# Machine Learning Based Intrusion Detection for IoT Botnet

Sikha Bagui, Xiaojian Wang, and Subhash Bagui

***Abstract*—In this article, we analyzed botnet traffic in an IoT environment using three machine learning classifiers: Logistic Regression, Support-Vector Machine and Random Forest. We classified each attack in each botnet for nine devices. We calculated the Accuracy, True Positive, False Positive, False Negative, True Negative, Precision, Recall, F1 score for each algorithm. We obtained impressive results (above 99%) using these three classifiers. We have a high attack detection rate. A brief analysis of the results is presented.**

***Index Terms*—Intrusion detection, machine learning, internet of things (IoT), botnet, logistic regression (LR), support vector machines (SVM), random forest (RF).**

## I. Introduction

A low estimate is that by 2025, the global worth of Internet of Things (IoT) devices will be $4 trillion dollars, and a high estimate is that by 2025, the global worth of IoT devices will be $11 trillion dollars [1]. With the development of IoT technologies, more and more devices have joined our lives, making security of systems an utmost concern. Many of the devices used in our everyday lives today, for example, smart phones, wearable devices, health monitoring devices, etc., generate vast amounts of private information, but have very little security, if any, built in. The internet is complex enough to secure, and these additional insecure IoT devices make the task of security even more challenging [1]. Botnets are able to infiltrate any internet connected device from smart watches and home smart kitchen appliances to corporate mainframes. Free availability of source code of IoT botnets like BASHLITE and Mirai have led to cyber attackers trying their hands at IoT malwares [1]. The IoT malware, Mirai, has actually inspired a renaissance of IoT malware and has been responsible for large scale DDos attacks [1]. The Mirai botnet and it's variants and imitators were basically a wake-up call to the industry to better secure IoT devices [2].

Botnets are typically created to infect as many devices as possible and complex botnets even self-propagate and update their behavior, finding and infecting devices automatically. Hence botnets are very difficult to detect [3]. Another reason why botnets are difficult to detect and contain is that they lurk on devices that do not significantly affect the performance of the device [3]. For example, a security camera may be part of an active botnet, but neither an average user nor a small business may be aware of this. Therefore, it is extremely important to identify botnets from the traffic of IoT devices.

In this paper, we use the dataset available in [4] to classify botnet traffic in the IoT environment. This dataset is real network traffic data, gathered from nine commercial IoT devices infected by two botnets, Mirai and BASHLITE. The data is analyzed using three classifiers, Logistic Regression (LR), Support Vector Machines (SVM) and Random Forest (RF), and classified by botnet, by attack, by device.

The rest of the paper is organized as follows: Section II presents the related works; Section III describes the dataset – the devices used, the attack categories and features; Section IV briefly presents the three classification algorithms used; Section V presents the results; Section VI presents the discussion; and Section VII presents the conclusions and future works.

## II. Related Works

In this section we grouped the work based on works done on intrusion detection systems and works done directly on IoT Botnet.

### A. Works on Intrusion Detection Systems

Several works have been done on intrusion detection systems. [5] designed fuzzy membership functions to solve dimensionality and anomaly mining, thereby reducing computational complexity and improving the computational accuracy of the classifier. [6] presented a dynamic coding mechanism, implementing a distributed signature based IDS in IP-USN (IP based ubiquitous sensor networks) and used Bloom filtering for signature matching. [7] designed and developed a virtual test platform to simulate a real network environment, deploying a signature-based Snort IDS for traffic monitoring and attack detection by mirroring the traffic to the server, and developing a stream-based IDS model using machine learning. They also implemented a flow-based anomaly detection model to overcome the limitations of the signature-based IDS. [8] designed a specification-based IDS for detecting a new type of threat - the topology attack. They proposed an IDS architecture using a network monitor backbone, and described its monitoring mechanisms through a RPL finite state machine. [9] developed a deep packet anomaly detection method that can be run on resource-constrained IoT devices, but can distinguish between normal and abnormal payloads.

Ref. [10] presented a DoS detection architecture for 6LoWPAN. This architecture integrated an IDS into the framework developed within the EU FP7 project ebbits. [11]

Manuscript received May 15, 2020; revised January 2, 2021. This work has been partially supported by the Askew Institute of the University of West Florida.

Sikha Bagui and Xiaojian Wang are with the Department of Computer Science, University of West Florida, Pensacola, FL 32514 USA (e-mail: bagui@uwf.edu; xw5@students.uwf.edu).

Subhash Bagui is with the Department of Mathematics and Statistics, University of West Florida, Pensacola, FL 32514 USA (e-mail: sbagui@uwf.edu).

doi: 10.18178/ijmlc.2021.11.6.1068

proposed an IDS framework for IoT based on 6LoWPAN, which included a monitoring system and a detection engine. SVELTE [12], primarily targeting routing attacks, used a host based IDS under 6LoWPAN environment. The goal of [13] was to detect DoS attacks and attack protocols for 6LoWPAN and CoAP communications and propose an IDS framework for detecting and preventing attacks in the internet integrated environment. An intrusion detection model based on node consumption analysis in 6LowPAN was proposed in [14]. Irregular energy consumption of the routing scheme in the 6LoWPAN grid and the sensor nodes were used to identify malicious attacks. A malicious pattern matching engine for lightweight security systems was proposed in [15]. Two novel techniques, assisted transfer and early decision making, were proposed to reduce performance degradation due to computational power and memory limitations.

Ref. [16] proposed an event-processing IDS architecture using Complex Event Processing (CEP) technology. [17] proposed an architecture that employs a Bayesian event prediction model that uses historical event data generated by the IoT cloud to calculate the probability of future events. Based on the characteristics of the secure cloud service system, [18] proposed a secure high-order clustering algorithm that quickly searches and finds a mixed cloud density peak. The client first uses homomorphic encryption to construct the encrypted object tensor with user data, uploads it to the cloud to fully implement the proposed protocol, returning the clustering results of a random number of perturbations to the client, to eliminate the perturbations.

Kalis [19], an adaptive knowledge-driven expert intrusion detection system, which can monitor various protocols without changing existing IoT software, is a comprehensive method for IoT intrusion detection.

