

User Terminals as Attackers: An Open Dataset Analysis of DDoS Attacks in 5G Networks

Maria Christopoulou¹, Apostolis Garos², Athina Vekraki¹, Dimitris Santorinaios¹, Ioannis Koufos¹, Sofia Karamitsiani¹, George Xilouris^{1, 3}, Michail-Alexandros Kourtis¹, Georgios Gardikis², Panagiotis Trakadas³

¹National Centre for Scientific Research "Demokritos", Agia Paraskevi, Greece

{maria.christopoulou, dsantorinaios, avekraki, ikoufos, skaramitsiani, akis.kourtis, xilouris}@iit.demokritos.gr

²Space Hellas S.A., Agia Paraskevi, Greece

{agaros, ggar}@space.gr

³National and Kapodistrian University of Athens, Greece

ptrakadas@pms.uoa.gr

Abstract—The 5th Generation (5G) of cellular networks, developed by the 3rd Generation Partnership Project (3GPP), aims to meet the growing demands for data and communication services. A key component of the 5G architecture is the Network Data Analytics Function (NWDAF), which enhances network performance and detects anomalies by analyzing real-time data. This paper focuses on detecting abnormal user behavior, specifically Distributed Denial of Service (DDoS) attacks, using a comprehensive dataset captured in a 5G testbed. We compare the Z-score method, a traditional statistical method, with machine learning models, including Decision Trees, Naive Bayes, kNN, and XGBoost. Our results demonstrate the improved performance of machine learning models in detecting anomalies in this context. Furthermore, we study the impact of various network features through Principal Component Analysis (PCA), while also employing the inherent explainability capability of Decision Trees to highlight the importance of features in distinguishing between benign and malicious traffic. This study provides valuable insights into DDoS detection in 5G networks, and the dataset is made publicly available to facilitate further research.

Index Terms—3GPP, 5G, Dataset, Detection, DDoS attack, Machine Learning, NWDAF

I. INTRODUCTION

The Fifth Generation (5G) of mobile networks represents the latest advancement in cellular technology, designed to address the increasing demands of the digital era through enhanced speed, latency, capacity, and reliability. This evolution introduced a transition to a software-driven architecture, emphasizing virtualization and cloud technologies, and adopted a Service-Based Architecture (SBA) for scalable and modular Network Functions (NFs). The Network Data Analytics Function (NWDAF), introduced in the 3rd Generation Partnership Project (3GPP) Release 15, marked a significant move towards data-driven and proactive network management. NWDAF's primary role is to gather and analyze data from across the 5G network, including performance metrics, user behavior, and network conditions, to optimize network performance, resource management, and user experience. Integrated within the SBA of 5G, the NWDAF communicates with other network functions through standardized interfaces, enabling comprehensive data collection and analytics services. It may

employ sophisticated data analytics and machine learning algorithms to anticipate network demands, identify anomalies, and suggest network configuration adjustments. The NWDAF can be used in a broad spectrum of applications [1], enhancing network security, optimizing quality of service, facilitating network slicing for varied service demands, and supporting edge computing for applications requiring low latency.

A. Related Work and Paper Contributions

In recent years, 5G security has garnered significant attention, with studies highlighting vulnerabilities of the 5G architecture. Ahmad et al. [2] explain that the 5G systems/deployments are being challenged by large swings of traffic that continuously test cyber resiliency. Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks are considered menacing threats for such infrastructures, directly affecting QoS and the availability of 5G Core (5GC) Network Functions (NFs). The authors in [3] describe how emerging DoS attacks are engineered and showcase their impact on experience degradation and energy inefficiency, by emulating traffic with a discrete-event synthetic-traces simulator. Considering the recent advances of Artificial Intelligence and Machine Learning (AI/ML), open datasets of high quality have become necessary for developing and training algorithms for detecting and mitigating attacks resulting from existing vulnerabilities.

In [4], the authors perform various reconnaissance, Denial of Service (DoS) and network reconfiguration attacks against 5G core interfaces, exploiting vulnerabilities of the open-source free5GC [7] implementation. The provided dataset includes network flow features and log files of 5GC NFs. In [5], the authors focus on attacks involving unauthorized control plane signaling targeting the Packet Forwarding Control Protocol (PFCP) -the protocol used for exchanging control messages between the Session Management Function (SMF) and the User Plane Function (UPF) NFs- utilizing the open-source Open5GS [8] 5GC implementation. The dataset includes TCP/IP network flow statistics and PFCP session status counters as part of the 5GC control plane signaling.

