Predicting Malicious Software in IoT Environment Based on Machine Learning and Data Mining Techniques

Abdulmohsen Alharbi, Md. Abdul Hamid, Husam Lahza
Department of Information Technology
Faculty of Computing and Information Technology
King Abdulaziz University, Jeddah 21589, Saudi Arabia

***Abstract*—The Internet of Things (IoT) enable the IoT to sense and respond using the power of computing to autonomously come up with the best solutions for any industry today. However, Internet of Things have vulnerabilities since it can be hacked by cybercriminals. The cybercriminals know where the IoT vulnerabilities are, such as unsecured update mechanisms and malware (Malicious Software) to attack the IoT devices. The recently posted IoT-23 dataset based on several IoT devices such as Philips Hue, Amazon Echo devices and Somfy door lock were used for machine learning classification algorithms and data mining techniques with training and testing for predictive modelling of a variety of malware attacks like Distributed Denial of Service (DDoS), Command and Control (C&C) and various IoT botnets like Mirai and Okiru. This paper aims to develop predictive modeling that will predict malicious software to protect IoT and reduce vulnerabilities by using machine learning and data mining techniques. We collected, analyzed and processed benign and several of malicious software in IoT network traffic. Malware prediction is crucial in maintaining IoT devices' safety and security from cybercriminals' activities. Furthermore, the Principal Component Analysis (PCA) method was applied to determine the important features of IoT-23. In addition, this study compared with previous studies that used the IoT-23 dataset in terms of accuracy rate and other metrics. Experiments show that Random Forest (RF) classifier achieved the predictive model produced classification accuracy 0.9714% as well as predict 8754 samples with various types of malware and obtained 0.9644% of Area Under Curve (AUC) which outperforms several bassline machine learning classification models.**

***Keywords*—Machine learning; internet of things; malware; predictive modeling; cyber threats**

I. INTRODUCTION

The Internet of Things are internet-connected devices that can transfer data over a network. Nowadays lots of cyber-attacks have increased, the cybercriminals seek to exploit or damage data, disrupt computer devices and network resources. The term cyber generally, defines computer devices, network, internet and information technology. [1] Cyber threat is a possibility to successful cyber-attack that aims to harm computer system or network, steal sensitive data and gain unauthorized access. However, the IoT are vulnerable in terms of security; cybercriminals use malware attacks such as DDoS, ransomware, and IoT botnet attack to disable systems and networks. Study by Gotsev et al. [2] used different machine learning models to evaluate the performance of Machine Learning (ML) for attack detection.

The researchers used all the features of IoT-23 dataset [3] which has 21 features and they detected different types of malware attack on IoT devices such as DDoS, Okiru, HorizontalPortScan and other IoT botnets. Similar study by Nicolas Stoian [4] focused on the security aspect of the IoT by investigating the usability of ML approaches on anomaly detection. In the research, the dataset has been split into 80% for training and 20% for testing for each ML algorithms. In results, the best ML algorithm is Random Forest with a weighted average precision of 100%.

Chunduri et al. [5] used multi class classification to detect IoT botnet malware. Their aim is to build a classifier to detect IoT botnet attack and to get the best accuracy possible by using machine learning classifiers. They have used Network Traffic Analysis Tool (Zeek) [6] that monitors all the traffic on network for malicious activity. The PCA method was applied to minimize features and maximize the accuracy rate by using ML classification algorithms such as Decision Tree (DT) K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Naïve Bayes (NB). The traditional approaches using static analysis method is complicated in terms of examining malicious software that exposes IoT to security breach risks. Therefore, it is important to improve the method by utilizing machine learning and data mining techniques to make it safer and more effective in predicting malicious software in IoT network traffic.

In this research the IoT-23 dataset [3] was used and collected, analyzed, processed to predict benign and several of malicious software in IoT network traffic. Furthermore, in results for prediction models, RF algorithm achieved highest accuracy for predicting malicious and benign in IoT network traffic. Section II is about background of cyber threats, the literature review is reviewed in Section III, Section IV is methodology, Section V describes performance evaluation of ML classification algorithms, Section VI describes experiments and results for validation of predictive models of malicious and benign in IoT network traffic and Section VII is about discussion of the results.

A. Problem Statement

Cybercriminals use malware attacks on IoT devices to hack and disable IoT devices. The attacks make IoT devices less efficient and cybercriminals can steal sensitive information and personal data. The cybercriminals know where the vulnerabilities of IoT devices are and exploit them through malware attacks on the devices. The existing studies focus on detecting threats and ignore the significance of predicting malware threats that increase IoT devices' vulnerabilities. Moreover, none of the studies included malicious software as a threat to security in IoT devices. The studies have focused on end-to-end devices or attack surfaces. This research aims to predict malicious software in IoT network traffic in order to protect IoT devices and reduce vulnerabilities by using ML and data mining techniques.

B. Importance of this Research

The harmful effects of malware into IoT environment are exposure of IoT to security breach risks and stealing of sensitive information and personal data. Therefore, we collected, analyzed and processed benign and several of malicious software in IoT network traffic of IoT-23 dataset [3]. Four types of malicious software were chosen out of IoT-23 and we considered those that have more significant effect on the IoT devices, which are DDoS, C&C and various botnets like Mirai and Okiru. Types of IoT devices in IoT-23 dataset are Echo device, Hue device and door lock device. It is extremely important to predict malware in IoT network traffic in order to protect IoT devices and reduce vulnerabilities by increasing the security level for IoT devices and to improve the environment and make it more motivational. This research will be using machine learning and data mining techniques to make it safer and more effective in predicting malicious software in IoT network traffic.

Our contributions in this work are:

- Minimize features and maximize the accuracy rate by using PCA method and ML classification algorithms.
- Propose prediction model to predict malicious and benign in IoT network traffic by using supervised learning.
- Increasing the security level for IoT to improve the environment and make it more motivational.

II. Background

This section provides related background information and context to explain the objectives and relevant field of research of this work. The concept, types, and IoT network activities of malicious software and botnet are discussed. Then, the concept of cybercriminals using malware attack is addressed. Finally, the most important malware attack that cause security risks on IoT environment is introduced.

