Designing an effective network-based intrusion-detecting system for 5G networks

Azamat Imanbayev[1,2,†], Ansar Jakupov[2,†], Yersultan Valikhan[2,†] and Roman Odarchenko[3,*,†]

[1] al-Farabi Kazakh National University, 71 al-Farabi ave., 050040 Almaty, Kazakhstan
[2] Kazakh-British Technical University, 59 Tole bi str., 050000 Almaty, Kazakhstan
[3] National Aviation University, 1 Liubomyra Huzara ave., 03680 Kyiv, Ukraine

Abstract

The rapid advancement of 5G networks brings unprecedented benefits including higher speeds, lower latency, and the ability to support a massive number of connected devices. These enhancements enable new applications and services across various sectors, such as healthcare, automotive, and smart cities, revolutionizing how these industries operate. Traditional security measures, which were designed for earlier generations of cellular networks, are often inadequate in addressing the sophisticated and dynamic nature of cyber threats targeting 5G networks. This paper presents the design and implementation of a network-based Intrusion Detection System (IDS) specifically tailored for 5G networks to address these new security challenges. The proposed IDS leverages advanced machine learning techniques to analyze network traffic in real time, accurately identifying and mitigating potential security threats. Our research highlights the architectural design of the IDS, its integration within the 5G core network, and its effectiveness in maintaining network security. The IDS is designed to operate in a distributed manner, with components deployed across various network segments to provide comprehensive coverage and timely threat detection. Through extensive testing and evaluation, we demonstrate the IDS's ability to enhance the security posture of 5G networks, ensuring robust protection against various cyber threats. This includes a detailed examination of the system's performance metrics, such as detection accuracy, false positive rate, and processing latency, which collectively underscores the system's efficiency and reliability in real-world 5G environments. Additionally, the research explores the integration of Network Function Virtualization (NFV) to deploy the IDS as a virtual network function within the 5G core. The use of NFV allows for rapid updates and reconfiguration of the IDS in response to evolving security threats, thereby enhancing its adaptability and resilience. By leveraging these technologies, the IDS can continuously learn and improve its detection capabilities, adapting to new attack vectors and strategies. By combining machine learning and NFV technologies, the IDS provides a scalable, flexible, and effective solution for safeguarding the next generation of telecommunications infrastructure. Future work will focus on further refining the IDS algorithms and exploring additional security measures to address emerging threats, ensuring the continuous protection of 5G infrastructures.

Keywords

5G network, intrusion detection system, network security, machine learning, real-time analysis, cybersecurity, network traffic analysis, security threats, 5G core network, network function virtualization

1. Introduction

The advent of 5G technology marks a significant milestone in the evolution of telecommunications, promising enhanced connectivity, reduced latency, and the capability to support a vast array of IoT devices. The deployment of 5G networks is expected to revolutionize various industries, including healthcare, automotive, and smart cities, by enabling new applications and services that require high-speed data transfer and real-time communication. Despite these advancements, the increased complexity and scalability of 5G networks present substantial security challenges. The architecture of 5G networks, which involves more extensive use of software and virtualization, introduces new vectors for cyber-attacks, making them more vulnerable to security breaches.

Traditional security measures, which were designed for earlier generations of cellular networks, often fall short in addressing the dynamic and sophisticated nature of cyber threats targeting 5G infrastructures. The move from hardware-based security solutions to Software-Defined Networking (SDN) and Network Function Virtualization (NFV) means that security protocols must evolve to address these new environments. The increased reliance on cloud

CSDP-2024: Cyber Security and Data Protection, June 30, 2024, Lviv, Ukraine
*Corresponding author.
†These authors contributed equally.
imanbaevazamat@gmail.com (A. Imanbayev); ersultan.valihan@gmail.com (Y. Valikhan); ansar.jakupov@gmail.com (J. Jakupov); odarchenko.r.s@ukr.net (R. Odarchenko)

0000-0003-3719-4091 (A. Imanbayev); 0009-0002-5950-4177 (Y. Valikhan); 0009-0003-4419-4206 (J. Jakupov); 0000-0002-7130-1375 (R. Odarchenko)

services and edge computing in 5G networks further complicates the security landscape, as data and applications are distributed across various locations, increasing the potential attack surface.

