I.J. Information Engineering and Electronic Business, **2025, 3, 1-13**
Published Online on June 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijieeb.2025.03.01

Classifying IoT Device's Traffic Traces Using Network Traffic Characteristics

Rajarshi Roy Chowdhury*
American International University-Bangladesh, Department of Computer Science, Dhaka, 1229, Bangladesh
E-mail: rajarshi@aiub.edu
ORCID iD: https://orcid.org/ 0000-0001-9235-3687
*Corresponding Author

Debashish Roy
Humber College, Faculty of Applied Science and Technology, Toronto, Canada
E-mail: debashish.roy@humber.ca
ORCID iD: https://orcid.org/0000-0002-1174-3868

Pg Emeroylariffion Abas
Universiti Brunei Darussalam, Faculty of Integrated Technologies, Gadong, BE1410, Brunei Darussalam
E-mail: emeroylariffion.abas@ubd.edu.bn
ORCID iD: https://orcid.org/0000-0002-7006-3838

Received: 03 November, 2024; Revised: 07 January, 2025; Accepted: 13 February, 2025; Published: 08 June, 2025

Abstract: The escalating proliferation of devices, including both Internet of Things (IoT) and non-IoT devices, has triggered a suite of emergent security challenges in cyberspace, such as accurate device identification and authentication. The wide array of device types, protocols, and usability exacerbates these challenges. While conventional addressing schemes such as the logical Internet Protocol addressing and physical Media Access Control addressing schemes are integral for communication, they are susceptible to spoofing attacks. Device fingerprinting can be used to address the issue of identifying devices and traffic types using only implicit identifiers such as network traffic characteristics. In this paper, supervised machine learning based a device fingerprinting model has been proposed for the classification of both IoT and non-IoT devices on three levels based on their communication traffic characteristics. A meticulous feature selection process, employing two attribute evaluators, identified a subset of twenty features crucial for generating unique fingerprints from a large set of features pool. Three publicly available datasets and two supervised classifiers were utilized for evaluation purposes. Experimental results illustrated that the proposed model attained a classification accuracy exceeding 99% in discerning between known and unknown traffic traces (Level-1) on both the UNSW IoT and D-Link IoT datasets using the Random Forest (RF) classifier, and 99.74% accuracy in classifying network traffic types (Level-2) on the UNSW dataset. Individual device identification (Level-3) proves equally robust, with the RF and J48 classifiers achieving 99.03% and 98.14% accuracies on the UNSW non-IoT and IoT datasets, respectively. These findings underscore the potential of the device fingerprinting model in enhancing network security. The model's robust classification capabilities across various datasets and identification levels make it a valuable asset in tackling modern security challenges in networked environments.

Index Terms: Internet of Things, Device Classification, Machine Learning, Network Traffic Traces, Traffic Features Analysis

1. Introduction

The Internet of Things (IoT) represents a significant technological evolution following the advent of the Internet. Originating in the 1960s with the Advanced Research Projects Agency Network (ARPANET), the structure of the Internet underwent refinement through the National Science Foundation Network (NSFNET) in the 1980s [1]. Since then, technological advancements have led to the rapid growth of the Internet, culminating in the emergence of IoT. Introduced by the Auto-ID center in 1999, the concept of IoT envisioned the identification of individual physical objects in the real world through the use of a globally unique identifier, facilitated by Radio Frequency Identification (RFID) [2]. The primary objective of IoT is to bridge the physical and digital worlds, offering various services over the

network [2]. The proliferation of heterogeneous IoT devices with different functionalities have given rise to new privacy and security challenges in cyberspace, including the need for accurate device identification [3, 4, 5], authentication [6, 7, 8, 9, 10], location tracking [8, 9, 10], and anomaly detection [11, 12, 13]. IoT devices, such as temperature sensors, door sensors, IP cameras, and smoke detectors, are constrained by limited memory, storage, processing power, and energy [14, 15, 16]. In contrast, general-purpose computing devices like laptops, tablets, and desktops are more powerful and capable of performing a wider range of functions; these are known as non-IoT devices [15].

Conventionally, network-connected devices are identified using Internet Protocol (IP) or Media Access Control (MAC) addresses, serving as unique identifiers essential for network communication. However, these addressing schemes have been demonstrated to be susceptible to straightforward manipulation through the application of networking knowledge and wide-available software, including MAC address randomization, IP address and MAC address spoofing attacks [17]. Therefore, identifying network-connected devices accurately can be challenging when relying on explicit identifiers like IP and MAC addresses. However, precise device identification is crucial for enhancing network security. Conversely, Device Fingerprinting (DFP) [17], a process of identifying both IoT (smart-plug, IP Camera, and door sensor) [6], and non-IoT (Laptop, Computer, and Tablet) [18, 19] devices solely based on their communication traffic characteristics, is more robust to manipulation. The communication traffic characteristics used include network packets [6, 18, 20], radio signal [21], and MAC frame [3, 22], forming the basis for generating unique identifier for device identification purposes. DFP can be classified into two categories based on network traffic capturing processes: active fingerprinting and passive fingerprinting. Notably, DFP has demonstrated significant efficacy in enhancing network security, largely attributed to its strong resilience against vulnerabilities such as node forging and masquerading attacks.

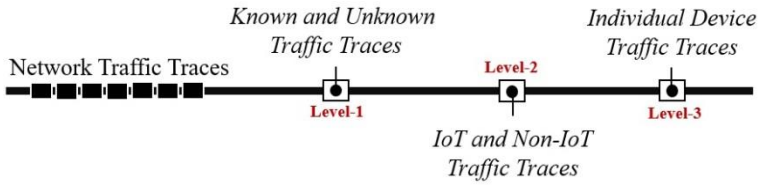

Fig. 1. Traffic traces classification levels.

An efficient DFP model based on Machine Learning (ML) has been proposed in this study, utilizing only network packet information, which has been captured passively for subsequent analysis. Network traffic traces (packets) can be captured using only low-cost hardware and software resources [34]. The proposed DFP model was designed to categorize network traffic traces on three levels, as illustrated in Figure 1. These levels encompass differentiation of known and unknown traffic traces (Level-1: known traffic traces refer to network traffic patterns or data that are well-understood and easily identifiable, whereas unknown traffic traces are those that are unfamiliar and not yet recognized), IoT and non-IoT devices traffic traces (Level-2), and individual device traffic traces (Level-3). A total of 82 packet header features are extracted from the IP network layer protocol, and both the Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) transport layer protocols. Two attribute evaluators were utilized to assess these features, with the top 20 ranked features selected as unique fingerprints. Subsequently, the proposed DFP model was evaluated on three publicly available datasets. Random Forest (RF) and J48 classifiers were employed for training and testing the model performances. Overall, the RF classifier demonstrated high classification performances, achieving 99.58% and 99.74% accuracies on the UNSW dataset for differentiating between known and unknown traffic traces (Level-1), and IoT and non-IoT devices traffic traces (Level-2), respectively. On Level-3 classification, the model attained over 99% and 97% accuracies in identifying individual devices network traffic traces of the individual devices on the non-IoT (UNSW) and IoT (UNSW) datasets, respectively. These findings highlight the significant potential of the DFP model to improve network security by offering precise and reliable device identification. The model's strong classification abilities across diverse datasets and identification scenarios underscore its value in addressing contemporary security challenges within smart home networked environments. All of these experiments were conducted using smart-home datasets. By leveraging its robust performance in different contexts, the proposed DFP model emerges as a crucial tool for enhancing security measures and safeguarding networks against emerging threats. Network administrators and cybersecurity professionals will benefit from this research through advanced tools that offer more accurate device identification and enhanced security measures. Moreover, organizations that implement smart-home technologies will experience improved protection against potential security threats.

While the proposed DFP model exhibited robustness in handling the three levels of classification tasks, it requires retraining upon enhance of a new device into the network. Importantly the study did not encompass tasks related to anomaly detection or the classification of malicious network traffic. The main contributions of this study are:

- Proposing a supervised ML-based DFP model that is not only able of differentiating between unknown and known network traffic traces but also distinguishing between traffic traces of non-IoT and IoT devices as well as classifying traffic traces of individual devices.