A. Cyber Threats

Nowadays lot of cyber-attacks have increased, the cybercriminals seek to exploit or damage data, disrupt computer devices and network resources. The term cyber generally, defines computer devices, network, internet and information technology [7]. Cyber-attack attribution is technique that tracks, identifies, and lays blame on the criminal of a cyber-attack or other hacking exploit. Cyberspace is the environment of the internet that involving a global of computer network or the internet to enable communications and data exchange activities. Cyber Threat Intelligence (CTI) generally is relying on the collection of information and its analysis with current or potential attacks that is threatening policy of the organizations [8]. Cybercriminals can be internal or external to the organization that is facing cyberattack. An attack on the computer devices, network or system performed by person who has authorization access is known as an insider attack. An attack that originates exposures from outside the organization and attempt to exploit IT equipment are known as external attacks [1]. Cyber threat is possibility to successful cyber-attack that aims to harm computer system or network, steal sensitive data and gain unauthorized access. Some top cyber threats are illustrated as following [9] [10] [11] [12] [13]:

B. Malware (Malicious Software)

Malware is computer code designed to disrupt and disable such as stealing sensitive data or taking control of computer system. Malware (Malicious Software) has remained the most common cyber threats since 2014. Approximately four million samples of malware on different devices are detected by security organizations in 2017. The increase of malware samples have escalated malware attacks.

C. Ransomware Attack

Ransomware is a type of malware, which restricts access to user files or a computer system till the victim pays a ransom. Ransomware is significant cybersecurity threat since it uses techniques to avoid detection system to attack legitimate users. Ransomware can be considered a part of malware and it has been evaluated as a separate threat, although it belongs to the malware category. Moreover, Ransomware is considered the most significant cyber-attacks nowadays.

D. Distributed Denial of Service (DDoS) Attack

It is a cyber-attack which is an attempt to compromise the availability of computer devices or network resources to make them unavailable to the legitimate or normal users. DDoS attack is aimed to send massive amount of superfluous requests in order to deny the server from responding to the valid requests immediately. Denial of Service (DoS) attack can damage the target that rely on an online presence, while DDoS attack strikes a target with several resources and is harder to stop DDoS attack.

E. Cyber Espionage

It is type of cyber-attack which is an act of obtaining confidential information without permission from the user of the information for economic, political, military or personal objectives. It includes utilization of the internet or a computer network over utilized proxy server, malicious software including Trojan horse and spyware. The targets of this attack are government and commercial sectors. Cybercriminals develop new tools and techniques to increase the number of attacks and the degree of damage caused to its victims.

F. IoT Botnet Attack

The Internet of Things (IoT) bot is a variant of a traditional botnet that contains a group of compromised computers, smart devices and sensors connected to the internet. IoT botnet attack is used by cybercriminals for causing damage such as financial and for illegitimate purposes in terms of control of malicious actors. Over 41% of all attacks are due to the vulnerabilities of the IoT devices and IoT botnet attack contribute approximately 13% total of attacks in various other information technology industries.

III. Related Work

Study by Gotsev et al. [2] used different machine learning models to evaluate the performance of ML for attack detection. They applied various ML classifiers such as Support Vector Machine, Random Forrest Naïve Bayes, Logistic Regression and Decision Tree. In the experiments, the researchers used all feature of IoT-23 dataset [3] which has 21 features. Furthermore, IoT-23 dataset contains labeled information of benign and malicious IoT network traffic. They detected different types of malware attack on IoT devices such as DDoS, Okiru, HorizontalPortScan and other. In testing results, DT and RF achieved highest accuracy detection which was 1.00% and LR classifier achieved 0.76% accuracy, SVM achieved 0.74% accuracy, while NB classifier had unsatisfied result and achieved 58% accuracy.

Similar study by Nicolas Stoian [4] focused on the security aspect of Internet of Things networks by investigating the usability of ML approaches of anomaly detection. The researcher used 14 features of IoT-23 dataset and applied statistical correlation to dataset in order to eliminate the data which was irrelevant to the label column. Furthermore, the research splitting the dataset into 80% for training and 20% for testing for each ML algorithms. In results, the best ML algorithm is Random Forest with a weighted average precision of 100%, another algorithm is AdaBoost with precision of 86% while Support Vector Machine has precision of 60% and Naïve Bayes with a weighted average precision of 76% Chunduri et al. [5] used multi class classification to detect IoT botnet malware. Their aim is to build a classifier to detect IoT botnet attack and to get the best accuracy possible by using machine learning classifiers. The researcher used IoT-23 dataset [3] which contains benign and malicious network traffic of IoT devices. They focused on six types of botnet attack which are Mirai, Bashlite, Torii, Hakai, Okiru and Muhstik, moreover they used Zeek (Network Traffic Analysis Tool) [6] that monitors all the traffic on network for malicious activity. Furthermore, the researchers selected 12 features of IoT-23 dataset; in results they applied ML classifiers to training and testing IoT-23 dataset. The best accuracy was achieved by RF 99.88%, GradientBoosting produced 99.36% and K-Nearest Neighbors achieved 96.14% while Support Vector Machine with 94.72% can be considered the least fit model. In study by Strecker et al. [14], the researchers compared the effectiveness machine learning classifiers based cyber security techniques on the IoT-23 dataset. They used seven features of the IoT-23 dataset and applied RF, SVM and KNN algorithms for IoT cyber security in 2021. Their result for malware detection, the highest accuracy is of RF 92.27%, the second-best accuracy is KNN 89.80% and SVM achieved 83.52%. In 2018, Mirsky et al. [15] built an Intrusion Detection System (IDS) with autoencoders for detection of online anomaly called Kitsune. The researchers have developed attribute extractor that consists in the following attribute categories which are Socket, Network Jitter, Host-MAC&IP and Channel. They have demonstrated on their results anomaly detection of Mirai botnet malware on IoT devices. In 2018, Meidan et al. [16] introduced a dataset called N-BaIoT for Bashlite and Mirai botnet malware that considering as Kitsune [17] attributes which have been implemented on nine different IoT devices. Ferrag et al. [18] have investigated the way seven contemporary Artificial Neural Networks (ANN) approaches perform training of the CICIDS-2018 and the BoT-IoT datasets. They have provided the details on overall accuracy, training time by using Deep Learning (DL) detection rate.

