IoTArgos: A Multi-Layer Security Monitoring System for Internet-of-Things in Smart Homes

Yinxin Wan, Kuai Xu, Guoliang Xue, Feng Wang
Arizona State University
Emails: {ywan28, kuai.xu, xue, fwang25}@asu.edu

***Abstract*—The wide deployment of IoT systems in smart homes has changed the landscape of networked systems, Internet traffic, and data communications in residential broadband networks as well as the Internet at large. However, recent spates of cyber attacks and threats towards IoT systems in smart homes have revealed prevalent vulnerabilities and risks of IoT systems ranging from data link layer protocols to application services. To address the security challenges of IoT systems in smart homes, this paper introduces IoTArgos, a multi-layer security monitoring system, which collects, analyzes, and characterizes data communications of heterogeneous IoT devices via programmable home routers. More importantly, this system extracts a variety of multi-layer data communication features and develops supervised learning methods for classifying intrusion activities at system, network, and application layers. In light of the potential zero-day or unknown attacks, IoTArgos also incorporates unsupervised learning algorithms to discover unusual or suspicious behaviors towards smart home IoT systems. Our extensive experimental evaluations have demonstrated that IoTArgos is able to detect anomalous activities targeting IoT devices in smart homes with a precision of 0.9876 and a recall of 0.9763.**

I. Introduction

Recent advances of embedded systems, wireless communications, cloud computing, and artificial intelligence have successfully driven the widespread deployment of IoT devices in millions of smart homes across the world [4, 7, 35, 36, 42]. However, the prevalent device, system, and application vulnerabilities and weakness of IoT systems [3, 15, 24, 27, 48] due to security design flaws, weak password management, vague trust management, lack of IoT security standards, and resource constraints for cryptographic functions have enabled cyber attacks to compromise and control millions of IoT devices for launching large scale distributed denial of service (DDoS) attacks, controlling all lights in a city, and breaking into residential homes or offices with smart but insecure locks [5, 22, 23, 32, 38, 41].

The broad attack vector across the entire IoT protocol stack against smart home IoT systems has created challenges for existing application-driven or device-specific security solutions [12, 26, 30, 47]. Thus, securing IoT systems in smart homes calls for security frameworks and standards that consider the weakness and vulnerabilities in all IoT protocol layers. Towards this end, this paper introduces IoTArgos, a security monitoring system that collects, analyzes, and characterizes *multi-layer* data communications of *all* IoT devices in smart homes via programmable home routers.

IoTArgos leverages home routers powered by OpenWrt, an embedded Linux operating system, to automatically collect TCP/IP-based network flow records via softflowd and nfcapd utilities and wireless packets captured by off-the-shelf wireless sniffers plugged into the routers. The combination of network flows and wireless packets offers a wide range of multi-layer features which capture behavioral patterns of data communications for heterogeneous IoT systems in smart homes and explain what, when, how, if, and why IoT devices communicate with other end systems including remote cloud servers or local IoT hubs in the same home.

In light of prevalent security threats towards IoT devices, we develop a two-stage machine learning (ML) based intrusion detection module in IoTArgos. This module explores supervised classification algorithms for detecting known attacks in the first stage, while relies on unsupervised anomaly detection algorithms in the second stage for capturing emerging zero-day attacks that are likely undetected by the classification stage.

To demonstrate the performance and benefit of the ML-based intrusion detection method, we replay a wide range of cyber attacks at different protocol layers against selected IoT systems in one smart home that deploys IoTArgos on an OpenWrt-supported Linksys WRT1900ACS home router equipped with one 1.6GHz dual-core processor and 512MB memory. Our extensive experimental results based on synthetic IoT data communication traffic demonstrate the effectiveness of the ML-based intrusion detection method in capturing known or new attack behaviors towards smart home IoT devices. Specifically, the two-stage method in IoTArgos, using a combination of random forest (RF) in the classification stage and principal component analysis (PCA) in the anomaly detection stage, achieves a high area under the curve (AUC) value of 0.9678 and 0.9876, 0.9763, 0.9818, 0.9819 of precision, recall, accuracy, $\mathcal{F}_1$ score in detecting IoT attacks.

The contributions of this paper are summarized as follows:

- We propose, design, implement, and evaluate IoTArgos, a multi-layer security monitoring system for characterizing data communications of heterogeneous IoT devices in smart homes and detecting a wide range of anomalous and intrusion activities towards IoT devices.
- We characterize the behavioral patterns of various IoT devices deployed in real-world smart homes based on a broad range of data communication and traffic features from multiple layers of TCP/IP protocol and IoT communication protocol.
- We develop an innovative two-stage ML-based intrusion

detection method for detecting a wide spectrum of attacks towards IoT devices, and run extensive experiments with synthetic IoT data traffic to demonstrate that the method is able to detect anomalous activities targeting IoT devices in smart homes with a precision of 0.9876 and a recall of 0.9763.

The remainder of this paper is organized as follows. Section II describes the background of smart home IoT security. Section III presents the system architecture and key components of our proposed IoTArgos security monitoring system, while Section IV describes the characterization of IoT communication patterns. Section V introduces the two-stage ML-based intrusion detection method, and Section VI demonstrates its performance. Section VII discusses related work, and Section VIII concludes this paper and outlines our future work.

II. Background

In this section, we first describe the rising deployment, applications, and services of IoT devices in millions of smart homes, and then discuss the existing vulnerabilities of heterogeneous and weakly-protected IoT devices and the challenges to secure them. Finally, we shed light on multi-layer behavioral fingerprints left by real-world IoT attacks and threats.

A. Internet-of-Things at Smart Homes

The recent years have witnessed the explosive deployment of Internet-connected IoT devices in smart homes. For example, as illustrated in Fig. 1, a smart home today can connect a wide range of IoT devices such as IP surveillance cameras, smart thermostats and smoke detectors, voice assistants, and air quality monitors via different communication protocols for home automation, home security and safety, energy efficiency, and healthcare.

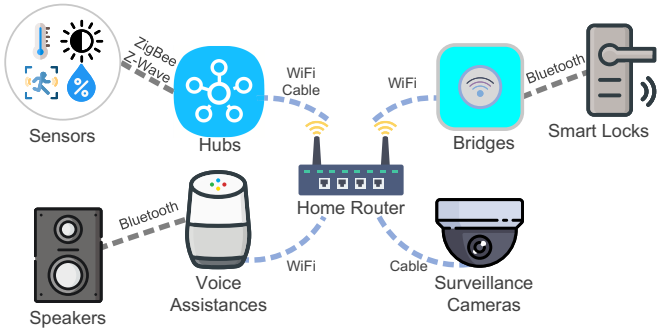

Fig. 1. An example smart home network with heterogeneous IoT devices.

Smart home IoT devices connect to the Internet either *directly* via running TCP/IP protocol stacks on themselves or *indirectly* via relying on a smart home platform such as Samsung's SmartThings [35], Google's Nest [18], and Amazon's Alexa [4]. In smart homes, the Internet-capable IoT devices such as smart TVs, IP surveillance cameras, Amazon Echo, and Google Home typically use wired cables or IEEE 802.11 wireless protocols (WiFi) to connect to home routers, while many embedded IoT devices such as Philips Hue smart bulb, Samsung multipurpose sensor, and August smart lock use low-power wireless protocols such as Zigbee, Bluetooth Low Energy (BLE), and Z-Wave to connect to the Internet via a smart gateway, hub, or bridge. For ease of presentation, we refer the Internet-capable IoT devices as *ic-IoT devices*, and the gateway-supported embedded IoT devices as *em-IoT devices*.

