Educational Administration: Theory and Practice
2024, 30(5), 13007-13024
ISSN: 2148-2403
https://kuey.net/

Research Article

Enhancing Security for NFV-Based IOT Networks through Machine Learning: A Comprehensive Review and Analysis

Sandeep N. Gite[1]*, Smita L. Kasar[2]

[1]*Maharashtra Institute of Technology, Aurangabad, Maharashtra, India, Email: sandeepgite37@gmail.com
[2]Maharashtra Institute of Technology, Aurangabad, Maharashtra, India, Email: smitakasar@gmail.com

***Corresponding author:** Sandeep N. Gite
*Email: sandeepgite37@gmail.com

Citation: Sandeep N. Gite, Smita L. Kasar (2024), Enhancing Security for NFV-Based IOT Networks through Machine Learning: A Comprehensive Review and Analysis, *Educational Administration: Theory and Practice*, 30(5), 13007-13024
Doi: 10.53555/kuey.v30i5.5656

ARTICLE INFO

Submitted- 15-04-2024
Reviewed- 04-05-2024
Acceptance- 22-05-2024
Published- 30-05-2024

ABSTRACT

The past several years have seen a notable surge in the adoption of Internet of Things devices, leading to the development of NFV-based IoT networks. However, security remains a significant concern in these networks, as they are vulnerable to various attacks. Machine learning is a promising solution for enhancing security in NFV-based IoT networks. This paper presents a comprehensive review and analysis of the latest research on security enhancement for NFV-based IoT networks using machine learning. The paper aims to identify the current state-of-the-art techniques used to enhance security in these networks and highlight the potential benefits of machine learning in this context. The integration of machine learning in NFV-based IoT networks can significantly enhance security against emerging threats. The absence of standards, affordable and efficient machine learning systems is only a few of the issues that must be resolved. One critical aspect is the development of effective machine learning systems capable of identifying malicious traffic and handling the vast attack surface and different attack vectors. In order to achieve this, the study examines how SDN and NFV are being adopted to protect IoT networks from new risks. This area of research holds great potential for improving the accuracy and efficiency of anomaly detection in NFV networks, keeping pace with rapidly evolving threats. In this regard, the paper provides a survey of machine learning-based algorithms for intrusion detection in NFV networks for enhancing security.

Kkeywords: IDS- Intrusion Detection System, NFV-Network Function Virtualization, BoT-IoT, recurrent neural networks (RNN-IDS), Software Defined Networking (SDN)

Introduction

The paper suggests exploring hybrid machine learning approach, incremental learning and ensemble methods and designing a framework that can detect anomalies and intrusions effectively across all functional environments of the NFV network [2]. Unsupervised techniques offer a broader resolution for intrusion detection and help identify outliers in the data that would be difficult to detect [2]. However, unsupervised approaches may have difficulty detecting advanced attacks in IoT networks [3]. Therefore, hybrid approaches combining supervised and unsupervised machine learning offer promise for intrusion detection in NFV networks, integrating the strengths of multiple techniques to enhance accuracy and efficiency in detecting anomalies in the complex NFV environment [3]. Supervised-based methods are more efficient and accurate in detecting anomalies than unsupervised-based methods [2]. However, they rely on labelled data that can be scarce in NFV environments and limit their effectiveness [3]. Researchers and network administrators need to develop methods that can automatically and accurately detect intrusions in the IoT network while identifying the source of the attack without affecting NFV functionality and quality of service [2]. These mechanisms should involve learning attack and normal profiles, abstracting security policies, and rethinking network security in IoT deployments [1]. Ultimately, the network needs to play a critical role in securing IoT deployments, and dynamic and context-aware enforcement capabilities are required to enhance security for

NFV-based IoT networks [1]. The following sections explore the relevant literature and discuss the key findings and conclusions of this review.

KronoDroid: Time-based Hybrid-featured Dataset for Effective Android Malware Detection and Characterization [1]

Technique: The current data sets only provide a static picture of a non-stationary event since they have not kept up with the growth of different types of Android malware. In the investigation of Android malware, the degenerative influence of the time variable, the literature on the effectiveness of machine learning-based classifiers has not received enough attention.

Crucial factors to take into account while developing Android malware detection systems that are more durable, robust, and dependable. This research combines a number of malware and benign data sources to produce a data set with 489 static and dynamic elements across a longer period of time. Two equally featured sub-datasets are generated by system calls, the dynamic feature source's quirks, are handled by employing an emulator and a real device. A new labelled hybrid-featured Android dataset covering the full 2008–2020 Android history, including timestamps for each data sample, is the primary product of this research. It takes into consideration the many dynamic data sources. There are 35,246 benign samples and 28,745 dangerous samples from 209 malware families in the emulator data set. 41,382 of the 36,755 innocuous software discovered in the real device data set are made up of 240 malware families.

The largest hybrid-featured Android dataset, KronoDroid, is now structured and publicly available. It is the only dataset with time stamped data, samples from more than 209 types of Android malware, taking into account the unique characteristics of dynamic sources.

Pros/Cons:

1. Information derived from hybrid analysis. More thorough information about the apps is provided by the mix of static and dynamic properties special techniques for time stamping in relation to time. A range of timestamp choices are considered and recommended.
2. It is feasible to learn from the growth of malware and develop machine learning algorithms that are more dependable, robust and long-lasting when time is taken into consideration. Labelled benign and harmful samples.
3. Applications from both categories may be applied to create efficient classifiers and conduct comprehensive application forensics investigations.
4. It was found that applications operating on real or simulated devices had dynamic data disparities.. For comparison and further investigation, the two equally feature-rich sub-datasets that comprise KronoDroid the emulator and the real device can be used.
5. A massive compilation of info. On both sub datasets, there are more than 28,000 samples in each class.

On machine learning effectiveness for malware detection in android OS using static analysis data [2]

Technique: While the Android operating system has undergone a number of security measures to increase its resilience, users are often unaware of whether an application has the potential to function as malware. Several techniques that employ static analysis data and machine learning to identify malware in an effort to solve this issue. Using static analysis data taken from the Drebin data set, system investigates the efficiency of supervised learning methods in this area. Additionally, it offers a brief overview of previous relevant studies within the field. Explains six popular classification methods in various setups with respect to their ability to identify Android malware and feature selection. Our test findings show a limited sample of features can be used to achieve excellent classification accuracy.

Pros/Cons: Different methods to measure the significance of the original features and author discuss significant works and methodologies. In this approach, we offer a comprehensive study in the field of machine learning for Android OS virus detection. The more comprehensive data collection, which is open and publicly accessible upon request, forms the foundation of our study. This research demonstrates that, even when app intrinsic qualities are the only factors taken into account, machine learning may be a potent foundation for the identification of dangerous apps. Therefore, to obtain high accuracy, (a) ML classifiers should be setup appropriately, and (b) the classification result is only affected by a small fraction of features. This issue has not been addressed in any earlier works. Furthermore, it is important to note that the majority of linked research projects do not make use of a high dimensional data collection or the entire Drebin data set. Specifically, research shows that, unlike the Drebin study, which reported accuracy as high as 94%, SVM may achieve total accuracy as high as 99%.