- Investigating individual network traffic features from the network and transport layers using two attribute evaluators: Info Gain and Gain Ratio, to identify a suitable subset of the top twenty features based on their significance for identifying network traffic traces.
- Evaluating the classification performance of the DFP model on three public datasets: the UNSW, IoT Sentinel, and D-Link IoT datasets, with the selected top twenty features.

The remainder of this paper is organized as follows: related works and existing DFP schemes are described in Section 2. Section 3 presents the proposed ML-based DFP model, datasets, and feature analysis process. Experimental outcomes from various datasets are provided and discussed in Section 4, and finally, Section 5 concludes the paper with the future direction of work.

2. Related Work

Researchers [6, 18, 23, 24, 31] have proposed various device fingerprinting models using ML or Deep Learning (DL) approaches by analyzing characteristics of network traffic. These models aim to classify IoT devices and distinguish between two different traffic types: IoT and non-IoT traffic. Unique fingerprints, serving as implicit identifiers, can be generated from different types of communication traffic, including network packets [18, 24, 25, 31], MAC frames [22, 26], and radio signals [27, 28, 29]. Many researchers have proposed DFP models that exclusively rely on network packet features due to their simplicity in capturing traffic traces compared to the use of radio signals. Table 1 provides a brief summary of the existing DFP models.

Miettinien et al. [6] proposed an automated DFP framework for identifying IoT devices by scrutinizing network traffic traces observed passively. Distinct device fingerprints were created by extracting 276-dimensional feature vectors from 12 consecutive network packets, encompassing information from the link, network, transport, and application layers protocols. These feature vectors were subsequently used to train individual ML models for each device type, enabling unique IoT device classification. The model attained an accuracy of 81.5% when considering the global ratio over a set of 27 smart IoT devices. Conversely, Aksoy and Gunes [25] proposed a DFP model, relying solely on a single TCP/IP packet header information for generating unique fingerprints. From a single packet, a total of 212 features were extracted, consisting of different layers protocols, including the network, transport, and application layers. The model attained an accuracy exceeding 82% on the IoT Sentinel dataset comprising 23 IoT devices.

Chowdhury et al. [24] introduced a DFP model by analyzing network traffic, with feature sets extracted from a single TCP/IP packet. The authors utilized different character-level metrics evaluators, including variability, stability, and suitability, for generating unique fingerprints, with the evaluation process involving the assessment a total of 218 features derived from packet header. Through the use of the UNSW and IoT Sentinel datasets, suitable subsets of 161 and 86 features, respectively, were identified. The model demonstrated commendable performances, achieving 83% accuracy on the IoT Sentinel dataset comprising 27 IoT devices and 97.78% accuracy on the UNSW dataset comprising 19 smart home IoT devices. Notably, the authors observed a decline in experimental performance when dealing with similar types of devices from the same manufacturers. The same researchers subsequently proposed an efficient DFP model by combining packet-level (20 features) and frame-level (4 features) features to improve classification performances [3]. The J48 classifier achieved an impressive accuracy of 99% on the D-Link IoT dataset, comprising 8 D-Link IoT devices.

Sivanathan et al. [19] introduced a supervised machine learning-based DFP model to classify IoT and non-IoT devices based on their communication traffic characteristics. From network packet flows (n number of packets) observed over a one-day period, 11 statistical features were derived, with the 11 features used to train and assess the model's performance. Their proposed DFP model achieved an accuracy exceeding 95% on the UNSW dataset, encompassing 21 smart-home IoT devices. Subsequently, the researchers enhanced the DFP model, utilizing only 8 statistical features derived from hourly traffic trace data [18]. This improved model attained over 99% accuracy on the UNSW dataset, which included 28 IoT devices. Hamad et al. [30] proposed a DFP model by analyzing a sequence of 20 to 21 for generating unique fingerprints. A comprehensive set of 67 statistical features, including packet size, header size, Inter Arrival Time (IAT), and TCP payload size, derived from both packet header and payload information, was utilized. The model exhibited notably high classification performance of the model was notably high, particularly when employing the random forest classifier with 100 estimators (decision trees), outperforming other machine learning classifiers. The DFP model demonstrated accuracy exceeding 90% when evaluated on the IoT Sentinel dataset containing 27 IoT devices [6].

Aneja et al. [32] proposed a DL based DFP model by utilizing 100 consecutive packets' IAT values, represented as plotting graphs, for generating fingerprints. These fingerprints were then used to train a Convolution Neural Network (CNN) model. The model demonstrated an accuracy of 86.7% on a private dataset with 2 devices. Another DFP model utilizing unique fingerprints in the form of a graph, generated by utilizing 1,000 consecutive network packets' IAT values has also been proposed [31]. The proposed model achieved an accuracy exceeding 97% on the GTID dataset comprising 58 devices [33]. The ResNet-50 architecture demonstrated superior classification performances when contrasted with the CNN-5L architecture, although the ResNet-50 CNN architecture required longer training and testing times.

Table 1. Existing ML/DL-based DFP approaches.

Source	Problem	No of Packets	Fingerprints	ML/DL	Dataset/Devices
[6]	Device Classification	12	Vector	ML-RF	*IoT Sentinel*^{27Devices}
[18]		1		ML-PART	*IoT Sentinel*^{23Devices}
[24]		1		ML-J48	*IoT Sentinel*^{27Devices} *UNSW*^{19Devices}
[3]		1			*D-Link IoT*^{8Devices}
[18]		$n^{Packets/Hour}$		ML-RF	*UNSW*^{28Devices}
[19]		$n^{Packets/Day}$			*UNSW*^{21Devices}
[30]		20-21			*IoT Sentinel*^{27Devices}
[31]		1000	Graph	DL-CNN	*GTID*^{58Devices}
[32]		100			*Private*^{2Devices}
[4]		100			*UNSW*^{21Devices}

Various device fingerprint models have been designed using either traditional ML algorithms [6, 24, 25] or DL algorithms [31, 32] for the different classification task, as summarized in Table 1. In these approaches, fingerprints were generated either using a large set network packets information or a single packet's information, leading to significant variations in classification performance. Some researchers have incorporated packet payload information in fingerprint generation, raising concerns about potential data privacy threats [30]. Consequently, a thorough analysis becomes essential to select an optimal subset of features from a vast pool of available features using a minimal number of network packets. In this study, network packets features have been analyzed deeply for generating distinctive fingerprints, aiming to enhance classification performance and streamline complexity. Network packet features have been focused on, due to the availability of network packets in both Local Area Network (LAN) and Wide Area Network (WAN), and the simplicity of capturing network traffic traces using affordable hardware devices [6, 34, 35].

3. Proposed Methodology

3.1. Network Traffic Traces

To assess the classification capabilities of the proposed device fingerprinting model, three publicly available datasets: UNSW [18], IoT Sentinel [6], and D-Link IoT [34] datasets, have been employed in this paper. Brief descriptions of these datasets are presented in Table 2.

Table 2. Publicly available datasets of both IoT and non-IoT devices.

Dataset	IoT	Non-IoT	Instances/Packets	Year	Source
IoT Sentinel	31	-	102,347	2017	[6]
UNSW (U-IoT)	21	-	1,276,657		[18]
UNSW (U-Non-IoT)	-	7	277,835		
D-Link IoT	12	-	333,864	2020	[34]

The UNSW (U-IoT), IoT Sentinel, and D-Link IoT datasets consist of communication traffic data 21, 31, and 12 different types of smart home IoT devices, including cameras, smart plugs, door sensors, printers, and home hubs IoT devices, respectively. On the other hand, the UNSW (U-Non-IoT) dataset consists of communication traffic data of only seven non-IoT devices, including laptops, tablets, and smartphones non-IoT devices, traffic traces. Miettinen et al. [6] recorded only set up communication traffic traces to obtain the IoT Sentinel dataset, whilst only normal/benign communication traffic traces over an extended period from IoT devices were recorded to obtain the D-Link IoT dataset [34]. Conversely, normal/benign communication traffic traces from both non-IoT and IoT devices were recorded to obtain the UNSW [18] dataset.