Potluri et al. [19] evaluates a Convolutional Neural Network (CNN) based network intrusion detection techniques. They used the NSL-KDD and the UNSW-NB15 datasets. These datasets are converted into an image such as format as part of the process. The researchers build the three layers of CNN to label for the attacks. The study is compared the GoogLeNet and ResNet50 with designed CNN approach that achieved the satisfying results, with accuracy rate achieved 91.14% on the NSL-KDD dataset and 94.9% on the UNSW-NB15 dataset. De La Torre Parra et al. [20] proposed a method for detecting attacks at the back-side and client end at the same time. The client's site uses a CNN model with micro security for the detection of DDoS, botnets, and phishing attacks. The authors designed a joint training method for minimizing the resource utilization for detection of attacks in IoT devices and maximized the usability of extracted features for using the back-end server. The scope of the study is limited to using the CNN model for detecting URL-based attacks aimed at the client's IoT device and the RNN–LTSM model at the back-end server for the detection of malware attacks.

The focus of the study by Pastor et al. [21] is to provide measures for the detection of these malwares using passive network-based monitoring. Network flow features were identified for this purpose according to relevancy, and they were used with deep learning models and machine learning models. The researchers used some algorithms i.e. C4.5 Random Forest (RF) and Deep Neural Networks (DNN) to compare their performance. The main aim was to monitor crypto mining and the detection of real-time flow. This was done through testing these models in complex scenarios using real servers and connections that were encrypted. Various features were employed to demonstrate the efficiency of these models against crypto mining.

Study by Li et al. [22] focused on Command and Control (C2) server that is employed by using a Domain Generation Algorithm (DGA) in order to generate communication between C2 and malware. This cannot be easily countered by using traditional methods like blacklisting. The researchers provide the framework of machine learning in order to deal with these threats. Real-time data were collected for one year using real traffic, and a deep learning model was proposed for

the classification of domains of DGA. Results showed an accuracy of 95.89%, 97.79%, 92.45% and 95.21% for framework classification, DNN model, clustering at the second level and HMM prediction, respectively.

The study by Sarker [23] presented the Cyberlearning for binary classification model in order to detect anomalies and classification of multi-class model of cyber-attacks. Features that are correlated to this were selected for an analysis of comprehensive nature. The empirical data on the effectiveness of this model was analyzed. This model takes the binary classification into account for evaluating the effectiveness in detecting anomalies and other cyber-attacks. The techniques for machine learning were employed. For the hidden layers, a security model that is based on an artificial neural network was presented, and the effectiveness for these was evaluated using various techniques. Security datasets NSL-KDD and UNSW-NB15 were examined to employ an experimental analysis. The findings were believed to provide a good reference to future research in the same field.

Another study by Li et al. [24] used a detection system called Significant Permission IDentification (SigPID). This system is designed with three levels without extracting the usage of Android permissions. These levels include pruning permission data to identify malicious apps and then classifying those malwares using only 22 significant permissions. These permissions are then compared with the baseline approach, and the final outcomes indicates the precision of up to 90% in F-measure, accuracy and recall as well. Their dataset contains 2000 malware and the SigPID is determined to have an effectiveness of 93.62 in the detection of dataset malware and effectiveness of 91.4% in detecting unknown malware. The researcher by Karanja et al. [25] used a novel approach towards analyzing and classifying malware. This is done using texture features and classical classifiers of machine learning that apply to the IoT malware. A low computation approach was employed by converting the malware binaries into images. This broke the environmental dependencies and platform barriers, considering the analysis of images is not limited to platforms. A 95% and 88% accuracy were achieved through a K-nearest neighbor and random forest classifier. The results showed that this method is applicable for real-time settings and can be employed for flagging off known IoT malware using preprocessed features of the image in known malware. D. Li and Q. Li [26] used a mixture of attacks that use multiple generative methods and yield adversarial malware with multiple manipulation sets. The adversarial training is used with a manipulated set with large cardinality. The robustness of malicious software detection against twenty-six evasion attacks is based on five methods using gradient-based, gradient-free, obfuscation, a mixture of attacks, and transfer the attack. The proposed methods improved the performance, but more research is required in the area of adversarial malware detection.

A. Data Mining Techniques

Data mining approaches in Internet of Things (IoT) systems are integrated to discover in terms of a range of well-established knowledge patterns such as supervised, unsupervised, semi-supervised, and statistical approaches. These data mining approaches enable classification, prediction and regression of upcoming streaming data to be able to be visualizing the knowledge and activate the sensors and actuators of the IoT systems. Numerous crucial data mining techniques are illustrated as following [27] [28] [29] [30]:

B. Classification

Classification in data mining is a popular technique that splits data points into various classes which assigns items in a collection to target categories or classes. It allows to organize dataset of all types, including complicated and massive dataset as well as small and simple ones.

C. Regression

Regression is a type of data mining technique utilized to predict numeric values given in a particular part of dataset. The most popular types of regression are linear and logistic regressions algorithms of machine learning. Furthermore, other types of regression can be performed depending on their performance on an individual dataset.

D. Prediction

Prediction in data mining is to predict the unknown values or outcomes. Prediction techniques in data mining discovers the correlation among dependent and independent variables and the correlation between independent variables. It predicts the identity of single variable based on the current description of some other related variable.

IV. Methodology

In methodology section, the author will show the chosen techniques and tools to implement the approach for predicting malicious and benign IoT network traffic by using machine learning and data mining techniques. Furthermore, The IoT dataset-23 was selected which has a labeled malicious and benign on IoT network traffic. In addition, data preprocessing was applied and used Principal Component Analysis (PCA) method for feature selection to make it suitable for a machine learning model. To achieve the goals of building a usable supervised machine learning model for predicting malicious and benign IoT network traffic, this research will be applying IoT-23 dataset [3] targeting Weka tools [31] for ML model and data Orange tools [32] for data mining techniques.

A. IoT Dataset Selection

A large dataset IoT-23 published in January 2020 [3] has been identified. IoT-23 consists of a labeled dataset with malicious and benign IoT network traffic, types of IoT devices in IoT-23 dataset i.e. Echo device, Hue device and door lock device. The IoT-23 dataset created by the Avast AIC (Artificial Intelligence and Cybersecurity) laboratory which is help for researchers to develop machine learning algorithms. IoT-23 dataset has twenty malicious captures executed from different IoT devices, in which 11 malware labels and one benign label have existed in IoT network traffic. Four types of malicious software have been chosen of IoT-23 and we consider those that have more significant effect on the IoT devices, which are DDoS, C&C and various botnets like Mirai and Okiru. Fig. 1 shows distribution of main labeled after preprocessing on IoT-23 dataset.

