Computers, Materials & Continua
DOI: 10.32604/cmc.2022.026581
Article

Tech Science Press

Detecting IoT Botnet in 5G Core Network Using Machine Learning

Ye-Eun Kim[1], Min-Gyu Kim[2] and Hwankuk Kim[2,*]

[1]Department of Electronics Information and System Engineering, Sangmyung University, Cheonan-si, Korea
[2]Department of Information Security Engineering, Sangmyung University, Cheonan-si, Korea
*Corresponding Author: Hwankuk Kim. Email: rinyfeel@smu.ac.kr
Received: 30 December 2021; Accepted: 02 March 2022

Abstract: As Internet of Things (IoT) devices with security issues are connected to 5G mobile networks, the importance of IoT Botnet detection research in mobile network environments is increasing. However, the existing research focused on AI-based IoT Botnet detection research in wired network environments. In addition, the existing research related to IoT Botnet detection in ML-based mobile network environments have been conducted up to 4G. Therefore, this paper conducts a study on ML-based IoT Botnet traffic detection in the 5G core network. The binary and multiclass classification was performed to compare simple normal/malicious detection and normal/three-type IoT Botnet malware detection. In both classification methods, the IoT Botnet detection performance using only 5GC's GTP-U packets decreased by at least 22.99% of accuracy compared to detection in wired network environment. In addition, by conducting a feature importance experiment, the importance of feature study for IoT Botnet detection considering 5GC network characteristics was confirmed. Since this paper analyzed IoT botnet traffic passing through the 5GC network using ML and presented detection results, think it will be meaningful as a reference for research to link AI-based security to the 5GC network.

Keywords: IoT botnet; 5G; B5G; malware; machine learning

1 Introduction

With the rapidly evolving network environment, the connection of IoT devices is rapidly changing from the individual wired internet (wi-fi) and 4G network connections to 5G networks. As the number of IoT devices connected to 5G mobile networks is increased, the importance of the security of IoT devices is also increasing. However, due to various security issues in IoT devices, IoT devices are exposed to security threats such as falsification, improper access, and information leakage [1,2]. Attackers exploit these IoT security vulnerabilities to create "IoT Botnet" of various malware types [3]. IoT Botnet infected with malware causes massive distributed denial-of-service (DDoS) attacks on networks connected to IoT [4,5]. Therefore, when the IoT device with weak security is connected to the 5G network, new security problem may occur in the 5G core network devices [6]. This leads to major

security concerns due to infringement of corporate or individual availability [7–9]. Therefore, studies on the IoT Botnet traffic analysis and detection in the field of mobile network are required [10,11].

5G has a service-based architecture (SBA) structure that segmented and virtualized into network function (NF) units. 5G end-to-end security framework of ITU-T SG17 includes SBA security in 5G core network security [12], and 3GPP is conducting research for the security of 5G SBA NFs [13]. With the development of AI technology, various studies are being conducted to native-AI in NFs. In particular, 5G has started to provide network designs suitable for AI services by introducing Network Data Analytics Function (NWDAF) to 5GC network to implement AI-based network automation and optimize related network functions [14,15]. Therefore, prior research that can be used to implement security functions using AI in core architecture of 5G and 6G networks is required.

Existing studies on the IoT Botnet detection have primarily focused on detecting IP packets of wired network occurred by IoT devices using machine learning (ML). However, the research on ML-based IoT Botnet detection methods and features in the 5G mobile network environment according to the evolution of the hyperconnected network environment is insufficient. Herein, a method to detect the IoT Botnet traffic based on ML in a 5G mobile network environment is studied, experimented results are explained, and future research is presented.

In this paper, the IoT Botnet detection was performed using packets collected from IoT devices and UPF in the 5G mobile network environment. The MedBIoT dataset [16] was used for the IP packet collected from the IoT devices. After replayed MedBIoT packets in the 5G mobile network simulated environment, the GTP-U packets collected by dumping directly from the UPF were used as the 5G mobile network packet. After this, the binary (benign/malicious) and multiclass (benign/bashlite/mirai/torii) were classified using the following ML algorithms: k-nearest neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF) and Stacking ensemble and the results were compared. The purpose of this study was as follows:

- ◯ Comparison of the detection performance in wired and 5G mobile network environments
- ◯ Realization of a model with the highest detection performance
- ◯ Establishing the necessity of study on the features of the IoT Botnet detection in 5GC
- ◯ Comparison of binary and multiclass classification results

The main contribution of this study can be summarized as follows: First, we conducted ML-based IoT botnet traffic detection study in the 5G core network paths. Second, binary and multiclass classification was performed to compare simple normal/malicious detection and normal/three type IoT Botnet malware detection. Previous studies focused on AI-based IoT botnet detection research in wired network environments, but this paper is meaningful as a reference for 5G security research by analyzing IoT Botnet traffic passing through the 5G core network and presenting detection results. Third, by comparing detection performance and feature importance in wired and 5G core network environments, the importance of feature research for IoT Botnet detection considering 5GC network characteristics was presented.

The subsequent sections of this paper are arranged as follows. Section 2 describes the 5G Network, GTP-U protocol, and previous works on IoT Botnet detection. Section 3 describes the Botnet detection methodology. Section 4 describes and compares the experimental results of the binary and multiclass classification for three experimental methods. Section 5 summarizes the study and presents future research directions.

2 Related Work

2.1 Overview of 5G Core Network and GTP Protocol

The structure of the 5G network consists of the 5G Radio Access Network (5G RAN) and 5GC, as shown in Fig. 1. The 5G RAN is composed of User Equipment (UE) and gNodeB (gNB) and provides a radio interface to the UE. The gNB provides the UE with the protocols for the Control and User Planes, General Packet Radio Service Tunneling Protocol (GTP). GTP is a tunneling protocol defined to deliver General Packet Radio Service (GPRS) within a mobile network. Mobile subscribers can access the Internet using UE while on the go by maintaining a connection to a Packet Data Network (PDN). GTP comprises consists of control plane (GTP-C), user plane (GTP-U), and charging (GTP' derived from GTP-C) traffic [17].

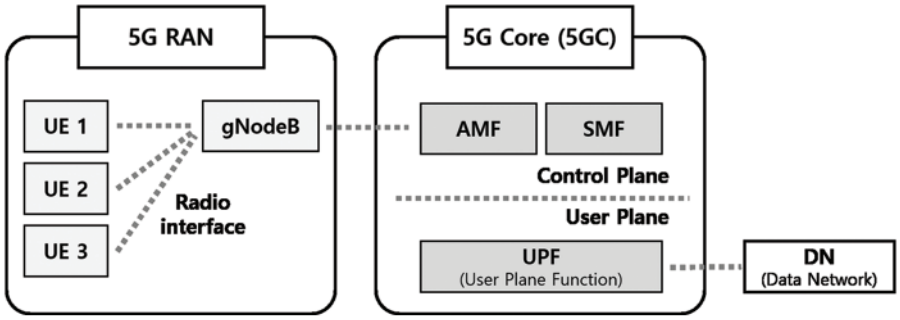


Figure 1: 5G network

5GC is composed of the Control and User Planes in charge of the control and data packet transmission, respectively. In the control plane, most mobility management is performed through Access and Mobility Management (AMF) and Session Management Function (SMF) in 5GC. AMF manages access and mobility in units of terminals through location service messages, and SMF allocates IP addresses to UEs and manages user plane services [18]. AMF and SMF provide services through SBA's single Service Based Interface (SBI). Each NF constituting SBA, network virtualization and software based network architecture, plays a role in controlling data transmission [19]. NWDAF, one of the 5G SBA configuration NFs, plays a key role as a functional entity that collects and analyzes various information on various network domains and uses them to provide analysis-based statistics and predictive insights to 5GC network functions such as Policy Control Function (PCF). ML algorithms linked NWDAF can utilize multiple network information collected through NWDAF to perform tasks such as mobility prediction and optimization, anomaly detection, prediction QoS, and data correlation [20]. The UPF of 5GC user plane is connected to the Data Network (DN) and handles the routing and forwarding of packets.