Towards improving detection performance for malware with a correntropy-based deep learning method [3]

Technique : Internet of Things (IoT) technologies are developing so quickly, it is becoming more and more important to detect and analyse malware in industrial applications of Cyber-Physical Systems, which use the IoT paradigm to provide a variety of services. The study of malware analysis and detection is now popular with several advanced machine learning approaches, including deep learning, and some progress has been made so far. But there are a few issues as well. For example, given the noise and outliers included in the malware datasets that are currently accessible, certain approaches are not robust enough. Therefore, there is still scope for improvement in the malware classification and detection accuracy. With this problem in mind, research suggests a novel approach that combines the deep learning model and correntropy. This recommended malware analysis and detection approach includes mixed correntropy because of its efficacy as a similarity metric in managing noisy, complicated datasets. In particular, it is incorporated into the popular Convolutional Neural Network deep learning model in order to recognise outlier features and reconstruct the network's loss function. Research describes in detail this method's design process. Moreover, the suggested technique is examined on a widely-used benchmark dataset as well as a practical malware dataset to confirm its learning efficiency.

Pros/Cons: In the context of an Internet of Things-enabled CPS application, the Android malware classifier is a challenging issue to detect and assess and this study attempts to address it. System offers a unique malware detection and analysis method based on the CNN deep learning classifier with the mixture correntropy-induced loss function, starting with the general analysis of prior work on the mixture correntropy and the CNN. Compared to other conventional classifiers, our mixture correntropy-based CNN deep learning model uses mixture correntropy, which is especially helpful for outlier learning problems, allowing the data to be handled more flexibly and steadily with improved robustness and classification accuracy. In the trials, the proposed method's classification performance is compared with other well-known approaches using benchmark datasets and real-world Android malware datasets. The experiment's results validate the study approach's effectiveness and efficiency in categorising malware data and photos.

Hybrid Model for Intrusion Detection Systems [4]

Technique: It is getting harder to detect any malicious activity right away in order to prevent the loss of important data and money due to the rising number of new attacks on the constantly expanding network traffic. One of the most popular methods for detecting network intrusions is anomaly-based detection. The accuracy varies depending on the dataset those strategies are tested on. This dataset typically does not accurately reflect network traffic. In light of this, this project entails improving KDD'99 (NSL-KDD) and analysing different ML algorithms used in IDS when tested on two datasets that are comparable to CICIDS2017. The goal of this study was to create a new hybrid intrusion detection system model after analysing various intrusion detection systems on both datasets. This novel hybrid technique combines decision tree and random forest algorithms using stacking scheme to achieve accuracy of 85.2% and precision of 86.2% for the NSL-KDD dataset and accuracy of 98% and precision of 98% for the CICIDS2017 dataset.

Pros/Cons: The results of the algorithmic analysis utilising CICIDS2017 demonstrated that this data may be utilised to test and train intrusion detection systems because it contains records of real network traffic, which aids in improving classifier training. Better detection accuracy, precision, and recall are the consequence. When tested against the CICIDS2017 dataset, the suggested hybrid model exhibits high accuracy, recall, and precision, comparable to the well-known techniques. This dataset accurately represents real-world network traffic, demonstrating the applicability of our concept for IDS implementation in real-world scenarios.

A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks [5]

Technique: Information security relies heavily on intrusion detection, where precise identification of different types of network attacks is crucial technology. In this work, author investigated the modelling of a intrusion detection system based on deep learning. In this research paper author suggested a deep learning method for intrusion detection through recurrent neural networks. Research look at how well the model performs in binary and multiclass classification and how different learning rates and neuron counts affect the performance of the proposed model. System evaluate it against ML techniques like J48, random forest, ANN, SVM and others that have been suggested by earlier researchers. The results of the experiment show that RNN-IDS is a very suitable model for building a classification model with high accuracy, and it outperforms traditional machine learning classification approaches in both binary and multiclass classification. The RNN-IDS model offers a fresh approach to intrusion detection research while also increasing intrusion detection accuracy.

Pros/Cons: Recurrent neural networks are composed of input, output, and hidden units; the hidden unit does most of the work. The RNN model is essentially a one-way information synthesis from the temporal concealment unit of the past to the timing hiding unit of the present, as well as a one-way information flow from the input units to the concealed units. The system retains end-to-end data and views concealed units as the network's general store. Research can conclude that the RNN embodies deep learning if we unwrap it. For supervised classification learning, one can employ RNNs.

Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT) [6]

Technique: In this study, researcher has examined eight popular data mining techniques. The R platform is used to run each simulation. In particular, he utilised R's H2O package to simulate DLANNs. Three authentic sensor datasets from the UCI data repository were used for the experiments29,30, and 31. Robot navigation, human activities, and body postures and movements are all classified using datasets that are gathered through the use of sensors and accelerometers. System pre-processed the datasets to make them appropriate for the

classifiers before simulating the methods. Since this was an early investigation, only included a portion of the datasets.
Pros/Cons: New data sets are brought about by the IoT paradigm, mostly through sensor device collection. One of the hardest things to do in data mining is to mine IoT data for this hidden knowledge. According to some experts, handling IoT data requires a fresh set of data mining methods. In this research, he looked at the suitability of DLANNs and other well-known data mining algorithms. Based on initial investigation, research can conclude that outcomes with comparatively higher accuracy can be obtained with C4.5, C5.0, ANNs, and DLANNS.

ANALYSIS OF SECURITY THREATS, ATTACKS IN THE INTERNET OF THINGS [7]

Technique: IoT is susceptible to several risks in its current state. Providing security assurance for data flow is one of the most important issues of IoT because every tier of the system exposes data to different types of attacks from attackers. There are layers in the Internet of Things, and each layer offers a service. The security needs vary throughout layers due to the distinct functions that each one fulfils. This study aims to investigate the many security and privacy concerns related to the Internet of Things. The paper provided a quick overview of several active security methods that function at different levels.
Pros/Cons: Researcher examined the various attack types and a few security issues, such as threats, based on this data. He used these to assess the different threats and their security implications, gaining insight into the extent of the IoT's data security vulnerabilities.

Lightweight Virtualization based security framework for Network Edge [8]

Technique: The increasing proliferation of IoT devices, from basic household sensors to intelligent industrial appliances, will make it more difficult to handle security requirements holistically. The -Saas- model, which provides on-demand security services, is receiving a lot of interest. Edge computing will be helpful in handling this by processing data locally. However, SECaaS solutions' overall performance may suffer as a result of edge nodes' diminished capabilities. The primary focus of this study is on how lightweight virtualization technologies may be used to offer virtualized security services in resource-constrained contexts. The primary emphasis of this analysis is to determine whether such a scenario is feasible and to evaluate its performance.
Pros/Cons: Orchestration of security services: Edge computing allows for the efficient distribution and coordination of several services among dispersed edge nodes in order to balance workloads. Furthermore, a number of edge nodes working together can offer improved security features. Regretfully, current orchestration systems are primarily tailored for use in data centres, and additional research efforts are required to address resource limitations and edge node dispersion.
Security of container virtualization: To achieve the required isolation between containers, container virtualization mostly depends on underlying kernel features. As a result, particular efforts ought to handle the pertinent security issues. Additionally, a sophisticated ecosystem of orchestration systems and container image repositories has sprung up around the Docker virtualization technologies. These complementing tools provide new security problems that extend beyond the traditional host domain

Flow Based Security for IoT Devices using an SDN Gateway [9]

Technique: In networked systems, the quantity of IoT-based devices is expected to rise at a nearly exponential rate, necessitating the need for a flexible and secure integration solution. The notion of Software Defined Networking facilitates network device configuration and administration being done centrally. This study proposes the use of an SDN gateway as a distributed mechanism to monitor traffic to and from Internet of Things devices. Then, this gateway has the ability to recognise unusual activity and take the necessary action. Research results indicate that, despite its ability to limit the number of installs that may happen in a second, the attack detection function is capable of successfully identifying and thwarting TCP and ICMP flood-based assaults.
Pros/Cons: The necessity for an adaptable and dynamic approach to Internet of Things security has been summarised in this study, along with the possibility of using an SDN-based gateway to solve the problem. An adaptable flow-based security system for Internet of Things devices has been developed using the Pox controller. Even though the first test results only cover a limited number of attacks blocked, they have successfully validated the strategy and offered a foundation for further development.