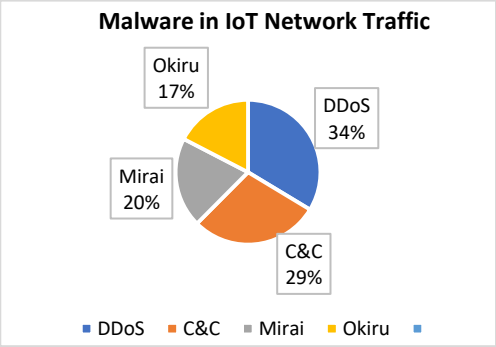

Fig. 1. Distribution of Main Labeled IoT-23 Dataset.

B. Feature Selection IoT-23 Dataset

After preprocessing IoT-23 dataset [3] the Principal Component Analysis (PCA) method was applied for feature selection. The IoT-23 dataset has 21 features and Weka tools [31] was used since it supports PCA method. After using PCA method, the IoT-23 dataset reduced to 18 features. The purpose of using PCA method is to find set of variables on IoT-23 dataset with less redundancy. Fig. 2 shows the main stages for feature selection using PCA method.

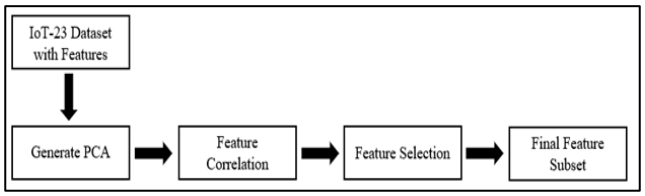

Fig. 2. Main Stages of the Feature Selection by PCA Method.

C. Principal Component Analysis

The PCA method was used for reducing the features of IoT-23 dataset. We eliminated these least important features for feature selection but do not lose original dataset completely. PCA method helps us to identify patterns in data of IoT-23 dataset based on the correlation among features. Furthermore, PCA method improved machine learning classification algorithms performance, removed correlated features and reduced overfitting by removing the unnecessary features in the IoT-23 dataset, which leads to minimizing features and maximizing the accuracy rate by using ML classification algorithms.

D. Supervised ML Models used for Prediction

In this research, the supervised learning model was used in terms of getting trained on a labelled dataset. A labelled IoT-23 dataset has two classes 0 and 1. 0 refers to benign while 1 refers to malicious, as binary classification we are predicting one of two classes in terms to know which features are malicious and benign of IoT network traffic. Fig. 3 shows prediction model for malicious and benign.

E. Weka Tools

Weka tools [31] is collection of machine learning algorithms for data mining tasks. It contains tools for data preprocessing, classification, clustering, regression and more. It is considered as an efficient tool for ML and data mining since it supports unsupervised and supervised ML algorithms.

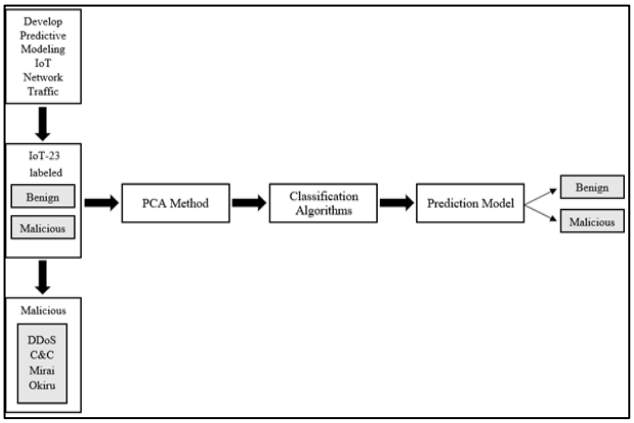

Fig. 3. Proposed Prediction Model.

F. Orange Tools

Orange tools [32] is an open-source platform to perform data analysis, machine learning and data mining Python scripting or visual programming. It contains prediction model that can predict the future based on previous attitude that happened before.

G. Classification Algorithms used in ML

Machine learning supervision was applied to train on a labelled IoT-23 dataset. A labelled dataset has two classes 0 and 1; 0 is benign and 1 is malicious. The labeled dataset is targeting to predict a packet which is malicious or benign. Furthermore, ML classification algorithms are applied since the labelled IoT-23 data has two classes. ML classification algorithms were used i.e. DT, RF, KNN, SVM and NB algorithms for prediction malicious and benign of IoT network traffic. Fig. 4 shows supervised ML classification algorithms used.

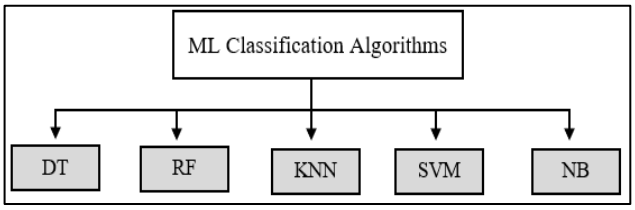

Fig. 4. Classification Algorithms used for Predication.

H. Training and Testing

For training and testing, the IoT-23 dataset was split into 70% for training and 30% for testing and 10-fold cross validation, to make it suitable for a machine learning models. Training or testing data is an approach to measure the accuracy of ML model. Moreover, machine learning classification algorithms were applied for training and testing model of ML.

V. Performance Evaluation

After training and validating ML models, some metrics are needed to identify the best model from a set of ML models. Our model of ML was developed to provide accurate prediction. Confusion matrix and evaluation metrics for classification model are used for predicting malicious and benign IoT network traffic. Additionally, four metrics used to

evaluate classification ML model which are accuracy, precision, recall and F1-score.

- **Accuracy** = $\frac{(TP+TN)}{(TP+FP+TN+FN)}$
- **Precision** = $\frac{TP}{(TP+FP)}$
- **Recall** = $\frac{TP}{(TP+FN)}$
- **F1-Score** = $\frac{2*(Precision*Recall)}{(Precision+Recall)}$

A. Confusion Matrix Model

Confusion Matrix Model (CMM) [33] [34] [35] [36] is applied to understand the performance of True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN). Furthermore, CMM is used to describe the performance of the classifier model of test data for which the true values are known. Confusion matrix was applied for prediction malicious and benign as shown in Table I.