The uplink flow through user plane of the 5G network traffic is shown in Fig. 2. The IP packet generated by the UE is forwarded to the gNB and transmitted to the DN. When gNB receives an IP packet, it encapsulates the packet in the GTP-U header and is further encapsulated inside the IP and UDP headers before being sent to the UPF. The encapsulated IP packet and GTP-U header are G-PDU (Protocol Data Unit). GTP-U header is composed of Message type, Tunnel Endpoint Identifier (TEID), Next Extension Header type, and QoS Flow Identifier (QFI). The Message type is fixed to decimal 255 (0xff), indicating that G-PDU, and the Next Extension Header type are fixed to a binary number "1000 0101" (0x85) that was the PDU session container. The QFI identifies the QoS flow that

the transmitted packet belongs. The TEID is a unique identifier assigned to each GTP user connection on each node.

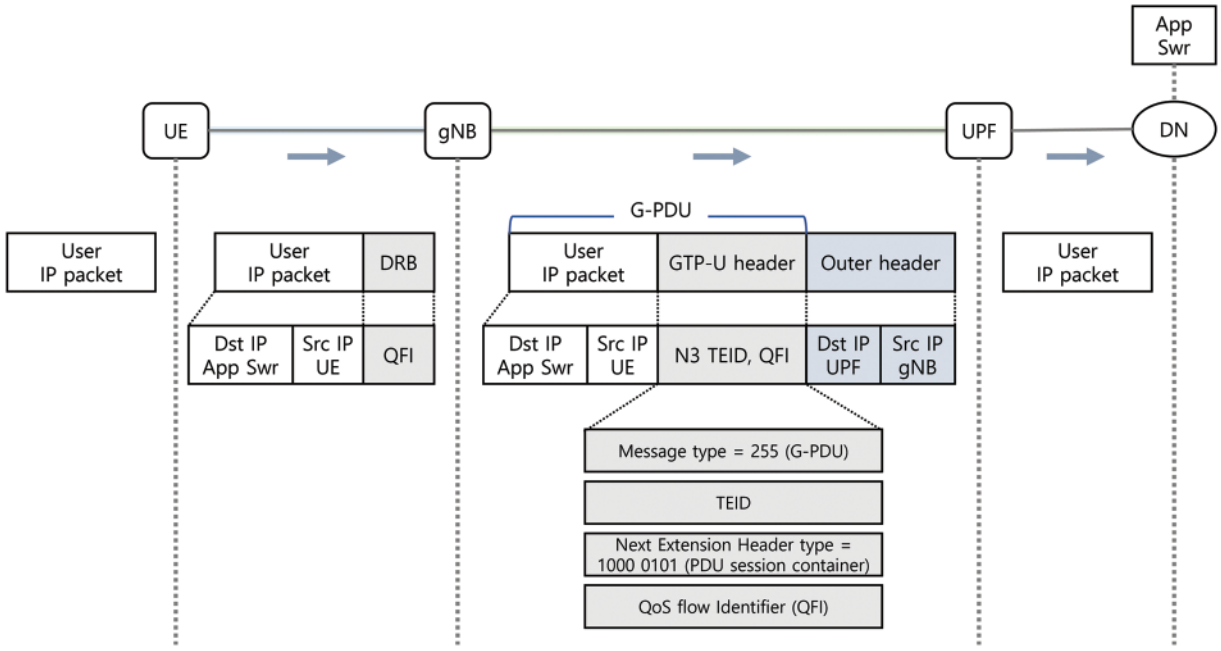


Figure 2: 5G network flow–Uplink [21]

2.2 Previous Work

Recently, the increased use of IoT devices in various fields has been accompanied by increased IoT device attacks. Accordingly, research on detecting abnormal behavior generated by IoT Botnets using machine learning is being actively conducted–Tab. 1.

Table 1: Summary of previous studies related to IoT Botnet detection with ML

Work	Techniques	Mobile Network	ML-based Detection
Stoian [22]	- Perform multi-class detection of IoT network anomalies with various ML algorithms (Random Forest, Naïve Bayes, ANN, SVM, AdaBoost).	X	O
Dutta et al. [23]	Proposes Stacking Ensemble (DNN, LSTM, Logistic Regression) app-roach method for classification performance improvement of network traffic outliers including IoT traffic.	X	O

(Continued)

Table 1: Continued

Work	Techniques	Mobile Network	ML-based Detection
Alzahrani et al. [24]	- Provides lightweight model (FastGRNN) applicable to IoT gateways able to multi-classify the IoT network traffic.	X	O
Soe et al. [25]	Proposes IoT Botnet detection framework with sequential architecture based on machine learning.	X	O
Hussain et al. [26]	- Proposes a universal feature set that performs well for detecting Botnet attacks regardless machine learning training dataset. - importance feature selection and feature frequency analysis	X	O
Saeedi [27]	- Study of IoT threats and vulnerabilities from a packet core perspective and proposes DDoS detection and mitigation method using ML (KNN, Decision Tree, Naïve Bayes, Logistic Regression) in the mobile core network	O (4G)	O
Salva-Garcia et al. [28]	- Proposes 5G-aware traffic filtering security framework to secure virtualized and multitenant 5G-based IoT traffic. - 5G IoT device packet analysis, Filtering Detection	O (5G)	X

Reference [22,23] proposed a model capable of improving IoT traffic detection performance in wired network environments using ML algorithms. [24,25] studied and proposed an ML-based IoT network anomaly detection model that can be applied to IoT gateways or IoT devices in wired networks. [26] performed a feature study to show good performance in detecting IoT traffic in wired networks regardless of the training dataset. As shown in [22–26], current studies related to IoT malicious traffic detection using ML are focused on detecting anomaly behavior in wired networks. In addition, since GTP data is not analyzed, there is a limit to applying it to the 5G core network.

As IoT devices connected to mobile networks increase, research to detect IoT abnormal behaviors occurring in mobile networks is also being gradually progressed [27,28]. [27] performed AI-based IoT traffic detection using GTP packets collected from Serving Gateway (SGW) of 4G mobile networks. Therefore, functions of 5GC were not considered in this study. [28] analyzed IoT anomaly traffic in the 5G mobile network environment and proposes a filtering-based IoT traffic security framework. As such, research on IoT malicious traffic in the 5G mobile network environment analyzes IoT traffic and even encapsulation-aware traffic filtering detection using it. Therefore, in this study, we determine

the difference in anomaly packet detection in wired networks and 5G mobile network environment using various machine learning algorithms.

3 Experimental Methods

This study consists largely of the Basic IoT Botnet Dataset and ML based IoT Botnet Detection steps. Basic IoT Botnet Dataset is associated with selecting and collecting the datasets used in the experiment. ML based IoT Botnet Detection is concerned with detecting IoT malicious packets using Machine Learning (ML) and performs Pre-Processing, Training, Binary Classification, and Model Evaluation. Each stage proceeds, as shown in Fig. 3, and is described in detail below.

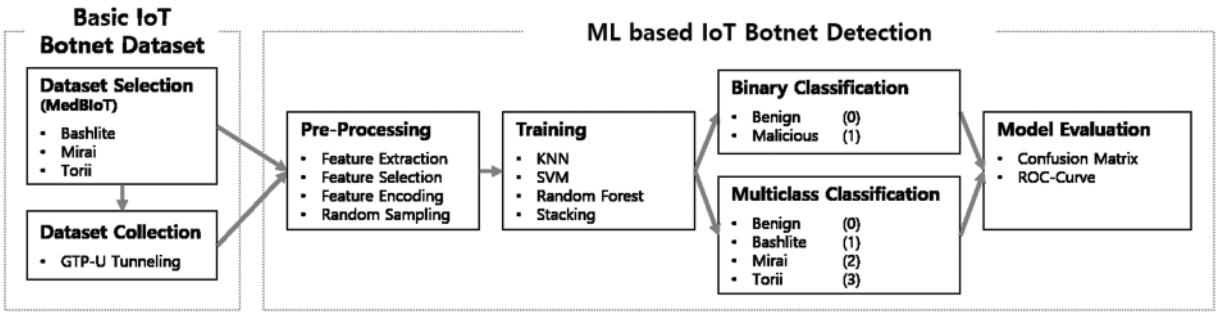


Figure 3: Methodology for IoT Botnet attacks detection

3.1 Basic IoT Botnet Dataset

3.1.1 Dataset Selection

This study was conducted using MedBIoT [16] created by the Center for Digital Forensics and Cyber Security at Tallinn University of Technology in 2020. [16] was performed to solve the lack of datasets for IoT Botnet detection. The MedBIoT dataset was collected by deploying real malware (bashlite, mirai, torii) to real and emulated IoT devices (ex. 83 devices) in a medium-sized network. [16] provides the dataset CSV containing the written extracted features and raw PCAP files. The MedBIoT dataset consists of about 4.3 and 12.54 million malicious and benign packets, respectively, as shown in Tab. 2.