3.2. Machine Learning Algorithms

In this study, two traditional ML classifiers: J48 and Random Forest (RF), were utilized to assess the proposed DFP model on different datasets. These tree-based classifiers are chosen due to their simplicity and resource efficiency, whilst DL algorithms, including CNN and Long Short-Term Memory (LSTM) networks, require large datasets (for training) and high computing resources. Drawing insights from existing research works in the literature [6, 18, 24, 25, 37], it has been observed that tree-based ML algorithms consistently achieve higher classification accuracy compared to rule-based and Bayesian theorem-based algorithms.

J48 (C4.5): The J48 classifier, which is based on the C4.5 classifier within the Weka tool, represents an enhanced iteration of the Iterative Dichotomiser 3 (ID3) algorithm [38, 39]. A decision tree resembling a tree structure for classification is produced by the classifier, where leaves and branches represent classification rules (or decision rules) and features, respectively [40]. Decision rules formulated based on if-then conditions to classify input data from a given dataset. The J48 classifier employs concepts from information theory, such as entropy, and utilizes tree pruning techniques to construct a decision tree for solving classification tasks. Each attribute or feature with high information

gain is selected in this process [41]. Known for its simplicity, the J48 classifier has found applications in various domains. In reference [42], the J48 algorithm was ranked 1st among the top 10 ML algorithms, and consequently, it has been utilized in data mining for diverse objectives, including malicious traffic identification [11, 43], IoT device classification [24, 25], and medical science [44]. Additionally, the algorithm also allows the investigation of a significant subset of features for classification tasks.

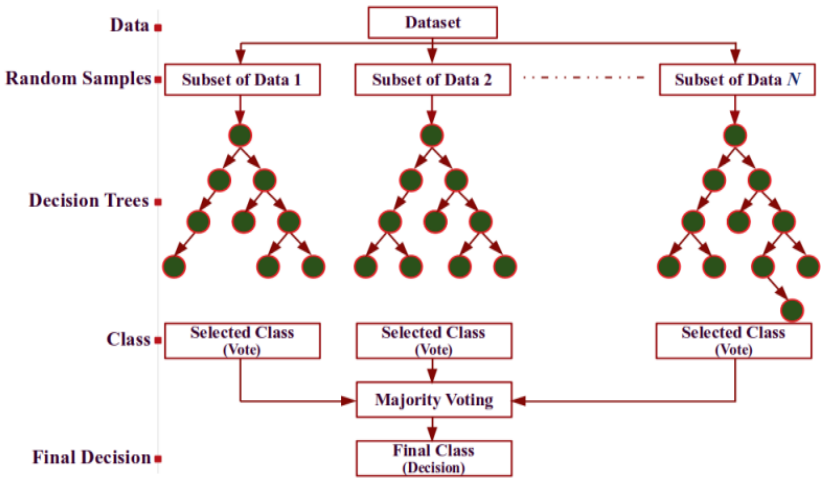

Fig. 2. Example of the Random Forest.

Random Forest (RF): RF is a classifier consisting of k number of decision tree-based classifiers (x, θ_n), n = 1, 2, 3...k. Each tree has a random vector (θ_n), and the input x is comprised of independent and identically distributed samples [45]. RF enhances classification accuracy by aggregating decisions from each tree, as presented in Figure 2. A voting mechanism is then used to predict the most popular class from the input data. The classifier incorporates bagging and boosting, contributing to improved accuracy. Bagging assists by using the random features selection process and estimating the generalization error, strength, and correlation of the ensemble of trees. RF stands out for its high execution speed and ability in reducing overfitting, making it an attractive choice for classification problems [46, 47]. Researchers have successfully utilized the classifier across diverse classification tasks, including device classification [18, 48, 49], anomaly detection [50], traffic classification [5, 51], disease classification [52], as well as IoT device identification [53, 54].

3.3. Proposed DFP Model

The outlined architecture for DFP is depicted in Figure 3. Both inbound and outbound network traffic traces originating from network-connected IoT and non-IoT devices, were captured on an Access Point (AP) for generating unique fingerprints. From existing works [5, 6, 18, 24, 25] it has been observed that device-originated communication traffic traces carry significant characteristics that can be used to classify devices as well as identify network traffic types. In this study, traffic traces originating from devices were filtered according to their individual device MAC addresses. Subsequently, they were stored in the Packet Capture (PCAP) file format for further processing.

TShark was utilized to extract a set of 82 features (or attributes), including *tcp.window_size*, *ip.proto*, *ip.len*, *udp.srcport*, and *tcp.dstport*, from each packet of different protocols: TCP, UDP, and IP, whilst each feature value consists of either numerical, null (or missing) or nominal (categorical data without any ordering or ranking) values. Subsequently, all the null or missing values were replaced with zero for ease of computation process. This information was then stored in a comma-separated values file format, to facilitate the analysis of network traffic characteristics. Individual packet information was annotated with labels corresponding to the 3 levels of identification: Level-1 for differentiating between known and unknown traffic types, Level-2 for differentiating between IoT and non-IoT traffic types, and Level-3 for differentiating traffic from individual devices. Subsequently, individual features were evaluated using two attribute evaluators: Gain Ratio (GR) and Info Gain (IG), to determine their significance. An average value was computed for each feature and these values were subsequently sorted in descending order to identify the top 20 features, which were made to serve as unique device fingerprints. Table 3 presents some of the significant features with GR, IG, and average values.

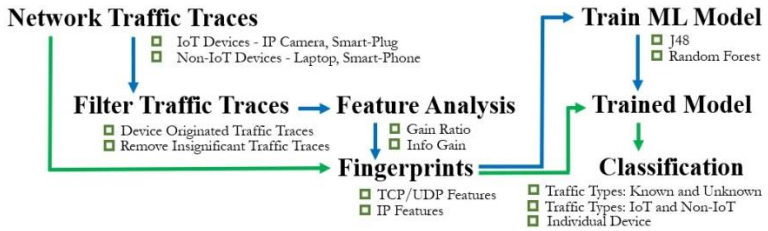

Fig. 3. The proposed device fingerprinting model.

Table 3. List of features with different attributes values.

Attributes/Features	GR	IG	Average	Sorted
ip.len	0.32214838	1.74606264	1.03410551	1
tcp.srcport	0.24463659	1.81798418	1.031310385	2
tcp.window_size	0.27744747	1.13436262	0.705905045	3
-	-	-	-	-
-	-	-	-	-
tcp.dstport	0.19025628	0.95853087	0.574393575	8
udp.checksum.status	0.05968236	0.00488872	0.03228554	24
-	-	-	-	-
tcp.checksum.status	0	0	0	39
tcp.analysis.acks frame	0	0	0	40

These fingerprints, representing the top 20 features, were then utilized for training a supervised ML model for the classification task, utilizing the labelled datasets during the training phase. The ML model underwent testing, with three levels of classification tasks performed to evaluate its performance: (1) classifications of traffic instances into one of two classes: known or unknown traffic traces (Level-1), (2) identification of network traffic instances into one of two classes: non-IoT or IoT traffic traces (Level-2), and (3) classifications of distinct devices traffic traces (Level-3).

3.4. Evaluation Metrics

The general objective of the different performance evaluation metrics is to assess the proposed DFP model classification performances on a test or unseen dataset. Two performance metrics were taken into account: accuracy and recall.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

These metrics were utilized to measure the effectiveness of the proposed DFP model employing various ML algorithms. The assessment metrics are computed through Equation 1 and Equation 2, respectively, where True Negative (*TN*) and True Positive (*TP*) represent the total count of correctly classified negative and positive instances (packets or traffic traces), respectively. Conversely, False Negative (*FN*) and False Positive (*FP*) indicate the total number of instances incorrectly classified as negative and positive, respectively.