A real-time hybrid intrusion detection framework, including an anomaly-based and specification-based intrusion detection module, is proposed in [20]. The anomaly-based intrusion detection agent, located in the root node, uses the unsupervised optimal path forest algorithm to predict the clustering model by using incoming packets. The specification-based intrusion detection agent in the router node analyzes the behavior of its host node and sends its local result through ordinary data packets to the root node. [21] proposed a new network intrusion detection method for IoT networks based on a conditional variational autoencoder with a specific architecture, which integrates intrusion tags.

### *B. Works on IoT Botnets Specifically*

Few works have also been done on detecting botnets on IoT devices. The authors of [22] proposed a host-based detection system based on one-class classifiers. Host based detection techniques can be considered less realistic for attacks on IoT botnets for various reasons including the fact that we would have to rely on the IoT manufacturers to install host-based anomaly detectors on the products. Also, given that IoT botnet attacks mutate at a very fast rate [2] and are becoming increasingly more and more complex by the day, some of these mutations will succeed in bypassing existing methods of early detection [23].

Ref. [24] used a one-class Support Vector Machine built with features such as CPU and memory usage to detect malicious activities. [25] proposed a deep learning-based botnet traffic analyzer called Botnet Traffic Shark (BoTShark) that uses only network transactions and is independent of deep packet inspection techniques to identify correlations between original features and new features in each layer of the autoencoder or CNN extracted in a cascaded manner. [26] proposed a state-of-the-art T-IDS, built on a novel randomized data partitioned learning model (RDPLM) relying on a compact network feature set and feature selection techniques, simplifying sub-spacing and multiple randomized meta-learning techniques. [27] analyzed the effectiveness of some community detection algorithms in detecting P2P botnets, especially with partial information. They showed that the approach can work with only about half of the nodes, reporting their communication graphs with only a small increase in detection errors. A method to detect compromised IoT devices included in a botnet is proposed in [28]. This method is based on logistic regression, which allows the estimation of the probability that a device initiating a connection is running a bot.

Ref. [29] empirically evaluates a network-based anomaly detection method which extracts behavior snapshots of the network and uses deep autoencoders to detect anomaly in network traffic from compromised IoT devices. [29] also presents a very good summary of IoT-related anomalies, botnets and malware attacks done by others.

While many of the previous works were on simulated data, in this paper we used real network traffic data, presented in [4], [29], to classify each attack in each botnet on each device using three classifiers, LR, RF and SVM.

## III. Dataset Description

The dataset used by this paper is from UCI's machine learning repository [4]. The data is divided into 10 attacks carried by 2 botnets, gafgyt and mirai. The 9 IoT devices are: Danmini Doorbell, Ecobee Thermostat, Ennio Doorbell, Philips B120N10 Baby Monitor, Provision PT 737E Security Camera, Provision PT 838 Security Camera, Samsung SNH 1011 N Webcam, SimpleHome XCS7 1002 WHT Security Camera, and SimpleHome XCS7 1003 WHT Security Camera.

Most of these devices were infected by both gafgyt and mirai, as can be seen in Tables I through VII; but Ennio Doorbell and Samsung SNH 1011 N Webcam was infected only by gafgyt and the Philips B120N10 Baby Monitor was infected only by Mirai.

Mirai is a kind of malware that can make a computing system running Linux a remotely controlled "zombie." This can lead to large-scale network attacks though Mirai's mainly infected IoT devices such as web cameras, routers, etc. Devices infected by Mirai continuously scan the IP address of the IoT device on the Internet. The default username and password are used to log in to the vulnerable devices, and then the Mirai software is injected. The Mirai botnet has five types of attacks: scan, ack, syn, udp, and udpplain. Scan does automatic scanning for vulnerable devices. Ack causes Ack flooding. Syn causes Syn flooding. UDP causes UDP flooding. UDPplain causes UDP flooding with fewer options,

optimized for higher PPS. [29]

Gafgyt (also known as BASHLITE) is a malware that infects Linux systems to initiate Distributed Denial of Service (DDoS) attacks. It mainly uses the Metasploit module to exploit known vulnerabilities in the WeMo UPnP protocol. The Gafgyt botnet also has five types of attacks: combo, junk, scan, udp, and tcp. Combo sends spam data and opens a connection to a specified IP address and port. Junk sends spam data. Scan scans the network for vulnerable devices. UDP causes UDP flooding. TCP causes TCP flooding. [29]

This dataset has 23 basic features [30] which can be categorized into the following attribute types: stream aggregation, time-frame and statistics extracted from packet streams.

Stream aggregation is composed of: (i) H stats, which summarizes the recent traffic from this packet's host (IP); (ii) MI stats, which summarizes the recent traffic from this packet's host (IP + MAC); (iii) HH stats, which summarizes the recent traffic going from this packet's host (IP) to the packet's destination host; (iv) HH_jit stats, which summarizes the jitter of the traffic going from this packet's host (IP) to the packet's destination host; (v) HpHp stats, which summarizes the recent traffic going from this packet's host+port (IP) to the packet's destination host+port.

Time-frame or the decay factor Lambda used in the damped window is: L5, L3, L1, L0.1 and L0.01. These statistics capture the recent history of the streams.

The statistics extracted from the packet streams are: (i) weight, which includes the weight of the stream (number of items observed in recent history); (ii) mean; (iii) standard deviation; (iv) radius, which is the root squared sum of the two streams' variances; (v) magnitude, which is the root squared sum of the two streams' means; (vi) cov, which is an approximated covariance between two streams; (vii) pcc, which is an approximated correlation coefficient between two streams. These features are extracted from a total of five time windows: 100ms, 500ms, 1.5sec, 10sec, and 1min, thus totaling 115 features. More details of each feature can be seen from [30].

The statistics are summarized from all of the traffic as follows [30]:

1) Originating from this packet's source MAC and IP address (denoted SrcMAC-IP).
2) Originating from this packet's source IP (denoted SrcIP).
3) Sent between this packet's source and destination IPs (denoted Channel).
4) Sent between this packet's source and destination TCP/UDP Socket (denoted Socket).