Table 2: The number of packets in MedBIoT

Number of packets	Traffic type	Number of devices
4,143,276	Bashlite	40
842,674	Mirai	25
319,139	Torii	12
12,540,478	Normal	83

Therefore, packets are captured using various IoT devices, and MedBIoT, containing twice as much normal to malicious packets, is used as the dataset for this study. MedBIoT's raw PCAP files are used as IoT device IP packets, and GTP-U tunneled packets are collected by replaying the raw PCAP file packets in the 5G simulation environment. Then, using ML, each classification of IoT device IP and GTP-U tunneling packets is performed and then compared.

3.1.2 Dataset Collection GTP-U Tunneling Packets

The Dataset Collection stage takes the GTP-U tunneled packets from the UPF in a 5G mobile network environment. This process was performed in the Open5GCore Rel.6 environment–Fig. 4. This experiment performed GTP tunneling using only UE and gNB simulations, and UPF among the components of Open5GCore.

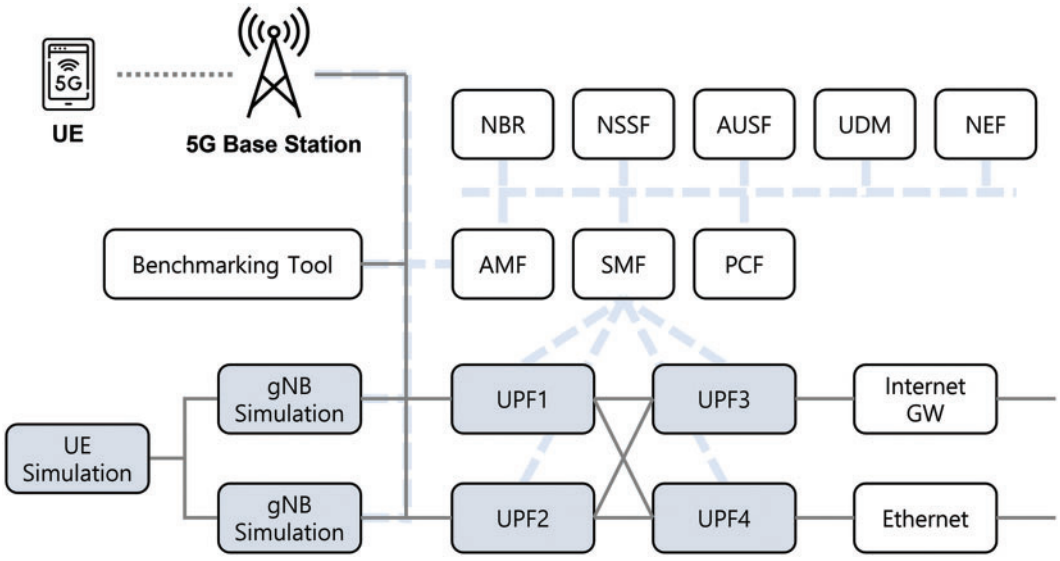


Figure 4: Open5GCore Rel.6–5G testbed [29]

First, the UE replayed the IP packets of the raw PCAP files of MedBIoT. The replayed packets pass the gNB into the GTP-U tunnel and delivered to the UPF. Then, GTP-U packets were collected by dumping packets using Wireshark in UPF. Only one fixed TEID appears because the packets are replayed at a specific UE and are captured at a specific UPF. Therefore, we changed TEID values to random TEID values according to the source and destination IPs/Ports of the IP packet of uplink or downlink.

3.2 ML Based IoT Botnet Detection Methodology

3.2.1 Pre-Processing

The pre-Processing of selected and collected datasets is performed using Feature Extraction and Selection, Feature Encoding, and Random Sampling. The data input for Pre-Processing are IP and GTP-U tunneling packets from MedBIoT and UPF, respectively. The features that have undergone Pre-Processing are converted into to CSV format and inputted into the ML algorithms.

Feature Extraction and Selection. MedBIoT dataset consist of original network packet as packet capture files (PCAPs). Therefore, we extracted header information from the IP and GTP-U tunneling packets of IoT devices and UPF using Tshark. Among them, the features per packet selected for input to the ML algorithm are shown in Tab. 3. And features per packet and "label features" for binary and multiclass classification are stored in CSV format.

Table 3: Feature selected per packet

Packet		Type of Feature	Raw Feature
IP Packet +GTP-U Tunneling Packet:	IP Packet (IoT Devices)	Packet Size	frame.len
		Ethernet Protocol	eth.src, eth.dst
		ARP Protocol	arp.src.proto ipv4, arp.dst.proto ipv4, arp.src.hw mac, arp.dst.hw mac, arp.opcode
		IP Protocol	ip.src, ip.dst, ipv6.src, ipv6.dst
		ICMP Protocol	icmp.type, icmp.code
		TCP Protocol	tcp.srcport, tcp.dstport
		UDP Protocol	udp.srcport, udp.dstport
	GTP-U Tunneling Packet (UPF)	Packet Size	frame.len
		IP Protocol	ip.src, ip.dst
		UDP Protocol	udp.srcport, udp.dstport
		GPRS Protocol	gtp.flags, gtp.message, gtp.length gtp.teid, gtp.ext_hdr.next

Feature Encoding. Encoding was performed on object-type features among 29 features (excluding the label feature) selected in Feature Selection. The IP address is converted to a decimal integer after removing ".". All object-type features (MAC address, GTP TEID, GTP extension next header) except for IP addresses were converted from hexadecimal strings to decimal integers. And the missing values were replaced with 0.

Random Sampling. We randomly sampled normal and malicious packets for experiments that change the normal/malicious packets count ratio. To save the state of each stage according to the count ratio, the state of the randomly sampled packets was stored, and the sampled packets were continuously added.

3.2.2 Training and Classification

Malicious packet detection was performed using four ML classification algorithms (KNN, SVM, RF, and Stacking Ensemble). The Regression Model of the Stacking Ensemble algorithm used KNN, SVM, and RF, and the Meta-Regressor used Logistic Regression. Each ML classification algorithm was implemented using python's Scikit-learn library.

The classification was performed in two ways: binary and multiclass. The binary classification classified packets into benign (0) and malicious (1). The multiclass classification classifies packets into benign (0) or three malicious packets: bashlite (1), mirai (2), and torii (3).

4 Experimental Results

This study performed the binary and multiclass classification using the IP packets of MedBIoT and GTP-U packets collected via the GTP-U tunneling of MedBIoT's IP packets. Then, the detection rates of both classification methods were compared. Each classification performed experiments and provided results under three experimental conditions in the following steps:

(1) Derived the optimal ratio of the number of benign packets to malicious packets for the experiment after comparing the difference in the detection performance with respect to the change in the ratio of the number of benign packets to malicious packets.
(2) Derived the optimal model and comparison of detection rates by each experimental method after performed the cross-validation using the optimal ratio of malicious to benign packets that derived in ①.
(3) Calculated the feature importance and confirmation of influence for each of three experimental methods (Fig. 5) in the experiment of ②.

② and ③ experiments were performed using three experimental methods as shown in Fig. 5 and the difference between the detection performance in IoT device and 5G mobile network UPF was presented. The three experimental methods presented in Fig. 5 are as follows:

a) Performed the classification using the IP packet of IoT device
b) Performed the classification using GTP-U packets collected from the UPF by replaying the IP packets of IoT device in 5G mobile network environment
c) Performed the classification using the IP packet of IoT device and the GTP-U packet of UPF together

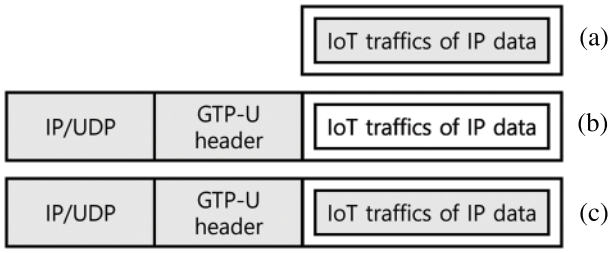


Figure 5: Packet structure in 5GC network experimental environment

4.1 Performance Evaluation Setup

This study was conducted in the following experimental environment: Windows 10 Platform. Intel Core i9-10980XE (3.0 GHz/24.75MB), DDR4 32GB PC4-25600 ∗ 4 (128GB), NVIDIA GeForce RTX 3090 24GB.