B. Security Challenges of Heterogeneous IoTs

As connected IoT devices in smart homes continue to grow in size and complexity, the security issue has become one of the top challenges in IoT research community. Securing IoT systems in smart home is a daunting task due to the heterogeneity of IoT systems, the prevalence of vulnerabilities in IoT devices and applications, as well as the broad attack vector across the entire IoT protocol stack. For example, the recent Mirai botnet [5, 23], formed by hundreds of thousands of IoT devices, launched an aggregated 600 Gbps DDoS attack towards Brian Krebs's security blog. Thousands of home IP cameras, as part of Mirai botnet, are remotely exploited by attackers via universal plug and lay (UPnP) enabled home routers which allow IoT devices behind network address translation (NAT) protection to automatically bind a service port for communicating with remote networked systems [38]. Another innovative attack [27, 32] discovers and leverages the software implementation bugs in the Zigbee light link (ZLL) protocol [48], to potentially control all the lights in a city via spreading IoT worms from a single infected bulb, i.e., patient zero, to all compatible IoT lights using their built-in Zigbee wireless connectivity and physical adjacency. Similarly, misbehaving and malicious smart home applications could explore the design flaws of smart home programming platforms such as Samsung SmartThings to gain over-privileged access and control of IoT systems and to launch event spoofing attacks e.g., triggering fake fire alarms [15, 47].

C. Multi-Layer Behavioral Fingerprint of IoT Attacks

Many cyber attacks towards IoT devices leave traffic and behavioral fingerprints at different TCP/IP layers and IoT protocol stacks. For example, the Mirai botnet employs three stages including infiltration, infection, and operation for scanning, controlling, and exploiting vulnerable IoT devices. These steps trigger Internet data traffic between the attacking devices and the UPnP-enabled home router as well as wireless packets between the router and wireless IP cameras or other IoT devices. Similarly, if an attacker with unauthorized access to an August smart lock account via a Web interface or a smart phone app remotely opens and closes the smart lock via August cloud services, the events will leave IP, WiFi, and Bluetooth data traffic between the cloud servers and the home router, the router and the August connect bridge, the bridge and the August smart lock, respectively. Therefore, the broad attack vector over the entire IoT protocol stack calls for a multi-layer security monitoring and analysis platform.

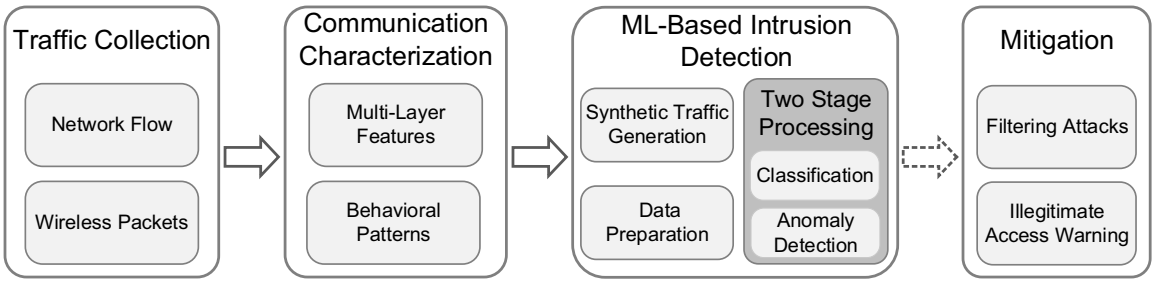

Fig. 2. The overall system architecture of IoTArgos.

III. IoTArgos System Overview and Design

In this section, we first present the system overview and architecture of IoTArgos. Subsequently, we discuss each key component of IoTArgos for monitoring and measuring IoT data communications via a multi-layer approach.

A. IoTArgos System Overview

IoTArgos is a multi-layer smart home security monitoring system for monitoring and analyzing data communications of IoT systems in smart homes and detecting and mitigating intrusions and anomalous activities. The design and implementation of IoTArgos are router-centered. Our intuition of developing IoTArgos on the home router originates from the network architecture of smart homes. Specifically, consumer-grade home routers serve as the residential gateway to route and forward data packets between internal IoT devices in smart homes and external cloud servers of IoT vendors or other remote servers on the Internet. The physical connection from IoT devices to the router is established either directly through wired cables or WiFi, or indirectly via a hub or bridge. In recent years, a number of research studies have explored the computation, storage, and bandwidth resources on commodity home routers to characterize end-to-end performance and troubleshoot performance anomalies and mis-configurations in home networks [2, 13, 45].

Fig. 2 illustrates the overall system architecture of IoTArgos, which consists of four key components: data collection, IoT communication characterization, ML-based intrusion detection, and real-time mitigation and defense. The primary objective of the *data collection* component is to configure and setup data collection instruments on programmable home routers, while the *IoT communication characterization* component is devoted to characterizing and profiling communication patterns of IoT devices. The *ML-based intrusion detection* component first explores supervised classification algorithms to classify known attacks of IoT data communications and then relies on unsupervised anomaly detection algorithms to detect unknown suspicious activities or zero-day attacks. Finally, the *real-time mitigation and defense* component, beyond the scope of this paper, is responsible for alerting home owners and automatically configuring firewall polices or deploying other defense mechanisms for filtering and mitigating anomalous activities towards or from IoT devices.

B. Key System Components

1) Multi-Layer IoT Data Collection: A key strength of our proposed router-centered security monitoring system lies in its flexibility of collecting all data communications at a centralized location. Such simple yet effective deployment and configuration is crucial for millions of regular home users to adopt the system, as existing application-driven or device-specific solutions require non-trivial skills and efforts for home users to manage and secure IoT devices with diverse operating systems and interfaces in smart homes. As the residential gateway of broadband home networks, many programmable home routers including bare-bone Raspberry Pi models have the computational resources and open-source packages to capture and store raw IP data packets and aggregated network traffic flows as well as the Ethernet and WiFi frames. Many battery-operated IoT devices in smart homes such as smart locks only communicate with low energy wireless protocols such as Zigbee and Bluetooth, which have been proven to be insecure by design. Therefore, in order to collect the entire data frames from heterogeneous IoT devices adopting different link layer protocols, we setup Texas Instrument CC2531 and CC2540 USB dongles for the capture of Zigbee and Bluetooth frames respectively. In this study, IoTArgos collects both network flow records and wireless packets of Zigbee, Bluetooth and WiFi protocols[1] at programmable home routers.