TABLE I
COMPARISON WITH EXISTING 5G-SPECIFIC ATTACKS DATASETS

Ref.	Description	Radio Access Data	5G Control Plane Signaling	Over-the-air Testbed
[4]	Various attacks against 5GC NFs over a simulated testbed	X	✓	X
[5]	PFCP DoS Attacks over a simulated testbed	X	✓	X
[6]	DoS Attacks against 5G testbed	X	X	✓
Our dataset	DDoS Attacks against 5G testbed	✓	✓	✓

In [6], the authors perform various DoS attacks -ICMP, UDP, SYN, HTTP floods- and scanning attacks, originated from a connected attacker to a 5G testbed, and provide a dataset focusing on data plane features.

Our contributions: Table I provides a comprehensive comparison between existing works and our paper. Unlike these works, in this study we combine both control plane signaling and radio access data, aligning with the relevant 3GPP Technical Specifications regarding abnormal user behavior. The paper describes the testbed setup, provides qualitative and quantitative information on the dataset recording methodology, and performs exploratory statistical analysis as a starting point for interested data scientists. To further contribute to the field of 5G security, we make the second version of this dataset publicly available at [9].

Paper organization: Section II provides an overview on abnormal User Equipment (UE) behavior detection, focusing on DDoS attacks and describes how NWDAF facilitates this use case. Section III describes the system setup for performing the attacks and the measurements. Section IV provides an overview of the dataset, model evaluation results of various machine learning and statistical models (z-score) used for anomaly detection in a 5G network. Section V discusses real-world use cases and discusses future work regarding the evolution of the presented dataset. Section VI concludes the paper.

II. ABNORMAL UE BEHAVIOR DATA ANALYTICS – THE NWDAF ROLE

The NWDAF can be leveraged to identify a group of UEs or a specific UE with abnormal behavior, i.e., being misused, hijacked, having malicious applications running on the UEs, or UEs which have been stolen, as reported in 3GPP TS 23.288 (Rel. 18) [10]. Such abnormal user be-

havior encompasses different use cases, termed “exceptions,” including unexpected long-live/large rate flows, unexpected wakeup, suspicion of DDoS attack, wrong destination address, too frequent service access, abnormal traffic volume or unexpected radio link failures. When there is an active subscription from a 5G NF (consumer), the NWDAF collects data from the 5GC NFs or the Operations-Administration-Maintenance (OAM) and performs analytics for abnormal user behavior detection. Table II presents a distilled version of Table 6.7.3.2-1 in 3GPP TS 23.288 (Rel. 18) [10]. The Table summarizes communication related data collected by the NWDAF in order to detect DDoS attacks against the 5G network. The consumer may also request additional UE behavior parameters from the Session Management Function (SMF), Access and Mobility Management Function (AMF) and Unified Data Management (UDM) to compare and identify behavior trends over time. In the case of DDoS attack detection, these parameters include the interval time for periodic communications, scheduled communication times of the UEs, the time the UE stays in Connection Management (CM) or Registration Management (RM) states, and the traffic profile (single or multi-packet transmissions) [10], [11].

After performing analytics, the NWDAF reports various metrics to the consumer. These metrics include the Exception, i.e., the type of abnormal behavior, the severity detected, the UE identifiers, and other UE behavioral metrics that the consumer has requested, depending on the exception.

TABLE II
5GC DATA COLLECTED BY THE NWDAF FOR DDOS ATTACK DETECTION [10]

Source	Data per Source
SMF	<ul style="list-style-type: none"> • Subscription Permanent Identifier (SUPI) • Internal Group ID • Single-Network Slice Selection Assistance Information (S-NSSAI) • Application ID • PDU Session ID • Session inactivity time • PDU Session Status • UE Session Behavior Trends • Metrics on UE communications
AMF	<ul style="list-style-type: none"> • Type Allocation Code (TAC) • UE locations (Cells where the UE enters) and timestamps • UE location trends • Metrics on UE Connection and State transitions (e.g., IDLE, CM, RM, Handover)
AF	<ul style="list-style-type: none"> • Generic Public Subscription Identifier (GPSI) • External Group ID
UPF	<ul style="list-style-type: none"> • Timestamps of communication beginning/end • DL/UL data rate and Traffic Volume