The IoT Botnet detection experiments proceeded with two classification methods (binary and multiclass classifications). Binary classification classified IoT Botnets into benign and malicious. Multiclass classification classified IoT Botnets into benign and three types of malware (bashlite, mirai, torii). Tab. 4 shows the detailed settings for each classification method. The total number of packets used in both classification methods is 40,000. In the experiments performed in Sections 4.2 and 4.3, a certain packet is selected and used among 40,000 packets. In each experiment, detection performance was evaluated using k-fold cross validation to prevent overfitting.

Table 4: Experiment setup by two classification methods

Classification type	Total input packets	Input packets count		Classification results	Using ML algorithm	Cross-validation
Binary classification	40,000	Benign Malicious	10,000 Bashlite:10,000 Mirai:10,000 Torii: 10,000	Benign (0) Malicious (1)	KNN SVM RF Stacking	cv = 7
Multiclass classification	40,000	Benign Bashlite Mirai Torii	10,000 10,000 10,000 10,000	Benign (0) Bashlite (1) Mirai(2) Torii(3)	KNN SVM, RF Stacking	cv = 7

The model evaluation for each experiment is evaluated by the ability to properly classify packets The possible results can be understood using the confusion matrix. In this study, accuracy, precision, recall, F1-score, receiver operating characteristic-area under the curve (ROC-AUC), and ROC-Curve were used for evaluation. The confusion matrix of multiclass classification was calculated by transforming the multiclass classification into the binary classification. Moreover, the ROC-AUC and ROC-Curve were obtained by macro-averaging after One-hot encoding of predictive values for each class in a One-*vs.*-One method. The detailed description of the evaluation index used for the experimental evaluation is shown in Tab. 5.

Table 5: Evaluation metrics description and calculation

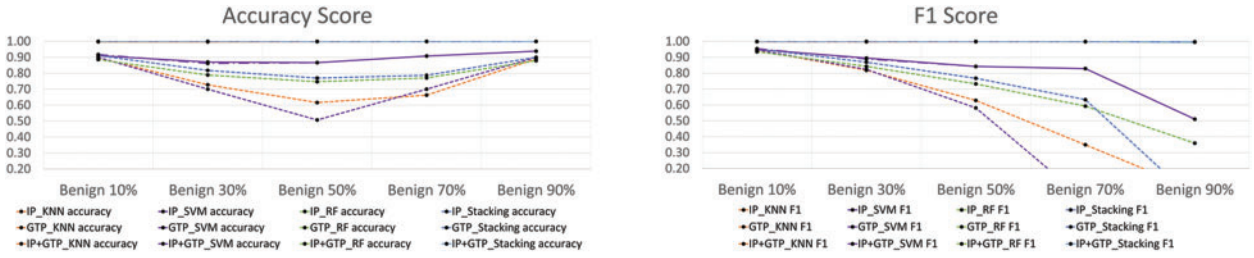
Evaluation metrics	Description	Calculation
Accuracy	The percentage of what the classification model correctly predicts the actual True as True and the actual False as False.	Accuracy=(TP+TN)/(TP+FP+FN+TN)
Precision	The percentage of what the actual true among those classified by predicting that the classification model is True.	Precision=TP/(TP+FP)
Recall	The percentage of what the classification model predicts as True out of what is actually True	Recall=TP/(TP+FN)
F1-score	Harmonic mean index of Recall and Precision	F1-score = 2∗(precision∗recall)/(precision+recall)
ROC-AUC	The area under the ROC-Curve. ROC-Curve: Plotting the True Positive Rate (TPR) against the False Positive Rate (FPR)	TPR=TP/(TP+FN) FPR=FP/(TN+FP)

4.2 Binary Classification

Here we classify the IP packet of IoT device and GTP-U packet of UPF into benign (0) and malicious (1) and compare the difference in detection performance by each experimental method in Fig. 5. First, the optimal ratio for the experiment was obtained by changing the ratio of the number of benign packets to malicious packets. Then, the performance detection rates for three experimental methods were compared using the optimal classification model score under the optimal ratio environment. Finally, confirmed the features having the high influence through a comparison of the feature importance in three experimental methods.

4.2.1 Experiments Related to Balanced Datasets (Benign vs. Malicious)

Experiment 4.2.1 was an experiment to find the optimal ratio of the number of benign to malicious packets that could compare the difference in detection performance through the binary classification of IP packets of IoT devices and GTP-U packets of the UPF. Thus, this experiment was performed by changing the ratio of the number of benign packets to malicious in a total of 40,000 randomly selected packets from 10% to 90%. The 3-malware benign and malicious packets provided by the experimented dataset MedBIoT were used at the same rate, and the binary classification was performed using the k-fold cross validation (cv = 7). The experimented results of the classification used the IP packet and GTP-U packet were shown in Fig. 6 and Tab. 6. As experimented result, when the ratio of the number of benign and malicious packets was 50%, it was confirmed that the score (ACC, F1 and AUC) had a median value not biased toward one side compared to when the ratio of the number of packets was 10% or 90%. Thus, we considered it as the optimal environment to compare the detection performance difference of binary classification less biasedly when the ratio of benign: malicious packets was 50%. The experimented result in 4.2.1 was used as the ratio of the number of benign packets to malicious in the subsequent 4.2.2 and 4.2.3 experiments.



(a) (b)

Figure 6: (Continued)

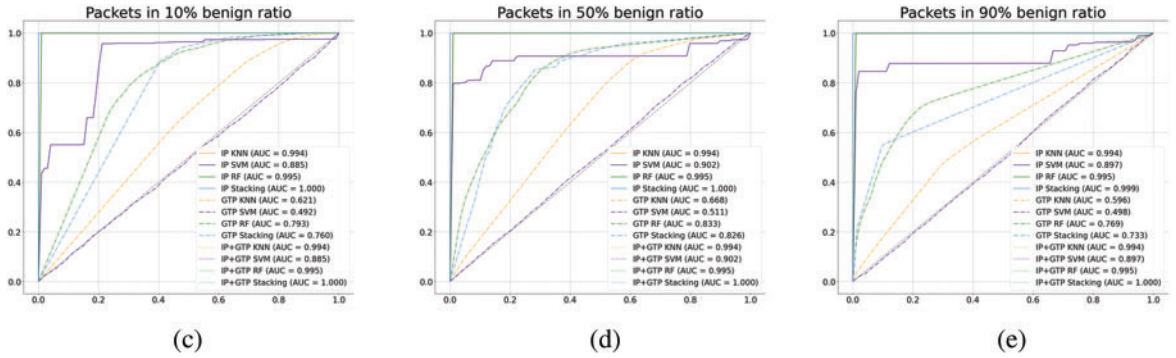


Figure 6: Benign classification experiment results changing benign packets count ratio–(a) Accuracy. (b) F1-score. (c) ROC-Curve that benign ratio 10%. (d) ROC-Curve that benign ratio 50%. (e) ROC-Curve that benign ratio 90%

Table 6: Binary classification score by benign packet count ratio–using IP and GTP-U packets