2) Characterizing IoT Data Communications: The collected multi-layer data by programmable home routers allow us to systematically characterize data communication behaviors of all IoT devices in smart homes, e.g., when, how, and why IoT devices in smart homes communicate with cloud servers, other remote networked systems, mobile applications that control the devices, home routers, and local IoT hubs. Specifically, IoTArgos profiles data communications of IoT devices with a broad range of basic *raw* communication and traffic features such as the IP address and domain name of remote end hosts, inter-packet arrival time, packet size, flow duration, source port, destination port, protocol, link layer protocol, as well as aggregated features such as the number and dynamics of remote hosts, and the dominant applications. These traffic features not only characterize IoT data communications, but also play a crucial role in the ML-based intrusion detection component.

[1]We are unable to capture Z-Wave wireless packets due to the unavailability of Z-Wave USB dongles for consumer-grade home routers.

3) ML-based Intrusion Detection: In general, we classify cyber attacks towards IoT systems as known and unknown attacks. The signature and patterns of known attacks are often public knowledge, while unknown attacks, e.g., new or zero-day attacks are typically not discovered or reported yet. Similar to traditional firewalls, detecting known intrusions and attacks often requires signature-based techniques or supervised machine-learning based methods which are often unable to uncover new attacks. In order to capture both known and unknown attacks, IoTArgos adopts a two-stage approach for ML-based intrusion detection: i) supervised classification stage, and ii) unsupervised anomaly detection stage. The first stage applies one supervised machine learning algorithm on the collected multi-layer data for classifying IoT attacks or normal IoT data communications, while the second stage applies one unsupervised anomaly detection algorithm to uncover anomalous behaviors that are not detectable by the supervised classification stage due to the unavailability of attack signatures or training data-sets.

4) Real-Time Mitigation and Defense: Once IoTArgos detects and identifies intrusion activities towards or from IoT devices in smart homes, the real-time mitigation and defense component will be triggered to take appropriate actions such as alerting the home owners via automated emails, disabling or disconnecting compromised IoT devices and their respective hubs, and configuring firewall polices on the home routers to filter the intrusion activities based on attack behavioral fingerprints.

IV. Characterizing IoT Data Communications

In this section, we first present how IoTArgos characterizes IoT data communications with features from multi-layer IoT communication protocols, and subsequently discuss the inherent and distinct communication patterns of diverse IoT devices deployed in real world smart homes.

A. Multi-Layer Feature Characterizations

We have deployed IoTArgos over 22 home networks across United States, Hong Kong SAR, and mainland China since August 2018. These smart homes house hundreds of IoT devices for a variety of purposes, among which we have observed 20 unique types of IoT devices, as summarized in Table I.

IoT devices in smart homes often exhibit various functions and diverse computation, storage, and communication capabilities. For example, smart TVs and Amazon Echo are often powered by electric power and have wired or wireless connections to home routers for Internet connections, while battery-operated motion and water leak sensors rely on low energy wireless communication protocols such as Zigbee, Z-Wave, or Bluetooth to connect with the specific bridges (also called hubs or gateways) which support both TCP/IP protocols and IoT wireless protocols and standards.

To characterize data communication for *all* IoT devices in smart homes with a *centralized* solution, IoTArgos explores a wide range of multi-layer features from TCP/IP-based network

TABLE I
The list of heterogeneous IoT devices observed in the 22 smart homes running IoTArgos.

IoT Device	Function	IoT Protocol	Type
Amazon Echo Dot	Voice Assistant	Bluetooth & WiFi	ic-IoT
Amazon Echo	Voice Assistant	Bluetooth & WiFi	ic-IoT
August Connect Bridge	Gateway	WiFi & Bluetooth	bridge
August Smart Lock Pro	Smart Lock	Bluetooth & Z-Wave	em-IoT
Google Home	Voice Assistant	WiFi & Bluetooth	ic-IoT
Philips Hue Smart Hub	Gateway	WiFi & Zigbee	hub
Philips Hue White	Smart Bulb	Bluetooth & Zigbee	em-IoT
Ring Video Doorbell	Doorbell	WiFi	ic-IoT
Samsung General Sensor	Sensor	Zigbee	em-IoT
Samsung Motion Button	Smart Button	Zigbee	em-IoT
Samsung Motion Outlet	Smart Outlet	Zigbee	em-IoT
Samsung Motion Sensor	Sensor	Zigbee	em-IoT
Samsung Hub	Gateway	Ethernet & WiFi Zigbee & Z-Wave	hub
Samsung Water Sensor	Snesor	Zigbee	em-IoT
Aeotec Multisensor 6	Sensor	Z-Wave	em-IoT
Aeotec Siren Gen5	Sensor	Z-Wave	em-IoT
TCL Smart TV	Smart TV	Ethernet & WiFi	ic-IoT
LG Smart TV	Smart TV	Ethernet & WiFi	ic-IoT
YI Home Camera	Camera	WiFi	ic-IoT
Reolink Camera	Camera	Ethernet & WiFi	ic-IoT

flow records extracted by *softflowd* and *nfcapd* packages installed on programmable home routers and wireless packets captured by wireless sniffers installed on the routers. The majority of IoT applications and services in smart homes have employed encryption in data communications, thus IoTArgos focuses on behavioral features such as the size, durations, protocols, and remote end systems of data communications.

B. Exploring Behavioral Patterns of IoT Devices

Mining and correlating multi-layer features of IoT data communications allow us to gain a deep understanding on behavioral patterns of these IoT devices in smart homes, which is a critical first step for securing these devices and detecting anomalies. For example, Fig. 3 shows the numbers of Zigbee wireless packets exchanged between a Samsung outlet and Samsung SmartThings hub as well as IP data packets between the hub and Samsung servers hosted on Amazon cloud over a 24-hour time window.

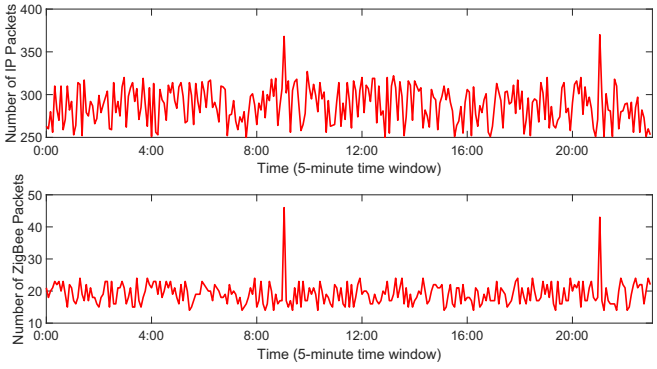

Fig. 3. The numbers of Zigbee wireless packets exchanged between a Samsung outlet and Samsung SmartThings hub as well as IP packets between the hub and Samsung could servers.

The objective of building behavioral patterns of IoT devices is to understand *what*, *when*, *how*, *if*, and *why* the devices communicate with other systems including their *local* bridges, hubs, or gateways in the same home and *remote* cloud servers.

Towards this end, IoTArgos extracts the *basic* traffic features of network flows and wireless packets originating from or destined to the IoT devices including source MAC address, destination MAC address, source IP address, destination IP address, source port, destination port, protocol, flow volumes in packets and bytes, flow duration, and then derives the *advanced* features such as inter-packet arrival time, average packet size, domain and autonomous system number (ASN) of cloud server. For example, Table II illustrates the observations of a number of features for six typical IoT devices over a 24-hour time window in one smart home equipped with IoTArgos. As shown in Table II, IoT devices exhibit distinct traffic patterns as evidenced by the behavioral fingerprints on remote ports, remote servers, traffic volumes and intensity.