As a response to these challenges, this paper proposes the development of a network-based Intrusion Detection System (IDS) specifically designed for 5G networks. The primary objective of this research is to create an IDS that can efficiently monitor and analyze network traffic, detect malicious activities, and respond to potential threats in real time. Unlike traditional IDS solutions that may struggle with the high throughput and low latency requirements of 5G, our proposed system leverages advanced machine learning algorithms to enhance its detection capabilities.

Machine learning techniques, particularly those involving anomaly detection, are well-suited to the dynamic environment of 5G networks. These techniques can learn from historical data to identify patterns indicative of normal and abnormal behavior, allowing the IDS to detect previously unknown threats. The integration of machine learning into the IDS framework enables continuous improvement in threat detection, as the system can adapt to new attack methods and strategies.

This study explores the integration of machine learning algorithms into the IDS to enhance its accuracy and effectiveness in identifying anomalies and cyber-attacks within the 5G environment. We present a detailed analysis of the IDS architecture, including its placement within the 5G core network, the data flow between network functions, and the methods used for real-time traffic analysis. The IDS is designed to operate in a distributed manner, with components deployed across various network segments to provide comprehensive coverage and timely threat detection.

Furthermore, the research examines the use of NFV to deploy the IDS as a virtual network function within the 5G core. This approach offers several advantages, including flexibility in deployment, scalability to handle varying network loads, and ease of integration with existing network infrastructure. The use of NFV also allows for rapid updates and reconfiguration of the IDS in response to evolving security threats.

In addition to the architectural design, we discuss the implementation of machine learning models for traffic analysis and threat detection. The models are trained on a diverse dataset of network traffic, including both benign and malicious flows, to ensure robust performance across different scenarios. We also address the challenges associated with data collection and labeling, as well as the strategies employed to mitigate these issues.

Finally, we present the results of extensive testing and evaluation of the proposed IDS. The evaluation includes performance metrics such as detection accuracy, false positive rate, and processing latency, demonstrating the system's effectiveness in real-world 5G environments. The findings indicate that our IDS can significantly enhance the security posture of 5G networks, providing robust protection against a wide range of cyber threats.

This research aims to contribute to the development of secure and resilient 5G networks by providing a comprehensive solution for network-based intrusion detection. By leveraging machine learning and NFV technologies, the proposed IDS offers a scalable, flexible, and effective approach to safeguarding the next generation of telecommunications infrastructure.

2. Related works

In the realm of 5G networks, various authors have proposed diverse options for implementing security mechanisms, with many solutions focusing on the creation of an IDS using machine learning methods [1].

The concept of an IDS is well-established in network security design, leading to a wide range of implementation options. One notable solution employs the MQTT protocol [2], particularly aimed at devices based on the Cellular Internet of Things (CIoT) concept [3]. This IDS module is integrated as a network function within the 5G virtual network core, where it analyzes input traffic duplicated from other network functions via the N4 network interface, linking the User Plane Function (UPF) and the Access and Mobility Management Function (AMF) [4].

It is also worth noting that the correct classification of attacks is one of the main criteria for ensuring security in 5G networks. So by the end of 2020, the European Union Agency for Cybersecurity (ENISA) published an updated report on threats to 5G networks. The report discusses new issues related to the security of networks and various processes. It also describes changes in 5G architecture and summarizes information from 5G standardization documents. [5].

The authors of this work [6] state that the current security system is not entirely effective, as it often detects malicious traffic only after or during an attack. This is why self-learning models for cybersecurity will be necessary in the future [7].

Moreover, the use of large volumes of data generated by 5G networks allows for the identification of abnormal network behavior, significantly contributing to the development of intelligent security mechanisms. The development and implementation of intrusion detection and prevention approaches based on artificial intelligence are essential components for ensuring the security of future 5G networks and enhancing existing security systems [8].