## IV. Classifiers

Three classifiers were used in this study: Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF).

LR is a machine learning classifier used to model the probability of a certain class. Though LR can also be extended to classifying several classes, in it's basic form, LR uses a logistic function to model a binary dependent variable.

SVM, relatively computationally inexpensive, is a supervised learning classifier mainly used for binary classification. In SVMs, we find the best hyperplane that divides the data into two categories and we generally have a low generalization error. The farther the data point from a decision boundary, the more confident we are about the prediction. The points separating the hyperplane are known as support vectors.

RF refers to a classifier that uses multiple trees to train and predict samples. Random forests establish a forest in a random way. After getting the forest, when a new sample is entered, each decision tree in the forest makes a separate judgment to see which class the sample should belong to (for the classification algorithm). The sample is predicted to be of the class to which it was classified the most times.

## V. Experimental Setup

Since we are classifying each attack in each botnet for each device, the data was grouped by device, by botnet and then by attack. Our initial results using the three classifiers, LR, SVM, and RF did not give us good performance, which was mainly due to the highly imbalanced nature of the data. To address this issue, we used an almost equal number of benign (normal) data as well as malicious data. The almost 50% of the benign data was randomly selected from the set of benign data and added to the malicious dataset before running the algorithms.

The data was then pre-processed using z-score normalization. Each of the classifiers (LR, SVM, and RF) were then used as binary classifiers on the normalized data and training and prediction was performed. 80% of the data was used for training and 20% for testing. Scikit Learn was used to run the classifiers.

## VI. Results

Eight metrics were used to evaluate and analyze the results: True Positive (TP) is actually positive, and the prediction is positive; False Positive (FP) is actually negative, and the prediction is positive; True Negative (TN) is actually negative, and the prediction is negative; False Negative (FN) is actually positive, and the prediction is negative; Accuracy, Precision, Recall and F1-score.

Accuracy is the ratio of the model's correct data (TP+TN) to the total data, given by:

$$\text{Accuracy} = (TP+TN)/(TP+FP+TN+F \quad (1)$$

Recall, also referred to as sensitivity, or Attack Detection Rate (ADR): This is the effectiveness of the model in identifying an attack, that is, for all positive cases (TP+FN) in the dataset, the positive cases (TP) correctly judged by the model, given by:

$$\text{Recall=sensitivity=ADR=TPR=TP/(TP+FN)} \quad (2)$$

Precision: This is the percentage of classified attack instances that are truly classified as attacks, that is, for all positive cases (TP+FP) judged by the model, the proportion of the real cases (TP).

$$\text{Precision}=TP/(TP+FP) \quad (3)$$

$$2/\text{F1-score}=1/\text{Precision}+1/\text{Recall} \quad (4)$$

F1-score: This is the relationship between precision and recall, given by:

The higher the F1-score, the more robust the classification model [24].

TABLE I: DANMINI DOORBELL RESULTS

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Danmini Doorbell | gafgyt | combo | LR | 0.99991 | 12035 | 2 | 0 | 9817 | 0.99983 | 0.99990 | 0.99992 |
| | | | SVM | 0.99945 | 12023 | 0 | 12 | 9819 | 1.00000 | 0.99950 | 0.99950 |
| | | | RF | 1.00000 | 12035 | 0 | 0 | 9819 | 1.00000 | 1.00000 | 1.00000 |
| | | junk | LR | 0.99968 | 5827 | 1 | 4 | 9892 | 0.99983 | 0.99961 | 0.99957 |
| | | | SVM | 0.99949 | 5823 | 0 | 8 | 9893 | 1.00000 | 0.99931 | 0.99931 |
| | | | RF | 0.99994 | 5830 | 0 | 1 | 9893 | 1.00000 | 0.99991 | 0.99991 |
| | | scan | LR | 0.99994 | 5878 | 0 | 1 | 10001 | 1.00000 | 0.99991 | 0.99991 |
| | | | SVM | 0.99937 | 5877 | 8 | 2 | 9993 | 0.99864 | 0.99943 | 0.99915 |
| | | | RF | 1.00000 | 5879 | 0 | 0 | 10001 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99984 | 21164 | 1 | 4 | 9916 | 0.99995 | 0.99986 | 0.99988 |
| | | | SVM | 0.99942 | 21150 | 0 | 18 | 9917 | 1.00000 | 0.99957 | 0.99957 |
| | | | RF | 0.99984 | 21164 | 1 | 4 | 9916 | 0.99995 | 0.99986 | 0.99988 |
| | | tcp | LR | 0.99989 | 18431 | 1 | 2 | 9904 | 0.99995 | 0.99990 | 0.99992 |
| | | | SVM | 0.99961 | 18422 | 0 | 11 | 9905 | 1.00000 | 0.99970 | 0.99970 |
| | | | RF | 0.99996 | 18432 | 0 | 1 | 9905 | 1.00000 | 0.99997 | 0.99997 |
| | mirai | ack | LR | 1.00000 | 20342 | 0 | 0 | 10007 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99993 | 20340 | 0 | 2 | 10007 | 1.00000 | 0.99995 | 0.99995 |
| | | | RF | 1.00000 | 20342 | 0 | 0 | 10007 | 1.00000 | 1.00000 | 1.00000 |
| | | scan | LR | 1.00000 | 21559 | 0 | 0 | 9888 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 1.00000 | 21559 | 0 | 0 | 9888 | 1.00000 | 1.00000 | 1.00000 |
| | | | RF | 1.00000 | 21559 | 0 | 0 | 9888 | 1.00000 | 1.00000 | 1.00000 |
| | | syn | LR | 1.00000 | 24459 | 0 | 0 | 9966 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99997 | 24458 | 0 | 1 | 9966 | 1.00000 | 0.99998 | 0.99998 |
| | | | RF | 1.00000 | 24459 | 0 | 0 | 9966 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99993 | 47606 | 0 | 4 | 9833 | 1.00000 | 0.99996 | 0.99996 |
| | | | SVM | 0.99990 | 47604 | 0 | 6 | 9833 | 1.00000 | 0.99994 | 0.99994 |
| | | | RF | 1.00000 | 47610 | 0 | 0 | 9833 | 1.00000 | 1.00000 | 1.00000 |
| | | udpplain | LR | 1.00000 | 16517 | 0 | 0 | 9789 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99996 | 16516 | 0 | 1 | 9789 | 1.00000 | 0.99997 | 0.99997 |
| | | | RF | 1.00000 | 16517 | 0 | 0 | 9789 | 1.00000 | 1.00000 | 1.00000 |