Using Machine Learning to Secure IoT Systems [10]

Technique: IoT has been expanding quickly over the past ten years, and by 2020, it's predicted that there will be over 25 billion devices linked together. One of the IoT's weaker areas during its expansion has been found to be security. An IoT network presents a number of issues when it comes to security implementation, two of which are the the system's heterogeneity and the overwhelming quantity of gadgets that need maintenance. with order to solve the challenges associated with protecting IoT devices, this research proposes the use of machine learning within an IoT gateway to aid with system protection. In order to identify irregularities in the data transmitted from the edge devices, it looks into utilising ANN in a gateway. Researcher really believes that this strategy can enhance IoT system security.

Pros/Cons: Securing these systems and devices is essential as several IoT devices are linked together. There is a security flaw in IoT at the moment that has to be fixed. It is important to secure the system overall, even though many people are focusing on vulnerabilities in access control and authentication. This research suggests using machine learning to identify anomalies or intrusions in an IoT system in order to do this. Researcher trained the network to identify erroneous data points by utilising neural networks. Lack of sufficient genuine data points caused us to struggle during our tests to build efficient neural networks. System was able to retrain algorithm to identify both legitimate and invalid data points by adding more data points that were invalid readings.

Security Function Virtualization for IoT Applications in 6G Networks [11]
Technique: Security Function Virtualization (SFV) is one of the key features anticipated for 6G. Similar to 5G networks' Network Function Virtualization (NFV), SFV offers fresh possibilities for enhancing security while cutting down on security overhead. Specifically, it offers a compelling solution for security-related compatibility problems. Because more IoT devices are anticipated to be connected to 5G and 6G networks, malware for these systems is becoming more and more popular among cybercriminals. To tackle these problems, this method suggests a security framework that uses SFV to softwarize security functions in order to increase user trust in IoT devices and stop malware from spreading. The results show that the proposed design may reduce the number of infected devices by 66% in just 10 seconds.
Pros/Cons: SFV is a new concept that claims to improve security and reduce the overhead of 7G networks. NFV is utilised to establish distinct networks for each category in order to isolate the devices, and a distributed ledger is employed to record the status of each device. Virtualized remote attestation techniques are used to ensure that the diverse collection of IoT devices is not incompatible with one another. Research demonstrated how the suggested system improves cost, scalability, flexibility, and malware control in addition to security. The outcomes showed that in just 10 seconds, the suggested framework was able to cut the number of infected devices by 66%.

Threat analysis of IoT networks Using Artificial Neural Network Intrusion Detection System [12]
Technique: As a paradigm, it is vulnerable to several serious infiltration concerns, nevertheless. In order to counter these dangers, an Artificial Neural Network is used in this paper's threat analysis of the Internet of Things. A supervised artificial neural network known as a multi-level perceptron is trained with internet packet traces and evaluated based on its resistance to Distributed Denial of Service attacks. The primary focus of this article is the categorization of threats and typical behaviour on an Internet of Things network. The ANN process is tested on a mock Internet of Things network. The experimental findings show that multiple DDoS/DoS attacks may be successfully detected with an accuracy of 99.4%.
Pros/Cons: In order to detect DDoS/DOS attacks on an Internet of Things network, system introduced a NN-based method in this study for intrusion detection. Part of the detection process involved categorising threat and normal patterns. Using a simulated Internet of Things network, the ANN model was validated and showed over 99% accuracy. In terms of true and false positive rates, it did well and was able to recognise various attack types with success.

Investigation on Intrusion Detection Systems in IoT [13]
Technique: Comfort and efficiency are key objectives in smart environments when it comes to improving the standard of living for people. Smart environment design is now feasible because to recent advancements in IoT technology. IoT-based systems put smart surroundings at risk for security breaches. In the age of the Internet of Things, big data analytics which may be utilised to identify patterns and anomalies in data is based on the data generated by connected devices.

Most cyber security systems use IDSs, which are utilised by many different approaches and architectures, to identify intrusions. Rather than analysing monitored events against a database of known intrusion experiences, anomaly-based intrusion detection systems learn the typical pattern of system behaviour and warn on unexpected events that occur. Because IoT's use modern information technology, smart grids have become the ideal intrusion target due to the utilisation of sensor devices to collect data from these environments.
Pros/Cons: The characteristics of every IDS technique were compiled in this report. This study uses deep learning design concepts to identify an efficient and effective intrusion detection solution IoT. Through an analysis of this fascinating and ever-evolving field of research, this review might potentially benefit security researchers greatly. It will support researchers searching for cutting-edge IDS solutions to manage communication-related IoT security.

Intrusion Detection System Using Recurrent Neural Networks and Attention Mechanism [14]
Technique: Sensitive data must be protected over the internet using specific algorithms that can identify even the smallest irregularities in the data. Since most of the data contains redundant information that is not useful for classifying data into normal and abnormal categories, it is more difficult to identify these abnormalities. Finding the best features is still a challenging endeavour. Researcher proposes an attention oriented RNN for intrusion detection in large scale networks, which will optimise the model to focus on the most relevant features

and train much quicker. The model is tuned to concentrate on the key characteristics for classification while utilising the full potential of the GRU network. The attention vector that is added to the network allows for the latter. The network's efficiency is assessed using a number of metrics, including F1-score, recall, accuracy, and precision. Researcher evaluates these metrics across the most recent classification algorithms to see how well our suggested model works in comparison to the existing methods.
Pros/Cons: This study proposes new intrusion detection architecture. The attention mechanism in the architecture is capable of concentrating on the best features to identify minute variations in network data. Using two widely used benchmark datasets, the system examined the suggested model and discovered that it performs much better in terms of accuracy, precision, recall, and F1-score.

A Novel SDNFV IoT Architecture Leveraging Softwarization Technology Services to Alleviate IoT Network Resource Restrictions [15]

Technique: An innovative architecture that integrates IoT with SDN and NFV was suggested as a way to lessen the restrictions on IoT resources by softwarization. The suggested architecture is then constructed and simulated using an IoT, SDN, and NFV-compatible mininet-IoT emulator. 6LoWPAN hosts are utilised as IoT devices, OVSwitch is used as a virtualization service, thus in the simulation, the RYU controller serves as an SDN controller. The results of the trial show that the SDNFVIoT Architecture outperforms the entry-level IoT system by a significant margin.
Pros/Cons: Integrating technology with the most significant is the Internet of Things step towards satisfying the vast demands of the smart environment. In this study, architecture for fusing IoT with SDN and NFV is proposed. As stated in the system, the suggested architecture consists of four primary levels, each of which has sub-layers representing the three technologies and the interactions between them. The three technologies are combined in the SDNFVIoT architecture that this study proposes to use. Additionally, it displays the protocol stack for every technology across all tiers and a proof of concept. The purpose of this work is to integrate SDN and NFV with the IoT

Evaluating Convolutional Neural Network for Effective Mobile Malware [16]

Technique: In this work, researcher examines the capacity of deep learning techniques to discriminate between real and fake Android samples. In order to do this, he developed a Convolutional neural network-based approach that uses dynamic analysis to analyse syscall instances. Utilising a recent dataset with 7100 real-world applications more than 3000 of which are malware the built deep learning classifiers were experimentally evaluated.
Pros/Cons: Because of the extensive usage of mobile devices and the growing volume of private and sensitive user data they contain, malware writers find it increasingly easy to create increasingly aggressive software specifically targeted at mobile platforms. In this, researcher suggests a deep learning-based Android malware detector. Specifically, he examines the application of Convolutional Neural Networks to develop a model assessing their efficacy in trustworthy and malware classification. Using 7100 real-world mobile apps, the system evaluates the proposed strategy; the accuracy ranges from 0.85 to 0.95.