Researchers also present threat models specific to the 5G ecosystem. In their studies, they develop an attack tree analysis methodology for examining service-oriented 5G architectures and conduct detailed vulnerability assessments, focusing on network function virtualization [9].

This article [10] proposes an intelligent identification system based on a programmable 5G architecture. In their study, the primary models are the Random Forest and the k-Nearest Neighbors method. Additionally, they incorporated boosting into their model, which provides excellent classification performance on their dataset.

On the other hand, if we return to security threats, this approach focuses on detecting and mitigating DOS/DDOS attacks [11]. This solution leverages a substantial set of training data and methodologies previously used in 4G network security [12]. Specific implementations emphasize the internal architecture of the IDS module, utilizing machine learning algorithms and neural networks. For instance, one solution employs a convolutional neural

network algorithm to detect suspicious traffic, achieving an accuracy of over 94% [1].

This paper proposes an optimal scheme for an IDS module based on Software-Defined Networks (SDN) for various types of devices, utilizing machine learning methods for enhanced detection capabilities.

One of the most promising approaches involves leveraging machine learning and artificial intelligence techniques to detect anomalies in 5G networks [1]. By analyzing large volumes of network traffic data in real time, these systems can identify suspicious behavior patterns indicative of potential security breaches. Additionally, deep learning models trained on extensive datasets enable Network Intrusion Detection (NID) systems to adapt to and learn from emerging threats, thereby improving detection accuracy and reducing false positives [1].

The integration of SDN and NFV technologies has significantly enhanced the deployment and scalability of NID systems in 5G environments. SDN separates network management from data forwarding functions, facilitating dynamic traffic analysis and policy enforcement. Meanwhile, NFV allows for the seamless creation of NID instances within virtualized network functions, promoting flexibility and efficiency.

Moreover, advances in hardware acceleration, such as FPGA-based packet processing and dedicated network processors, have empowered NID systems to meet the stringent performance requirements of 5G networks without compromising detection capabilities. These hardware solutions enable high-speed packet inspection and deep analysis with minimal impact on network latency and throughput.

By incorporating these advanced techniques and technologies, the proposed IDS scheme aims to provide a robust, scalable, and efficient security solution for 5G networks.

In the field of 5G security, one of the first significant studies was initiated by the European Union. For instance, in the second half of 2019, the Network and Information Security Directive (NIS) released a report evaluating the risks of 5G mobile networks [16]. Subsequently, the group published an important document on a toolkit for mitigating cybersecurity risks in 5G networks [13].

By the end of 2020, the European Union Agency for Cybersecurity (ENISA) published an updated report on threats to 5G networks. The report discusses new issues related to the security of networks and various processes. It also describes changes in 5G architecture and summarizes information from 5G standardization documents. [14].

The authors of this work [15] state that the current security system is not entirely effective, as it often detects malicious traffic only after or during an attack. This is why self-learning models for cybersecurity will be necessary in the future [16].

Moreover, the use of large volumes of data generated by 5G networks allows for the identification of abnormal network behavior, significantly contributing to the development of intelligent security mechanisms. The development and implementation of intrusion detection and prevention approaches based on artificial intelligence are essential components for ensuring the security of future 5G networks and enhancing existing security systems [17].

Researchers also present threat models specific to the 5G ecosystem. In their studies, they develop an attack tree analysis methodology for examining service-oriented 5G architectures and conduct detailed vulnerability assessments, focusing on network function virtualization [18].

This article [19] proposes an intelligent identification system based on a programmable 5G architecture. The advantages and similarities of this work with ours lie in the fact that they also classify traffic. In their study, the primary models are the Random Forest and the k-Nearest Neighbors method. Additionally, they incorporated boosting into their model, which provides excellent classification performance on their dataset.

Similar approaches are described in [20–21].

3. Internal and external design of the network-based intrusion detection system for 5G

The objective of this research is to design a network-based intrusion detection system for 5G networks to monitor outgoing traffic, identify and capture potential impacts in real time, and classify them using data analysis.