TABLE I. CONFUSION MATRIX

Confusion Matrix	Predicted Class: Benign	Predicted Class: Malicious
Actual Class: Benign	True Positive (TP)	False Negative (FN)
Actual Class: Malicious	False Positive (FP)	True Negative (TN)

B. Receiver Operating Characteristic (ROC) Curve

Receiver Operating Characteristic (ROC) curve was applied to show in a graphical way the trade-off between clinical sensitivity and specificity. The x-axis is false positive rate and the y-axis is true positive rate. Two metrics used to evaluate a ROC curve, Area Under the Curve (AUC) if equals 0.70% the model will be able to distinguish between true positive class and false positive class.

- **Sensitivity** = $\frac{TP}{(TP+FN)}$
- **Specificity** = $\frac{TN}{(TN+FP)}$

VI. EXPERIMENTS AND RESULTS

In the experiments, will present the results of the predicting malicious and benign of IoT network traffic by using various supervised machine learning classification algorithms. Four types of malicious software have predicted for each class which are DDoS, C&C, Mirai and Okiru. Performance model is evaluated using accuracy, precision, recall and F1-score. The IoT-23 dataset has split into 70% for training and 30% for testing and 10-fold cross validation, to make it suitable for a machine learning model. Experiments are done on prepared feature of IoT-23 dataset [3] using ML classification algorithms such as DT, RF, KNN, SVM and NB. This chapter also presents comparison of evaluation metrics for predictive model of the proposed technique with existing studies, for predicting malicious and benign IoT network traffic. The implementation of the ML model by Weka tools [31] and data mining techniques by Orange tools [32].

A. Malicious and Benign of IoT Network Traffic

The performance evaluation is computed using four metrics which are accuracy, precision, recall and F1-score. For training and testing by supervised learning (SL) the number of samples of malicious is 8K samples while benign is 43K samples as shown in Fig. 5 and Fig. 6. ML classification algorithms used i.e. DT, RF, KNN, SVM and NB for validation of malicious and benign in IoT network.

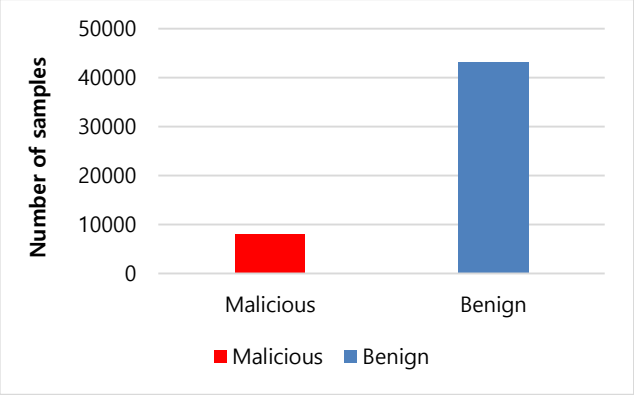

Fig. 5. Number of Samples Malicious and Benign.

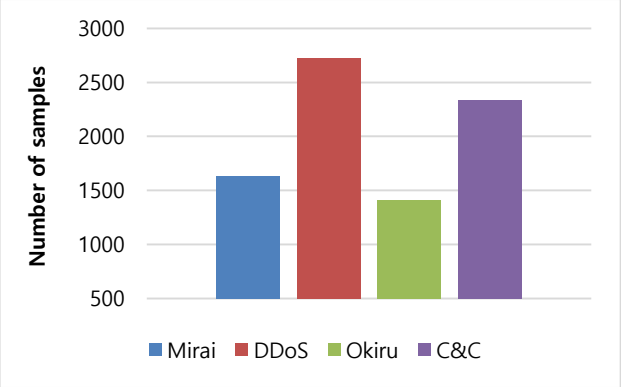

Fig. 6. Number of Samples each Type of Malware.

B. Performance Evaluation of ML Algorithms

Four metrics were applied for evaluating model of machine learning, the best results of ML classifiers algorithms were Random Forest and Support Vector Machine. Other ML algorithms obtained satisfying results, Table II shows performance evaluation of ML algorithms. Fig. 7 shows comparison performance metrics of ML algorithms, Fig. 8 comparison of ML algorithms using True Positive Rate (TPR).

TABLE II. PERFORMANCE EVALUATION OF ML ALGORITHMS

Classifier	Accuracy	Precision	Recall	F-1 Score
DT	0.9567	0.9580	0.9556	0.9553
RF	0.9848	0.9855	0.9850	0.9855
KNN	0.9674	0.9770	0.9773	0.9770
SVM	0.9840	0.9845	0.9850	0.9845
NB	0.9479	0.9668	0.9480	0.9544

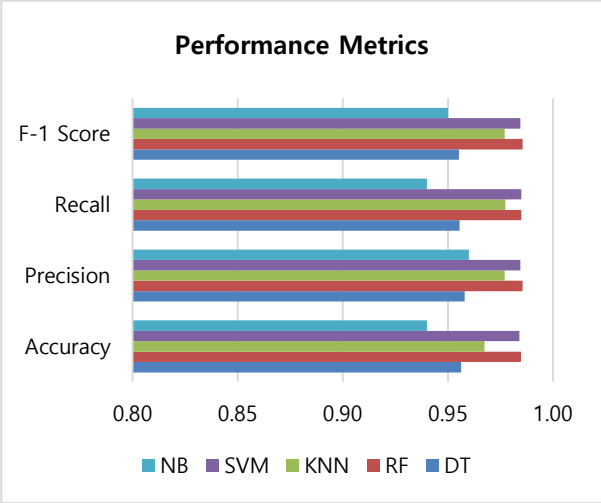

Fig. 7. Comparison Performance Metrics of ML Algorithms.

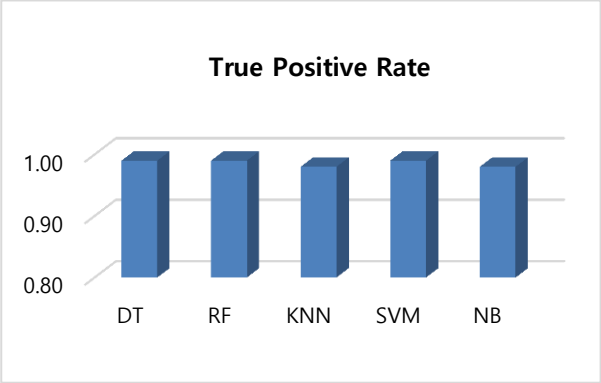

Fig. 8. Comparison of ML Algorithms using TPR.