The multi-layer data communication features allow IoTArgos to establish a behavioral profile to summarize the communication patterns of IoT devices over time. Such behavioral profiles and communication patterns, if captured in a controlled smart home laboratory environment, can serve as a baseline of normal IoT communications. To identify multi-layer features of IoT data communications, we rely on permanent physical layer addresses of IoT devices and their hubs as well as the temporal sequences to associate network flow records with wireless packets for generating *augmented* network flow records which summarize individual conversations or events of IoT applications and services, e.g., remotely opening an August smart lock via data communications among the lock owner's smart phone, August cloud server, the smart home router, and the August connect bridge.

The monitoring and analysis on IoT behavioral patterns and dynamics by IoTArgos have led to many interesting observations. For example, some cloud-based IoT devices like voice assistants and smart hubs often have long or periodical connections for exchanging low-bandwidth heartbeat signals with cloud servers. However, once a certain event happens, e.g., the smart lock is opened or the voice assistant is requested to play a Spotify song, the underlying data traffic features exhibit dramatic changes. Such observations suggest that IoTArgos provides a uniform and consistent framework based on behavioral patterns to tell what is happening to heterogeneous IoT devices in smart homes, and more importantly to protect these devices via proactive securing monitoring and reactive real-time mitigation and defense.

In summary, IoTArgos characterizes and profiles IoT data communications with a broad range of features from multi-layer IoT protocol stacks. Mining these multi-layer features also leads us to discover distinct communication patterns of IoT devices in smart homes. These features and patterns provide critical insights and valuable inputs for exploring ML algorithms to detect intrusion activities and anomalous behaviors towards or from IoT devices.

V. Machine-Learning Based Intrusion Detection

A number of recent studies and surveys have reported that the prevalent and exploitable vulnerabilities of IoT systems in smart homes have enabled the attackers to compromise and control thousands of IoT systems for launching large scale DDoS attacks or listening to the private conversations in hacked smart voice assistants. Hence, detecting intrusion activities and anomalous behaviors of IoT systems in smart homes is an important design goal of IoTArgos.

In this section, we present our two-stage ML-based intrusion detection design in IoTArgos, which relies on the supervised classification algorithms in the first stage to classify known attacks, and explores the unsupervised anomaly detection in the second stage to detect unknown or zero-day attacks towards IoT devices in smart homes.

A. First Stage: Supervised Classification

The intuition of our proposed two-stage intrusion detection strategy is driven by the diversity and complexity of the existing and potential attacks and attacks towards heterogeneous IoT devices in smart homes. Traditional signature-based detection methods or emerging ML-based classifications approaches are very efficient for classifying attacks whose signatures are available or whose prior instances are captured and labelled. However, such methods often have challenges in recognizing zero-day attacks that are created by attackers via exploiting newly discovered or exposed vulnerabilities from one or more types of IoT devices.

Therefore, as a first step, IoTArgos explores supervised classification algorithms to detect and filter a subset of attacks via training a suite of classification algorithms and selecting the *desired* algorithm that balances the intrusion detection performance and system consumption such as CPU and memory cost on resource-constrained home routers. Specifically, we choose five well-known and computationally lightweight classification algorithms including k-nearest neighbors (k-NN), logistic regression (LR), naïve bayes (NB), RF, and support vector machine (SVM).

B. Second Stage: Unsupervised Anomaly Detection

As several prior studies on Internet intrusion and anomaly detection have pointed out, cyber attackers often employ new attack techniques thanks to the newly discovered zero-day vulnerabilities in the compromised systems, they often exhibit similar behavioral patterns, e.g., such as port scanning, penetration testing, and brute-force password attempts as existing attacks. Given the likely new attacks that are misclassified as "normal" in the first stage, IoTArgos develops the second-stage with anomaly detection algorithms for identifying new attacks from the remaining "normal" data communications. In this stage, we also select and evaluate computationally lightweight anomaly detection algorithms such as clustering-based local outlier factor (CBLOF), fast angle-based outlier detection (FastABOD), feature bagging (FB), isolation forest (IForest), local outlier factor (LOF), and PCA.

Therefore, the goal of the second stage is to uncover anomaly behaviors that are not detectable by the supervised intrusion classification technique due to the unavailability of attack signatures or training data-sets or the emerging new

TABLE II
DISTINCT BEHAVIORAL FEATURES OF A SAMPLE SET OF IOT DEVICES OVER 24-HOUR TIME WINDOW.

Feature	Amazon Echo	Google Home	Samsung Hub	Hue Bridge	August Lock Bridge	Reolink Camera
Remote ports	53, 80, 123, 137, 443	53, 443, 5228, 5353	53, 443	80, 123, 443	80, 443	11004, 10995, 51097, 51223
Remote servers	352 unique IP from Amazon and NTP servers	27 unique IP from Google and DNS servers	66 unique IP from Amazon	24 unique IP from Amazon, Akamai CDN, and Google	8 unique IP from Amazon and Akamai CDN	6 unique IP from Amazon and Verizon networks
ASNs	AS16509 (Amazon), AS7806 (Binary Net)	AS15169 (Google)	AS16509 (Amazon)	AS20940 (Akamai), AS15169 (Google), AS16509 (Amazon)	AS20940 (Akamai), AS16509 (Amazon)	AS14618 (Amazon), AS22394 (Cellco)
Dominant ASN	AS16509	AS15169	AS16509	AS20940	AS16509	AS22394
Dominant service	HTTP	HTTPS	HTTPS	HTTPS	HTTPS	UDP
Total packets	141k	151k	86k	104k	116k	94k
Packets per 5 min.	490	527	302	364	406	326
Total bytes	46MB	25MB	19MB	26MB	22MB	33MB
Bytes per 5 min.	161KB	88KB	66KB	91KB	77KB	117KB
Total flows	5,464	920	6,768	765	20	9
Flows per 5 min.	19	3	24	3	0	0

vulnerabilities or weakness of IoT systems to be discovered by attackers. In other words, each data communication with various multi-layer features of IoT systems in smart homes will be initially inspected by the the first-stage classification algorithms in IoTArgos. If the classifier reports normal, the record will go through the second stage anomaly detection model.

VI. PERFORMANCE EVALUATIONS

We have implemented the IOTArgos system and deployed the system across 22 real-world smart home networks. In this section, we present results of our extensive performance evaluations along with our key observations. We first describe our experiment setup, synthetic IoT data communication generations, and evaluation metrics. Subsequently, we systematically evaluate the performance of our proposed two-stage intrusion detection technique.

A. Experiment Setup and Synthetic IoT Traffic Generation

To detect and classify attacks and intrusion activities towards IoT devices in smart homes, we built a simple yet effective ML-based intrusion detection component into the IoTArgos system. To demonstrate the performance, benefit, and feasibility of our approach, the IoTArgos system not only collects the normal multi-layer data communications of IoT devices in distributed smart homes in real-time, but also captures the simulated attack traffic towards selected IoT devices. To comprehensively simulate the existing cyber attacks against smart home IoT systems, we referred to recent studies [3, 5, 31, 32] on security attacks and threats towards IoT devices and simulate and replay a wide range of cyber attacks, as summarized in Table III, across multiple layers of IoT communication protocol stack.