We need a set of input data to continue working with our module. That is, all traffic coming from the network of nodes (gNodeB) will be sent for processing to the virtual network functions of the 5G core (Evolved Packet Core), where we will have the IDS module. (Fig. 1)

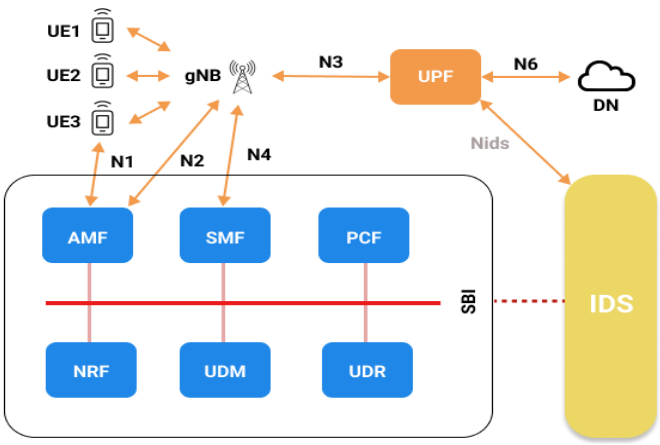

Figure 1: 5G Core network with IDS

So, since the IDS module will be located inside the 5G network core as a network function, the first step will be to register it as a network function in a special Network Function Repository (NRF). The NRF function, in turn, can take into account factors such as load potential, accessibility, and location. Fig. 2 shows an example diagram of how network function registration would occur.

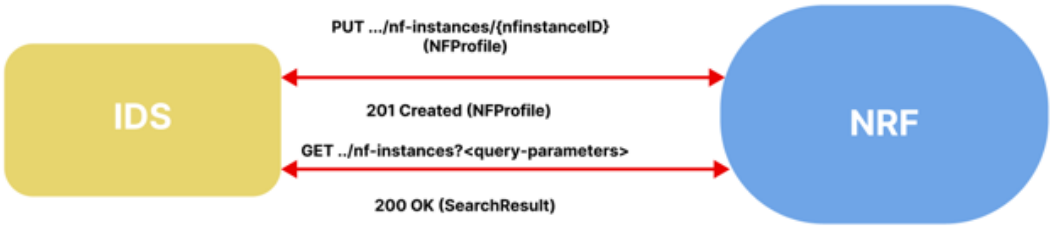

Figure 2: Registration between IDS and NRF

According to Fig. 1, the UPF function will be responsible for forwarding incoming traffic to the IDS function. So that data packets can be duplicated and forwarded to IDS using UPF, it has been connected to this function. Due to this IDS can take the necessary actions to analyze and detect potential attacks in real time. The IDS function will also have connectivity through SBI with features such as AMF and SMF for quick response and detection when suspicious traffic arrives.

If the traffic is classified as an attack, it becomes necessary to apply certain measures to alert and partially prevent the attack. This issue in the module will be dealt with by a function that will notify other network functions that can communicate with each other using the SBI (Service-based interface) interface via the HTTP/2 protocol (Fig. 1).

Given the anticipated high load of the system, driven by numerous connected devices and a substantial influx of data, scalability and fault tolerance are paramount considerations in its design. Introducing a new network function into an established 5G core implementation necessitates a careful examination of its impact.

It's essential to highlight that the IDS module will receive input data through duplication from the UPF function, streamlining its integration into the system. This approach avoids creating unnecessary dependencies between the IDS as the data recipient and the UPF as the data sender.
In essence, the implementation of the IDS module serves as a new feature that enhances the existing functionality without introducing complexities or dependencies. This streamlined integration process ensures seamless operation and facilitates the system's scalability and fault tolerance.

It is also important to note the importance of monitoring for timely response from those responsible for the stability of the system where data is exchanged. For these purposes, the module will include the collection of metrics on the volume of incoming traffic, classification, and prediction of a network attack (Fig. 3).

The primary functionality of the IDS module is to receive input traffic, process it, and transmit it to a service that will analyze the traffic to identify potential security threats and, if necessary, classify the threat type.