Experimental method	Algorithm	Score	Benign ratio				
			10%	30%	50%	70%	90%
IP Packet: (a)	KNN	Accuracy	99.715%	99.692%	99.822%	99.877%	99.937%
		F1	99.841%	99.780%	99.822%	99.795%	99.685%
	SVM	Accuracy	90.985%	87.010%	86.637%	90.760%	93.842%
		F1	94.980%	89.520%	84.218%	82.884%	51.073%
	Random forest	Accuracy	99.942%	99.955%	99.955%	99.945%	99.945%
		F1	99.968%	99.959%	99.955%	99.908%	99.723%
	Stacking ensemble	Accuracy	99.958%	99.967%	99.958%	99.908%	99.723%
		F1	99.977%	99.976%	99.958%	99.847%	99.54%
GTP-U Tunneling Packet: (b)	KNN	Accuracy	89.760%	72.737%	61.650%	66.272%	89.005%
		F1	94.512%	81.816%	62.829%	34.925%	6.6061%
	SVM	Accuracy	94.737%	82.353%	58.125%	0.0000%	0.0000%
		F1	90.000%	70.000%	50.695%	70.000%	90.000%
	Random forest	Accuracy	88.560%	78.857%	74.637%	77.047%	87.887%
		F1	93.512%	84.298%	73.236%	59.259%	35.924%
	Stacking ensemble	Accuracy	91.108%	81.775%	76.967%	78.708%	90.000%
		F1	95.272%	86.922%	76.761%	63.327%	0.0000%
IP+GTP-U Packets: (c)	KNN	Accuracy	99.772%	99.83%	99.847%	99.875%	99.922%
		F1	99.873%	99.878%	99.847%	99.791%	99.61%
	SVM	Accuracy	91.695%	86.137%	86.637%	90.76%	93.827%
		F1	95.394%	88.944%	84.218%	82.884%	51.02%
	Random forest	Accuracy	99.922%	99.927%	99.922%	99.915%	99.922%
		F1	99.957%	99.948%	99.922%	99.858%	99.609%
	Stacking ensemble	Accuracy	99.967%	99.9%	99.925%	99.942%	99.95%
		F1	99.981%	99.929%	99.93%	99.903%	99.749%

4.2.2 Experiments Related to Optional ML Algorithm

Experiment 4.2.2, confirmed the difference in detection performance for each algorithm in a low bias environment that the benign to malicious packets ratio was set as 50% derived experiment 4.2.1. Then, confirmed the optimal model for the binary classification, and compared the detection rates for each of the three experimental methods. These experiments were collected the benign and malware packets of three types of malware (bashlite, mirai, torii) at the same rate and performed by the k-fold cross-validation (cv = 7). The experimented results in 4.2.2 used the three experimental methods in Fig. 5 were shown in Fig. 7 and Tab. 7. In the experimental methods (a) and (b), the Stacking algorithm showed the highest accuracy (approximately 99.958% and 99.924% respectively), the F1-score (approximately 99.958% and 99.925% respectively), and the ROC-AUC (approximately 99.958% and 99.925% respectively). In the experimental method (c), Stacking showed the highest accuracy (approximately 76.967%) and the F1-score (approximately 76.761%), and showed the second highest ROC-AUC (approximately 76.967%). In the detection case used only GTP-U packets collected from the UPF of 5GC, it was reduced by approximately 22.99% compared to the detection used IP packets of the wired network. However, the IoT Botnet detection using both GTP-U packet and IP packet features showed similar performance with a difference of approximately 0.03% compared to the detection using IP packets. Through this, when performed the IoT Botnet detection by the binary classification in 5GC environment, it was confirmed that the same performance as that in a wired network can be obtained when the GTP-U packet feature and IP packet feature are used together.

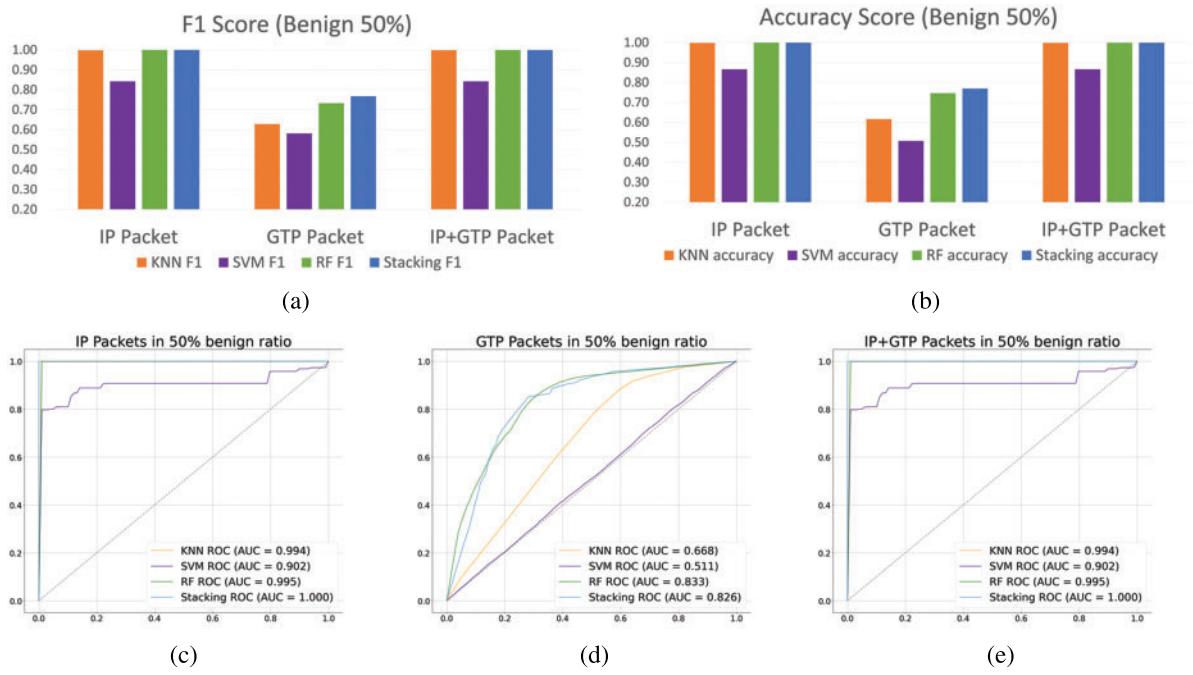


Figure 7: IoT botnet detection algorithm performance comparison by 3 experimental methods for binary classification–(a) Accurac. (b) F1-score. (c) ROC-Curve that using IP packets. (d) ROC-Curve that using GTP-U packets. (e) ROC-Curve of IP and GTP-U packets

Table 7: Binary classification Score by algorithm - benign to malicious packets count ratio is 50%

Experimental method	Score	Algorithm			
		KNN	SVM	RF	Stacking
IP Packet: Experimental method (a)	Accuracy	99.822%	86.637%	99.955%	99.958%
	F1	99.822%	84.218%	99.955%	99.958%
	ROC-AUC	99.388%	90.161%	99.495%	99.958%
GTP-U Packet: Experimental method (b)	Accuracy	61.650%	50.695%	74.637%	76.967%
	F1	62.829%	58.125%	73.236%	76.761%
	ROC-AUC	66.788%	51.120%	83.280%	76.967%
IP + GTP-U Packets: Experimental method (c)	Accuracy	99.847%	86.637%	99.922%	99.924%
	F1	99.847%	84.218%	99.922%	99.925%
	ROC-AUC	99.420%	90.166%	99.495%	99.925%

4.2.3 Experiments Related to Feature Importance in 5GC

In this section, the feature importance was calculated for the binary classification for the three experimental methods in Fig. 5 used the Stacking Ensemble and the optimal algorithm derived in 4.2.2, and the impact of IP packet features was confirmed–Tab. 8. Results showed that 3 of the 10 features used in experimental method (b) used only GTP-U packets had importance values. Among them, the TEID feature showed the highest importance at approximately 0.590397, it was related to the IPs and ports of the source and destination of the IP packet. Experimental methods (a) and (c) using IP packets of a wired network were used 18 and 28 features, respectively, and the features related to the source and destination addresses showed the high importance. This showed that the IP packet features had a large influence in the IoT Botnet detection through binary classification.