C. Prediction Model for Malware of IoT Network Traffic

After training and testing ML model by Weka tools [31] Orange tools [32] was used for validating performance in terms of predicting malicious and benign in IoT network traffic. The results show RF algorithm is one the best accurate prediction mehods, this is due to the Classification Accuracy (CA) achieved 0.9714% while SVM algorithm obtained 0.7284% and we consider it obtained an inaccurate prediction, DT algorithm achieved 0.9141%, KNN obtained 0.9378% and NB obtained 0.8455%. As shown in Table III the number of samples for predicting malicious and benign in IoT network traffic. Fig. 9 shows a comparison of ML classifiers for predictive model accuracy.

TABLE III. VALIDATION A PREDICTION MODEL OF ML CLASSIFIERS

Classifier	Malicious	Benign
DT	12420	38854
RF	8754	42520
KNN	10733	40541
SVM	30753	20521
NB	21861	29413

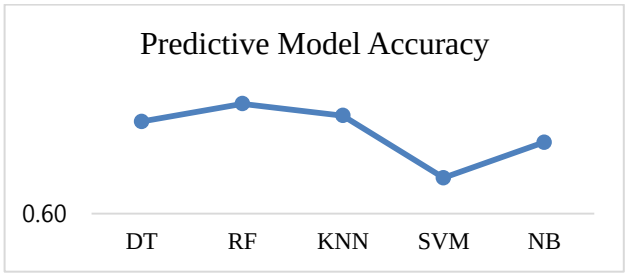

Fig. 9. ML Classifiers for Predictive Model Accuracy.

D. The Important Features for PCA Data Analysis

In Weka tools [31] Principal Component Analysis was used to reduce the features of the IoT-23 dataset [3]. We eliminate these least important features for feature selection but we do not lose original dataset completely. PCA method helps us to identify patterns in data of IoT-23 dataset based on the correlation among features. Furthermore, PCA method improved machine learning classification algorithms performance, removed correlated features and reduced overfitting by removing the unnecessary features in the IoT-23 dataset, which leads to minimizing features and maximizing the accuracy rate by using ML classification algorithms. PCA method considered the important features of IoT-23 dataset which are Ts, Uid, ID_orig.h, ID_orig.p, ID_resp.h, ID_resp.p, Proto, Service, Duration, Resp_bytes, Conn_state, Local_orig, Local_resp, Missed_bytes, History, Orig_pkts, Resp_pkts and Tunnel_parents. Fig. 10 shows a scatter plot after applied PCA method on IoT-23 dataset.

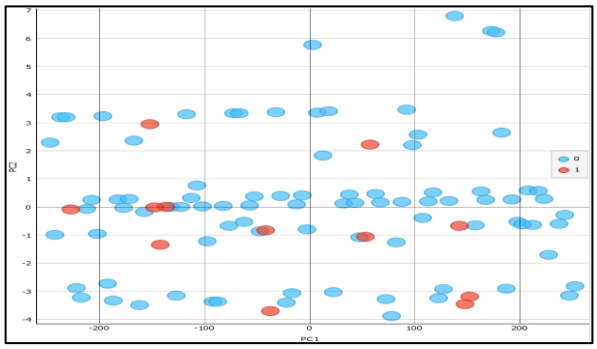

Fig. 10. Scatter Plot of Label IoT-23 Dataset.

E. Comparison with Previous Studies for Evaluating ML Models

The comparison of the machine learning algorithms with existing behavioural-based IoT-23 dataset is showed in Table IV. Gotsev et al. [2] used DT and RF classifiers and achieved 1.00% with four metrics. Nicolas Stoian [4] employed AdaBoost algorithm and it did not perform well as it achieved only 0.87% accuracy. Chunduri et al. [5] used RF and GBM algorithms and obtained highest accuracy rate, RF produced 0.9988% and GBM produced 0.9936%. Strecker et al. [14] obtained of RF classifier 0.9227% accuracy. Our model of ML produced the accuracy 0.9848% using RF classifier. In addition, model of ML achieved satisfied results for all ML classification algorithms with all metrics i.e. AUC, accuracy, precision, recall and f-1 score.

F. Confusion Matrix

After applying different ML classifiers, confusion matrix was applied i.e. True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN). The x-axis describes predicted label and the y-axis describes true label. The results show as per DT classifier, 12999 predicted a packet is benign and it actually is, 1712 predicted a packet is malicious and it actually is not, 667 predicted a packet is benign but it actually is not, 4 predicted a packet is malicious but it actually is, as well as the other ML classifiers i.e. RF, KNN, SVM and NB shown in Table V.

TABLE IV. COMPARISON WITH PREVIOUS STUDIES FOR EVALUATING ML MODEL

Study	Model	AUC	Accuracy	Precision	Recall	F-1 Score
[2] (Gotsev et al. 2021)	Naïve Bayes	-	0.58	0.75	0.58	0.51
	Support Vector Machine	-	0.74	0.70	0.74	0.70
	Logistic Regression	-	0.76	0.74	0.76	0.73
	Decision Tree	-	1.00	1.00	1.00	1.00
	Random Forest	-	1.00	1.00	1.00	1.00
[4] (Nicolas Stoian 2020)	Support Vector Machine	-	0.67	0.60	0.67	0.59
	Naïve Bayes	-	0.23	0.27	0.38	0.10
	Artificial Neural Network	-	0.66	0.71	0.66	0.52
	Random Forest	-	0.84	0.88	0.85	0.84
	Adaptive Boosting	-	0.87	0.86	0.87	0.83
[5] (Chunduri et al. 2021)	K-Nearest Neighbors	0.9568	0.9614	-	-	-
	Random Forest	0.9960	0.9988	-		
	Support Vector Machine	0.9400	0.9472	-	-	-
	Gradient Boosting Machine	0.9867	0.9936	-	-	-
[14] (Strecker et al. 2021)	K-Nearest Neighbors	0.8982	0.8990	0.8982	0.8971	0.9280
	Random Forest	0.9193	0.9227	0.9193	0.9330	0.9393
	Support Vector Machine	0.8352	0.8352	0.8352	0.8298	0.8559
Our Study	Decision Tree	0.9422	0.9567	0.9580	0.9556	0.9553
	Random Forest	0.9644	0.9848	0.9855	0.9850	0.9855
	K-Nearest Neighbors	0.9583	0.9674	0.9770	0.9773	0.9770
	Support Vector Machine	0.9628	0.9840	0.9845	0.9850	0.9845
	Naïve Bayes	0.9161	0.9479	0.9668	0.9480	0.9544