4. Results and Discussion

4.1. Hardware and Software Details

The proposed ML-based DFP model, designed for the classification of devices (both IoT and non-IoT devices) and network traffic traces, was evaluated on a Windows-based system. This evaluation was conducted using the benchmark open-source Java-based data mining software, Waikato Environment for Knowledge Analysis (WEKA) [39]. WEKA, a cutting-edge state-of-the-art software, is versatile and applicable for various tasks, including data pre-processing, classification, regression, clustering, and visualization tasks. A laptop was equipped with the WEKA software, operating within the Windows Operating System (OS) environment.

4.2. Known and Unknown Traffic Types Classification

The effectiveness of the proposed device fingerprinting model was assessed in distinguishing between known and unknown traffic instances. This evaluation utilized two well established supervised ML classifiers: J48 and RF, and was conducted on two public datasets: the D-Link IoT and UNSW (U-IoT). Figure 4 depicts the Confusion Matrix (CM) representing the classification performance of the proposed model in differentiating between known and unknown traffic from the D-Link IoT and UNSW IoT datasets, referred to as Level-1. The model exhibited high classification

accuracy with both the J48 and RF classifiers. Specifically, the J48 achieved a maximum accuracy of 99.60% on the UNSW IoT dataset, whereas the RF classifier achieved its highest accuracy of 99.94% on the D-Link IoT dataset.

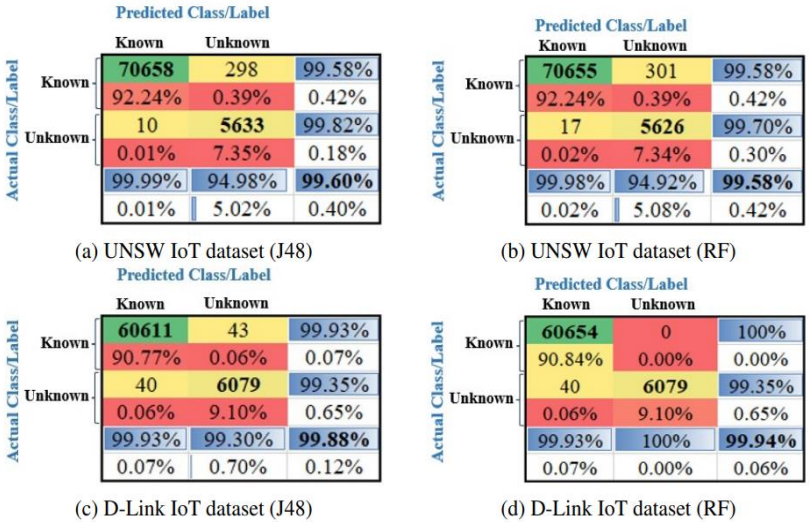

Fig. 4. Confusion Matrices of the proposed device fingerprinting (DFP) model in differentiating between Unknown and Known traffic traces: (a) UNSW IoT dataset (J48), (b) UNSW IoT dataset (RF), (c) D-Link IoT dataset (J48), and (d) D-Link IoT dataset (RF).

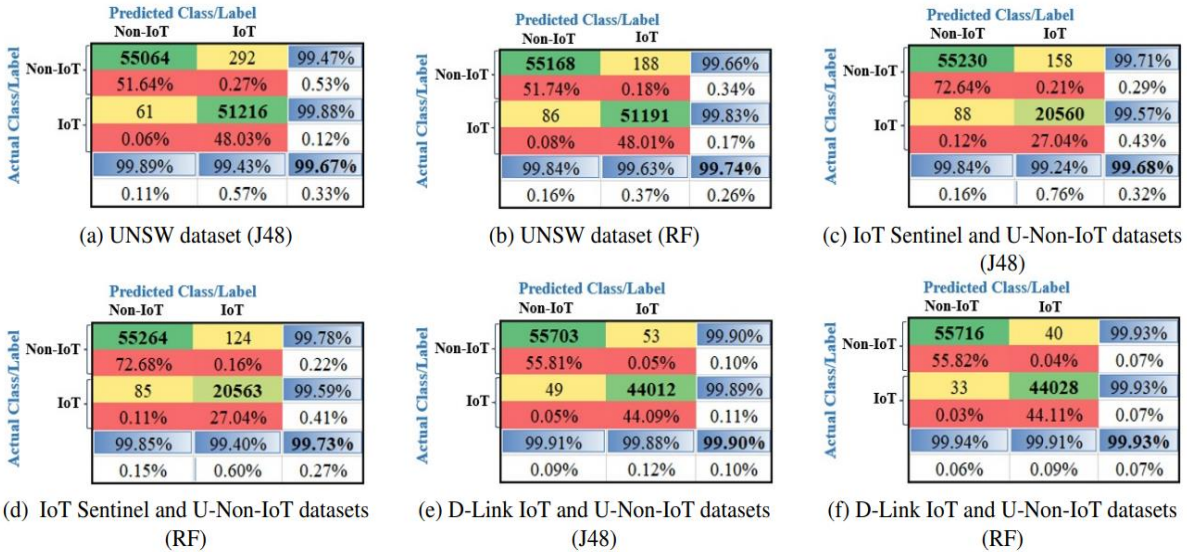

Fig. 5. Confusion Matrices of the proposed DFP model for distinguishing between IoT and non-IoT devices traffic traces: (a) UNSW dataset (J48), (b) UNSW dataset (RF), (c) IoT Sentinel and U-Non-IoT datasets (J48), (d) IoT Sentinel and U-Non-IoT datasets (RF), (e) D-Link IoT and U Non-IoT datasets (J48), and (f) D-Link IoT and U-Non-IoT datasets (RF).

Comparing between the distinct classifiers on the D-Link IoT dataset, both the J48 and RF classifiers correctly identified 6,079 out of 6,119 unknown instances as known traffic instances. However, the RF classifier managed to correctly identify all known instances, with the J48 classifier failing to identify 43 of the 60,654 known instances. On the UNSW IoT dataset, the J48 classifier performed slightly better, correctly identifying 70,658 out of 70,956 known instances and 5,633 out of 5,643 unknown instances, as compared to the RF classifier. Overall, the proposed DFP model demonstrated generally higher classification accuracy using the RF classifier, as depicted in Figures 4(b) and 4(c).

4.3. IoT and Non-IoT Devices Traffic Types Classification

In this study, three publicly accessible datasets: UNSW, IoT Sentinel, and D-Link IoT datasets, were utilized to assess the performance of the proposed DFP model in distinguishing between IoT and non-IoT traffic traces. The results, illustrated in Figure 5, indicate an impressive overall accuracy exceeding 99%. Notably, the RF classifier consistently outperforms the J48 classifier across all cases.

As shown in Figure 5(f) on the combined D-Link IoT and U-Non-IoT dataset, the RF classifier correctly identified 55,716 non-IoT instances and 44,028 IoT instances, while making only 40 false identifications of non-IoT as IoT and 33 false identifications of IoT as non-IoT. On the UNSW dataset, the RF classifier achieved the highest accuracy at 99.74% (binary classification), as depicted in Figure 5(b). In contrast, the J48 classifier exhibited a lower classification

performance of 99.67%; falsely identifying 292 non-IoT traffic instances as IoT and 61 IoT traffic instances as non-IoT. These findings underscore the superior classification capabilities of the RF classifier in distinguishing between IoT and non-IoT instances across diverse datasets.

4.4. Individual Device Classification Performances

The proposed model was assessed across three separate datasets: IoT Sentinel, and UNSW (both U-IoT and U-nonIoT), to evaluate its performance in classifying individual devices (IoT and non-IoT). Experimental findings demonstrate that the model consistently achieved higher accuracy using all the datasets, as presented in Figure 6. Specifically, the J48 classifier demonstrated accuracies of 91.55% and 98.14% using the IoT Sentinel and UNSW IoT datasets, respectively, whilst its performance exhibited a slight decline of 2.42% and 0.16% when utilizing the same datasets with the RF classifier.