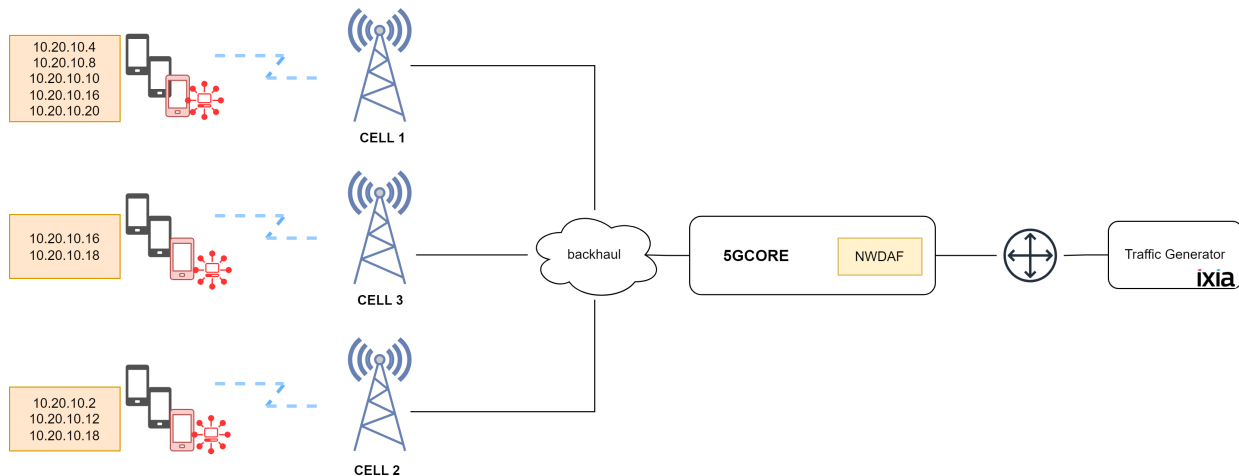


Fig. 1. Testbed setup. Handovers were performed during the recording, resulting in some connections changing cells.

III. SYSTEM SETUP AND ATTACKS DESCRIPTION

Following the 3GPP Technical specifications described in Section II, this paper focuses on DDoS attacks, recording and providing openly control signaling metrics from the 5G network that could infer such an attack by malicious UEs. Figure 1 illustrates the testbed topology which comprises 3 cells with a total of 9 UEs connected to the same core network. The 5G network is implemented by the Amarisoft Callbox Mini solution, and we further employ two more cells using the Amarisoft Classic, that also hosts the 5G core [12].

The setup utilizes a broad set of UE devices comprising a set of smart phones (Huawei P40), microcomputers (Raspberry Pi 4 - Waveshare 5G Hat M2), industrial 5G routers (Industrial Waveshare 5G Router), a WiFi-6 mobile hotspot (DWR-2101 5G Wi-Fi 6 Mobile Hotspot) and a CPE box (Waveshare 5G CPE Box). All UEs are being operated by subsidiary hosts which are responsible for the traffic generation, occurring from scheduled communications times. All identifiers are artificially generated and do not represent real personal data. We identify each UE through its 'imeisv' ID, that corresponds to the device in use, due to vendor implementation, that uses the same IMSI for all UEs.

Table III summarizes the details regarding the traffic profiles of all UEs and the date and time of the attacks. The attacks selected are as follows:

- **SYN Flood:** Exploits the TCP handshake by sending numerous SYN packets with spoofed IPs, causing half-open connections and leading to denial of service.
- **ICMP Flood:** Overwhelms the target with ICMP Echo Requests, forcing the system to respond.
- **UDP Fragmentation:** Sends a large volume of fragmented UDP packets, exhausting the target's resources during reassembly.
- **DNS Flood:** Sends high volumes of DNS queries to overwhelm DNS servers, affecting domain name resolution and access to online services.

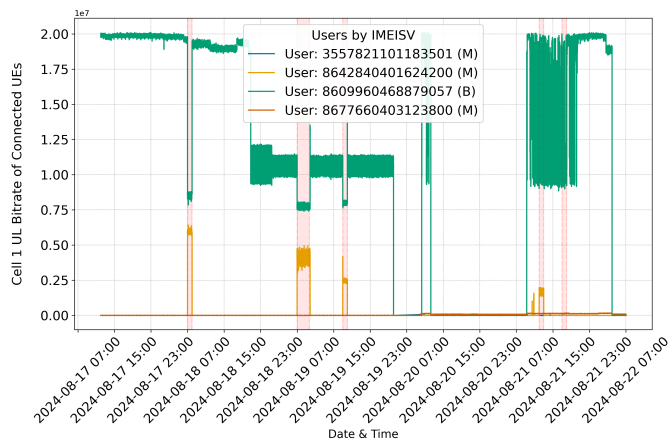


Fig. 2. UL Bitrate of UEs in Cell 1 (M refers to Malicious UE and B to Benign UE.)

- **GTP Flood:** Targets the GTP protocol in mobile networks by sending excessive GTP packets.

IV. DATASET PROCESSING AND EVALUATION OF RESULTS

A. The Dataset

The dataset is recorded through a data collector interfacing with the 5G network, gathering data regarding UEs, gNBs, and the Core Network. The data are pre-processed into three separate .csv files: "amari_ue_data.csv", "enb_counters.csv", and "mme_counters.csv". Due to space limitations, we focus on 'amari_ue_data.csv', which contains features related to UE identification, IP addressing, bearer, and cell information, as well as aggregated bitrates. A complete description of the features can be found in [9].