Table 8: Top 10 feature importance by 3 experimental methods for binary classification

Feature importance rank	IP Packet: (a)		GTP-U Packet: (b)		IP Packet + GTP-U Packet: (c)	
Number of features used	18		10		28	
Rank 1	eth.src	0.306732	gtp.teid	0.590397	eth.src	0.305114
Rank 2	ip.dst	0.203758	frame.len	0.299844	ip.dst	0.203945
Rank 3	eth.dst	0.197681	gtp.length	0.109760	eth.dst	0.194969
Rank 4	ip.src	0.163671	ip.src	0	ip.src	0.164341
Rank 5	frmae.len	0.068601	ip.dst	0	frame.len	0.063545
Rank 6	tcp.srcport	0.026958	udp.srcport	0	tcp.srcport	0.026928
Rank 7	tcp.dstport	0.012855	udp.dstport	0	tcp.dstport	0.014897
Rank 8	udp.srcport	0.012562	gtp.flags	0	udp.srcport	0.012045

(Continued)

Table 8: Continued

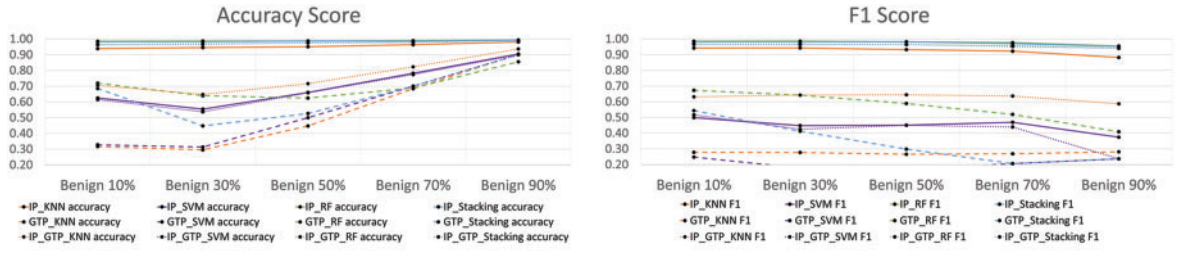
Feature importance rank	IP Packet: (a)		GTP-U Packet: (b)		IP Packet + GTP-U Packet: (c)	
Rank 9	arp.dst.proto_ipv4	0.003589	gtp.message	0	gtp.length	0.005374
Rank 10	udp.dstport	0.002011	gtp.ext_hdr.next	0	arp.dst.proto_ipv4	0.002834

4.3 Multiclass Classification

Section 4.3 is an experiment to multi-classify benign (0) and three types of malicious into bashlite (1), mirai (2) and torii (3) by using the IP packet of IoT device and the GTP-U packet of UPF, and to compare the difference in detection performance by each experimental method mentioned in Fig. 5. The same as experiment process with Section 4.2, the optimal ratio of the number of packets for the experiment was found, the detection rates for each of the three experimental methods through the optimal model were compared, and then the features with high influence were confirmed through the feature importance comparison.

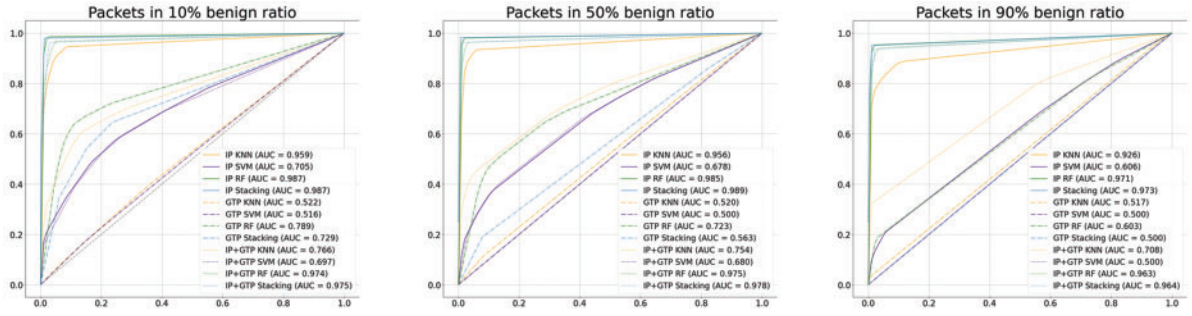
4.3.1 Experiments Related to Balanced Datasets (Benign vs. Malicious)

Experiment 4.3.1 is an experiment to find the optimal ratio of the number of benign to malicious packets which could compare the difference in detection performance through the multiclass classification of the IP packet of IoT device and the GTP-U packet of UPF. As same with experiment 4.2.1, this experiment was performed by changing the ratio of the number of benign packets to malicious for a total of 40,000 randomly selected packets from 10% to 90%. The benign and malicious packets of the 3-malware provided by the experimental dataset MedBIoT were used at the same rate, and the multiclass classification was performed using the k-fold cross validation (cv = 7). The experimented results of the classification used the IP packet and GTP-U packet were as shown in Fig. 8 and Tab. 9. As experimented results, it was confirmed that, as same in the binary classification, the score (ACC, F1, AUC) had an unbiased intermediate value when the ratio of the number of benign and malicious packets was 50%. Thus, we considered the ratio of benign: malicious packets as 50% as the optimal environment for less biased comparison of the detection performance difference of the multiclass classification. The experimented result in 4.3.1 was used as the ratio of the number of benign packets to malicious in the followed experiments in 4.3.2 and 4.3.3



(a) (b)

Figure 8: (Continued)



(c) (d) (e)

Figure 8: Multiclass classification experiment results changing benign packets count ratio–(a) Accuracy. (b) F1-score. (c) ROC-Curve that benign ratio 10%. (d) ROC-Curve that benign ratio 50%. (e) ROC-Curve that benign ratio 90%

Table 9: Multiclass classification score by benign packet count ratio–using IP and GTP-U packets

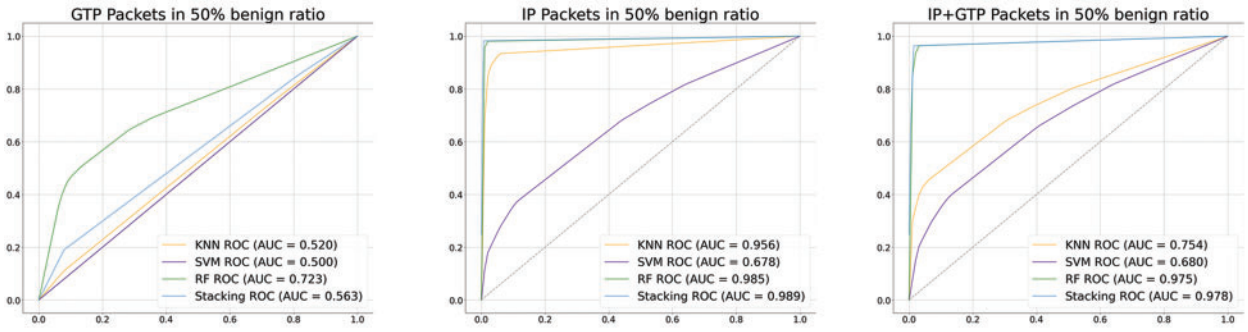
Experimental method	Algorithm	Score	Benign ratio				
			10%	30%	50%	70%	90%
IP Packet: (a)	KNN	Accuracy	93.900%	94.532%	95.020%	96.382%	98.072%
		F1	94.211%	94.280%	93.231%	92.232%	88.234%
	SVM	Accuracy	62.510%	55.440%	65.995%	78.210%	90.392%
		F1	49.899%	44.803%	45.095%	46.927%	37.414%
	Random	Accuracy	98.487%	98.617%	98.630%	98.930%	99.332%
	Forest	F1	98.528%	98.564%	98.131%	97.643%	95.488%
	Stacking	Accuracy	98.258%	98.316%	98.716%	98.691%	99.325%
	ensemble	F1	98.247%	98.248%	98.231%	97.044%	95.388%
GTP-U Tunneling	KNN	Accuracy	31.684%	29.534%	44.722%	68.422%	90.254%
		F1	27.780%	27.691%	26.556%	26.872%	28.123%
Packet: (b)	SVM	Accuracy	32.754%	31.332%	50.000%	70.000%	90.000%
		F1	24.758%	17.738%	16.666%	20.588%	23.684%
	Random	Accuracy	71.880%	63.955%	62.477%	68.847%	85.527%
	forest	F1	67.293%	64.176%	58.864%	51.960%	40.897%
	Stacking	Accuracy	68.533%	44.758%	52.650%	70.000%	89.975%
	ensemble	F1	54.329%	41.290%	29.832%	20.588%	23.685%
IP+GTP-U	KNN	Accuracy	70.518%	64.758%	71.61%	82.285%	93.64%
Packets: (c)		F1	63.141%	64.27%	64.431%	63.658%	58.695%
	SVM	Accuracy	61.487%	53.755%	62.830%	77.398%	90%
		F1	51.682%	42.529%	44.976%	43.918%	23.684%
	Random	Accuracy	96.478%	96.858%	97.458%	98.11%	99.138%
	Forest	F1	96.735%	96.692%	96.418%	95.686%	94.42%
	Stacking	Accuracy	96.241%	96.6%	97.5%	97.925%	99.1%
	Ensemble	F1	96.51%	96.426%	96.397%	95.22%	93.993%