For each type of attacks, we wrote dedicated scripts and followed the state-of-the-art penetration test procedure to replay the attacks with varying instances and intensity towards selected IoT devices. In order to replay link layer attacks against devices running the Zigbee and Bluetooth protocols [14, 27, 32, 34, 48], we rely on the widely-used HackRF One transceiver, a software defined radio (SDR) device capable of transferring and receiving radio signals ranging from 1MHz to 6GHz, and run the open-source radio frequency monitoring and injecting tools such as GNU Radio and Scapy-radio for customizing the frequency or the type of the link layer frames.

For collecting the *normal* IoT data communication, we built a smart home sandbox in a laboratory environment that deploys the IoTArgos system on a OpenWrt-supported Linksys WRT1900ACS home router equipped with 1.6GHz dual-core processor and 512MB memory. In the smart home laboratory, all IoT devices are configured behind the NAT-enabled router. In addition, we disabled UPnP from the security setting on the home router to prevent outside attackers in close proximity from directly targeting all IoT devices in the smart home sandbox. The smart home sandbox, consisting of a variety of IoT devices and the programmable home router, has been continuously running for over six months. We consider the data collected by the router in the sandbox during these six months as the *normal* IoT data communications, and combine with the simulated attacks to generate large scale *synthetic* IoT communication data-sets. We eventually built a labelled data-set consisting of over 6 million normal network flow records and over 300 thousand attack flows for our evaluation experiment. Once acquiring the synthetic IoT data communications consisting of normal IoT data communications and simulated attacks, we evaluate the performance and cost of a suite of ML models for determining the most optimal model to balance the intrusion detection quality and the cost of computational and memory resources.

B. Evaluation Metrics

In our synthetic IoT data traffic, we label the simulated attacks as *positive* and normal traffic flows and wireless packets as *negative*. During the performance evaluation of ML-based algorithms, the correctly detected attacks are denoted

TABLE III
THE LIST OF SIMULATED ATTACKS TOWARDS IOT DEVICES IN SMART HOMES.

Category	Attack Strategy	Description
Scanning Attacks	host scanning [5, 43]	identifying IoT devices and scanning for vulnerabilities
	port scanning [5, 31]	
	nexpose scanning [3]	
	nessus scanning [3]	
Flooding Attacks	HTTP flooding [5]	application-layer
	DNS flooding [5]	application-layer
	GRE-IP flooding [5]	application-layer
	UDP flooding [5]	volumetric
	UDP plain flooding [5]	volumetric
	SYN flooding [5]	TCP state exhaustion
	ACK flooding [5]	TCP state exhaustion
	IP AH flooding [5]	IPSec
	IP ESP flooding [5]	IPSec
Brute Force Attacks	SSH brute force [31]	brute forcing
	Telnet brute force [5]	user credentials
Data Link Layer Attacks	ACK spoofing [27, 32]	fake scan response
	blind attack [27]	malicious identify request
	DoS attack [27]	junk traffic
	force re-pairing [34]	manipulate pairing request

by true positive (TP), while the attacks detected as normal scenarios are considered as false negative (FN). Similarly, true negative (TN) refers to the cases when normal IoT data communications are recognized as normal, while false positive (FP) represents the cases when normal IoT communications are incorrectly detected as attacks.

In addition to the TP, FN, TN, FP measures in the widely-used confusion matrix, we also evaluate ML algorithms with precision, recall, accuracy and $\mathcal{F}_1$ score. Specifically, the precision $\mathcal{P}$ is calculated as $\mathcal{P} = \frac{TP}{TP+FP}$, while the recall $\mathcal{R}$ is derived as $\mathcal{R} = \frac{TP}{TP+FN}$. The accuracy metrics $\mathcal{A}$, reflecting the overall correct detection as attacks or normal scenarios, is calculated as $\mathcal{A} = \frac{TP+FN}{TP+FP+FN+TN}$, while the $\mathcal{F}_1$ score, balancing the precision and recall, is derived as $\mathcal{F}_1 = 2 \times \frac{\mathcal{P} \times \mathcal{R}}{\mathcal{P}+\mathcal{R}}$.

C. Evaluation of the First Stage

In the first stage for IoT intrusion detection, we select five widely-used classification algorithms, i.e., k-NN, LR, NB, RF, and SVM, and apply k-fold cross validation with k set as 10 to evaluate and compare their performance. Table IV illustrates accuracy, precision, recall, and $\mathcal{F}_1$ score of the five classification algorithm. As shown in Table IV, all classification algorithms achieve over 0.90 across all metrics except the precision and $\mathcal{F}_1$ score for NB classification.

TABLE IV
THE PERFORMANCE METRICS OF DETECTING IOT INTRUSIONS WITH FIVE CLASSIFICATION ALGORITHMS.

Model	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{F}_1$
k-NN	0.9833	0.9878	0.9649	0.9762
LR	0.9573	0.9718	0.9076	0.9386
NB	0.9195	0.9413	0.8292	0.8817
RF	0.9858	0.9893	0.9703	0.9797
SVM	0.9707	0.9806	0.9365	0.9580

While all these classification algorithms are very effective, the effectiveness of these algorithms depends on the complete knowledge of the attacks. In reality, there are always the possibility of *new* and *unknown* attacks. In order to study the impact of *new* and *unknown* attacks, we conducted a study in which we *intentionally* treat certain types of attacks as "normal" data in the training phase. As a result, the performance of all classification algorithms decreases significantly due to the existence of *new attacks*. Table V shows the decreased performance metrics of detecting IoT intrusions with five classification algorithms with 25% types of IoT attacks with the lowest instances in the synthetic data-set as new or unknown during model training process. For example, the accuracy, precision, recall and $\mathcal{F}_1$ score of RF classification drops 0.1088, 0.1004, 0.0950, 0.0977, respectively. This simple study demonstrates that *the classification algorithm alone, albeit efficient in detecting IoT attacks with high quality labelled data-sets, is not sufficient enough to detect the rising new attacks and threats towards IoT systems in smart homes*.

TABLE V
THE PERFORMANCE METRICS OF DETECTING IOT INTRUSIONS WITH FIVE CLASSIFICATION ALGORITHMS WITH 25% TYPES OF IOT ATTACKS AS NEW OR UNKNOWN DURING MODEL TRAINING PROCESS.

Model	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{F}_1$
k-NN	0.8945	0.8958	0.8781	0.8869
LR	0.8388	0.8620	0.8388	0.8502
NB	0.7802	0.8668	0.7636	0.8119
RF	0.8770	0.8889	0.8753	0.8820
SVM	0.8766	0.8886	0.8642	0.8762

D. The Benefits of the Second Stage

Anomaly detection algorithms have been extensively studied to identify unknown attacks or anomalies towards networked systems. Thus, we introduce anomaly detection algorithms in IoTArgos as the second stage for discovering those "new" types of IoT attacks that are incorrectly detected as normal data communications by the classification algorithm in the first stage. Considering the rare nature of the new or zero-day IoT attacks, we only consider the types of IoT attacks with the lowest instances from our labelled data-set as the "new" types of attacks in the experiments.