The 'enb_counters.csv' includes 107 features, focusing on cell-level metrics such as bitrates, user ratios, and control plane signaling. The 'mme_counters.csv' provides 61 features on the Non-Access Stratum (NAS) with insights on PDU session status and UE context. In this paper, we define an anomaly as

TABLE III
TRAFFIC PROFILES OF MALICIOUS AND BENIGN UES

Attack/Benign Traffic	Date	Time (UTC)	IPs	IMEISV
SYN FLOOD	18.08.2024	7:00 - 8:00	10.20.10.2 10.20.10.4	8642840401612300 8642840401624200
ICMP FLOOD	19.08.2024	7:00 - 9:41	10.20.10.2 10.20.10.4	
UDP FRAG	19.08.2024	17:00 - 18:00	10.20.10.2 10.20.10.4	
DNS FLOOD	21.08.2024	12:00 - 13:00	10.20.10.2 10.20.10.4	
GTPU FLOOD	21.08.2024	17:00 - 18:00	10.20.10.2 10.20.10.4 10.20.10.6 10.20.10.8 10.20.10.10	8642840401612300 8642840401624200 8642840401594200 8677660403123800 3557821101183500
Skype-1050p YouTube 4K	Start Date: 17.08.2024 End Date: 21.08.2024	Start Time (17.08): 16:00 End Time (21.08): 20:00	10.20.10.12 10.20.10.16 10.20.10.18 10.20.10.20	8628490433231150 8609960480859050 8609960480666910 8609960468879050

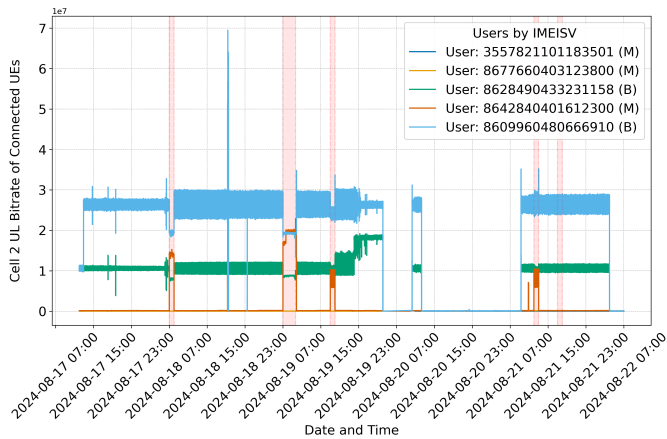


Fig. 3. UL Bitrate of UEs in Cell 2 (M refers to Malicious UE and B to Benign UE.)

an observation from a different probability distribution than the majority, whereas an outlier deviates within the same distribution without implying a shift.

Table V compares the performance of the Z-Score method with the machine learning models across various metrics, including test accuracy, precision, recall, and F1-score. Considering that this is an imbalanced dataset, the accuracy can be misleading, as it can be impacted by the majority class (i.e., the benign traffic). Therefore, we provide the recall, precision and F1-score as additional metrics for a more comprehensive characterization of the models' performances.

B. Principal Component Analysis (PCA) Results

We applied PCA to visualize the relationship between benign and malicious traffic, using the first two principal

TABLE IV
TOP FEATURES FOR PRINCIPAL COMPONENTS (PC1 AND PC2)

PC1 (Principal Component 1)	PC2 (Principal Component 2)
dl_bitrate (0.303)	pusch_snr (0.380)
ul_tx (0.307)	ul_mcs (0.348)
ul_retx (0.267)	eprc (0.330)
ul_bitrate (0.265)	dl_tx (0.141)
dl_tx (0.292)	ul_tx (0.082)

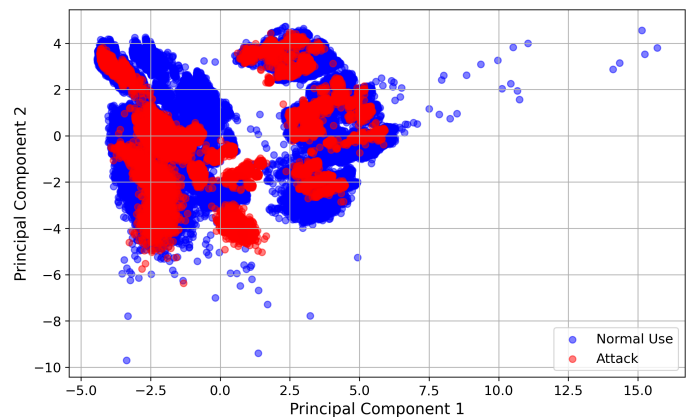


Fig. 4. PCA Components of the Dataset

components (PC1 and PC2) that capture significant variance. As a result, we selected a subset of the features to train the AI/ML models, as described in Section IV.