TABLE II: ECOBEE THERMOSTAT RESULTS

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ecobee Thermostat | Gafgyt | combo | LR | 0.99992 | 10681 | 0 | 1 | 2543 | 1.00000 | 0.99995 | 0.99995 |
| | | | SVM | 0.99924 | 10672 | 0 | 10 | 2543 | 1.00000 | 0.99953 | 0.99953 |
| | | | RF | 1.00000 | 10682 | 0 | 0 | 2543 | 1.00000 | 1.00000 | 1.00000 |
| | | junk | LR | 0.99977 | 6049 | 0 | 2 | 2634 | 1.00000 | 0.99983 | 0.99983 |
| | | | SVM | 0.99931 | 6045 | 0 | 6 | 2634 | 1.00000 | 0.99950 | 0.99950 |
| | | | RF | 1.00000 | 6051 | 0 | 0 | 2634 | 1.00000 | 1.00000 | 1.00000 |
| | | scan | LR | 0.99975 | 5426 | 1 | 1 | 2694 | 0.99982 | 0.99972 | 0.99982 |
| | | | SVM | 0.99926 | 5421 | 0 | 6 | 2695 | 1.00000 | 0.99945 | 0.99945 |
| | | | RF | 1.00000 | 5427 | 0 | 0 | 2695 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99983 | 20949 | 2 | 2 | 2628 | 0.99990 | 0.99957 | 0.99990 |
| | | | SVM | 0.99915 | 20933 | 2 | 18 | 2628 | 0.99990 | 0.99919 | 0.99952 |
| | | | RF | 0.99996 | 20950 | 0 | 1 | 2630 | 1.00000 | 0.99998 | 0.99998 |
| | | tcp | LR | 0.99977 | 18975 | 3 | 2 | 2647 | 0.99984 | 0.99938 | 0.99987 |
| | | | SVM | 0.99945 | 18967 | 2 | 10 | 2648 | 0.99989 | 0.99936 | 0.99968 |
| | | | RF | 0.99995 | 18976 | 0 | 1 | 2650 | 1.00000 | 0.99997 | 0.99997 |
| | Mirai | ack | LR | 1.00000 | 22714 | 0 | 0 | 2566 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99992 | 22712 | 0 | 2 | 2566 | 1.00000 | 0.99996 | 0.99996 |
| | | | RF | 0.99996 | 22713 | 0 | 1 | 2566 | 1.00000 | 0.99998 | 0.99998 |
| | | scan | LR | 0.99973 | 8693 | 2 | 1 | 2565 | 0.99977 | 0.99955 | 0.99983 |
| | | | SVM | 0.99956 | 8691 | 2 | 3 | 2565 | 0.99977 | 0.99944 | 0.99971 |
| | | | RF | 0.99991 | 8693 | 0 | 1 | 2567 | 1.00000 | 0.99994 | 0.99994 |
| | | syn | LR | 1.00000 | 23345 | 0 | 0 | 2639 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99992 | 23343 | 0 | 2 | 2639 | 1.00000 | 0.99996 | 0.99996 |
| | | | RF | 0.99996 | 23344 | 0 | 1 | 2639 | 1.00000 | 0.99998 | 0.99998 |
| | | udp | LR | 0.99982 | 30242 | 2 | 4 | 2671 | 0.99993 | 0.99956 | 0.99990 |
| | | | SVM | 0.99982 | 30240 | 0 | 6 | 2673 | 1.00000 | 0.99990 | 0.99990 |
| | | | RF | 1.00000 | 30246 | 0 | 0 | 2673 | 1.00000 | 1.00000 | 1.00000 |
| | | udpplain | LR | 0.99990 | 17464 | 2 | 0 | 2631 | 0.99989 | 0.99962 | 0.99994 |
| | | | SVM | 0.99995 | 17463 | 0 | 1 | 2633 | 1.00000 | 0.99997 | 0.99997 |
| | | | RF | 1.00000 | 17464 | 0 | 0 | 2633 | 1.00000 | 1.00000 | 1.00000 |

The results for 7 of the IoT devices are shown in Tables I-VII respectively (we could not show results of all nine due to space limitations). The results are presented for: Danmini Doorbell, Ecobee Thermostat, Ennio Doorbell, Philips B120N10 Baby Monitor, Provision PT 737E Security Camera, Provision PT 838 Security Camera and Samsung SNH 1011 N Webcam. These results are shown by botnet, for each attack for each device, on the three classifiers: LR, SVM

and RF. Tables I-VII compare the accuracy and other statistical metrics of the three classification models, LR, SVM and RF, for 7 of the devices for the different attack types.

Fig. 1 and Fig. 2 present the classification accuracy of the other two devices, SimpleHome XCS7 1002 WHT Security Camera and SimpleHome XCS7 1003 WHT Security Camera. The classification accuracy is compared by classifier, LR, SVM and RF, by attack.