The intuition and rationale of introducing anomaly detection algorithms lie in our observations that all outlier IoT data communications are likely anomalous and suspicious activities towards IoT systems since these data communications likely deviate from common and normal IoT data communication patterns. In this study, we select six algorithms i.e., CBLOF, FastABOD, FB, IForest, LOF, and PCA which are widely used for anomaly or outlier detections, and run each of these algorithms against all the "normal" IoT data communications to generate distinct clusters of various sizes for grouping similar patterns or to assign anomaly scores for each communication flow.

TABLE VI
PERFORMANCE METRICS OF COMBINING RF CLASSIFICATION AND THE SECOND STAGE FOR DETECTING KNOWN AND NEW IOT ATTACKS.

Model	$\mathcal{A}$	$\mathcal{P}$	$\mathcal{R}$	$\mathcal{F}_1$
CBLOF	0.9757	0.9798	0.9661	0.9729
FastABOD	0.9241	0.9365	0.9212	0.9288
FB	0.9467	0.9521	0.9353	0.9436
IForest	0.9876	0.9897	0.9750	0.9823
LOF	0.9444	0.9500	0.9342	0.9420
PCA	0.9818	0.9876	0.9763	0.9819

Table VI illustrates the performance of combining RF classification with the second-stage anomaly detection algorithms for detecting new IoT intrusions that are misclassified as

normal communications in the first stage. To systematically evaluate the benefit of adding anomaly detection algorithms in the second stage, we run all the 30 combinations of the (classification, anomaly detection) pair on our labelled dataset of synthetic IoT data communications. Table VII shows the substantial improvement on AUC of combining two-stages over the first classification stage alone, where each entry shows the result of the corresponding (classification, anomaly detection) pair.

We observe from Table VII that running anomaly detection after classification improves accuracy in all cases, and the improvement is significant except for a few combinations. For example, applying PCA anomaly detection after running the RF classification improves 11.36% on the AUC metric, i.e., from an AUC of 0.8691 in the first stage to the final AUC of 0.9678. To have a better view of the improvement, we use Fig. 4 to illustrate both the average improvement and maximum improvement of AUC via combining anomaly detection stage and classification for all 30 combinations. All of the models have significant AUC increment on average and the isolation forest model managed to improve the raw naïve bayes model's AUC over 16%.

In summary, these experimental results demonstrate that our proposed two-stage intrusion detection algorithm is very effective and has significant advantages over classification alone. In addition, we have run a series of experiments with varying ratios, e.g., 95%/5%, 90%/10%, 85%/15%, 80%/20%, 75%/25%, and 70%/30% of normal and attack traffic flows. All the experiments show similar performances of our proposed two-stage ML-based method in detecting simulated attacks towards smart home IoT devices. Given the overall performance and robustness of our proposed method, it is safe to expect significant improvement of the two-stage algorithm when new/better classification or anomaly detection algorithms

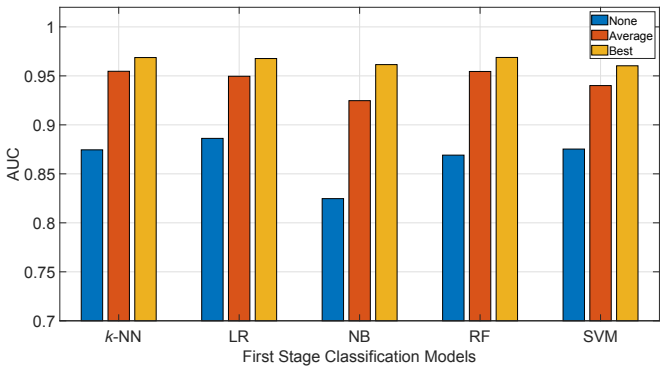

Fig. 4. The average and best AUC improvement by combining two-stages over the first classification stage alone.

VII. Related Work

Smart home IoT security has been of recent interest to many researchers due to the wide deployment of IoT devices in home networks. Research effort has been devoted to empirically study the current status of IoT deployment, to discover and raise awareness of the potential vulnerabilities and their growing security implications, to design more secure IoT application frameworks, and to understand the behavior of IoT components and further detect anomaly.

State of Home IoT Deployment and Security: A recent SoK paper [3] categorizes the home IoT security research into device, mobile application, cloud endpoint, and communication, describes the attacks and mitigation, and proposes recommendations to stakeholders for each category. In addition, the paper evaluates the security properties of 45 home IoT devices and their applications. Most recently, [24] carries out a large-scale empirical analysis of IoT devices in real-world homes, covering 83 million devices in 16 million homes, and presents methodologies to identify the types of IoT devices in home networks and provides their regional distributions. The paper also analyzes the status of services and weakness in the IoT devices and the scanning behavior of smart homes. As the first paper investigating the security problems of Smart Home IoT applications, [15] identifies over 50% of the applications on Samsung's SmartThings platform with serious over-privilege problems. [17] discusses the similarities and differences between IoT security research and classic IT security research from hardware, system software, network, and application layers. [28] provides a comprehensive survey on security issues and defense mechanisms for various IoT applications in smart homes, vehicles, cities and buildings.

Attack Techniques: The popularity of home IoT devices has introduced a new spectrum of attacks towards all the components in smart homes. [3] compiles a comprehensive list of attacks towards different types of IoT devices, popular services supported by the devices, weak trust management and weak credentials, mobile application development, and communication channels. With an example attack where worms automatically spread over a large area among physically adjacent lamps in a chain reaction using only the standard Zigbee wireless interface, [32] investigates a new attack paradigm where IoT devices with ad hoc networking capabilities can spread malware to their physically adjacent neighbors bypassing the Internet. In addition, vulnerable IoT devices have been recently exploited to form high profile botnets. [5][23] provide a detail insight of Mirai botnet, which is a high profile DDoS threat sourced from hundreds of thousands of IoT devices, and how the insecurity of IoT devices contributes to the growth of this largest ever botnet. [20] discusses Hajime, which is the latest botnet consisting of IoT devices managed in a peer-to-peer fashion.

Application Security: [9] gives a thorough analysis of smart home IoT applications' security and privacy issues. SmartAuth [39] is a framework that identifies required permissions for IoT applications running on platforms like SmartThings and Apple Home. SAINT [8] is a static information flow tracking and analysis tool for evaluating privacy risks in IoT implementation. Their results show that 60% SmartThings market apps include sensitive data flow. Soteria [10] presents a static analysis system for validating whether an IoT app or environment is secure and operates correctly by automat-

TABLE VII
THE AUC IMPROVEMENT BY COMBINING TWO-STAGES OVER THE FIRST CLASSIFICATION STAGE ALONE.