TABLE V. CONFUSION MATRIX FOR EACH ML CLASSIFIERS

DT Classifier		
True Label	Benign	Malicious
Benign	12999	4
Malicious	667	1712
	Predicted Label	

RF Classifier		
True Label	Benign	Malicious
Benign	12950	53
Malicious	180	2199
	Predicted Label	

KNN Classifier		
True Label	Benign	Malicious
Benign	12748	255
Malicious	246	2133
	Predicted Label	

SVM Classifier		
True Label	Benign	Malicious
Benign	12890	61
Malicious	210	2221
	Predicted Label	

NB Classifier		
True Label	Benign	Malicious
Benign	12871	155
Malicious	239	2117
	Predicted Label	

VII. Discussion of the Results

The IoT-23 dataset has approximately 160K rows and 21 features from 20 malware traffic captured from different IoT devices i.e. Echo device, Hue device and door lock device in which 11 malware labels and one benign label have existed in IoT network traffic. This study found the RF classifier to be the best performing; produced an AC 0.9714% and AUC achieved 0.9644%. For predicting malicious software over the IoT network traffic, all ML algorithms were predicting well except SVM algorithm this is due to AC produced was 0.7284%. DT algorithm predicted 12420 of malware which predict DDoS C&C, Mirai and Okiru; as well as the other ML algorithms have predicted malware i.e. RF, KNN, SVM and NB. Moreover, PCA method helps to improve ML performance and decrease overfitting by removing the unnecessary features in the IoT-23 dataset in order to improve accuracy rate for prediction. The IoT-23 dataset was split into 70% for training and 30% for testing and 10-fold cross validation, to make it suitable for a machine learning model. This study had two limitations. First, the types of malicious software in the IoT-23 dataset is limited. However, as discussed previously, four types of malicious software have been chosen on the IoT-23 dataset. However, we consider these types of malicious software selected have more significant effect in IoT environment. Second, after applied predictive modelling, malicious software cannot be prevented on the IoT devices. This is because the experiments were performed by machine learning and data mining techniques for predictive modeling without preventing tools for malicious software such as Intrusion Detection System (IDS).

VIII. Conclusion and Future Work

In this research, ML classification algorithms and data mining techniques were used for predictive modeling for validation of prediction of malicious and benign in IoT network traffic. Types of malware and IoT botnet used in this study for predicting are DDoS, C&C and various IoT botnet like Mirai and Okiru. The PCA method was applied to determine the important features of IoT-23 dataset and this study has been compared with previous studies that used the IoT-23 dataset in terms of accuracy rate and other metrics. We achieved better accuracy rate of ML classification i.e. KNN, SVM and NB. The highest accuracy rate for models of ML is RF classifier which produced 0.9844% and SVM classifier produced 0.9840%. For prediction model of malicious and benign in IoT network traffic, RF algorithm obtained the best accurate predictive model and achieved AC 0.9714% and predicted 8754 samples of various types of malware such as DDoS, C&C and various IoT botnet like Mirai and Okiru. In future work, the researchers will extensively understand the behavior of various types of malware attacks in IoT. Furthermore, will study these types of malware attacks against machine learning algorithms-based IDS. Also will investigate and evaluate IDS to prevent malicious software in network traffic.

References

[1] R. Saxena and E. Gayathri, "Cyber threat intelligence challenges: Leveraging blockchain intelligence with possible solution," Mater. Today Proc., p. S2214785321045752, Jul. 2021, doi: 10.1016/j.matpr.2021.06.204.

[2] L. Gotsev, M. Dimitrova, B. Jekov, E. Kovatcheva, and E. Shoikova, "A Cybersecurity Data Science Demonstrator: Machine Learning in IoT Network Security," p. 6, 2021.

[3] "Sebastian Garcia, Agustin Parmisano, & Maria Jose Erquiaga. (2020). IoT-23: A labeled dataset with malicious and benign IoT network traffic. doi: 10.5281/zenodo.4743746.

[4] N.-A. Stoian, "Machine Learning for Anomaly Detection in IoT networks: Malware analysis on the IoT-23 Data set," p. 10.

[5] H. Chunduri, T. Gireesh Kumar, and P. V. S. Charan, "A Multi Class Classification for Detection of IoT Botnet Malware," in Computing Science, Communication and Security, vol. 1416, N. Chaubey, S. Parikh, and K. Amin, Eds. Cham: Springer International Publishing, 2021, pp. 17–29. doi: 10.1007/978-3-030-76776-1_2.

[6] Zeek Network Security Monitor (2019). https://docs.zeek.org/en/current/intro/.

[7] M. S. Abdullah, A. Zainal, M. A. Maarof, and M. Nizam Kassim, "Cyber-Attack Features for Detecting Cyber Threat Incidents from Online News," in 2018 Cyber Resilience Conference (CRC), Putrajaya, Malaysia, Nov. 2018, pp. 1–4. doi: 10.1109/CR.2018.8626866.

[8] U. Noor, Z. Anwar, T. Amjad, and K.-K. R. Choo, "A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise," Future Gener. Comput. Syst., vol. 96, pp. 227–242, Jul. 2019, doi: 10.1016/j.future.2019.02.013.

[9] K. Alieyan, M. M. Kadhum, M. Anbar, S. U. Rehman, and N. K. A. Alajmi, "An overview of DDoS attacks based on DNS," in 2016 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Oct. 2016, pp. 276–280. doi: 10.1109/ICTC.2016.7763485.

[10] A. Alzahrani, A. Alshehri, R. Alharthi, H. Alshahrani, and H. Fu, "An Overview of Ransomware in the Windows Platform," in 2017 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, Dec. 2017, pp. 612–617. doi: 10.1109/CSCI.2017.106.