4.3.2 Experiments Related to Optimal ML Algorithm

In experiment 4.3.2, the benign: malicious packets count ratio was set as 50% and confirmed the difference in detection performance for each algorithm in an environment with less bias. Through this, the optimal model for multiclass classification was confirmed, and the detection rates were compared for each of the three experimental methods mentioned in Fig. 5. As same in 4.2.2 experiments, these experiments were collected the benign and malware packets of three malware (bashlite, mirai, and torii) at the same rate and performed by the k-fold cross validation (cv = 7). The experimented results of 4.3.2 used the three experimental methods mentioned in Fig. 5 were as shown in Fig. 9 and Tab. 10. In the experimental method (a), the Stacking showed the highest accuracy (approximately 98.716%), F1-score (approximately 98.231%), and ROC-AUC (approximately 98.904%). In the experimental method (b), the RF showed the highest accuracy (approximately 62.477%), F1-score (approximately 58.864%), and ROC-AUC (approximately 72.265%). In the experimental method (c), Stacking showed the highest accuracy (approximately 97.5%) and the ROC-AUC (approximately 97.785%), and showed the second highest F1-score (approximately 96.396%). This shows that the multiclass classification using IP packets has great performance when using the Stacking algorithm, and the multiclass classification using only GTP packets has great performance when using the RF algorithm. When compared the accuracy of the best algorithm for each experimental method, in the detection case used only GTP-U packets collected from the UPF of 5GC, the accuracy was decreased by approximately 36.71% compared to the detection used the IP packets of wired network. However, the IoT Botnet detection used both GTP-U packet and IP packet features showed similar performance to the multiclass classification used the IP packet with an accuracy reduction of approximately 1.231%. Through this, when performed the multiclass classification on the benign and individual IoT Botnet malware in 5GC environment, it was confirmed that the performance similar to that in a wired network could be obtained when the GTP-U packet and IP packet features were used together as in the binary classification.

(a) (b)

Figure 9: (Continued)



(c) (d) (e)

Figure 9: IoT Botnet detection algorithm performance comparison by 3 experimental methods for multiclass classification–(a) Accuracy. (b) F1-score. (c) ROC-Curve that using IP packets. (d) ROC-Curve that using GTP-U Packets. (e) ROC-Curve that using IP and GTP-U packets

Table 10: Multiclass classification score by algorithm–benign to malicious packets count ratio is 50%

Experimental method	Score	Algorithm			
		KNN	SVM	RF	Stacking
IP Packet: Experimental method (a)	Accuracy	95.020%	65.995%	98.630%	98.716%
	F1	93.231%	45.095%	98.131%	98.231%
	ROC-AUC	95.603%	67.828%	98.504%	98.904%
GTP-U Packet: Experimental method (b)	Accuracy	44.722%	50.000%	62.477%	52.650%
	F1	26.556%	16.666%	58.864%	29.832%
	ROC-AUC	51.959%	50.000%	72.265%	56.278%
IP Packet + GTP-U Packet: Experimental method (c)	Accuracy	71.610%	65.830%	97.457%	97.500%
	F1	64.430%	44.975%	96.418%	96.396%
	ROC-AUC	75.389%	68.013%	97.496%	97.785%

4.3.3 Experiments Related to Feature Importance in 5GC

In Experiment 4.3.3, the feature importance was calculated for the multiclass classification by using the optimal algorithm for each of the three experimental methods derived from 4.3.2 and confirmed the influence of IP packet features–Tab. 11. As experimented results, in the experimental method (b) of multiclass classification used only GTP-U packets, three features had importance values. Among them, the TEID feature related to the IPs and ports of the source and destination of IP packet showed the highest importance by 0.570975. Experimental methods (a) and (c) used the IP packets of wired networks showed the high importance for features related to IP and port addresses. This shows that IP packet features have a large influence on IoT Botnet detection of multiclass classification in 5GC as in binary classification.

Table 11: Top 10 feature importance by 3 experimental methods for multiclass classification

Feature importance rank	IP Packet: (a)		GTP-U Packet: (b)		IP Packet + GTP-U Packet: (c)	
Number of features used	18		10		28	
Rank 1	ip.dst	0.227764	gtp.teid	0.570975	ip.dst	0.227183
Rank 2	tcp.srcport	0.218623	gtp.length	0.216603	tcp.srcport	0.213114
Rank 3	ip.src	0.205318	frame.len	0.212422	ip.src	0.204700
Rank 4	tcp.dstport	0.165589	ip.src	0	tcp.dstport	0.161656
Rank 5	udp.dstport	0.099744	ip.dst	0	udp.dstport	0.099483
Rank 6	frame.len	0.081292	udp.srcport	0	frame.len	0.060910
Rank 7	udp.srcport	0.001636	udp.dstport	0	frame.len (gtp)	0.014244
Rank 8	icmp.type	0.000034	gtp.flags	0	gtp.teid	0.010706
Rank 9	eth.dst	0	gtp.message	0	gtp.length	0.006485
Rank 10	eth.src	0	gtp.ext_hdr.next	0	udp.srcport	0.001398

4.4 Comparison of Binary and Multiclass Classification Results

In this section, the results of the binary classification and multiclass classification performed under the same conditions are compared. Both classification experiments were performed in an environment where the ratio of benign and malicious packets was 50% by using approximately 40,000 packets. Tab. 12 shows the scores and feature importance top five for the algorithms that showed the best performance for each experimental method of each classification.

Table 12: Comparison of the best algorithm, score, and feature importance of classification

Classification type		Binary classification			Multi classification		
Experimental method		(a)	(b)	(c)	(a)	(b)	(c)
Best algorithm		Stacking	Stacking	Stacking	Stacking	RF	Stacking
Top score	ACC	99.958%	76.967%	99.924%	98.716%	62.477%	97.500%
	F1-score	99.958%	76.761%	99.925%	98.231%	58.864%	96.396%
	ROC-AUC	99.958%	76.967%	99.925%	98.904%	72.265%	97.785%
Feature importance Rank	Rank 1	eth.src	gtp.teid	eth.src	ip.dst	gtp.teid	ip.dst
	Rank 2	ip.dst	frame.len	ip.dst	tcp.srcport	gtp.length	tcp.srcport
	Rank 3	eth.dst	gtp.length	eth.dst	ip.src	frame.len	ip.src
	Rank 4	ip.src		ip.src	tcp.dstport		tcp.dstport
	Rank 5	frmae.len		frame.len	udp.dstport		udp.dstport

In binary classification, the Stacking ensemble algorithm showed the best performance. In multiclass classification, in the experimental methods (a) and (c) using IP packets showed the best performance in Stacking, and in the experimental methods (b) using GTP-U packets, RF showed the best performance. And, in the multiclass classification that required more sophisticated classification (normal/bashlite/mirai/torii), it could be confirmed that all scores were slightly reduced (maximum approximately 23.315%, minimum approximately 0.054%) than the binary classification. In particular, detection of IoT Botnet using only GTP-U packets decreased by up to 22.99% (binary classification) and 36.71% (multiclass classification) compared to detection using IP packets together. Also, the experimental method (b) used the GTP-U packet in both classifications (binary/multiclass) had the feature importance value of high TEID, and the next, features related to the packet length had high feature importance. Experimental methods (a) and (c) showed the high importance for features related to the source and destination addresses in both classifications (binary/multiclass). As a result of the experiment, it was confirmed that the GTP-U packet could not fully reflect the characteristics of the IP packet in both binary and multiclass classification. Therefore, in order to perform IoT Botnet detection in 5GC, feature research that can reflect the characteristics of both IP packet and GTP-U packet is required.