Model	None	CBLOF	FastABOD	FB	IForest	LOF	PCA
k-NN	0.8745	0.9669(+10.57%)	0.9171(+4.87%)	0.9558(+9.30%)	0.9684 (+10.73%)	0.9517(+8.83%)	0.9687(+10.77%)
LR	0.8862	0.9648(+8.87%)	0.9123(+2.94%)	0.9389(+5.95%)	0.9677(+9.20%)	0.9591(+8.23%)	0.9548(+7.74%)
NB	0.8247	0.9160(+11.07%)	0.9049(+9.72%)	0.9183(+11.35%)	0.9615(+16.59%)	0.9238(+12.02%)	0.9239(+12.03%)
RF	0.8691	0.9688(+11.47%)	0.9212(+5.99%)	0.9457(+8.81%)	0.9676(+11.33%)	0.9558(+9.98%)	0.9678(+11.36%)
SVM	0.8753	0.9395(+7.33%)	0.9086(+3.80%)	0.9370(+7.05%)	0.9603(+9.71%)	0.9412(+7.53%)	0.9540(+8.99%)

ically extracting a state model from a SmartThings IoT app and applying model checking to identify property violations. IoTSan [29] uses model checking to reveal interaction level flaws by identifying events that can lead the system to unsafe states. FlowFence [16] proposed a framework that splits application codes into sensitive and non-sensitive modules and orchestrates the execution through opaque handlers. SIFT [25] is a safety-centric programming platform which leads to more robust and reliable IoT apps.

Behavior Modeling and Intrusion Detection: [47] proposes a third-party defender which monitors the smart home side-channel traffic and detects misbehavior in smart apps such as privileged accesses and event spoofing. Their approach leverages wireless fingerprints to detect mis-behaviors in a resource-constraint IoT environment. [1] demonstrates a multi-stage privacy attack to identify device types, states and activities by passively observing the encrypted wireless traffic. [43] investigates the current encryption status of four popular medical devices by capturing and analyzing network traffic and retrieving clear-text information, which reveals sensitive medical conditions and behaviors. [46] discusses the active learning approach for detecting intrusion targeting IoT devices. [6, 26, 44, 45] model the device behavior at network, transport and application layers, while [19, 21, 37] model device behavior with linker layer and physical layer traffic.

A recent position paper [40] summarizes the attacks and security functions in device, network, and service layers, and introduces a cross layer framework to connect and bridge the gap between different layers. Another relevant paper [12] demonstrates smart home devices are vulnerable to attacks from malicious mobile apps running on authorized phones and implements and evaluates a HanGuard system where the home router enforces role based access control between mobile apps and IoT devices with the help of a user-space monitoring app running on the mobile phone. Their adoption of router for enforcing security policy is similar to our work. [11] proposes to leverage smart home applications on activity recognition, health monitoring, and automation for detecting abnormal home and user behavior in the homes. In contrast to these research effort, our proposed IoTArgos system focuses on characterizing IoT data communications with multi-layer behavioral features, and detecting intrusion and anomalous activities towards IoT with a two-staged ML-based method.

VIII. CONCLUSIONS AND FUTURE WORK

The recent high-profile cyber attacks towards vulnerable and insecure IoT devices have highlighted the prevalent security threats towards millions of smart homes and the great risks of data and user privacy. These attacks call for systematic approaches for protecting IoT devices from the broad attack vector which spans the entire IoT protocol stacks due to design flaws of rapidly developed and deployed protocols, weak credential management, and lack of cryptographic functions on resource-constrained IoT devices. As a first step of securing IoT devices in smart homes, this paper designs, develops, and evaluates IoTArgos, a multi-layer security monitoring system on programmable home routers. Based on data captured from hundreds of IoT devices in real-world smart homes, IoTArgos characterizes and models data communication behaviors of heterogeneous IoT devices with a broad range of communication and traffic features. To detect intrusions towards IoT devices, IoTArgos develops a two-stage method to first explore supervised classification algorithms for identifying known attacks based on trained labelled data-sets and then rely on unsupervised anomaly detection algorithms for capturing emerging attacks without prior attack labels or signatures. Our extensive experiments based on synthetic IoT data traffic with normal communications collected from a smart home sandbox and simulated attacks have shown the two-stage method is very effective in detecting a wide range of IoT attacks.

Our future work will focus on implementing and deploying IoTArgos across a large number of smart homes and small businesses to monitor the security of IoT systems in these edge networks and correlating the security monitoring of distributed homes for discovering coordinated and large scale cyber attacks towards IoT devices with similar vulnerabilities. We plan to further investigate the mitigation techniques for efficiently identifying and filtering attacks. Our efforts will be also centered on balancing the trade-off between accuracy and false possible rate and looking deep into the data to differentiate between the legitimate and illegitimate accesses of the IoT devices.

ACKNOWLEDGMENT

This research was supported in part by NSF grants 1816995, 1717197, and 1704092. The information reported here does not reflect the position or the policy of the funding agency.

REFERENCES

[1] A. Acar, H. Fereidooni, T. Abera, A.-K. Sikder, M. Miettinen, H. Aksu, M. Conti, A.-Z. Sadeghi, and A.-S. Uluagac, "Peek-a-Boo: I See Your Smart Home Activities, Even Encrypted!" *arXiv preprint arXiv:1808.02741*, 2018.

[2] B. Aggarwal, R. Bhagwan, T. Das, S. Eswaran, T. Padmanabhan, and G. Voelker, "NetPrints: Diagnosing Home Network Misconfigurations Using Shared Knowledge," in *Proc. of USENIX NSDI*, 2009.

[3] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose, "SoK: Security Evaluation of Home-Based IoT Deployments," in *Proc. of IEEE S&P*, 2019.

[4] Amazon, "Alexa Skills Kit," https://developer.amazon.com.

[5] M. Antonakakis, et al., "Understanding the Mirai Botnet," in *Proc. of USENIX Security*, 2017.

[6] B. Bezawada, M. Bachani, J. Peterson, H. Shirazi, I. Ray, and I. Ray, "Behavioral Fingerprinting of IoT Devices," in *Proc. of ACM ASHES*, 2018.

[7] C. Cao, W. Gong, W. Dong, J. Yu, C. Chen, and J. Liu, "Network Measurement in Multihop Wireless Networks with Lossy and Correlated Links," in *Proc. of IEEE INFOCOM*, 2018.

[8] Z.-B. Celik, L. Babun, A.-K. Sikder, H. Aksu, G. Tan, P. McDaniel, and A.-S. Uluagac, "Sensitive Information Tracking in Commodity IoT," in *Proc. of USENIX Security*, 2018.

[9] Z.-B. Celik, E. Fernandes, E. Pauley, G. Tan, and P. McDaniel, "Program Analysis of Commodity IoT Applications for Security and Privacy: Challenges and Opportunities," *arXiv preprint arXiv:1809.06962*, 2018.

[10] Z.-B. Celik, P. McDaniel, and G. Tan, "Soteria: Automated IoT Safety and Security Analysis," in *Proc. of USENIX ATC*, 2018.

[11] J. Dahmen, D. Cook, X. Wang, and H. Wang, "Smart Secure Homes: A Survey of Smart Home Technologies that Sense, Assess, and Respond to Security Threats," *Journal of Reliable Intelligent Environments*, vol. 3, no. 2, pp. 83–98, 2017.

[12] S. Demetriou, N. Zhang, Y. Lee, X. Wang, C. Gunter, X. Zhou, and M. Grace, "HanGuard: SDN-driven Protection of Smart Home WiFi Devices from Malicious Mobile Apps," in *Proc. of ACM WiSec*, 2017.