Key features contributing to PC1 include dl_bitrate (0.303), ul_bitrate (0.265), and retransmission rates, such as ul_retx (0.267) and dl_retx (0.222), indicating that traffic volume and retransmissions are important in distinguishing between attack

TABLE V
PERFORMANCE COMPARISON OF DIFFERENT MODELS ON 5G DATASET

Model	Test Accuracy	Precision (0.0)	Recall (0.0)	F1-Score (0.0)	Precision (1.0)	Recall (1.0)	F1-Score (1.0)	W. Avg Precision	W. Avg F1-Score	Training Time (s)
Z-Score Method	0.9186	0.96	0.96	0.96	0.33	0.31	0.32	0.33	0.32	0.45
Decision Tree	0.9700	0.97	1.00	0.98	0.90	0.58	0.70	0.97	0.97	3.13
Decision Tree K-folds	0.9700	0.97	1.00	0.98	0.90	0.58	0.70	0.97	0.97	13.24
Decision Tree K-folds w/ SMOTE	0.9127	0.99	0.92	0.95	0.40	0.82	0.54	0.95	0.93	41.13
Naive Bayes	0.8676	0.94	0.91	0.93	0.11	0.16	0.13	0.89	0.88	0.27
Naive Bayes K-folds	0.8676	0.94	0.91	0.93	0.11	0.16	0.13	0.89	0.88	2.44
Naive Bayes K-folds w/ SMOTE	0.7343	0.95	0.76	0.84	0.08	0.34	0.13	0.89	0.80	10.25
XGBoost	0.9969	1.00	1.00	1.00	0.98	0.97	0.97	1.00	1.00	2.67
XGBoost K-folds	0.9969	1.00	1.00	1.00	0.98	0.97	0.97	1.00	1.00	10.66
XGBoost K-folds w/ SMOTE	0.9969	1.00	1.00	1.00	0.96	0.99	0.99	1.00	1.00	32.01
KNN	0.9997	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.10
KNN K-folds	0.9997	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	345.84
KNN K-folds w/ SMOTE	0.9181	0.99	0.92	0.95	0.42	0.87	0.57	0.96	0.93	658.90

and normal traffic. Features like PUSCH SNR (0.379 for PC2) and UL MCS (0.348 for PC2) reflect the signal quality during high-load scenarios.

As illustrated in Fig. 4, there is a noticeable overlap between malicious and benign traffic, which suggests that the malicious traffic shares similar characteristics with the benign traffic. For example, flooding traffic increases the uplink bitrate (ul_bitrate), a pattern that can also occur during benign traffic transmission. This correlation between features makes it more challenging to fully separate the two classes. To address these challenges, we evaluate AI/ML models like Decision Trees in section IV-C, that can handle such correlations, as they are capable of capturing non-linear relationships between features.

C. Evaluation of Z-Score and Machine Learning Models

We evaluated several machine learning models, including KNN, XGBoost, Decision Tree, and Naive Bayes, using cross-validation and SMOTE to address class imbalance. We also used stratified splits for maintaining the class distribution in the training sets, considering that this dataset is imbalanced. The Z-Score method was also employed as a baseline for comparison. The training time evaluations were conducted on a workstation equipped with an Intel Core i7-7700 processor (3.60 GHz).

The **Z-Score** method provided a baseline approach for anomaly detection in this study, as a statistical technique for flagging outliers. To account for potential non-stationarity in the dataset, first-order differencing was applied to selected numerical features, ensuring that time-dependent fluctuations did not skew the results. Following this, the numerical features were normalized using a MinMax scaler, which transformed all feature values to a range between -1 and 1. This normalization helped ensure that all features contributed equally to the Z-Score calculation, avoiding dominance by features with larger magnitudes.

In the Z-Score method, any data points with Z-Scores exceeding a threshold of 7.0 were flagged as anomalies. This threshold was determined through hyperparameter tuning to optimize the trade-off between detecting true anomalies and minimizing false positives. A higher threshold was chosen to reduce the likelihood of flagging points from the distribution's tail, which may naturally exhibit extreme values without representing genuine anomalies. The Z-Score method

performed well in scenarios where the dataset adhered to a relatively normal distribution, offering a test accuracy of 91.9%. However, when focusing on malicious traffic detection (class 1), the method demonstrated precision of 0.33 and recall of 0.31, indicating a significant number of false negatives.

While the Z-Score method provided an acceptable balance between precision and recall for anomaly detection, it missed a substantial number of true anomalies—6,438 instances in total—demonstrating its limitations in more complex datasets. The method works well for capturing basic outliers in a relatively simple data distribution, but it struggles with dynamic network traffic of our dataset. Although the Z-Score method is computationally efficient and easy to implement, its inability to effectively capture the complexities of the data, particularly for malicious traffic, makes it less suitable for high-precision tasks in 5G networks.