SEIRS epidemic model with delay for transmission of malicious objects in computer network [17]

Technique: The computer network's SEIRS of malicious items is developed, with a constant high fatality rate for contaminated nodes and a constant death rate from causes other than malicious object attacks. When a node in a computer network dies, it is similar to state that it is isolated from the network, which can still transmit harmful items even when anti-malware software is running continuously. A collection of integro-differential equations make up the model. A node that has been recovered from the infected class experiences a brief immunity acquisition with probability p (0 6 p 6 1) and a probability of 1 p when it dies due to a malicious object attack. The free equilibrium of malicious objects is examined, and the threshold parameter is used to represent how stable the results are.
Pros/Cons: In a computer network with continuous latent and immune phases, a fluctuating populace, A model for the SEIRS pandemic harmful object transmission has been created. Diekmann and Heesterbeek's SIR epidemic models have been extended according to SEIR type (A) with constant latent period s and constant exposure time x. According to this model, upon a node's removal from the infected class, it either dies or recovers temporarily, acquiring temporary immunity with probability p (0 6 p 6 1) from the attack of malicious objects. Yan and Liu [10] assumed that a node's recovery from the infected class would result in a probability of acquiring permanent immunity.

UNSW-NB15: A Comprehensive Data set for Network Intrusion Detection systems [18]

Technique: One of the primary research issues in this area is the lack of a comprehensive network-based data collection that can reflect modern network traffic situations, a variety of low-footprint incursions, and depth-structured information about the network traffic. Ten years ago, benchmark data sets KDD98, KDDCUP99, and NSLKDD were created for the purpose of evaluating research efforts related to network intrusion detection systems. Several recent studies have shown that, in the current network security environment, these data sets are not representative of network traffic and modern low-footprint assaults.

This study explores the creation of a UNSW-NB15 data set in an effort to solve the problem of network benchmark data sets not being easily accessible. This data set's network traffic assault activities are a cross between the real, contemporary normal and the contemporary manufactured ones. The UNSWNB15 data set's features are produced using both cutting-edge and established methodologies. You can use this link to access the data set, which is open for study.
Pros/Cons: This paper shows that the existing benchmark datasets are not adequate to represent the modern nature of network traffic and attack scenarios. UNSWNB15 is constructed in the UNSW cyber security lab's synthetic environment. The primary IXIA tool utilised in the synthetic environment has enabled the ability to provide a modern representation of both synthetic abnormal network traffic and the real, modern normal. UNSW-NB15 illustrates nine main attack families with the IXIA Perfect Storm tool.
49 features have been generated using Argus, Bro-IDS tools, and twelve algorithms covering aspects of network packets. However, the existing benchmark data sets such as KDD98, KDDCUP99, and NSLKDD only found a limited amount of attacks and outdated packet data.

Deep neural architectures for large scale android malware analysis [19]

Technique: This speed problem can be addressed by machine learning, which automates the classification process. While there have been numerous attempts to detect Android malware using conventional machine learning approaches, there hasn't been a good enough attempt to employ the more recent DL models in this system. In order to identify Android malware from a sizable dataset of more than 55 GB of Android malware, researches used a variety of deep learning models in this research. In order to offer insights into the dataset and assess its effectiveness in contrast to deep learning-based models, he also used Bayesian machine learning to this issue area. System demonstrates that by employing these models, researcher can attain superior outcomes in comparison to the most advanced methods.
Pros/Cons: Researcher has provided a thorough description of a DL application for Android malware analysis in this research. He presented the empirical findings from several deep learning models. He also investigated a different strategy called Bayesian machine learning, which produces models that are much easier to understand, because all of these models greatly hinder the interpretation of their findings. He gave a thorough explanation of our experiment's findings and showed how study findings equal or even exceed the most accurate state-of-the-art outcomes when they are applied to a sizable dataset of Android malware.

Classification of Android Apps and Malware Using Deep Neural Networks [20]

Technique: Malware that targets mobile devices is a common issue in today's world. In essence, malware detection is a software classification problem that uses data from programme analysis. This study looks at the effectiveness of DNNs in the classification of Android applications using system API-call sequences. DNNs' capacity to pick up intricate and adaptable traits could result in the prompt and efficient detection of malware. Researchers designed a Convolutional Neural Network for sequence classification and conducted malware detection and software functionality grouping experiments in order to assess and compare our CNN with classifications by recurrent neural networks and other n-gram based techniques. LSTM and CNN both performed noticeably better than n-gram based techniques. Surprisingly, CNN performs far better than the LSTM, which is thought to be the best option for sequential data.
Pros/Cons: In this research he demonstrated the utility of Convolution Neural Networks, a type of Deep Neural Network, for software categorization. This could result in malware classification and detection that is both successful and efficient. This study is one of the first attempts to construct DNNs for Android application/malware classification. The outcomes of our experiments demonstrated that DNNs outperformed the conventional n-gram-based approaches by a significant margin.
Although this research showed that DNNs are a useful tool for classifying Android apps, there is still much space for improvement. Specifically, researcher has concentrated within an application's core code, on the API-call sequences. Researcher thought that an improved software classifier can be created by combining data from various parts or facets of an application.

A SDN-IoT Architecture with NFV Implementation [21]

Technique: SDN with cloud computing and NFV, is expected to become a key enabler that will drastically change how network operators design and make money off of their infrastructure. However, the IoT is revolutionising how cyberspace and physical space interact, having a profound effect on day-to-day living. The issue's astounding breadth and the new, strict criteria for dependability, performance, and security will necessitate the development of new methodological and engineering techniques in order for these technologies to be effective.
In order to tackle the new difficulties posed by the Internet of Things, A simple and broad SDN-IoT architecture with NFV implementation is proposed in this study. It provides particular recommendations on where and how to use SDN and NFV techniques. The architecture's flexibility, which opens up new possibilities for the quick deployment of software-enabled global services, will spur innovation in the IoT space. The business viewpoint is also looked at in the research, where SDN and NFV are seen as enablers of new value-added services that can be added to the existing infrastructure. By employing rapidly deployable services across the value chain, this approach increases revenue prospects.

Pros/Cons: SDN and NFV are revolutionising the telecom sector, and they will serve as the foundation for a huge number of businesses in the future. System specified when and how to use NFV and SDN within an IoT "framework." Scalability and mobility challenges in IoT networks were addressed with a general-purpose, straightforward SDN-IoT architecture combined with NFV. In an IoT context, being dynamic, scalable, and elastic is made feasible by virtualizing the IoT gateway. By integrating SDN and NFV into the current infrastructure, research also examines the business side of things. It is clear that this strategy presents more chances for income generation along the services value chain.