TABLE III: ENNIO DOORBELL RESULTS

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ennio Doorbell | Gafgyt | combo | LR | 1.00000 | 10617 | 0 | 0 | 7806 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99973 | 10612 | 0 | 5 | 7806 | 1.00000 | 0.99976 | 0.99976 |
| | | | RF | 1.00000 | 10617 | 0 | 0 | 7806 | 1.00000 | 1.00000 | 1.00000 |
| | | junk | LR | 0.99993 | 5848 | 0 | 1 | 7931 | 1.00000 | 0.99991 | 0.99991 |
| | | | SVM | 0.99964 | 5844 | 0 | 5 | 7931 | 1.00000 | 0.99957 | 0.99957 |
| | | | RF | 0.99993 | 5848 | 0 | 1 | 7931 | 1.00000 | 0.99991 | 0.99991 |
| | | scan | LR | 0.99970 | 5660 | 0 | 4 | 7780 | 1.00000 | 0.99965 | 0.99965 |
| | | | SVM | 0.99933 | 5660 | 5 | 4 | 7775 | 0.99912 | 0.99933 | 0.99921 |
| | | | RF | 1.00000 | 5664 | 0 | 0 | 7780 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99958 | 20778 | 6 | 6 | 7817 | 0.99971 | 0.99947 | 0.99971 |
| | | | SVM | 0.99937 | 20770 | 4 | 14 | 7819 | 0.99981 | 0.99941 | 0.99957 |
| | | | RF | 0.99993 | 20782 | 0 | 2 | 7823 | 1.00000 | 0.99995 | 0.99995 |
| | | tcp | LR | 0.99982 | 20379 | 3 | 2 | 7744 | 0.99985 | 0.99976 | 0.99988 |
| | | | SVM | 0.99940 | 20367 | 3 | 14 | 7744 | 0.99985 | 0.99946 | 0.99958 |
| | | | RF | 0.99989 | 20380 | 2 | 1 | 7745 | 0.99990 | 0.99985 | 0.99993 |

TABLE IV: PHILIPS_B120N10 BABY MONITOR RESULTS

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Philips_B120N10 Baby Monitor | mirai | ack | LR | 1.00000 | 18078 | 0 | 0 | 35195 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99998 | 18077 | 0 | 1 | 35195 | 1.00000 | 0.99997 | 0.99997 |
| | | | RF | 1.00000 | 18078 | 0 | 0 | 35195 | 1.00000 | 1.00000 | 1.00000 |
| | | scan | LR | 0.99995 | 20771 | 0 | 3 | 34999 | 1.00000 | 0.99993 | 0.99993 |
| | | | SVM | 0.99995 | 20771 | 0 | 3 | 34999 | 1.00000 | 0.99993 | 0.99993 |
| | | | RF | 0.99998 | 20773 | 0 | 1 | 34999 | 1.00000 | 0.99998 | 0.99998 |
| | | syn | LR | 1.00000 | 23598 | 0 | 0 | 35076 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99998 | 23597 | 0 | 1 | 35076 | 1.00000 | 0.99998 | 0.99998 |
| | | | RF | 1.00000 | 23598 | 0 | 0 | 35076 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99996 | 43378 | 0 | 3 | 35074 | 1.00000 | 0.99997 | 0.99997 |
| | | | SVM | 0.99992 | 43375 | 0 | 6 | 35074 | 1.00000 | 0.99993 | 0.99993 |
| | | | RF | 1.00000 | 43381 | 0 | 0 | 35074 | 1.00000 | 1.00000 | 1.00000 |
| | | udpplain | LR | 1.00000 | 16162 | 0 | 0 | 35048 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99994 | 16159 | 0 | 3 | 35048 | 1.00000 | 0.99991 | 0.99991 |
| | | | RF | 1.00000 | 16162 | 0 | 0 | 35048 | 1.00000 | 1.00000 | 1.00000 |

TABLE V: PROVISION_PT_737E SECURITY CAMERA RESULTS

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Provision_ PT_737E Security Camera | Gafgyt | combo | LR | 0.99996 | 12375 | 0 | 1 | 12331 | 1.00000 | 0.99996 | 0.99996 |
| | | | SVM | 0.99972 | 12369 | 0 | 7 | 12331 | 1.00000 | 0.99972 | 0.99972 |
| | | | RF | 0.99996 | 12375 | 0 | 1 | 12331 | 1.00000 | 0.99996 | 0.99996 |
| | | junk | LR | 0.99979 | 6188 | 0 | 4 | 12419 | 1.00000 | 0.99968 | 0.99968 |
| | | | SVM | 0.99968 | 6186 | 0 | 6 | 12419 | 1.00000 | 0.99952 | 0.99952 |
| | | | RF | 0.99984 | 6189 | 0 | 3 | 12419 | 1.00000 | 0.99976 | 0.99976 |
| | | scan | LR | 0.99995 | 5837 | 0 | 1 | 12453 | 1.00000 | 0.99991 | 0.99991 |
| | | | SVM | 0.99989 | 5836 | 0 | 2 | 12453 | 1.00000 | 0.99983 | 0.99983 |
| | | | RF | 0.99995 | 5837 | 0 | 1 | 12453 | 1.00000 | 0.99991 | 0.99991 |
| | | udp | LR | 0.99946 | 20750 | 14 | 4 | 12465 | 0.99933 | 0.99934 | 0.99957 |
| | | | SVM | 0.99970 | 20748 | 4 | 6 | 12475 | 0.99981 | 0.99970 | 0.99976 |
| | | | RF | 0.99997 | 20753 | 0 | 1 | 12479 | 1.00000 | 0.99998 | 0.99998 |
| | | tcp | LR | 0.99961 | 20928 | 10 | 3 | 12392 | 0.99952 | 0.99953 | 0.99969 |
| | | | SVM | 0.99982 | 20927 | 2 | 4 | 12400 | 0.99990 | 0.99982 | 0.99986 |
| | | | RF | 1.00000 | 20931 | 0 | 0 | 12402 | 1.00000 | 1.00000 | 1.00000 |
| | Mirai | ack | LR | 0.99988 | 12199 | 1 | 2 | 12340 | 0.99992 | 0.99988 | 0.99988 |
| | | | SVM | 0.99992 | 12199 | 0 | 2 | 12341 | 1.00000 | 0.99992 | 0.99992 |
| | | | RF | 1.00000 | 12201 | 0 | 0 | 12341 | 1.00000 | 1.00000 | 1.00000 |
| | | scan | LR | 0.99997 | 19349 | 0 | 1 | 12437 | 1.00000 | 0.99997 | 0.99997 |
| | | | SVM | 0.99997 | 19349 | 0 | 1 | 12437 | 1.00000 | 0.99997 | 0.99997 |
| | | | RF | 1.00000 | 19350 | 0 | 0 | 12437 | 1.00000 | 1.00000 | 1.00000 |
| | | syn | LR | 0.99992 | 13212 | 0 | 2 | 12366 | 1.00000 | 0.99992 | 0.99992 |
| | | | SVM | 0.99973 | 13210 | 3 | 4 | 12363 | 0.99977 | 0.99973 | 0.99974 |
| | | | RF | 1.00000 | 13214 | 0 | 0 | 12366 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99998 | 31282 | 0 | 1 | 12398 | 1.00000 | 0.99998 | 0.99998 |
| | | | SVM | 0.99995 | 31281 | 0 | 2 | 12398 | 1.00000 | 0.99997 | 0.99997 |
| | | | RF | 1.00000 | 31283 | 0 | 0 | 12398 | 1.00000 | 1.00000 | 1.00000 |
| | | udpplain | LR | 1.00000 | 11338 | 0 | 0 | 12429 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99996 | 11337 | 0 | 1 | 12429 | 1.00000 | 0.99996 | 0.99996 |
| | | | RF | 1.00000 | 11338 | 0 | 0 | 12429 | 1.00000 | 1.00000 | 1.00000 |