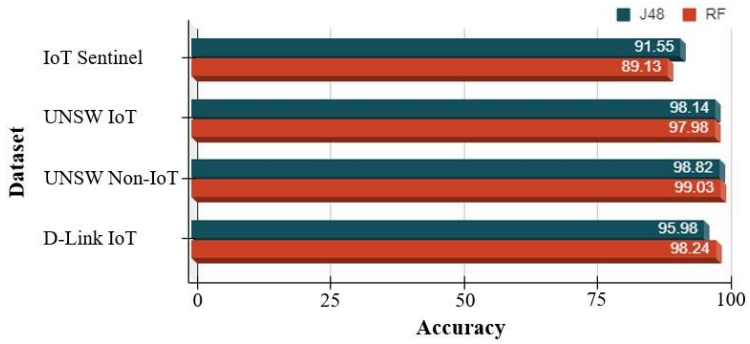

Fig. 6. Accuracy of the proposed model across varied datasets.

Conversely, the RF classifier achieved 99.03% and 98.24% accuracies using the UNSW non-IoT and D-Link IoT dataset, respectively, marginally surpassing the J48 classification performances. The classification performances on the IoT Sentinel and D-Link IoT datasets exhibited significant variation, despite the similarity in the characteristics of both datasets. This could be attributed to the similarity of these datasets, which include comparable types of IoT devices from the same manufacturer.

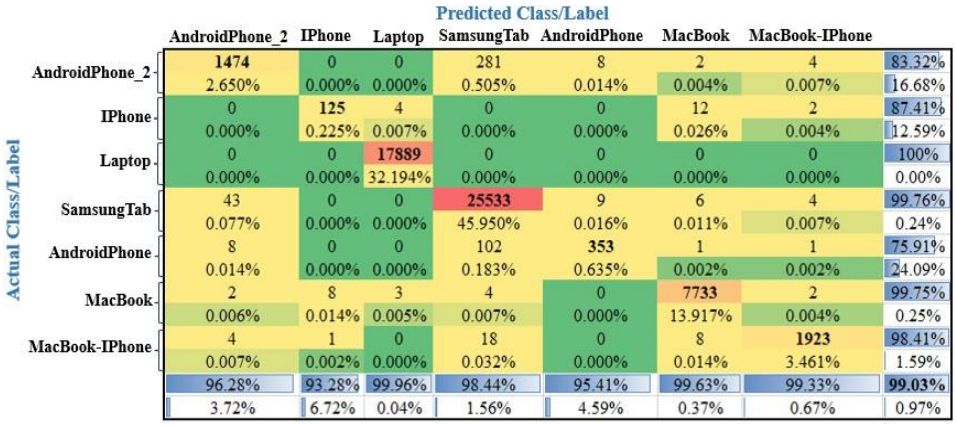

Fig. 7. Confusion Matrix of the proposed device fingerprinting model utilizing UNSW non-IoT dataset with the RF classifier.

From Figure 7 on the UNSW non-IoT dataset using RF classifier, it can be seen that a combined total of 537 traffic instances from the AndroidPhone_2, IPhone, SamsungTab, AndroidPhone, MacBook and MackBook-IPhone non-IoT devices were incorrectly classified as different devices. The model was able to correctly identify all of the 17,889 instances from the Laptop, despite incorrectly identifying 7 instances from other devices as coming from the Laptop. Notably, the Android devices, specifically AndroidPhone_2 and AndroidPhone, demonstrated lower classification performances compared to other devices. Overall, the RF classifier exhibited an accuracy of 99.03% on the UNSW non-IoT dataset, surpassing the J48 classifier's performance by 0.21%.

Figures 8 and 9 depict the confusion matrix of the proposed DFP model for Level-3 classification task on the UNSW IoT and D-Link IoT datasets, respectively. In Figure 8, it is evident that out of 21 devices, four devices: LiFXSmartBulb, NESTalarm, SmartThings and WithingsSmartscale, attained 100% recall, with 12 devices attaining recall levels surpassing 98% on the UNSW IoT dataset. BelkinWemoSwitch exhibited the lowest classification performance, with a recall of 83.61%. Overall, the proposed model achieved accuracy exceeding 98% using the J48 classifier, whereas the RF classifier demonstrated a slightly lower performance at 97.98%.

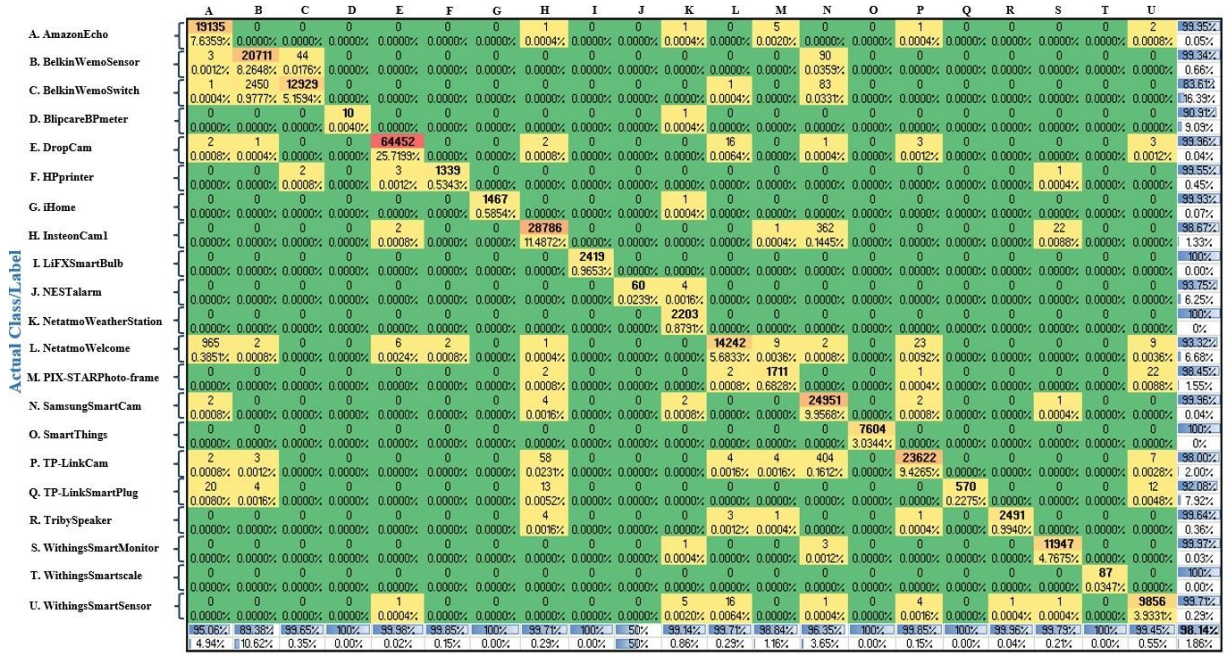

Fig. 8. Confusion Matrix of the proposed device fingerprinting model using UNSW IoT dataset with the J48 classifier.