A Multi-Perspective malware detection approach through behavioural fusion of API call sequence [22]

Technique: The behavioural difference between calling sequences that are harmful and those that are not was demonstrated by our behavioural models. Researcher constructed many relational viewpoint models that describe process behaviours based on that distinction. He tested method on the Windows and Android platforms to demonstrate its originality. These tests showed that suggested approach detected harmful samples that were not yet visible with minimal false positive rates and a high degree of precision. These model's average detection accuracy for the hidden samples used for testing malware on Windows and Android is 0.997 and 0.977, respectively. Furthermore, researcher suggested a unique method of API indexing that considers contextual similarities. Additionally, he proposed a new expressive form that visualises the sequence of API calls. As a result, we decided to include a confidence metric in our model categorization. Additionally, he created a behavioural heuristic that was successful in detecting malicious API request sequences that were mimicking or deceitful.

Pros/Cons: (1) The analytical procedure is made more difficult by the vast array of API functions. (2) The majority of the time, writers of malware use imitation techniques to make their dangerous programmes appear legitimate. To conceal their harmful behaviour, they thus create a significant number of API calls that match to typical events. Such a dishonest mindset increases ambiguity and complicates the analysis process. (3) Preventing exploratory attacks from compromising internal model processes.

Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset [23]

Technique: As Internet of Things systems proliferate, malevolent entities have begun to target them. Realistic investigation and protection countermeasures must be created in order to handle this. Network forensics and intrusion detection systems are two examples of such countermeasures. A representative and well-organized dataset is therefore essential for training and confirming the legitimacy of the systems. Despite the fact that there are multiple networks, most of the time not much is disclosed regarding the scenarios using botnets that were used. This study proposes a new dataset named Bot-IoT. It includes simulated and real-world IoT network traffic as well as several kinds of attacks.

Additionally, research offers a practical test bed environment to mitigate the shortcomings of current datasets, which include incomplete network information capture, imprecise labelling, and a range of recent and sophisticated attacks. System compares the BoT-IoT dataset to other datasets and assesses the dataset's trustworthiness for forensic purposes using various statistical and machine learning techniques. The foundation for botnets identification across IoT-specific networks is provided by this work.

Pros/Cons: The new dataset, Bot-IoT, is presented in this study. It includes a variety of attack traffic types frequently employed by botnets, in addition to regular Internet of Things and other network traffic. Using a realistic test bed as the development platform, this dataset was labelled, indicating an attack flow as well as the attack's category and subcategory for potential multiclass classification needs. More characteristics were developed in order to increase the prediction ability of the classifiers trained on this model. Statistical analysis produced a subset of the original dataset made up of the top 10 characteristics. Ultimately, four measures were particularly utilised to compare the dataset's validity i.e. fall-out, precision, recall and accuracy. Research contends that improved outcomes could be obtained with additional model optimisation.

On the Race of Worms, Alerts, and Patches [24]

Technique: This research offers a methodical approach for assessing how well automatic patching solutions operate. Researcher applies it to measure the rate at which patches or alerts must spread in order to contain worms. Concerns regarding scalability and trust motivate research into a hierarchical system where network hosts are organised into subnets, each of which has a patch server. Patches are distributed to end hosts inside subnets following verification and to super hosts via an overlay linking them. The analytical framework's unique abstraction of a minimal broadcast curve allows for the support of several overlays. Filtering of scans across subnets is also supported. System designers can use the quantitative estimations provided by the framework to help them with the sizing of patching systems that operate automatically. The outcomes are computed theoretically and confirmed via modelling.

Pros/Cons: According to our calculations, (i) patching is only going to be successful provided that the ratio of the proportion of the host population that is not patched and the rate of worm infection to patching rate are not excessively high. (ii) The concept of minimum broadcast curve may prove useful (iii) Worm-like patch distribution works well, provided there is no host quarantining and the verification and installation time is not too long.

IoT Security Techniques Based on Machine Learning [25]
Technique: Identity-based attacks, such as spoofing and Sybil attacks, can be identified and prevented by IoT devices by using authentication [8]. Unauthorised users cannot access IoT resources due to access restriction. IoT devices can access the processing and storage power of edge devices and servers for computationally demanding and latency-sensitive processes made possible by safe offloading techniques. Malware detection shields Internet of Things (IoT) devices from malware including viruses, worms, and Trojan horses that can compromise privacy, drain battery life, and impair network performance.
Pros/Cons: Partial state observation: Current reinforcement learning (RL)-based security approaches count on each learning agent to be aware of the correct state at all times and to assess the instant reward for every action taken. Furthermore, the agent has to put up with the poor strategies, particularly in the early stages of learning. Nevertheless, IoT devices typically struggle to determine the network and attack state precisely, and they must prevent a security catastrophe brought on by a poor policy at the start of the learning process. Overhead in computation and communication: A lot of the current ML-based security techniques include high overhead in computation and communication, necessitating a lot of training data and a laborious feature-extraction procedure.

Distributed denial-of-service attack mitigation in network functions virtualization-based 5G networks using management and orchestration [26]
Technique: The Internet of Things is predicted to connect billions of devices with the help of fifth generation technology. All sorts of cyber attacks will, however, eventually target Internet of Things devices. The term DDoS assault refers to the most prevalent type. NFV offers a significant potential to offer the advantage of elasticity and low-cost solutions for defending 5G networks in order to resist such threats. In light of this, the study suggests a novel defence against DDoS assaults in 5G NFV networks. The suggested method uses virtual machines from an intrusion prevention system to intercept the queries. IPS uses management and orchestration to dynamically deploy its virtual machines based on the amount of DDoS traffic, distributing the load. Experiments are carried out in a real 5G NFV environment constructed with 5G NFV environment tools in order to assess the efficacy of the process. To the best of our knowledge, this is the first experimental testing of an NFV-based method for 5G network DDoS mitigation in a real-world NFV environment. The experimental findings confirm that the suggested approach is capable of efficiently mitigating DDoS attacks.
Pros/Cons: By include packet entropy in decision mechanisms in further work, researcher can strengthen the validity of system's mitigation strategy for a real-world implementation. Later, other parameters might be included in a later version of MANO, and additional policies might be introduced later on. For example, dynamic resource management has been the main emphasis of this article in order to maximise the elasticity and efficacy of NFV.
It is believed that the IPS can fully identify the flood attacks in the test cases. Container-based virtualization is one of the solutions that are planned in case the number of "scaled out" IPS VMs exceeds capacity and causes a decline in QoS. The primary benefit of Kubernetes is its integrated auto scaling mechanism. Therefore, to administer these scaling policies, MANO is not required. Still, there aren't many Telco NFV solutions now using Kubernetes.

Virtualized Network Function Orchestration System and Experimental Network Based QR Recognition for a 5G Mobile Access Network [27]
Technique: In addition to being able to establish adaptable network architecture and deployment, this system should enable connectivity between network devices. This system gives access networks more of an emphasis. Researcher tested if it is feasible for new services to be adopted quickly and if it is possible to manage network resources effectively by conducting trials with various user situations services being created and turned on within a network. The suggested approach relies on mesh networking and Bluetooth transfer technology to establish automated connections across network workstations. It also makes use of Docker Form, a container virtualization technology, to configure and manage important features. Furthermore, the system is equipped with a Docker platform-based clustering and recovery mechanism for network function.
Pros/Cons: A network orchestration system built on network virtualization technologies was presented in this study. Docker-based virtual network function was built up and open-source hardware was employed as the network for actual performance testing. This demonstrated how service availability was maintained by shifting network functions according on node CPU usage. Furthermore, a quicker method of returning to the pre-reset environment was suggested in this research to address the restart issue that arises when the network is reset. This enhanced the structure of the container, which allowed it to have the same continuity in the system as virtual machines, despite the latter's inability to sustain service. This research also proposed an enhanced network orchestration system user service and evaluated the performance of a two-dimensional QR code recognition technique.