## VII. Discussion

From the statistical results we observe that the best performance is given by the RF classifier, followed by the LR. But, for the Provision_PT_737E Security Camera and the Provision_TP_838 Security Camera, SVM performs better than LR for the UDP attack. Though RF and LR perform better than SVM overall, the SVM results are only very slightly lower than RF and LR. In terms of attacks, we can say that the udp attack, of the gafgyt botnet, had a slightly lower classification rate than most other attacks. It would be difficult to say which attack had the best classification rate overall – most of the classification results were very good. Couple reasons for the good classification results might be: (i) the flow is expressed very finely and pre-processed using z-score normalization; and (ii) all features were collected in five time windows, and this data was pretty consistent for all time windows. As future work it might be good to see if all five different time windows are necessary and which features are really important for this classification.

TABLE VI: Provision_PT_838 Security Camera Results

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Provision_PT_838 Security Camera | Gafgyt | combo | LR | 0.99987 | 11656 | 0 | 4 | 19549 | 1.00000 | 0.99983 | 0.99983 |
| | | | SVM | 0.99958 | 11647 | 0 | 13 | 19549 | 1.00000 | 0.99944 | 0.99944 |
| | | | RF | 0.99990 | 11657 | 0 | 3 | 19549 | 1.00000 | 0.99987 | 0.99987 |
| | | junk | LR | 0.99996 | 5890 | 0 | 1 | 19626 | 1.00000 | 0.99992 | 0.99992 |
| | | | SVM | 0.99988 | 5888 | 0 | 3 | 19626 | 1.00000 | 0.99975 | 0.99975 |
| | | | RF | 0.99992 | 5889 | 0 | 2 | 19626 | 1.00000 | 0.99983 | 0.99983 |
| | | scan | LR | 0.99988 | 5721 | 0 | 3 | 19659 | 1.00000 | 0.99974 | 0.99974 |
| | | | SVM | 0.99980 | 5719 | 0 | 5 | 19659 | 1.00000 | 0.99956 | 0.99956 |
| | | | RF | 0.99988 | 5721 | 0 | 3 | 19659 | 1.00000 | 0.99974 | 0.99974 |
| | | udp | LR | 0.99929 | 20942 | 26 | 3 | 19664 | 0.99876 | 0.99927 | 0.99931 |
| | | | SVM | 0.99973 | 20941 | 7 | 4 | 19683 | 0.99967 | 0.99973 | 0.99974 |
| | | | RF | 0.99995 | 20943 | 0 | 2 | 19690 | 1.00000 | 0.99995 | 0.99995 |
| | | tcp | LR | 0.99965 | 17848 | 9 | 4 | 19720 | 0.99950 | 0.99966 | 0.99964 |
| | | | SVM | 0.99976 | 17847 | 4 | 5 | 19725 | 0.99978 | 0.99976 | 0.99975 |
| | | | RF | 0.99992 | 17849 | 0 | 3 | 19729 | 1.00000 | 0.99992 | 0.99992 |
| | Mirai | ack | LR | 0.99990 | 11700 | 0 | 3 | 19600 | 1.00000 | 0.99987 | 0.99987 |
| | | | SVM | 0.99990 | 11700 | 0 | 3 | 19600 | 1.00000 | 0.99987 | 0.99987 |
| | | | RF | 1.00000 | 11703 | 0 | 0 | 19600 | 1.00000 | 1.00000 | 1.00000 |
| | | scan | LR | 0.99990 | 19334 | 0 | 4 | 19784 | 1.00000 | 0.99990 | 0.99990 |
| | | | SVM | 0.99990 | 19334 | 0 | 4 | 19784 | 1.00000 | 0.99990 | 0.99990 |
| | | | RF | 0.99997 | 19337 | 0 | 1 | 19784 | 1.00000 | 0.99997 | 0.99997 |
| | | syn | LR | 1.00000 | 12361 | 0 | 0 | 19712 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 1.00000 | 12361 | 0 | 0 | 19712 | 1.00000 | 1.00000 | 1.00000 |
| | | | RF | 1.00000 | 12361 | 0 | 0 | 19712 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 1.00000 | 31541 | 0 | 0 | 19884 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 0.99998 | 31540 | 0 | 1 | 19884 | 1.00000 | 0.99998 | 0.99998 |
| | | | RF | 1.00000 | 31541 | 0 | 0 | 19884 | 1.00000 | 1.00000 | 1.00000 |
| | | udpplain | LR | 1.00000 | 10751 | 0 | 0 | 19709 | 1.00000 | 1.00000 | 1.00000 |
| | | | SVM | 1.00000 | 10751 | 0 | 0 | 19709 | 1.00000 | 1.00000 | 1.00000 |
| | | | RF | 1.00000 | 10751 | 0 | 0 | 19709 | 1.00000 | 1.00000 | 1.00000 |