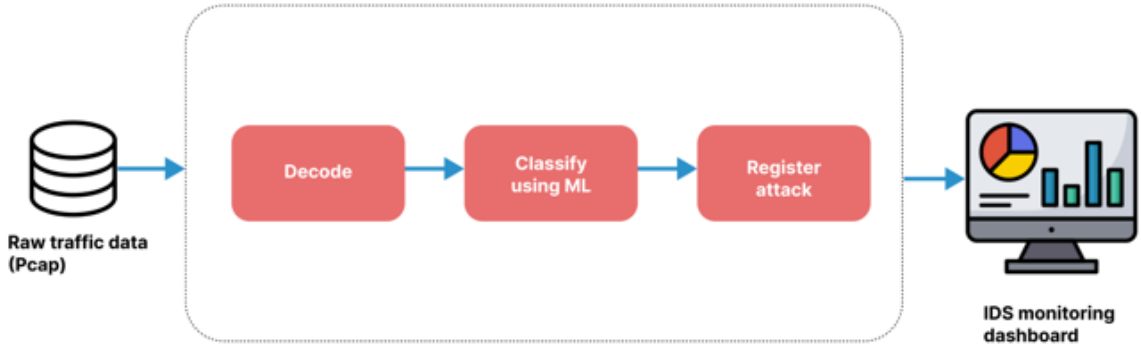

Figure 3: Internal scheme of IDS

Once our IDS was able to measure the distance in front of it, the problem of alerting other network functions residing on the common bus would be solved. We need this to pre-suppress network traffic. Through SBI interfaces, messages will be sent with changes in the data where the attack comes from and the classification of the threat. It is expected that the Session Management Function (SMF), upon receiving such an alert from our IDS module, will automatically remove the attacker's resources and terminate the established PDU session.

The AMF performs the registration blocking procedure and sets the registration state to RM-DEREGISTERED for the device user (UE).

In this context, the UE stops storing location or routing information for the UE, so the UE becomes available to the AMF. However, some parts of the UE context may still be stored in the UE and AMF, for example to avoid authentication procedures during registration of each procedure.

These measures will be very useful in the implementation of another security module in 5G networks, the main goal of which will be to prevent any threat (IPS).

4. Development of a model for a network-based intrusion detection system for 5G

Dataset

In this study, models were created and evaluated that are capable of identifying malicious traffic. In this section, we will introduce the dataset, and the models that were tested for binary classification, compare the obtained results, and choose the best model.

One of the first problems we encountered was the lack of necessary data; there are practically no open datasets with malicious traffic on the 5G network available on the Internet, which made our research difficult in terms of testing on various data. However, students from the American University provided access to generated 5G attack traffic, on which our model was built.

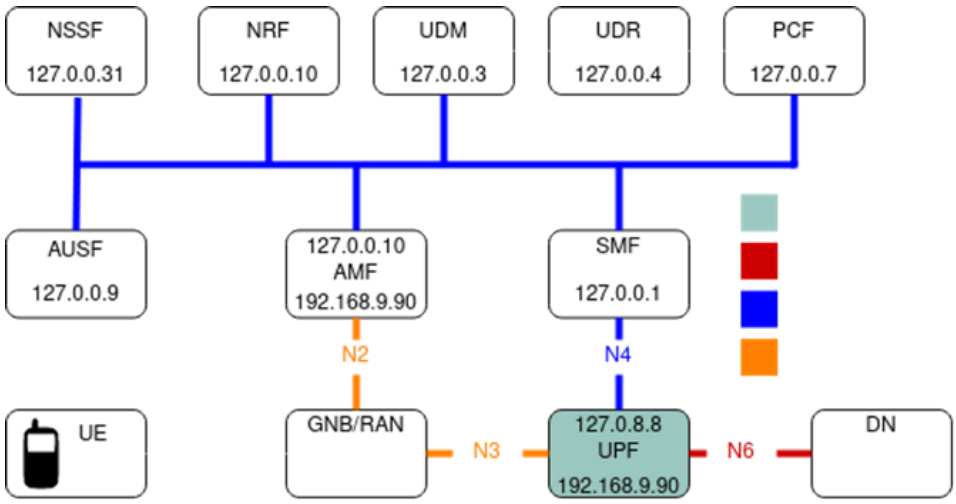

Figure 4: Traffic collection

Fig. 4 illustrates the process of collecting data on malicious traffic. The set of files containing malicious traffic consists of 10 types, each generated by attacking specific parts of the core. For example, the AMF attack type is conducted by requesting information from one of the core blocks. The issue is that this attack appears benign from the inside.