We also tested the **Interquartile Range (IQR)** method, which identifies anomalies by analyzing the range of the middle 50% of data, defined by the first (Q1) and third (Q3) quartiles. A factor of 5.5 was applied to extend the IQR range for anomaly detection, optimizing sensitivity while attempting to minimize false positives. While the IQR method achieved perfect recall (1.0) -meaning all actual anomalies were detected- it suffered from a low precision (0.065) and accuracy (0.112)- indicating a high number of false positives, with 135,665 normal instances flagged as anomalies, meaning that further threshold refinements are needed to enhance precision. Therefore, we use the Z-Score as the baseline method for comparison with ML models.

In comparison to the Z-Score method, the machine learning models demonstrated higher accuracy, precision, and recall. **KNN**, configured with 10 nearest neighbors, achieved a test accuracy of 99.97% and perfect precision, recall, and F1-scores across both classes. However, KNN relies on distance-based classification, making it vulnerable to overfitting in imbalanced datasets. New data points tend to be classified based on their proximity to the majority class (benign traffic), leading to poor generalization for the minority class (malicious traffic) [13]. When *KNN K-folds w/ SMOTE* was employed to address this imbalance, the performance remained high but at a much higher computational cost, with training time increasing to 658.90 seconds.

TABLE VI
CONFUSION MATRICES, FALSE POSITIVE RATES (FPR), AND FALSE
NEGATIVE RATES (FNR) FOR EACH MODEL

Model	Confusion Matrix	FPR (%)	FNR (%)
Decision Tree	$\begin{bmatrix} 128228 & 560 \\ 3551 & 4867 \end{bmatrix}$	0.43	42.18
Decision Tree K-folds	$\begin{bmatrix} 128228 & 560 \\ 3551 & 4867 \end{bmatrix}$	0.43	42.18
Decision Tree SMOTE	$\begin{bmatrix} 118287 & 10501 \\ 1478 & 6940 \end{bmatrix}$	8.15	17.56
Naive Bayes	$\begin{bmatrix} 117701 & 11087 \\ 7077 & 1341 \end{bmatrix}$	8.61	84.07
Naive Bayes K-folds	$\begin{bmatrix} 117701 & 11087 \\ 7077 & 1341 \end{bmatrix}$	8.61	84.07
Naive Bayes SMOTE	$\begin{bmatrix} 97930 & 30858 \\ 5594 & 2824 \end{bmatrix}$	23.96	66.45
XGBoost	$\begin{bmatrix} 128579 & 209 \\ 214 & 8204 \end{bmatrix}$	0.16	2.54
XGBoost K-folds	$\begin{bmatrix} 128579 & 209 \\ 214 & 8204 \end{bmatrix}$	0.16	2.54
XGBoost SMOTE	$\begin{bmatrix} 128467 & 321 \\ 100 & 8318 \end{bmatrix}$	0.25	1.19
KNN	$\begin{bmatrix} 128764 & 24 \\ 13 & 8406 \end{bmatrix}$	0.02	0.15
KNN K-folds	$\begin{bmatrix} 128764 & 24 \\ 13 & 8406 \end{bmatrix}$	0.02	0.15
KNN SMOTE	$\begin{bmatrix} 118609 & 10179 \\ 1065 & 7353 \end{bmatrix}$	7.90	12.65

Similarly, **XGBoost** showcased a test accuracy of 99.69%, offering precision, recall, and F1-scores of 1.00 across multiple configurations. As a gradient boosting algorithm, XGBoost iteratively refines its predictions by minimizing classification errors, which makes it effective in various applications. However, like KNN, XGBoost is prone to overfitting when applied to imbalanced datasets, as it tends to focus heavily on optimizing for the majority class [13]. Adjustments such as tuning the `scale_pos_weight` parameter can help address this issue by increasing the weight of the minority class, but improper tuning may still lead to a skewed model. Despite this, *XGBoost K-folds w/ SMOTE* maintained high performance, at the expense of increased training time to 32.01 seconds.