TABLE VII: Samsung SNH1011N Webcam Results

| Device | Botnet | Attack | Algorithm | Accuracy | TP | FP | FN | TN | Precision | Recall | F1 score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SamsungSNH1011N Webcam | gafgyt | combo | LR | 0.99995 | 11740 | 1 | 0 | 10423 | 0.99991 | 0.99995 | 0.99996 |
| | | | SVM | 0.99982 | 11736 | 0 | 4 | 10424 | 1.00000 | 0.99983 | 0.99983 |
| | | | RF | 1.00000 | 11740 | 0 | 0 | 10424 | 1.00000 | 1.00000 | 1.00000 |
| | | junk | LR | 0.99988 | 5712 | 0 | 2 | 10377 | 1.00000 | 0.99982 | 0.99982 |
| | | | SVM | 0.99956 | 5707 | 0 | 7 | 10377 | 1.00000 | 0.99939 | 0.99939 |
| | | | RF | 0.99988 | 5712 | 0 | 2 | 10377 | 1.00000 | 0.99982 | 0.99982 |
| | | scan | LR | 0.99994 | 5504 | 0 | 1 | 10465 | 1.00000 | 0.99991 | 0.99991 |
| | | | SVM | 0.99950 | 5497 | 0 | 8 | 10465 | 1.00000 | 0.99927 | 0.99927 |
| | | | RF | 1.00000 | 5505 | 0 | 0 | 10465 | 1.00000 | 1.00000 | 1.00000 |
| | | udp | LR | 0.99972 | 22080 | 6 | 3 | 10465 | 0.99973 | 0.99965 | 0.99980 |
| | | | SVM | 0.99948 | 22071 | 5 | 12 | 10466 | 0.99977 | 0.99949 | 0.99962 |
| | | | RF | 0.99994 | 22081 | 0 | 2 | 10471 | 1.00000 | 0.99995 | 0.99995 |
| | | tcp | LR | 0.99977 | 19565 | 4 | 3 | 10415 | 0.99980 | 0.99973 | 0.99982 |
| | | | SVM | 0.99967 | 19560 | 2 | 8 | 10417 | 0.99990 | 0.99970 | 0.99974 |
| | | | RF | 0.99977 | 19561 | 0 | 7 | 10419 | 1.00000 | 0.99982 | 0.99982 |

From these results we can also note a very high attack detection rate, well over 99% in most cases and even 100% in many cases, mostly using the RF algorithm. The Damini Doorbell and Provision_PT_838 Security Camera had 100% ADR using the other algorithms too, mostly in the Mirai botnet. All three algorithms also had a very high precision and F1 scores (one or very close to one) for almost all of the attacks.

We present the graphical results of classification accuracy for the SimpleHome Security Camera and SimpleHome_XCS71003WHT Security Camera. From these two figures too, we can observe that, on the average, RF

performed the most consistently, LR performed the second best and SVM performed the least consistently, though the classification accuracy of all three algorithms were very high.

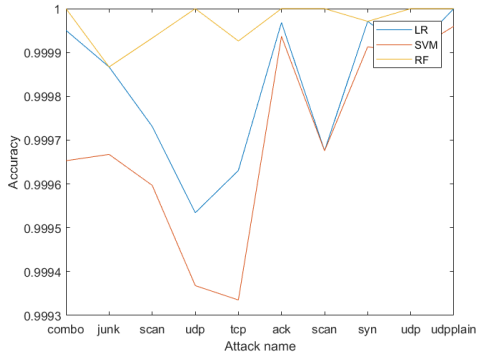


Fig. 1. Classification accuracy for simplehome security camera.

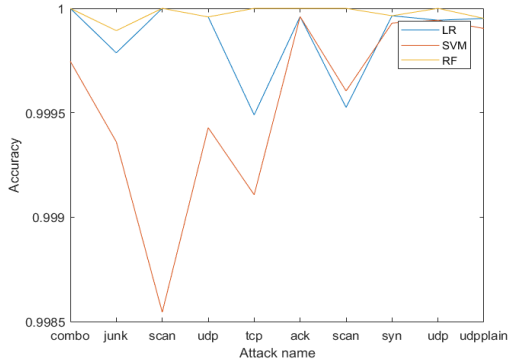


Fig. 2. Classification accuracy for simplehome_XCS71003WHT security camera.

## VIII. Conclusions and Future Works

Though the results, by botnet, for each attack on each device, for all three classifiers, show very high ADRs and classification accuracy (over 99%) with regard to determining whether an IoT device is attacked by a particular botnet, we can say that, on the average, the RF algorithm performed the best and SVM performed the lowest of the three algorithms. The high F1 scores show the robustness of three algorithms used.

This being an initial study, we used all the features in the dataset. As a follow-up study, it would be good to do feature selection and see which of the features perform the best for each attack for each device. A detailed study of the features would also be useful information. For example, it would be interesting to see if each of the attacks on the security cameras had similar characteristics or each of the attacks on the doorbells had similar characteristics, etc. This would be helpful in determining how to handle and prevent future attacks.

## Conflict of Interest

There is no conflict of interest to report.

## Author Contributions

Dr. Sikha Bagui helped in designing the research plan and the write-up of the paper. Xiaojian also helped in the research plan, worked on the programming and on the write-up of the paper. Dr. Subhash Bagui provided the statistical guidance.

## Acknowledgements

This work is partially supported by the Askew Institute of the University of West Florida.

## References

[1] K. Angrishi, "Turning Internet of Things (IoT) into Internet of Vulnerabilities (IoV): IoT botnets," *ArXiv:1702.03681*, pp. 1-17, 2017.

[2] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80-84, 2017.

[3] What is a Botnet? [Online]. Available: https://www.pandasecurity.com/mediacenter/security/what-is-a-botnet/

[4] Detection_of_IoT_botnet_attacks_N_BaIoT Data Set. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_N_BaIoT

[5] R. K. Gunupudi, M. Nimmala, N. Gugulothu, and S. R. Gali, "Clapp, a self constructing feature clustering approach for anomaly detection, *Future Gener. Comput. Syst.*, vol. 74, pp. 417-429, 2017.

[6] S. O. Amin *et al.*, "A novel coding scheme to implement signature based IDS in IP based sensor networks," in *Proc. IFIP/IEEE International Symposium on Integrated Network Management-Workshops*, 2009.

[7] A. Abubakar and B. Pranggono, "Machine learning based intrusion detection system for software defined networks," in *Proc. Seventh International Conference on Emerging Security Technologies (EST)*, 2017, pp. 138-143.