[11] M. Libicki, "The coming of cyber espionage norms," in 2017 9th International Conference on Cyber Conflict (CyCon), Tallinn, May 2017, pp. 1–17. doi: 10.23919/CYCON.2017.8240325.

[12] L. L. Dhirani, E. Armstrong, and T. Newe, "Industrial IoT, Cyber Threats, and Standards Landscape: Evaluation and Roadmap," Sensors, vol. 21, no. 11, p. 3901, Jun. 2021, doi: 10.3390/s21113901.

[13] A. P. Namanya, A. Cullen, I. U. Awan, and J. P. Disso, "The World of Malware: An Overview," in 2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud), Barcelona, Spain, Aug. 2018, pp. 420–427. doi: 10.1109/FiCloud.2018.00067.

[14] S. Strecker, R. Dave, N. Siddiqui, and N. Seliya, "A Modern Analysis of Aging Machine Learning Based IoT Cybersecurity Methods," J. Comput. Sci. Appl., vol. 9, no. 1, pp. 16–22, Oct. 2021, doi: 10.12691/jcsa-9-1-2.

[15] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection," ArXiv180209089 Cs, May 2018, Accessed: Mar. 02, 2022. [Online]. Available: http://arxiv.org/abs/1802.09089.

[16] Y. Meidan et al., "N-BaIoT—Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders," IEEE Pervasive Comput., vol. 17, no. 3, pp. 12–22, Jul. 2018, doi: 10.1109/MPRV.2018.03367731.

[17] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," J. Inf. Secur. Appl., vol. 50, p. 102419, Feb. 2020, doi: 10.1016/j.jisa.2019.102419.

[18] B. Yan and G. Han, "Effective Feature Extraction via Stacked Sparse Autoencoder to Improve Intrusion Detection System," IEEE Access, vol. 6, pp. 41238–41248, 2018, doi: 10.1109/ACCESS.2018.2858277.

[19] S. Potluri, S. Ahmed, and C. Diedrich, "Convolutional Neural Networks for Multi-class Intrusion Detection System," in Mining Intelligence and Knowledge Exploration, vol. 11308, A. Groza and R. Prasath, Eds. Cham: Springer International Publishing, 2018, pp. 225–238. doi: 10.1007/978-3-030-05918-7_20.

[20] G. De La Torre Parra, P. Rad, K.-K. R. Choo, and N. Beebe, "Detecting Internet of Things attacks using distributed deep learning," J. Netw.

Comput. Appl., vol. 163, p. 102662, Aug. 2020, doi: 10.1016/j.jnca.2020.102662.

[21] A. Pastor et al., "Detection of Encrypted Cryptomining Malware Connections With Machine and Deep Learning," IEEE Access, vol. 8, pp. 158036–158055, 2020, doi: 10.1109/ACCESS.2020.3019658.

[22] Y. Li, K. Xiong, T. Chin, and C. Hu, "A Machine Learning Framework for Domain Generation Algorithm-Based Malware Detection," *IEEE Access*, vol. 7, pp. 32765–32782, 2019, doi: 10.1109/ACCESS. 2019.2891588.

[23] I. H. Sarker, "CyberLearning: Effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks," Internet Things, vol. 14, p. 100393, Jun. 2021, doi: 10.1016/j.iot.2021 .100393.

[24] J. Li, L. Sun, Q. Yan, Z. Li, W. Srisa-an, and H. Ye, "Significant Permission Identification for Machine-Learning-Based Android Malware Detection," IEEE Trans. Ind. Inform., vol. 14, no. 7, pp. 3216–3225, Jul. 2018, doi: 10.1109/TII.2017.2789219.

[25] E. M. Karanja, S. Masupe, and M. G. Jeffrey, "Analysis of internet of things malware using image texture features and machine learning techniques," Internet Things, vol. 9, p. 100153, Mar. 2020, doi: 10.1016/j.iot.2019.100153.

[26] D. Li and Q. Li, "Adversarial Deep Ensemble: Evasion Attacks and Defenses for Malware Detection," IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 3886–3900, 2020, doi: 10.1109/TIFS.2020.3003571.

[27] M. M. Gaber et al., "Internet of Things and data mining: From applications to techniques and systems," WIREs Data Min. Knowl. Discov., vol. 9, no. 3, May 2019, doi: 10.1002/widm.1292.

[28] H. A. Madni, Z. Anwar, and M. A. Shah, "Data mining techniques and applications — A decade review," in 2017 23rd International Conference on Automation and Computing (ICAC), Huddersfield, United Kingdom, Sep. 2017, pp. 1–7. doi: 10.23919/IConAC .2017.8082090.

[29] F. Ali, D. Bhatt, T. Choudhury, and A. Thakral, "A Brief Analysis of Data Mining Techniques," in 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, Dec. 2019, pp. 752–758. doi: 10.1109/ICCIKE47802.2019.9004252.

[30] I. Batra, S. Verma, and K. Janjua, "Performance Analysis of Data Mining Techniques in IoT," in 2018 4th International Conference on Computing Sciences (ICCS), Jalandhar, Aug. 2018, pp. 194–199. doi: 10.1109/ICCS.2018.00039.

[31] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

[32] Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python, Journal of Machine Learning Research 14(Aug): 2349−2353.

[33] P. Bedi et al., "Detection of attacks in IoT sensors networks using machine learning algorithm," Microprocess. Microsyst., vol. 82, p. 103814, Apr. 2021, doi: 10.1016/j.micpro.2020.103814.

[34] A. Sivanathan, H. Habibi Gharakheili, and V. Sivaraman, "Managing IoT Cyber-Security Using Programmable Telemetry and Machine Learning," IEEE Trans. Netw. Serv. Manag., vol. 17, no. 1, pp. 60–74, Mar. 2020, doi: 10.1109/TNSM.2020.2971213.

[35] H. Naeem et al., "Malware detection in industrial internet of things based on hybrid image visualization and deep learning model," Ad Hoc Netw., vol. 105, p. 102154, Aug. 2020, doi: 10.1016/j.adhoc. 2020.102154.

[36] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust Intelligent Malware Detection Using Deep Learning," IEEE Access, vol. 7, pp. 46717–46738, 2019, doi: 10.1109/ACCESS.2019.2906934.