Decision Tree models, on the other hand, provided a more balanced approach to imbalanced data. Decision Trees rely on splitting criteria like Gini Impurity and Information Gain, which ensure that both benign and malicious traffic are considered during the tree-building process [13]. The *DecisionTreeClassifier* used in this study (with `max_depth=5`, `min_samples_split=10`, and `min_samples_leaf=2`) performed consistently well, achieving a test accuracy of 97.00%. Precision and recall for benign traffic were particularly high, but the recall for malicious traffic (class 1) was lower at 0.58. When SMOTE was applied, the recall improved significantly to 0.82, while maintaining a balanced performance across all metrics. Decision Trees also benefit from being less prone to overfitting compared to KNN and XGBoost, especially when the tree

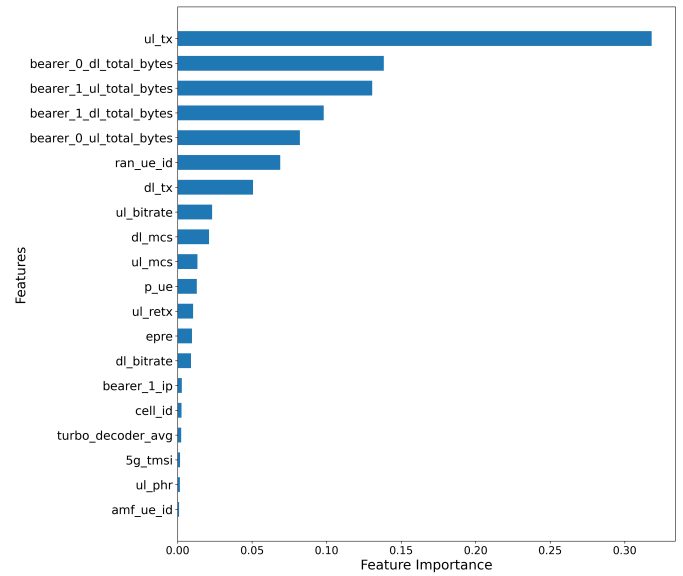


Fig. 5. Feature importance produced by Decision Trees

depth is carefully controlled, and pruning is applied.

Naive Bayes did not properly handle the complexities of the dataset. This model assumes conditional independence among features, an assumption that is violated in this dataset where features such as signal quality and retransmission rates are correlated. Naive Bayes produced a test accuracy of 86.76%, but its precision and recall for malicious traffic remained below 0.16, even after applying SMOTE. These results indicate that Naive Bayes, while computationally efficient, may not be proper algorithm for detecting anomalies in this complex dataset.

In Cybersecurity, the False Positive Rate (FPR) measures how often benign activities are misclassified as threats, while the False Negative Rate (FNR) reflects the likelihood of threats going undetected. Table VI presents the FPR and FNR for the tested models. The Decision Tree variants demonstrated balanced performance; without SMOTE, the Decision Tree had an FPR of 0.43% but a high FNR of 42.18%, indicating limitations in detecting malicious traffic. With SMOTE, the FNR dropped to 17.56%, though the FPR increased to 8.15%, highlighting a trade-off between sensitivity and specificity. KNN and XGBoost achieved low FPR (KNN: 0.02%, XGBoost: 0.16%) and FNR (KNN: 0.15%, XGBoost: 2.54%), with slightly higher FNRs in SMOTE scenarios. These results indicate that XGBoost and Decision Trees are a promising solution in anomaly detection in this context, with tuning needed to address potential overfitting.

Finally, we investigated the inherent explainability of the Decision Trees. As illustrated in Fig. 5, uplink transmission features, particularly *ul_tx*, play a dominant role in distinguishing between normal and attack traffic. Other critical features include both uplink and downlink traffic volumes associated with different bearers, such as *bearer_0_dl_total_bytes* and *bearer_0_ul_total_bytes*, highlighting the importance of traffic

volume in identifying anomalies (consistent with the features suggested by the 3GPP in DDoS attack detection scenarios [10]). Modulation and coding scheme features, like *ul_mcs* and *dl_mcs*, along with *ul_bitrate*, also show moderate importance, indicating their role in adapting network conditions during attack scenarios. Meanwhile, network quality features, such as *pusch_snr* and *cqi*, highlight signal degradation during high-traffic loads. In contrast, features like *imei_sv* and *initial_ta* have minimal impact, suggesting their limited relevance in detecting anomalies in this context.

V. USE CASES AND FUTURE WORK

Considering the discussion in Section II, the NWDAF can analyze patterns of network traffic and user behavior to detect anomalies that may indicate security threats or network issues. The applicability of 5G in various use cases, such as logistics, maritime industry, and smart cities, can provide detection of abnormal events. We discuss some potential applications below:

Logistics: The NWDAF can support fleet management and warehouse operations in 5G-based logistics operations. The NWDAF can monitor the communication patterns from a fleet of trucks. Anomalies such as a vehicle communicating more frequently than expected or not adhering to the expected communication schedule may indicate a hijacking or deviation from the planned route. In addition, in smart warehouses, the NWDAF can track data traffic related to automated systems and IoT sensors [14], [15], where sudden increases in traffic might suggest system malfunction or an attempt to overload the network as part of a cyber-attack.