Non-malicious traffic was collected from YouTube video views, conferences in Microsoft Teams, website visits, and file downloads and uploads. In total, the dataset consists of 120,000 unique records, approximately 100,000 of which are normal traffic and 20,000 are malicious.

Processing files containing the required traffic data requires approximately 96 gigabytes of RAM and a couple of hours for each file. However, by using the sniff() function, we reduced the data processing workload, avoiding the need to store every data packet in memory.

Model

Fig. 5 represents a methodological scheme describing the overall pipeline of the conducted work. The data preprocessing process involves several stages.

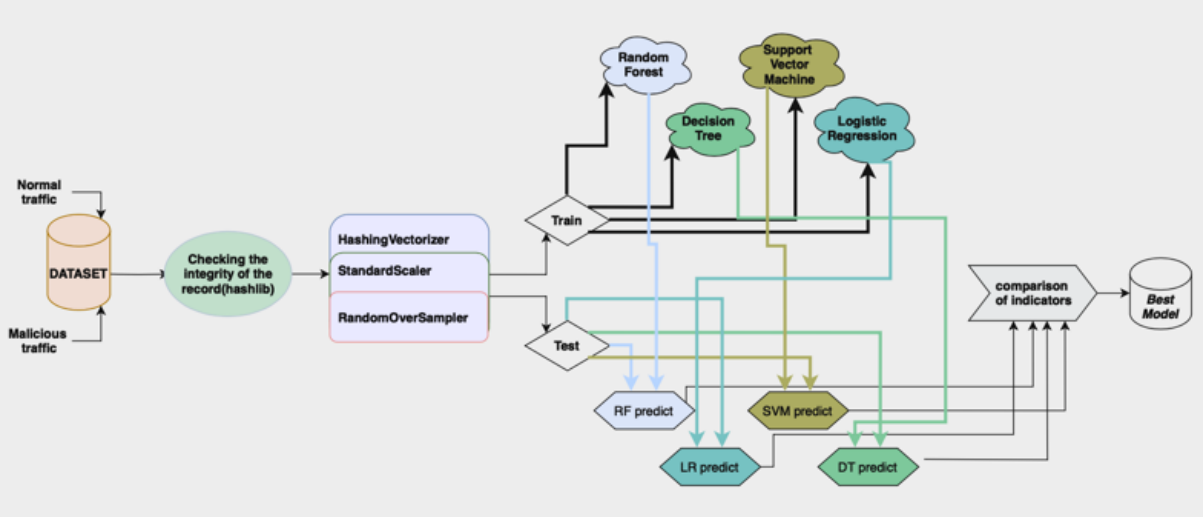

Figure 5: Methodology

The first stage includes data integrity verification, where it is necessary to ensure that the traffic indeed utilizes hash encoding. For this verification, we use the built-in Python library—hashlib. The output provides a result indicating whether the encoding matches or not.

As mentioned earlier, the number of unique records with normal traffic is five times greater than the number of records with malicious traffic. To address issues with uneven class distribution, we utilized random generation methods such as RandomOverSampler to augment the data.

The next stage involved translating our encodings into a computer-understandable language. We used the HashingVectorizer library for vectorizing our records, which will be required for model training.

The outcome was testing various machine learning models for binary classification. These models include RandomForestClassifier, LogisticRegression, DecisionTree, and SupportVectorMachine.

Results

We tested 4 development paths and arrived at the following results.

To evaluate and compare our models, we will use the following metrics: F1—score, and accuracy.