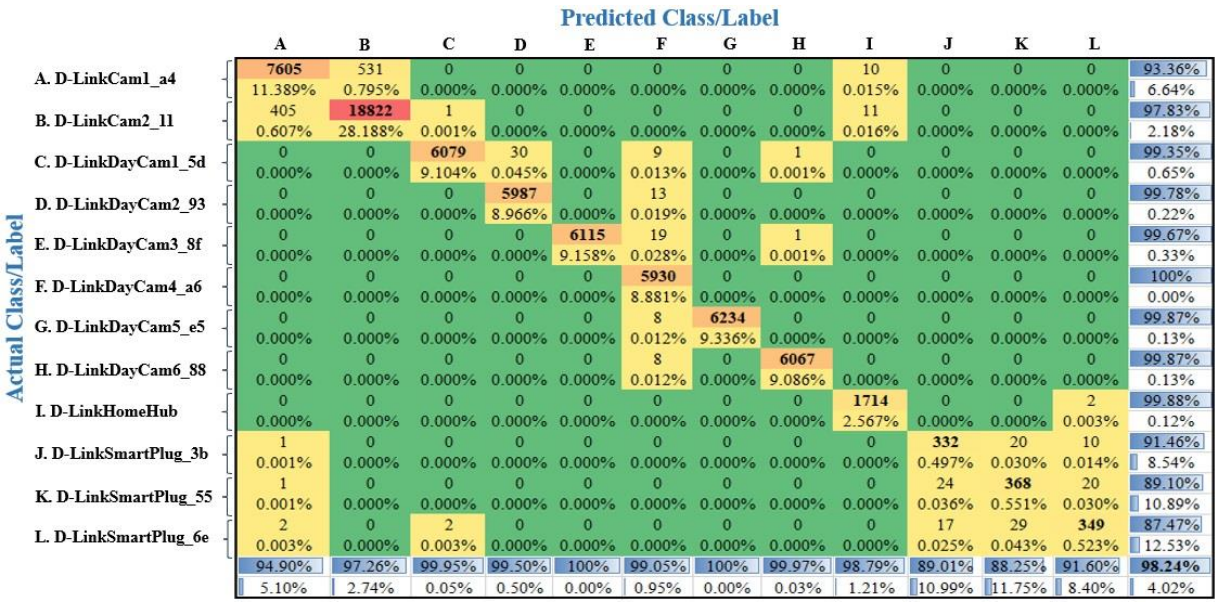

Fig. 9. Confusion Matrix of the proposed DFP model using D-Link IoT dataset with the RF classifier.

The proposed DFP model was also evaluated on the D-Link IoT dataset, aiming to classify individual IoT devices. All devices within this dataset are sourced from a singular manufacturer (such as D-Link), encompassing four unique device types. As illustrated in Figure 9, the model demonstrated individual recall of over 91% in most cases, with the exception of traffic traces coming from the two D-Link Smart Plug devices (D-LinkSmartPlug_3b and D-LinkSmartPlug_55), which exhibited recall of 89.10% and 87.47%, respectively. Overall, the RF classifier attained 98.24% accuracy on the D-Link IoT dataset, while the J48 classifier achieved a slightly lower accuracy of 95.98% accuracy. Throughout the experiment, a total of 1,175 instances (541 + 417 + 40 + 13 + 20 + 0 + 8 + 8 + 2 + 31 + 45 + 50) from the various devices, including DLinkCam1_4, D-LinkCam1_11, D-LinkDayCam1_5d, D-LinkDayCam2_11, D-LinkDayCam3_8f, D-LinkDayCam4_a6, D-LinkDayCam5_e5, D-LinkDayCam6_88, D-LinkHomeHub, D-LinkSmartPlug_3b, D-LinkSmartPlug_55 and D-LinkSmartPlug_6e, were incorrectly classified as different types of IoT devices. Meanwhile, 65,602 instances accurately classified as the correct types of device instances. Notably, instances from the D-LinkDayCam4_a6 devices achieved a perfect 100% classification accuracy, whilst the lowest performance at 87.47% recall, was observed for instances from the LinkSmartPlug_6e IoT device. The proposed device fingerprinting model demonstrated a 91.55% accuracy on the IoT Sentinel dataset, when employing the J48 classifier. Conversely, the RF classifier achieved a slightly lower accuracy of 89.13%. Notably, overall performances witnessed a decline, attributed to the presence of similar types of devices within the dataset.

Table 4 provides a comparative summary of various DFP approaches available in the literature, against the proposed DFP method for classifying network traffic instances. An 81.5% accuracy in identifying IoT devices was reported in reference [6], with a substantial number of features used for generating a unique fingerprint.

Table 4. Contrasting the proposed DFP scheme with existing DFP approaches.

Source	Problem	Dataset	Devices	Packet(s)	Feature	Algorithm	Performance
[6]	Device Classification	IoT Sentinel	27	12	23	RF	81.5%
[25]		IoT Sentinel	23	1	212	PART	82%
[24]		IoT Sentinel UNSW	27 19	1	161 86	J48	83.35% 97.78%
[3]		D-Link IoT	8	1 + 1 MF.	24	J48	99%
[18]		UNSW	28	$n^{Packets/Hour}$	8	RF	99%
[19]		UNSW	21	$n^{Packets/Day}$	11	RF	95%
[30]		IoT Sentinel	27	20-21	67	RF	90.3%
[31]		GTID	58	1000	1	CNN	97.5%
[32]		Private	2	100	1	CNN	86.7%
[4]		UNSW	21	100	2	CNN	98.49%
*	Device Cls[a,b,c,d] IoT vs non-IoT[e] Known vs Unknown[f]	IoT Sentinel[a] UNSW[c,e] D-Link IoT[b,e,f]	31[a] 21[c, U-IoT] 7[d,e,U-nIoT] 12[b,e,f]	1	20	J48[a,c] RF[b,d,e,f]	91.55%[a] 98.24%[b] 98.14%[c] 99.03%[d] 99.93%[e] 99.94%[f]

Note: Random Forest – RF, The proposed DFP model – *, IEEE 802.11 Mac Frame – MF, Classification - Cls, Convolutional Neural Networks – CNN.

Meanwhile, references [25] and [24] achieved accuracies of 82% and 83.35%, respectively, in classifying individual IoT devices within the IoT Sentinel dataset, by utilizing 212 [25] and 161 [24] features from a TCP/IP packet, respectively. In references [18] and [19], fingerprints were generated utilizing hourly and day-wise traffic analysis, respectively, using statistical features to identify individual devices. On the other hand, a network packet header, and a probe request frame (IEEE 82.11 MAC Frame) information were utilized in reference [3] for fingerprint generation, achieving 99% accuracy on the D-Link IoT dataset. In reference [30], the proposed DFP model utilized both packet header and payload information for fingerprint generation, despite data privacy threat from using payload information. References [31, 32] and [4], employed extensive packet information to generate fingerprints in the form of graphs for individual identification of network-connected devices, achieving accuracies of 97.5%, 86% and 98.49% on the GTID, private and UNSW IoT datasets, respectively. Comparatively, the proposed ML-based DFP model, which employs a meagre single packet header information, attained over 91%, 98.24%, 98.14% and 99.03% accuracy on the IoT Sentinel, D-Link IoT, UNSW IoT and UNSW non-IoT datasets, respectively.

5. Conclusion

The growing prevalence of diverse IoT devices with varying functionalities, combined with traditional non-IoT devices, has introduced complex security and privacy challenges. These issues include device classification, anomaly detection, and device management within IoT networks. Accurate device identification plays a crucial role in addressing and mitigating these challenges. This paper highlights the crucial importance of device identification and network traffic classification, based on inherent network traffic characteristics, to effectively address these challenges. The proposed DFP model offers three levels of classification tasks and employs state-of-the-art supervised ML classifiers, including RF and J48. The model was rigorously tested on three publicly available datasets: IoT Sentinel, D-Link IoT, and UNSW, which include both non-IoT and IoT datasets. Experimental results indicate that the proposed DFP model achieves an accuracy exceeding 99% in distinguishing between unknown and known traffic traces (Level-1) on both D-Link IoT and UNSW IoT datasets when utilizing the RF classifier. Additionally, it achieved an impressive 99.74% accuracy in classifying network traffic types (Level-2) on the UNSW dataset. For the identification of individual devices (Level-3), the RF and J48 classifiers demonstrated accuracies of 98.14% and 99.03%, respectively, on the UNSW IoT and UNSW non-IoT datasets. Notably, across the various experiments, the RF classifier consistently outperforms the J48 classifier, demonstrating higher overall accuracy. However, this study has some key limitations, including testing the model in a real-world deployment, handling a large dataset with different device types (both IoT and non-IoT) and retraining the model when a new device join the network.

Looking ahead, future research should delve into a more comprehensive analysis of network traffic features to further enhance classification performance across all three levels. Additionally, it is required to perform adversarial analysis for evaluating the model's resilience against spoofing and evasion attacks. A detailed analysis of advanced DL approaches, including transformers, Generative Adversarial Networks (GANs), and LSTM architectures, is also necessary to evaluate classification performance in real-world scenario.

Acknowledgment

The authors are profoundly grateful to the Department of Computer Science, American International University-Bangladesh, Bangladesh and the Faculty of Integrated Technologies, Universiti Brunei Darussalam, Brunei for supporting this research work.