[13] L. DiCioccio, R. Teixeira, and C. Rosenberg, "Impact of Home Networks on End-to-End Performance: Controlled Experiments," in *Proc. of ACM SIGCOMM Workshop on Home Networks*, 2010.

[14] J. Dunning, "Taming the Blue Beast: A Survey of Bluetooth Based Threats," *IEEE Security & Privacy*, vol. 8, no. 2, 2010.

[15] E. Fernandes, J. Jung, and A. Prakash, "Security Analysis of Emerging Smart Home Applications," in *Proc. of IEEE S&P*, 2016.

[16] E. Fernandes, J. Paupore, A. Rahmati, D. Simionato, M. Conti, and A. Prakash, "FlowFence: Practical Data Protection for Emerging IoT Application Frameworks," in *Proc. of USENIX Security*, 2016.

[17] E. Fernandes, A. Rahmati, K. Eykholt, and A. Prakash, "Internet of Things Security Research: A Rehash of Old Ideas or New Intellectual Challenges?" *IEEE Security & Privacy*, vol. 15, no. 4, pp. 79–84, 2017.

[18] Google, "Nest, Create a Connected Home," https://nest.com/.

[19] T. Gu and P. Mohapatra, "BF-IoT: Securing the IoT Networks via Fingerprinting-Based Device Authentication," in *Proc. of IEEE MASS*, 2018.

[20] S. Herwig, K. Harvey, G. Hughey, R. Roberts, and D. Levin, "Measurement and Analysis of Hajime, a Peer-to-peer IoT Botnet," in *Proc. of NDSS*, 2019.

[21] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "IoT Devices Fingerprinting Using Deep Learning," in *Proc. of IEEE MILCOM*, 2018.

[22] Y. Jia, Y. Xiao, J. Yu, X. Cheng, Z. Liang, and W. Z., "A Novel Graph-based Mechanism for Identifying Traffic Vulnerabilities in Smart Home IoT," in *Proc. of IEEE INFOCOM*, 2018.

[23] G. Kambourakis, C. Kolias, and A. Stavrou, "The Mirai botnet and the IoT Zombie Armies," in *Proc. of IEEE MILCOM*, 2017.

[24] D. Kumar, K. Shen, B. Case, D. Garg, G. Alperovich, D. Kuznetsov, R. Gupta, and Z. Durumeric, "All Things Considered: An Analysis of IoT Devices on Home Networks," in *Proc. of the USENIX Security*, 2019.

[25] C. Liang, B. Karlsson, N. Lanet, F. Zhao, J. Zhang, Z. Pan, Z. Li, and Y. Yu, "SIFT: Building an Internet of Safe Things," in *Proc. of IPSN*, 2015.

[26] M. Miettinen, S. Marchal, I. Hafeez, N. Asokan, A.-R. Sadeghi, and S. Tarkoma, "IoT SENTINEL: Automated Device-Type Identification for Security Enforcement in IoT," in *Proc. of IEEE ICDCS*, 2017.

[27] P. Morgner, S. Mattejat, Z. Benenson, C. Muller, and F. Armknecht, "Insecure to the Touch: Attacking ZigBee 3.0 via Touchlink Commissioning," in *Proc. of ACM WiSec*, 2017.

[28] A. Mosenia and M. Jha, "A Comprehensive Study of Security of Internet-of-Things," *IEEE Trans. on Emerging Topics in Computing*, vol. 5, no. 4, pp. 586 – 602, 2016.

[29] D.-T. Nguyen, C. Song, Z. Qian, S.-V. Krishnamurthy, E.-J.-M. Colbert, and P. McDaniel, "IotSan: Fortifying the Safety of IoT Systems," in *Proc. of CoNEXT*, 2018.

[30] J. Obermaier and M. Hutle, "Analyzing the Security and Privacy of Cloud-based Video Surveillance Systems," in *Proc. of ACM IoTPTS*, 2016.

[31] Y.-M.-P. Pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, and C. Rossow "IoTPOT: Analysing the Rise of IoT Compromises," in *Proc. of USENIX WOOT*, 2015.

[32] E. Ronen, C. O'Flynn, A. Shamir, and A.-O. Weingarten, "IoT Goes Nuclear: Creating a ZigBee Chain Reaction," in *Proc. of IEEE S&P*, 2017.

[33] E. Ronen and A. Shamir, "Extended Functionality Attacks on IoT Devices: The Case of Smart Lights ," in *Proc.of IEEE EuroS&P*, 2016.

[34] M. Ryan, "Bluetooth Smart: The Good, the Bad, the Ugly, and the Fix," 2013, https://lacklustre.net/bluetooth/bluetooth_smart_good_bad_ugly_fix-mikeryan-blackhat_2013.pdf.

[35] Samsung, "SmartThings," https://www.smartthings.com/.

[36] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[37] S. Siby, R. Maiti, and N. Tippenhauer, "IoTScanner: Detecting Privacy Threats in IoT Neighborhoods," in *Proc. of ACM IoTPTS*, 2017.

[38] V. Sivaraman, D. Chan, D. Earl, and R. Boreli, "Smart-Phones Attacking Smart-Homes," in *Proc. of ACM WiSec*, 2016.

[39] Y. Tian, N. Zhang, Y.-H. Lin, X.-F. Wang, B. Ur, X. Guo, and P. Tague, "SmartAuth: User-Centered Authorization for the Internet of Things," in *Proc. of the USENIX Security*, 2017.

[40] A. Wang, A. Mohaisen, and S. Chen, "XLF: A Cross-layer Framework to Secure the Internet of Things (IoT)," in *Proc. IEEE ICDCS*, 2019.

[41] Q. Wang, W. Hassan, A. Bates, and C. Gunter, "Fear and Logging in the Internet of Things," in *Proc. of NDSS*, 2018.

[42] R. Want, B. Schilit, and S. Jenson, "Enabling the Internet of Things," *Computer*, vol. 48, no. 1, pp. 28 – 35, 2015.

[43] D. Wood, N. Apthorpe, and N. Feamster, "Cleartext Data Transmissions in Consumer IoT Medical Devices," in *ACM Workshop on IoT S&P*, 2017.

[44] L. Xiao, X. Wan, X. Lu, Y. Zhang, and D. Wu, "IoT Security Techniques Based on Machine Learning: How Do IoT Devices Use AI to Enhance Security?," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 41 – 49, 2018.

[45] K. Xu, Y. Wan, G. Xue, and F. Wang, "Multidimensional Behavioral Profiling of Internet-of-Things in Edge Networks," in *Proc. of IEEE/ACM IWQoS*, 2019.

[46] K. Yang, J. Ren, Y. Zhu, and W. Zhang, "Active Learning for Wireless IoT Intrusion Detection," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 19 – 25, 2018.

[47] W. Zhang, Y. Meng, Y. Liu, X. Zhang, Y. Zhang, and H. Zhu, "HoMonit: Monitoring Smart Home Apps from Encrypted Traffic," in *Proc. of ACM CCS*, 2018.

[48] T. Zillner and S. Strobl, " ZigBee Exploited: The Good, the Bad and the Ugly," in *Proc. of the DeepSec Conferences In Depth Security*, 2015.