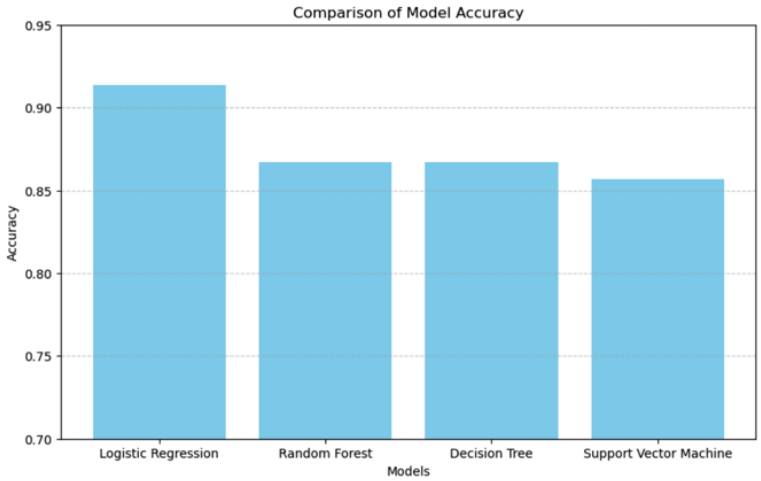

Figure 6: Accuracy comparison

In the figure above, we can observe that among the models, there is a favorite in terms of prediction accuracy based on the overall metric. We could conclude that this model suits our needs the best. However, let's take a look at other metrics as well.

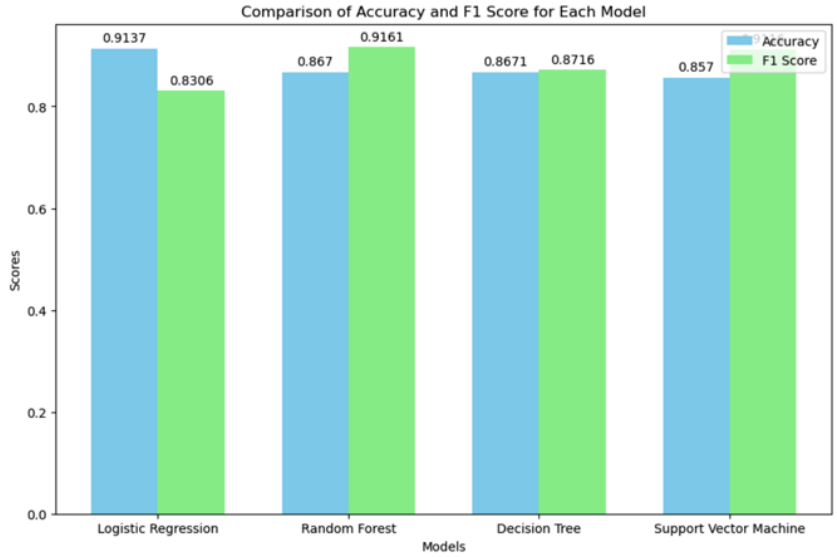

Figure 7: F1-score comparison

If we now look at the comparison of our metrics, namely accuracy and F1-score, the conclusions are not straightforward. The impact of class imbalance greatly affects the logistic regression model, as its F1 score is much lower than that of the other models, despite our efforts to balance the classes. Due to the similarity of the generated values, the model struggles to correctly identify the attack class.

If we go further and look at what values we have for each class, then it becomes more and more clear.

Table 1
F1 comparison for each class.

Class	Model			
	RFC	LR	DT	SVM
attack	0.87	0.66	0.80	0.86
normal	0.96	0.99	0.94	0.96

Based on all the aforementioned metrics and indicators, we can say that the Random Forest model performs the best in binary classification of these classes. However, the SVM

model is only a few points behind, meaning we can use the SVM model with an error only 0.01 higher. Nevertheless, it is worth noting that the SVM model takes significantly more time to train, which leads us to prefer the RFC model.

