provides the **NPT!** results for each component and the tool as a whole. The first trial does not have a BLE component, so that trial's timing is not considered in **CITIoT!**'s total average, but is included for the Wi-Fi units' averages.

The mean **NPT!** for the BLE component of **CITIoT!** operating alone is 24.1 seconds per 25,000 packets. On average, one hour of captured BLE traffic resulted in about 441.75 thousand packets and took the BLE component of the tool approximately 7 minutes to process. The standard deviation of .2 indicates that the individual

trials' **NPT!** did not vary considerably.

The average **NPT!** for the Wi-Fi component of **CITIoT!** is 22.0 seconds per 25,000 packets. On average, one hour of captured Wi-Fi traffic resulted in about 417.73 thousand packets and took the Wi-Fi component of the tool approximately 6 minutes to process. The standard deviation of .7 indicates that the individual trials' **NPT!** did not vary considerably.

The average **NPT!** presented for **CITIoT!** as a whole assumes the BLE and Wi-Fi components are operating simultaneously. On average, the **NPT!** is seconds per 25,000 packets.

1.3.2 HDU!.

1.4 Chapter Summary

Bibliography