References

[1] Glowniak j. History, structure, and function of the internet. Seminars in Nuclear Medicine. 1998; 28(2): 135-144. doi: 10.1016/S0001-2998(98)80003-2.
[2] Madakam S, Ramaswamy R, Tripathi S. Internet of Things (IoT): A Literature Review. Journal of Computer and Communications. 2015; 3(5): 1-10. doi: 10.4236/jcc.2015.35021.
[3] Chowdhury RR, Che A, Abas PE. Packet-level and IEEE 802.11 MAC frame-level analysis for IoT device identification. Turkish Journal of Electrical Engineering and Computer Sciences. 2022; 30(5): 1-18. doi: 10.55730/1300-0632.3915.
[4] Chowdhury RR, Che A, Abas PE. A Deep Learning Approach for Classifying Network Connected IoT Devices Using Communication Traffic Characteristics. Journal of Network and Systems Management. 2023; 31(1). doi: 10.1007/s10922-022-09716-x.
[5] Chowdhury RR, Che A, Abas PE. Internet of things: Digital footprints carry a device identity. AIP Conference Proceedings 2643. 2023; 40003. doi: 10.1063/5.0111335.
[6] Miettinen M, Marchal S, Hafeez I, Asokan N, Sadeghi AR et al. IoT SENTINEL: Automated Device-Type Identification for Security Enforcement in IoT. In: 2017 IEEE 37th International Conference on Distributed Computing Systems; Atlanta, GA, USA; 2017. pp. 2177-2184. doi: 10.1109/ICDCS.2017.283.
[7] Song Y, Huang Q, Yang J, Fan M, Hu A, Jiang Y. IoT device fingerprinting for relieving pressure in the access control. ACM International Conference Proceeding Series. 2019. pp. 1-8. doi: 10.1145/3321408.3326671.
[8] Jeong YS. An Efficient IoT Healthcare Service Management Model of Location Tracking Sensor. Journal of Digital Convergence. 2016; 14(3): 261–267. doi: 10.14400/jdc.2016.14.3.261.
[9] Ramnath S, Javali A, Narang B, Mishra P, Routray SK. IoT based localization and tracking. IEEE International Conference on IoT and its Applications (ICIOT 2017). 2017. doi: 10.1109/ICIOTA.2017.8073629.
[10] Aume C, Andrews K, Pal S, James A, Seth A, Mukhopadhyay S. TrackInk: An IoT-Enabled Real-Time Object Tracking System in Space. Sensors. 2022; 22(2): 1–15. doi: 10.3390/s22020608.
[11] Hasan M, Islam MM, Zarif MII, Hashem MMA. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things 2019; 7: 100059. doi: 10.1016/j.iot.2019.100059.
[12] Meidan Y, Bohadana M, Mathov Y, Mirsky Y, Shabtai A, Breitenbacher D, Elovici Y. N-BaIoT-Network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Computing. 2018; 17(3): 12–22. doi: 10.1109/MPRV.2018.03367731.
[13] Nascita A, Cerasuolo F, Monda DD, Garcia J, Montieri A, Pescape A. Machine and Deep Learning Approaches for IoT Attack Classification. IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2022; 1-6. doi: 10.1109/INFOCOMWKSHPS54753.2022.9797971.
[14] Garcia-Morchon O., Kumar S. Sethi M., Ericsson. Internet of Things (IoT) Security: State of the Art and Challenges (draft-irtf t2trg-iot-seccons-16). Datatracker. 2018; URL:https://datatracker.ietf.org/doc/draft-irtf-t2trg-iotseccons/16/.
[15] Chowdhury RR, Abas PE. A survey on device fingerprinting approach for resource-constraint IoT devices: Comparative study and research challenges. Internet of Things (Netherlands). 2022; 20(1). doi: 10.1016/j.iot.2022.100632.
[16] Saleh M., Jhanjhi N., Abdullah A., Saher R. Proposing Encryption Selection Model for IoT Devices Based on IoT Device Design. 23rd International Conference on Advanced Communication Technology. 2022; doi: 10.23919/ICACT53585.2022.9728914.
[17] Xu Q, Zheng R, Saad W, Han Z. Device Fingerprinting in Wireless Networks: Challenges and Opportunities. IEEE Communications Surveys & Tutorials 2015; 18(1): pp. 94-104. doi: 10.1109/COMST.2015.2476338.
[18] Sivanathan A, Gharakheili HH, Loi F, Radford A, Wijenayake C et al. Classifying IoT Devices in Smart Environments Using Network Traffic Characteristics. IEEE Transactions on Mobile Computing 2018; 18(8): 1745-1759. doi: 10.1109/TMC.2018.2866249.
[19] Sivanathan A, Sherratt D, Gharakheili HH, Radford A, Wijenayake C et al. Characterizing and classifying IoT traffic in smart cities and campuses. In: 2017 IEEE Conference on Computer Communications Workshops; Atlanta, GA, USA; 2017. pp. 559-564. doi: 10.1109/INFCOMW.2017.8116438.
[20] Chowdhury RR, Idris AC, Abas PE. Internet of Things Device Classification using Transport and Network Layers Communication Traffic Traces. International Journal of Computing and Digital Systems. 2022; 12(1): 2210–142. doi: 10.12785/ijcds/120144.
[21] Wang X, Zhang Y, Zhang H, Li Y, Wei X. Radio Frequency Signal Identification Using Transfer Learning Based on LSTM. Circuits Syst Signal Process. 2020; 39(11): 5514–5528. doi: 10.1007/s00034-020-01417-7.
[22] Robyns P, Bonne B, Quax P, Lamotte W. Noncooperative 802.11 MAC Layer Fingerprinting and Tracking of Mobile Devices. Security and Communication Networks 2017; 2017. doi: 10.1155/2017/6235484.
[23] Sivanathan A. IoT Behavioral Monitoring via Network Traffic Analysis. 2020. V. abs/2001.10632. URL: https://api.semanticscholar.org/CorpusID:210943061.
[24] Chowdhury RR, Aneja S, Aneja N, Abas PE. Network Traffic Analysis based IoT Device Identification. In: BDIOT 2020: Proceedings of the 2020 the 4th International Conference on Big Data and Internet of Things; Singapore; 2020. pp. 79–89. doi: 10.1145/3421537.3421545.