5. Conclusions

In conclusion, the implementation of a robust network-based IDS is imperative for the security and integrity of 5G networks. Our proposed IDS, which leverages advanced machine learning techniques, has proven effective in real-time detection and mitigation of security threats. The integration of this IDS within the 5G core network not only enhances its security capabilities but also ensures minimal disruption to network performance. The research findings demonstrate that the IDS can adapt to the evolving threat landscape, providing a scalable and efficient solution for protecting 5G networks. Future work will focus on further refining the IDS algorithms and exploring additional security measures to address emerging threats, ensuring the continuous protection of 5G infrastructures.

References

[1] S. Gnanasivam, D. Tveter, N. Dinh, Performance Evaluation of Network Intrusion Detection Using Machine Learning, IEEE (2024).

[2] T. Le, et al., 5G-IoT-IDS: Intrusion Detection System for CIoT as Network Function in 5G Core Network, in: IEEE Global Communications Conference (2023) 4773–4778, doi: 10.1109/GLOBECOM54140.2023.10437158.

[3] T. Moges, et al., Cellular Internet of Things: Use cases, technologies, and future work, Internet of Things 24 (2023). doi: 10.1016/j.iot.2023.100910.

[4] A. Imanbayev, et al., Research of Machine Learning Algorithms for the Development of Intrusion Detection Systems in 5G Mobile Networks and Beyond, Sensors 22 (2022) 9957. doi: 10.3390/s22249957.

[5] ENISA, The Heat Is Online. Threat Landscape for 5G Networks Report (2020). URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-for-5g-networks

[6] Y. Siriwardhana, et al., Robust and Resilient Federated Learning for Securing Future Networks, Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit) (2022) 351–356.

[7] Y. Siriwardhana, et al., AI and 6G Security: Opportunities and Challenges, Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit), IEEE (2021) 616–621.

[8] R. Santos, et al., Machine Learning Algorithms to detect DDoS Attacks in SDN, Concurrency and Computation: Practice and Experience 32(16) (2020).

[9] R. Na, et al., 5G Mobile Network Slicing for THz Services, IEEE 2nd 5G World Forum (5GWF) (2019).

[10] J. Li, Z. Zhao, R. Li, Machine Learning-Based IDS for Software-Defined 5G Network, IET Networks 7 (2017) 53–60.

[11] G. Iashvili, et al., Intrusion Detection System for 5G with a Focus on DOS/DDOS Attacks, in: 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (2021) 861–864. doi: 10.1109/IDAACS53288.2021.9661021.

[12] S. Park, et al., Threats and Countermeasures on a 4G Mobile Network, Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (2014) 538–541. doi: 10.1109/IMIS.2014.79.

[13] NIS cooperation group, The Heat Is Online. EU Coordinated Risk Assessment of the Cybersecurity of 5G Networks (2019). URL: https://digital-strategy.ec.europa.eu/en/news/eu-wide-coordinated-risk-assessment-5g-networks-security

[14] N. C. Group, The Heat Is Online. Cybersecurity of 5G Networks EU Toolbox of Risk Mitigating Measures (2020).

[15] ENISA, The Heat Is Online. Threat Landscape for 5G Networks Report (2020). URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-for-5g-networks

[16] Y. Siriwardhana, et al., Robust and Resilient Federated Learning for Securing Future Networks, Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit) (2022) 351–356.

[17] Y. Siriwardhana, et al., AI and 6G Security: Opportunities and Challenges, Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit), IEEE (2021) 616–621.

[18] R. Santos, et al., Machine Learning Algorithms to detect DDoS Attacks in SDN, Concurrency and Computation: Practice and Experience 32(16) (2020).

[19] R. Na, et al., 5G Mobile Network Slicing for THz Services, in: IEEE 2nd 5G World Forum (5GWF) (2019).

[20] O. Solomentsev, et al., Data Processing in Case of Radio Equipment Reliability Parameters Monitoring, Proceedings - 2018 Advances in Wireless and Optical Communications, RTUWO (2018) 219–222.

[21] O. Solomentsev, et al., Signal processing in case of radio equipment technical state deterioration, in: Signal Processing Symposium, SPSympo (2015).