[8] A. Le, J. Loo, Y. Luo, and A. Lasebae, "Specification-based IDS for securing RPL from topology attacks," in *Proc. IFIP Wireless Days (WD)*, 2011.

[9] D. Summerville, K. Zach, and Y. Chen, "Ultra-lightweight deep packet anomaly detection for Internet of Things devices," in *IEEE Proc. the 34th International Performance Computing and Communications Conference (IPCCC)*, 2015, pp. 1-8.

[10] P. Kasinathan, C. Pastrone, and M. Spirito, "Denial-of-Service detection in 6LoWPAN based internet of things," in *Proc. IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2013, DOI: 10.1109/WiMOB.2013.6673419.

[11] P. Kasinathan, G. Costamagna, and H. Khaleel, "An ids framework for internet of things empowered by 6lowpan," in *Proc. the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 2013, pp. 1337-1340.

[12] S. Raza, L. Wallgren, and T. Voigt, "SVELTE: Real-time intrusion detection in the Internet of Things," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2661-2674, 2013.

[13] J. Granjal, J. M. Silva, and N. Lourenco, "Intrusion detection and prevention in CoAP wireless sensor networks using anomaly detection," *Sensors*, vol. 18, no. 8, pp. 2445, 2018.

[14] T. Lee, C. Wen, L. Chang, H. Chiang, and M. Hsieh, "A lightweight intrusion detection scheme based on energy consumption analysis in 6LowPAN," *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*, vol. 260, Springer, Netherlands, pp. 1205-1213, 2014.

[15] D. Oh, D. Kim, and W. Ro, "A malicious pattern detection engine for embedded security systems in the Internet of Things," *Sensors*, vol. 14, no. 12, pp. 24188-24211, 2014.

[16] J. Chen, and C. Chen, "Design of complex event- processing IDS in internet of things," in *Proc. IEEE Sixth International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 2014.

[17] B. Karakostas, "Event prediction in an IoT environment using nave bayesian models," *Procedia Comput. Sci.*, vol. 83, pp. 11-17, 2016.

[18] Y. Zhao, L. T. Yang, and J. Sun, "A secure high-order CFS algorithm on clouds for industrial internet-of-things," *IEEE Trans. Industr. Inf.*, vol. 14, no. 8, pp. 3766-3774, 2018.

[19] D. Midi, A. Rullo, A. Mudgerikar, and E. Bertino, "Kalis: A system for knowledge-driven adaptable intrusion detection for the Internet of Things," in *Proc. the IEEE 37th International Conference on Distributed Computing Systems*, 2017.

[20] H. Bostani, and M. Sheikhan, "Hybrid of anomaly-based and specification-based IDS for Internet of Things using unsupervised OPF based on MapReduce approach," *Computer Communications*, vol. 98, pp. 52-71, 2017.

[21] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, "Conditional variational autoencoder for prediction and feature

recovery applied to intrusion detection in IoT," *Sensors*, vol. 17, no. 9, 2017.
[22] V. H. Bezerra, V. G. T. DaCosta, S. B. Junior, R. S. Miani, and B. B. Zarpelao, "One-class classification to detect botnets in IoT devices," in *Proc. 18th Brazilian Symposium on Information and Computational Systems Security*, 2018, pp. 43-56.
[23] E. Bertino and N. Islam, "Botnets and internet of things security," *Computer*, vol. 50, no. 2, pp. 76-79, 2017.
[24] M. Guller, *Big Data Analytics with Spark*, APress, 2015.
[25] S. Homayoun, M. Ahmadzadeh, S. Hashemi, and A. Dehghantanha, "Botshark: A deep learning approach for botnet traffic detection," *Computer Science*, pp. 137-153, 2018.
[26] O. Y. Al-Jarrah, O. Alhussein, P. D. Yoo, S. Muhaidat, K. Taha, and K. Kim, "Data randomization and cluster-based partitioning for botnet intrusion detection," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1796-1806, 2016.
[27] H. P. Joshi, M. Bennison, and R. Dutta, "Collaborative botnet detection with partial graph information," in *Proc. 2017 IEEE 38th Sarnoff Symposium*, 2017.
[28] A. O. Prokofiev, Y. S. Smirnova, and V. A. Surov, "A method to detect Internet of Things botnets," in *Proc. 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2018.
[29] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai, and Y. Elovici, "N-BaIoT: Network-based detection of IoT botnet attacks using deep Autoencoders," *IEEE Pervasive Computing*, vol. 13, no. 9, pp. 1-8, 2018.
[30] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of Autoencoders for online network intrusion detection," presented at Network and Distributed Systems Security Symposium (NDSS), San Diego, USA, February 18-21, 2018.



**Sikha Bagui** is a professor and Askew fellow in the Department of Computer Science, at The University West Florida, Pensacola, Florida. Dr. Bagui is active in publishing peer reviewed journal articles in the areas of database design, data mining, big data and big data analytics, and machine learning. Dr. Bagui has worked on funded as well unfunded research projects and has numerous peer reviewed publications. She has also co-authored several books on database and SQL. Bagui also serves as an associate editor and is on the editorial board of several journals.

**Xiaojian Wang** received her bachelor's degree from Taiyuan University, China, in 2017, majoring in internet of things. She completed her master's degree in computer science from the University of West Florida, Pensacola, Florida, USA, in December, 2019. She also graduated with a master's degree from Taiyuan University of Technology in China in July 2020, majoring in computer science and technology. Her research interests are in the areas of IoT security and data analytics.

**Subhash C. Bagui** received his B.Sc. in statistics from University of Calcutta, M. Stat. from Indian Statistical Institute and Ph.D. from University of Alberta, Canada. He is currently a University distinguished professor at the University of West Florida. He has authored a book titled, "Handbook of Percentiles of Non -central t-distribution", and published many high quality peer reviewed journal articles. He is currently serving as associate editors/ editorial board members of several statistics journals. His research interests include nonparametric classification and clustering, statistical pattern recognition, machine learning, central limit theorem, and experimental designs. He is also a fellow of American Statistical Association (ASA) and Royal Statistical Society (RSS).