[25] Aksoy A, Gunes MH. Automated IoT Device Identification using Network Traffic. In: ICC 2019 - 2019 IEEE International Conference on Communications; Shanghai, China; 2019. pp. 1-7. doi: 10.1109/ICC.2019.8761559.
[26] Dalai AK, Jena SK. WDTF: A Technique for Wireless Device Type Fingerprinting. Wireless Personal Communications: An International Journal. 2017; 97(2): 1911–1928. doi: 10.1007/s11277-017-4652-y.
[27] Qing G, Wang H, Zhang T. Radio frequency fingerprinting identification for Zigbee via lightweight CNN. Physical Communication 2021; 44: 101250. doi: 10.1016/j.phycom.2020.101250.
[28] Chen L, Zhao C, Zheng Y, Wang Y. Radio Frequency Fingerprint Identification Based on Transfer Learning. 2021 IEEE/CIC International Conference on Communications in China (ICCC). 2021; 81–85. doi: 10.1109/ICCC52777.2021.9580203.
[29] Tian Q, Lin Y, Guo X, Wen J, Fang Y, Rodriguez J, Mumtaz S. New Security Mechanisms of High-Reliability IoT Communication Based on Radio Frequency Fingerprint. IEEE Internet Things Journal. 2019; 6(5): 7980–7987. doi: 10.1109/JIOT.2019.2913627.
[30] Hamad SA, Zhang WE, Sheng QZ, Nepal S. IoT Device Identification via Network-Flow Based Fingerprinting and Learning. In: 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering; Rotorua, New Zealand; 2019. pp. 103–111. doi: 10.1109/TrustCom/BigDataSE.2019.00023.
[31] Aneja S, Aneja N, Bhargava B, Chowdhury RR. Device fingerprinting using deep convolutional neural networks. International Journal of Communication Networks and Distributed Systems 2022; 28(2): 171-198. doi: 10.1504/IJCNDS.2022.121197.
[32] Aneja S, Aneja N, Islam MS. IoT Device fingerprint using deep learning. Proceedings - 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS 2018) 2019; 174–179. doi: 10.1109/IOTAIS.2018.8600824.
[33] Radhakrishnan SV, Uluagac AS, Beyah R. GTID: A Technique for Physical Device and Device Type Fingerprinting. IEEE Transactions on Dependable and Secure Computing 2015; 12(5): 519-532. doi: 10.1109/TDSC.2014.2369033.
[34] Chowdhury RR, Aneja S, Aneja N, Abas PE. Packet-level and IEEE 802.11 MAC frame-level network traffic traces data of the D-Link IoT devices. Data in Brief 2021; 37: 107208. doi: 10.1016/j.dib.2021.107208.
[35] Wu P, Lu Z, Zhou Q, Lei Z, Li X, Qiu M. Bigdata logs analysis based on seq2seq networks for cognitive Internet of Things. Future Generation Computer Systems. 2019; 90: 477–488. doi: 10.1016/j.future.2018.08.021.
[36] Hameed A, Violos J, Leivadeas A. A Deep Learning Approach for IoT Traffic Multi-Classification in a Smart-City Scenario. IEEE Access. 2022; 10:21193-21210. doi: 10.1109/ACCESS.2022.3153331.
[37] Suroso DJ, Rudianto A, Arifin M, Hawibowo S. Random forest and interpolation techniques for fingerprint-based indoor positioning system in un-ideal Environment. International Journal of Computing and Digital Systems. 2021; 10(1). doi: 10.12785/IJCDS/100166.
[38] Quinlan JR. C4.5: programs for machine learning. Publisher: Morgan Kaufmann; San Francisco, CA, United States; 1993.
[39] Witten IH, Frank E, Hall MA, Pal CJ. Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques. Publisher: Morgan Kaufmann; San Francisco, CA, United States; 2016.
[40] Bouzida Y, Cuppens F. Neural networks vs. decision trees for intrusion detection. IEEE International Conference. 2006; 2394–2400; URL:https://api.semanticscholar.org/CorpusID:17067528.
[41] Moustafa N, Turnbull B, Choo K. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. IEEE Internet Things Journal. 2019; 6(3) 4815–4830. doi: 10.1109/JIOT.2018.2871719.
[42] Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q et al. Top 10 algorithms in data mining. Knowledge and Information Systems 2008; 14(1): 1–37. doi: 10.1007/s10115-007-0114-2.
[43] Anthi E, Williams L, Slowinska M, Theodorakopoulos G, Burnap P. A Supervised Intrusion Detection System for Smart Home IoT Devices. IEEE Internet Things Journal. 2019; 6(5): 9042–9053. doi: 10.1109/JIOT.2019.2926365.
[44] Ortega J, Resureccion MR, Natividad LR, Bantug ET, Lagman AC, Lopez SR. An Analysis of Classification of Breast Cancer Dataset Using J48 Algorithm. International Journal of Advanced Trends in Computer Science and Engineering. 2020; 9(1.3): 475-480. doi:10.30534/ijatcse/2020/7591.32020.
[45] Breiman L. Random forests. Machine Learning. 2001; 45: 5–32. doi: 10.1023/A:1010933404324.
[46] Ho TK. Random decision forests. 3rd international conference on document analysis and recognition. 1995; 278–282. doi: 10.1109/ICDAR.1995.598994.
[47] Mishra AK, Ratha BK. Study of Random Tree and Random Forest Data Mining Algorithms for Microarray Data Analysis. International Journal on Advanced Electrical and Computer Engineering (IJAECE). 2016; 3(4): 5–7. URL: http://www.irdindia.in/journal ijaece/pdf/vol3 iss4/2.pdf
[48] Ganesan E, Hwang IS, Liem AT, Ab-Rahman MS. Sdn-enabled fiwi-iot smart environment network traffic classification using supervised ml models. Photonics. 2021; 8(6). doi: 10.3390/photonics8060201.
[49] Nguyen-An H, Silverston T, Yamazaki T, Miyoshi T. Entropy-based IoT devices identification. APNOMS 2020 - 2020 21st Asia-Pacific Network Operations and Management Symposium: Towards Service and Networking Intelligence for Humanity. 2020; 73–78. doi: 10.23919/APNOMS50412.2020.9236963.
[50] Yousefnezhad N, Madhikermi M, Framling K. MeDI: Measurement-based Device Identification Framework for Internet of Things. 16th International Conference on Industrial Informatics, INDIN 2018. 2018; 95–100. doi: 10.1109/INDIN.2018.8472080.
[51] Nguyen-An H, Silverston T, Yamazaki T, Miyoshi T. IoT Traffic: Modeling and Measurement Experiments. IoT. 2021; 2(1): 140–162. doi: 10.3390/iot2010008.
[52] Yang L, Wu H, Jin X, Zheng p, Hu S, Xu X, Yu W, Yan j. Study of cardiovascular disease prediction model based on random forest in eastern China. Scientific Reports. 2020; 10(1): 1–8. doi: 10.1038/s41598-020-62133-5.
[53] Chowdhury RR, Idris AC, Abas PE. Identifying SH-IoT devices from network traffic characteristics using random forest classifier. Wireless Networks. 2023; 1-15 doi: 10.1007/s11276-023-03478-3.
[54] Chowdhury RR, Idris AC, Abas PE. Device identification using optimized digital footprints. IAES International Journal of Artificial Intelligence. 2023; 12(1):232-240 doi: 10.11591/ijai.v12.i1.pp232-240.

Authors' Profiles

Rajarshi Roy Chowdhury, earned his Bachelor's degree in Computer Science from the University of Wollongong (UOW), Australia in 2011, followed by a Master's degree in Computer Science from Universiti Sains Malaysia (USM), Malaysia in 2012. He completed his Ph.D. in Systems Engineering in 2023 from the Faculty of Integrated Technologies (FIT), Universiti Brunei Darussalam (UBD), Brunei. Currently, he is working as an Assistant Professor in Department of Computer Science, American International University-Bangladesh (AIUB), Bangladesh. Prior to joining AIUB, he held academic positions at Sylhet International University (SIU) from November 1, 2012 to January 31, 2023 and at Metropolitan University (MU), Bangladesh, from February 1, 2023 to December 31, 2023. His research interests include networking, Internet of Things (IoT), data mining, and machine learning (ML).

Debashish Roy is a multifaceted expert at the intersection of Machine Learning (ML), Computer Network Security (CNS), and Education. With a distinguished academic background encompassing a Ph.D. and M.Sc. in ML and CNS, he possesses a comprehensive understanding of both the theoretical foundations and practical applications within these fields. Over the past two decades, he has honed his expertise not only through rigorous academic study but also through extensive teaching experience, guiding students through the complexities of these disciplines and fostering a passion for innovation. Beyond the classroom, he has demonstrated remarkable versatility and leadership in software development, serving as both a team leader and research project lead. Currently, he serves as a Professor at Humber College Institute of Technology & Advanced Learning, Toronto, ON, Canada. His research interests include network security, Internet of Things (IoT), data mining, and machine learning.

Pg Emeroylariffion Abas received his B.Eng. Information Systems Engineering from Imperial College, London in 2001, before obtaining his Ph.D. Communication Systems in 2005 from the same institution. He is now working as an Assistant Professor in System Engineering, Faculty of Integrated Technologies, Universiti Brunei Darussalam. His present research interests include data analytics, security of info-communication systems, communication and power, signal processing and design of photonic crystal fiber in fiber optics communication.

How to cite this paper: Rajarshi Roy Chowdhury, Debashish Roy, Pg Emeroylariffion Abas, "Classifying IoT Device's Traffic Traces Using Network Traffic Characteristics", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.17, No.3, pp. 1-13, 2025. DOI:10.5815/ijieeb.2025.03.01