5 Conclusion

In this study, the IoT Botnet packets were detected, results were compared and studied for their detection efficiency through the ML-based binary and multiclass classification in the mobile network environment. In order to build the dataset, after collecting the IP packets of IoT devices, GTP-U Tunneling packets were collected by replaying them in the 5G simulation environment. Experiments for both classification methods (binary/multiclass) were proceeded in three stages for the three experimental methods (IoT device IP packet, UPF GTP-U packet, and IP packet + GTP-U packet). In the first experiment, it was confirmed that the environment with the least bias was when the ratio of the number of benign to malicious packets was 50%. In the second experiment, IoT Botnet detection was performed using binary and multiclass classification. In the binary classification, the stacking ensemble algorithm showed the best detection rates with accuracies of approximately 99.958%, approximately 76.967%, and approximately 99.924% for each experimental method. Further, in the multiclass classification, when the experimental methods (a) and (c) were the Stacking algorithm, showed the highest accuracy (approximately 98.716%, approximately 97.5%), and when the experimental methods (b) was the RF, showed the highest accuracy (approximately 62.477%). In both classification methods, when IoT Botnet detection is performed using only GTP-U packets in the 5GC environment, the accuracy decreased by about 22.99% (binary) and about 36.71% (multiclass) compared to detection using IP packets. It showed that it was difficult to detect IoT Botnets in the 5GC environment using only the GTP-U packets for both binary and multiclass classification. Then, it was confirmed that the GTP-U packet and IP packet should be used together for the IoT Botnet detection in 5GC environment to obtain performance detection similar to that in wired network environment. In the third experiment, we confirmed that features related to IP packets show high importance for both binary and multiclass classification through feature importance comparison. Through this experiment, the importance of feature research for IoT Botnet detection considering 5GC network characteristics was confirmed for IoT Botnet detection in 5GC network path.

There are various ways to utilize the ML-based security function experimented in this paper. In fact, 6G has an AI-enabled architecture structure, and NWDAF is being discussed as a method. Before linking the ML-based security function with NWDAF, this paper conducted ML-based experiments using data collected from UPF and confirmed the possibility of NWDAF security function through

AI-native. Therefore, we plan to study how to link ML-based security functions to NWDAF through NWDAF structure and procedure analysis in the future. In addition, the experiment proposed in the paper performed IoT Botnet detection using features extracted from individual IoT Botnet packets. Therefore, we plan to conduct a feature research so that IoT Botnet traffic for sampled traffic can be detected with good performance using IoT Botnet feature that reflects the characteristics of the 5GC environment.

Funding Statement: This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2021-0-00796, Research on Foundational Technologies for 6G Autonomous Security-by-Design to Guarantee Constant Quality of Security)

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. Alizadeh, K. Andersson and O. Schelen, "A survey of secure internet of things in relation to blockchain," *Journal of Internet Services and Information Security (JISIS)*, vol. 10, no. 3, pp. 47–75, 2020.
[2] M. Zhong, Y. Zhou and G. Chen, "Sequential model based intrusion detection system for IoT servers using deep learning methods," *Sensors*, vol. 21, no. 4, pp. 1113, 2021.
[3] M. Wazzan, D. Algazzawi, O. Bamasaq, A. Albeshri and L. Cheing, "Internet of things botnet detection approaches: Analysis and recommendations for future research," *Applied Sciences*, vol. 11, no. 12, pp. 5713, 2021.
[4] M. A. Haq and M. A. R. Khan, "DNNBIoT: Deep neural network-based botnet detection and classification," *Computers, Materials & Continua (CMC)*, vol. 71, no. 1, pp. 1729–1750, 2021.
[5] R. Doshi, N. Apthorpe and N. Feamster, "Machine learning DDoS detection for consumer internet of things devices," in *2018 IEEE Security and Privacy Workshops (SPW)*, San Francisco, pp. 29–35, 2018.
[6] H. Kim, "5G core network security issues and attack classification from network protocol perspective," *Journal of Internet Services and Information Security (JISIS)*, vol. 10, no. 2, pp. 1–15, 2020.
[7] A. Abhishta, W. V. Heeswijk, M. Junger, L. J. M. Nieuwenhuis and R. Joosten, "Why would we get attacked? an analysis of attacker's aims behind DDoS attacks," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, vol. 11, no. 2, pp. 3–22, 2020.
[8] J. H. Park and J. H. Park, "Blockchain security in cloud computing: Use cases, challenges, and solutions," *Symmetry*, vol. 9, no. 8, pp. 164, 2017.
[9] M. Kolomeets, A. Chechulin and I. Kotenko, "Bot detection by friends graph in social networks," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, vol. 12, no. 2, pp. 141–159, 2021.
[10] S. Nowaczewski and W. Mazurczyk, "Securing future internet and 5G using customer edge switching using DNSCrypt and DNSSEC," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, vol. 11, no. 3, pp. 87–106, 2020.
[11] G. Choudhary, J. Kim and V. Sharma, "Security of 5G-mobile backhaul networks: A survey," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, vol. 9, no. 4, pp. 41–70, 2018.
[12] X. Yang, "5G security in ITU-T SG17," in *ETSI Security Week 2018*, International Telecommunication Union (ITU), France, pp. 1–19, 2018.
[13] 3GPP, "Security architecture and procedures for 5G system," 3GPP TS 33.501 v17.4.1, Jan. 2022.
[14] 3GPP, "System System architecture for the 5G system (5GS)," 3GPP TS 23.501 v17.3.0, Dec. 2021.
[15] J. Wu, R. Li, X. An, C. Peng, Z. Liu *et al.,* "Toward native artificial intelligence in 6G networks: System design, architectures, and paradigms," arXiv preprint arXiv:2103.02823, pp. 1–7, 2021.

[16] A. G. Manzanares, J. M. Galindo, H. Bahsi and S. Nomm, "Medbiot: Generation of an IoT botnet dataset in a medium-sized IoT network," in *6th Int. Conf. on Information Systems Security and Privacy (ICISSP)*, Valletta, Malta, pp. 207–218, 2020.
[17] 3GPP, "General packet radio service (GPRS); GPRS tunneling protocol (GTP); across the Gn and Gp interface," 3GPP TS 29.060 v17.1.0, Sep. 2021.
[18] L. Bonati, M. Polese and S. D'Oro, "Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead," *Computer Networks*, vol. 182, pp. 1–28, 2020.
[19] Q. Duan, "Intelligent and autonomous management in cloud-native future networks-a survey on related standards from an architectural perspective," *Future Internet*, vol. 13, no. 2, pp. 42, 2021.
[20] A. Ghosh, A. Maeder, M. Baker and D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, pp. 127639–127651, 2019.
[21] "Netmanias, "5G traffic flow," [Online]. Available: https://www.netmanias.com/en/post/oneshot/14104/5g/5g-traffic-flow, 2019.
[22] N. A. Stoian, "Machine learning for anomaly detection in IoT networks: Malware analysis on the IoT-23 data set," B.S. thesis, University of Twente, Netherlands, 2020.
[23] V. Dutta, M. Choras, M. Pawlicki and R. Kozik, "A deep learning ensemble for network anomaly and cyber-attack detection," *Sensors*, vol. 20, no. 16, pp. 4583, 2020.
[24] H. Alzahrani, M. Abulkhair and E. Alkayal, "A Multi-class neural network model for rapid detection of IoT botnet attacks," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 11, no. 7, pp. 668–696, 2020.
[25] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto and K. Sakurai, "Machine learning based IoT-botnet attack detection with sequential architecture," *Sensors*, vol. 20, no. 16, pp. 1–15, 2020.
[26] F. Hussain, S. G. Abbas, U. U. Fayyaz, G. A. Shah, A. Toqeer *et al.,* "Towards a universal features set for IoT botnet attacks detection," in *2020 IEEE 23rd Int. Multitopic Conf. (INMIC)*, Bahawalpur, Pakistan, pp. 1–6, 2020.
[27] K. Saeedi, "Machine learning for DDoS detection in packet core network for IoT," M.S. thesis, Dept. Computer Science, Eletrical and Space Engineering, Luleå University of Technology, Swedish, 2019.
[28] P. Salva-Garcia, J. M. Alcaraz-Calero, Q. Wang, J. B. Bernabe and A. Skarmeta, "5G NB-IoT: Efficient network traffic filtering for multitenant IoT cellular networks," *Security and Communication Networks*, vol. 2018, pp. 21, 2018.
[29] Fraunhofer FOKUS, "Open5Gcore, fundamental 5g core network functionality," [Online]. Available: https://www.open5gcore.org/, 2021.