Maritime industry: In this case, the NWDAF can support automated vessel tracking and port logistics. The NWDAF can analyze communication patterns from ships to detect anomalies in message frequency or data volume, which might indicate distress or unusual behavior. Furthermore, analyzing data traffic for logistics operations in ports can help identify unusual patterns that might suggest unauthorized access to systems or potential cyber-attacks on port 5G infrastructure.

Smart Cities: Some potential applications include traffic management systems and utilities monitoring. The NWDAF can analyze traffic flow data from a variety of sensors. A sudden change in the pattern could indicate a traffic incident or manipulation of the traffic control systems. Also, smart utility meters that suddenly show increased or irregular data transmission might be indicative of tampering or technical issues.

To further develop the dataset, we plan to continuously release newer versions towards more realistic use scenarios. This dataset evolution encompasses use case specific user data traffic and mobility patterns, application-level QoE metrics, UE behavior trends, as well as gradually increasing the number of connected devices to examine scalability and network resiliency issues.

VI. CONCLUSION

In this paper, we have presented a novel dataset collected from an operational 5G testbed, which includes detailed in-

formation on both the radio access network and control plane signaling. Our study has been guided by 3GPP standards for identifying abnormal behaviors, with a specific focus on detecting DDoS attacks from malicious user equipment. To support ongoing 5G security research, we are releasing this dataset to the public. In the future, we plan to enhance the dataset with more scenarios that reflect real-life use cases and examine the results of the models in real-time conditions.

ACKNOWLEDGMENT

This work has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the EU Horizon Europe programme PRIVATEER under Grant Agreement No. 101096110. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the EU or SNS JU.

REFERENCES

- [1] P. Gkonis et al., "Leveraging Network Data Analytics Function and Machine Learning for Data Collection, Resource Optimization, Security and Privacy in 6G Networks," *IEEE Access*, pp. 1–1, 2024, doi: 10.1109/ACCESS.2024.3359992.
- [2] I. Ahmad, S. Shahabuddin, N. Kumar, J. Okwuibe, A. Gurtov, and M. Ylianttila, "Security for 5G and Beyond," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3682–3722, 2019, doi: 10.1109/COMST.2019.2916180.
- [3] R. Ettiane, A. Chaoub, and R. Elkouch, "Toward securing the control plane of 5G mobile networks against DoS threats: Attack scenarios and promising solutions," *Journal of Information Security and Applications*, vol. 61, p. 102943, Sep. 2021, doi: 10.1016/j.jisa.2021.102943.
- [4] C. Coldwell et al., "Machine Learning 5G Attack Detection in Programmable Logic," in *2022 IEEE Globecom Workshops (GC Wkshps)*, Sep. 2022, pp. 1365–1370. doi: 10.1109/GCWkshps56602.2022.10008647.
- [5] G. Amponis et al., "Threatening the 5G core via FPCP DoS attacks: the case of blocking UAV communications," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, p. 124, Dec. 2022, doi: 10.1186/s13638-022-02204-5.
- [6] Y. Siriwardhana, "5G-NIDD: A Comprehensive Network Intrusion Detection Dataset Generated over 5G Wireless Network." *IEEE*, Dec. 02, 2022. Accessed: Feb. 04, 2024. [Online]. Available: <https://iee-dataport.org/documents/5g-nidd-comprehensive-network-intrusion-detection-dataset-generated-over-5g-wireless>
- [7] "free5GC." Accessed: Oct. 07, 2024. [Online]. Available: <https://free5gc.org/>
- [8] "https://open5gs.org/." Accessed: Oct. 07, 2024. [Online]. Available: <https://open5gs.org/>
- [9] National Centre of Scientific Research "Demokritos", & Space Hellas (Greece), "NCSR-DS-5GDDoS: 5G Radio and Core metrics containing sporadic DDoS attacks." Zenodo, Oct. 07, 2024. doi: 10.5281/zenodo.13900057.
- [10] 3GPP TS 23.288 (V18.4.0), "Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 18)," Dec. 2023.
- [11] 3GPP TS 23.502 (V18.4.0), "Procedures for the 5G System (5GS); Stage 2 (Release 18)," Dec. 2023.
- [12] Amarisoft. Accessed: Oct. 07, 2024. [Online]. Available: <https://www.amarisoft.com/>
- [13] Chen, W., Yang, K., Yu, Z., et al. (2024). A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review*, 57, 137. <https://doi.org/10.1007/s10462-024-10759-6>
- [14] Kavre, M., Gaddekar, A., & Gadhadre, Y. (2019). Internet of Things (IoT): A Survey. *2019 IEEE Pune Section International Conference (PuneCon)*, Pune, India, 1-6. <https://doi.org/10.1109/PuneCon46936.2019.9105831>
- [15] Ghasempour, A. (2019). Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges. *Inventions*, 4(1), 22. <https://doi.org/10.3390/inventions4010022>