A Novel Intrusion Detection Approach Using Machine Learning Ensemble for IoT [28]

Technique: IoT has grown more and more relevant because it may be used in a variety of settings. The IoT's usage of a scalable and flexible infrastructure is another factor contributing to its influence. Cybercriminals have been drawn to the Internet of Things due to its widespread and varied application in recent years. They discourage current and new stakeholders by taking advantage of the open-source IoT framework's weaknesses caused by the lack of strong and uniform security procedures. In order to stop malevolent actors from gaining access to the Internet of Things network and its accessories, the authors suggest a binary classifier technique that was created from ML ensemble method. The model's positive class performance metrics yielded 98.27% accuracy, 96.40% precision, and 95.70% recall. The suggested model's efficacy against cyber attacks is demonstrated by the simulation results, which qualify it for important Internet of Things applications.

Pros/Cons:

• An emphasis on mobile network devices, such as Internet of Things-enabled IDS platforms. • A binary classifier based on machine learning that distinguishes between harmful and benign network traffic in order to offer network security. The model underwent testing, validation, and training using the ensemble approach of random forest and GBM with a hyper parameter optimizer, utilising the "CSE-CIC-IDS2018-V2" dataset. SQL injection and insertion attacks were used to demonstrate the performance test.

Efficient Detection of DDoS Attacks Using a Hybrid Deep Learning Model with Improved Feature Selection [29]

Technique: Research introduced a powerful hybrid deep learning model that is improved by the feature selection method. Effective use of the proposed deep learning methodology has already been demonstrated in a number of applications, such as extremist association identification, rumour classification and intent detection. After developing an x2 test for feature selection, research classified DDoS assaults using a CNN + BiLSTM hybrid model. Thus, the proposed method might use CNN and BILSTM layers in conjunction with optimal feature selection to forecast the results of DDoS attacks from data.

Pros/Cons: Using benchmark data, the authors created a feed-forward DL-based method for predicting DDoS attacks. Through the use of a DL technique called a Feed Forward DL classifier, it attempted to predict DDoS attacks. Based on benchmark data, traditional DL classifiers might not offer a useful method for anticipating DDoS attacks. For many reasons, such as poor choosing predictors, using machine learning classifiers after classical feature sets, it is challenging to use benchmark data to forecast upcoming DDoS attacks.

Network Intrusion Detection Model Based on CNN and [30]

Technique: The three main stages of the proposed network intrusion detection method, known as the CNN-GRU model, are as follows: first, original data is transformed into numerical features and normalized; next, the dataset is balanced by the ADRDB algorithm; finally, the features are extracted by the RFP algorithm and ultimately transformed into a grey-scale map; Second, there is the training phase, when the Convolutional Block Attention Module (CBAM) uses residuals to give various weights to the features in the pre-processed data. The CNN module first extracts the spatial features, and then it combines Average pooling and Max pooling to further aggregate the spatial data. Afterwards, several GRU units extract the temporal features.

Pros/Cons: Incomplete feature extraction and generic multi-classification effects are common problems with traditional intrusion detection methods. This paper suggests an intrusion detection approach that combines gated recursive units and Convolutional neural networks to overcome these issues. The ADRDB and RFP algorithms are used by the model to address the issues of data set imbalance and feature redundancy. Comprehensive and sufficient feature learning is then achieved by combining CNN and GRU, and an attention module is added to assign different weights to the features, thereby lowering overhead and enhancing model performance. Based on the NSL-KDD dataset, UNSW_NB15 dataset, and CIC-IDS2017 dataset, the suggested model's accuracy is 99.69%, 86.25%, and 99.65%, respectively. Moreover, the accuracy is 99.65%, 86.92%, and 99.63%.

Using feature selection analysis experiments, hybrid model versus single model comparison experiments, feature extraction method comparison experiments, pooling method comparison experiments, and performance analysis experiments on the dataset, this paper shows that the model has a strong feature extraction capability, high detection accuracy, and low false alarm rate when dealing with large-scale high-dimensional network data. It also greatly improves the detection effect for a few classes, which offers promising potential real-time applications for intrusion detection systems.

IoT Intrusion Detection Using Machine Learning with a Novel High Performing Feature Selection Method [31]

Technique: Intrusion Detection Systems are vital instruments for self-defence against a range of cyber attacks. However, functional and physical variety presents major hurdles for IoT IDS systems. Utilizing these features and traits for IDS self-protection is difficult and impractical due to these IoT characteristics. This study proposes and implements a novel feature selection and extraction method for anomaly-based intrusion detection systems. The method starts with selecting and extracting pertinent characteristics in a range of ratios utilizing two entropy-based techniques (information gain and gain ratio.

On IoTID20, system approach produced 11 and 28 relevant features, respectively, using the intersection and union. On NSL-KDD, it produced 15 and 25 relevant features, respectively. Researcher also contrasted methodology with that of other cutting-edge research. The comparison shows that, with a very high classification accuracy of 99.98%, our model is more capable and superior.

Pros/Cons: The contributions of this paper include that we:

• Provide a filter-based approach to optimize the FS process utilizing the IG and GR methods, which combine many approaches to extract only the most crucial information. • Use the intersection and union theory concepts of mathematical sets to create a hybrid feature selection strategy (referred to as hybrid here since it combines two filter-based feature ranking techniques, IG and GR that extract the minimum and maximum of the best relevant features. There are two feature selection modules in the suggested approach. The most pertinent features from the previous phase are chosen by the first module using the intersection rule. While using the union rule, the second module fulfils the same function as the first. These modules choose the most relevant and optimal features, which are subsequently supplied to ML classifiers in the subsequent stage for the ensemble.

Thus, our hybrid offers an easy-to-use, practical, efficient, yet effective methodology that performs better than previous approaches while requiring less training time. Offering comprehensive experimental results to shed light on the suggested strategy as a practical and all-encompassing IoT ecosystem IDS solution methodology.

A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks [32]

Technique: Information security relies heavily on intrusion detection, and the critical element here is the ability to precisely identify different types of network intrusions. In this study, researcher investigated the modelling of an IDS using deep learning, and suggested a deep learning methodology for intrusion detection through the use of RNN-IDS. On the benchmark data set, researcher compared it with those of J48, artificial neural networks, random forests, support vector machines, and other machine learning techniques suggested by other researchers. According to the experimental findings, RNN-IDS performs better than typical ML classification methods in both binary and multiclass classification, making it an excellent choice for developing classification models with high accuracy.

Pros/Cons: Input, output, and hidden units are the components of recurrent neural networks; the hidden unit completes the majority of the work. Researcher can think of hidden units as the network's overall store, retaining end-to-end data. Upon dissecting the RNN, research discovered that deep learning is embodied in it. One possible method for supervised classification learning is to employ RNNs. The input of the hidden layer is influenced not only by the output of the input layer but also by the output of the final hidden layer.

Intrusion Detection Framework for the Internet of Things Using a Dense Random Neural Network [33]

Technique: A DnRaNN's central concept is taken from the human brain's architecture. Dense cell clusters can be found in several significant regions of the human brain. These clusters could include a range of distinct cells or similar cells. Through synapses and dendrites, these clusters can facilitate significant communication due to the density of their arrangement. Thus, in order to construct a DnRaNN, a mathematical model of dense clusters is created that can represent synapses as well as soma-to-soma contacts. Four cluster hidden layers, one output layer, and one input layer make up this model. Dense clusters provide a framework for multilayer design, where a limited quantity of compact nuclei is present in each layer. Through synapse and soma-to-soma contacts, every nucleus has communicating cells with one another in a fully integrated architecture. The classic MLFF architecture serves as the foundation for the communication framework between the concealed layers. An input layer provides the excitation signal to the first layer's nuclei.

Pros/Cons: The main contributions of this article are listed below. 1) It suggests using a brand-new dense random neural network for IoT network intrusion detection and categorization. 2) A thorough evaluation of the suggested strategy is carried out using the ToN_IoT, a new generation IoT security dataset. 3) Several performance assessment measures, such as accuracy, precision, recall, and F1 score, are used to assess the DnRaNN's performance in binary class and multiclass settings. 4) Lastly, it provides a thorough analysis of how well the suggested approach performs in comparison to a few well-known ML/DLbased intrusion detection systems.

Random-Forests-Based Network Intrusion Detection Systems [34]

Technique: It is impractical to expect security breaches to be entirely prevented by the security technology now in use. Intrusion detection is therefore a crucial part of network security. Nevertheless, a large number of IDSs in use today are rule-based, which limits their ability to identify new incursions. Furthermore, encoding rules require a lot of time and heavily rely on the awareness of known invasions. Thus, system provide novel, structured frameworks for anomalous, hybrid network-based intrusion detection systems that leverage the data mining algorithm known as random forests. The random forests technique automatically builds patterns of incursions over training data for abuse detection. Subsequently, intrusions are identified by comparing network activity with the patterns.

The random forests algorithm's outlier detection technique detects novel intrusions in anomaly detection. The outlier detection method identifies outliers associated with the patterns after the random forests algorithm builds the patterns of network services. By integrating the benefits of both anomaly detection and heavy usage, the hybrid detection system enhances detection performance.

Pros/Cons: The following is a summary of this paper's main contributions.
1) Provide fresh, methodical frameworks for network ID that use the random forests technique. Up till now, automated intrusion detection has not been implemented using the random forests approach. 2) To enhance the IDS's efficacy in misuse detection, use sampling strategies and feature selection algorithms. The detection rate of minority intrusions is increased by the sampling procedures. The total detection performance is enhanced by the features selection technique. 3) Use a novel approach to anomaly detection in service-based unsupervised outlier detection. The program generates outliers associated with the constructed patterns by constructing patterns of network services. Attack-free training data, which are hard to come by in real-world network environments, are not required for the suggested approach. 4) Integrate anomaly and misuse detection. The earlier IDSs' overall performance is enhanced by the combo.

Machine Learning for Detecting Anomalies and Intrusions in Communication Networks [35]
Technique: As a result of their increasing sophistication, cyber attacks are become harder to identify. An essential component of cyber security is the use of efficient and effective machine learning techniques to identify anomalies and breaches in networks. Different machine learning algorithms have been used to assist in identifying network users' malevolent intents. In order to categorize known network incursions, research assessed the effectiveness of recurrent neural networks as well as the Broad Learning System with its extensions in this work. By extending the network structure and employing different subsets of input data, the methods can be utilized to create generalized models. The models' performance is assessed using the following metrics: training time, F-Score, accuracy, and chosen features.
Pros/Cons: The recently suggested approaches made use of a feature selection algorithm together with varied numbers of mapped features and groups of mapped features with and without cascades. There are variants for incremental learning for both techniques. A process for identifying network intrusions was also covered. According to performance evaluation, BLS with improvement node cascades needed a lot more time to train.
Observe that no retraining of the model was required in the gradual BLS scenario. Larger mapped feature counts, mapped feature groups, and enhancement nodes all necessitated more memory and longer training times, as was to be expected. Additionally, a single experiment with integrated steps for feature selection and model generation is used to construct the VFBLS and VCFBLS models. By assigning fewer features produced from the input data, these models' training times may be shortened.

An efficient combined deep neural network based malware detection framework in 5G environment [36]
Technique: An effective framework based on hybrid deep neural networks is proposed for Android malware detection, which can distinguish between malware and benign apps with speed and accuracy.
• Pre-detection uses the random forest in conjunction with quickly derived permission features to identify key traits. Pre-detection results can increase the detection effectiveness of big samples and decrease the number of samples to be evaluated in the deep detection phase.
• Real-world data set experiments show how effective the suggested DLAMD is. Pre-detection and deep detection phases were the subjects of experiments, and both phases' accuracy was above 93%. In comparison to a single step, the accuracy is improved by roughly 2%–3% in the experiment of the entire detection framework. There is an approximate 10%–20% increase in accuracy when compared to typical machine learning methods in comparison trials.
Pros/Cons: (1) The smali file that was retrieved through decompilation contains Dalvik instructions. The opcode is derived from Dalvik instructions included in smali files, including "return-void," "invoke-direct," and "iput-object." Each of these opcodes represents a distinct operation and has sequence information according to appearance order. An APK can be thought of as an opcode sequence made up of other opcodes. Currently, identifying dangerous and benign apps using opcode features and examining the various ways that related instruction sequences manifest on their labels have changed the problem of malware identification.
(2) There are hundreds of opcodes, and some will show up more than once in various locations. Thus, it makes sense to think of representing APK behaviour with opcode sequences. On the other hand, an opcode sequence that is too long will introduce redundant information and lower detection efficiency. As of right now, it is possible to break down the issue of malware detection using opcode features into two smaller issues. The inefficiency brought on by length dependence depending on opcode characteristics is one sub problem.

Prediction and Detection of FDIA and DDoS Attacks in 5G Enabled IoT [37]

Technique: A Markov stochastic process is used for the behavioural analysis in order to forecast and identify attacks. By introducing several security ranges with different thresholds, the stochastic modelling seeks to determine the behaviour of IoT gadgets within the framework according to their log file. The latter keeps track of every interaction that takes place between the system and IoT devices. Researcher believes that any gadget is capable of carrying out the three fundamental elementary tasks of reading, updating, and deleting. One of these simple tasks can be maliciously carried out by a device to create an attack. A denial-of-service (DDoS) assault, for example, happens when a device overloads the system with reading activity. Security assaults are divided into two groups: damaging and harmless attacks. An innocuous attack may unintentionally occur from a gadget that is handled incorrectly. But a malicious gadget is the one behind a destructive attack.

Pros/Cons: Our suggestion is a stochastic model that utilizes a device's past evolution to determine its current state. A measurement sifting policy based on range is employed to depict the possible states of every instrument. In order to categorize each device into five classifications system will analyze each device's activity. This stochastic model's primary concept is to introduce five classes AHL, AHF, SHL, MHL, SHF, and MHF each of which has six threshold values, signifying in that order authentic harmless, authentic harmful, suspicious harmless, suspicious harmful, and malicious harmful. From there, system can determine which class each device belongs to. The thresholds in our work are fixed numbers that are determined by taking the historical log profile's number of activities and dividing it by two.

End-to-end malware detection for android IoT devices using deep learning [38]

Technique: It is imperative to provide effective techniques for Android malware detection given the recent dramatic growth in this type of malware. End-to-end malware detection techniques that do not involve human expert interaction are necessary. Two deep learning-based end-to-end techniques for Android malware detection are presented in this research. The suggested approaches' end-to-end learning process gives them an advantage over the detection techniques now in use.

The suggested approaches can reach detection accuracy of 93.4% and 95.8%, respectively, according to experiments. Suggested approaches are more appropriate for use on Android IoT devices as consume fewer resources than the existing methods.

Pros/Cons: The main contributions are as follows:

For the first time, we pre process Android application classes.dex files into fixed-size sequences using the two re sampling techniques. DLMs can be trained directly with the pre processed sequences. Two deep learning models for malware detection are suggested. A detection accuracy of 93.4% may be attained by the first model, named DexCNN, and 95.8% by the second model, named DexCRNN.

Identifying the attack surface for IoT [39]

Technique:

- Partitioning the IoT infrastructure into trust zones is one method of locating the attack vectors. In order to simulate attack paths, security threats, and vulnerabilities exploited at the device and network levels, second research goal is to dissect IoT architecture in trust zones as per this research. • Researcher think that rather than concentrating on a few key applications, such as the healthcare industry, IoT security should be viewed from a broader perspective.
- Finding device-level vulnerabilities and indicating which security flaws are introduced at the network level as a result of the use of IoT devices in a business is our fourth research challenge.
- Every weakness in an IoT network introduces a penetration point or penetration points and raises one or more security issues. Consequently, creating links between various attacks and the exploited vulnerabilities is our sixth research challenge. In other words, specify how an attacker can take advantage of a weakness to carry out a security risk and harm the system.

Pros/Cons:

In order to reduce the likelihood that a threat would materialize, research also look at which security measures should be implemented on the vulnerabilities that have been found. The device-threat mapping we provide is systems fourth and most important contribution; it helps identify the security settings that are best suited to defend a device against one or more attacks. Research established a 1:1 and 1:n link between security risks and vulnerabilities at the device level for this mapping.

IMCFN: Image-based malware classification using fine-tuned Convolutional neural network architecture [40]

Technique: Demonstrated their efficacy in picture classification for malware binary detection. Research suggests an innovative approach for multiclass classification issues that departs from the current methods. Suggested technique transforms the unprocessed malware binaries into colour images, which the optimized CNN architecture uses to recognize and locate malware families.

It also shows that the colored malware dataset outperformed the gray scale malware photos in terms of accuracy. System evaluated IMCFN's performance against Google's InceptionV3, ResNet50, and VGG16 architectures. Research discovered that our approach requires minimal run-time to detect malware with hidden code, obfuscated malware, and malware family variants. Our approach is resistant to simple obfuscation techniques that hackers frequently employ to mask malware, like packing and encryption.

Pros/Cons:

• Using the back-propagation technique to apply a clever fine-tuning methodology to malware fingerprint photos. The approach of data augmentation is employed to enhance the IMCFN algorithm's performance and manage dataset imbalance. The majority of actual applications require less processing overhead and faster classification, therefore only fine-tuned layers FC2, FC1 and Block 5 are used.
• Using image normalization to find potentially obfuscated or packed malware that contains known dangerous code inside of a program that has been normalized.
• Performing a thorough investigation on big datasets using a variety of state-of-the-art deep learning and classical machine learning architectures in order to assess the effectiveness of our suggested architectures in handling sizable datasets of novel malware variations.
• Testing this model with 25 malware families from the Mailing dataset to see how well it predicts the durability of an obfuscated malware attack.

Effective detection of mobile malware behaviour based on explainable deep neural network [41]
Technique: The property and privacy of users are seriously at risk due to the explosive increase in new mobile virus strains. Deep neural networks have demonstrated excellent accuracy in detecting malicious communications, according to recent studies. On the other hand, a DNN functions similarly to a "black box" in that its architecture conceals any information about how it operates. To get around this problem, the system provides a technique that uses a deep neural network's rules to identify malicious network traffic.
Next, we build a single hidden-output tree to symbolize the rules that are retrieved from the relationship between each hidden layer's output and the neural network's output. Ultimately, the outputs of the hidden layers are used as a bridge to fuse these trees into a single rule tree. By contrasting our approach with other cutting-edge techniques The results of this experiment demonstrate that our approach, which only uses the first nine packets' packet size as a feature, achieves high accuracy and provides a good interpretability of the deep neural network's performance in detecting malicious traffic.

Pros/Cons:

The system provides a novel approach to DNN rule extraction that guarantees high accuracy and improves the deep neural network's interpretability. To detect malware, the system implemented the suggested DNN rule extraction approach on network traffic. The system then compared the method's performance with three cutting-edge technologies and four well-known machine-learning algorithms. This approach performs better, as evidenced by experimental findings that show 98.55% accuracy, 0.9793 precision, 0.9827 recall, and 0.9804 F-Measure. Research developed an online FPGA-based virus detection solution for high-speed network environments.

Conclusion

From this literature survey, we reached to the conclusion that there is a need to create a hybrid machine learning-based approach for detecting viruses and malware. Improving the accuracy of attack detection in Internet of Things networks is the research's main challenge. Furthermore, one of the difficult tasks in the detection phase is classifying the assault in a big IoT network. Therefore, to detect malware and virus attacks, implement suitable defences, and offer a patching mechanism, a new paradigm is needed.

The primary goal of the research will be to create a hybrid machine-learning algorithm that works with the majority of Internet of Things devices. This model will predict NFV malware attacks so that appropriate countermeasures may be implemented. Updates are released via the service by the virtualized system as soon as the malware virus is discovered. All these algorithms as working separately find one parameter to get intrusion detection of IoT network. With an integrated approach, we will get multiple parameters to classify data and accordingly execute appropriate countermeasures against intrusion detection systems. Distribution of updates is enabled via an NFV service-based infrastructure that is distance-bound. The system manages a broad range of viruses and adware that are encountered.

The integration of machine learning in NFV-based IoT networks presents a promising approach to enhance network security against emerging threats. This study highlights the challenges that need to be addressed when implementing such mechanisms, including the need for standardization, cost-effectiveness, and effective machine learning systems. To achieve this, the mechanisms should involve learning attack and normal profiles, abstracting security policies, and rethinking network security in IoT deployments. While this study provides a comprehensive review of the current state of research in this area, there are still several limitations and gaps that need to be addressed. For example, there is a need for more empirical studies to validate the effectiveness of these mechanisms in real-world scenarios. Additionally, the lack of standardization and cost-effectiveness may limit the widespread adoption of these mechanisms. Therefore, future research should focus on addressing these challenges to enable the full potential of machine learning in enhancing security in NFV-based IoT networks. Overall, this study contributes to the ongoing advancement of knowledge in the field and provides valuable insights for researchers and practitioners working on enhancing security in IoT networks.

Research Funding:
No Financial support
Competing interests:
The authors declare that none of the work reported in this study could have been influenced by any known competing financial interests or personal relationships.

Ethical approval

This manuscript was ethically vetted by my research guide before being accepted for publication in this journal.

Acknowledgments
The author wishes to extend his sincere appreciation to the supervisor for all of his help and assistance during this research.

References

1. Alejandro Guerra-Manzanares,Hayretdin Bahsi, SvenNomm, "KronoDroid: Time-based Hybrid-featured Dataset for Effective Android Malware Detection and Characterization", ELSEVEIR-computers & security, 110 (2021) 102399.
2. Vasileios Syrris, Dimitris Geneiatakis "On machine learning effectiveness for malware detection in Android OS using static analysis data", Elseveir-Journal of Information Security and Applications Volume 59, June 2021, 102794.
3. Xiong Luo, Jianyuan Li ,Weiping Wang , Yang Gao, Wenbing Zhao, "Towards improving detection performance for malware with a correntropy-based deep learning method ", Digital Communications and Networks Volume 7, Issue 4, November 2021, Pages 570-579.
4. Baha Rababah, Srija Srivastava Depar, "Hybrid Model For Intrusion Detection Systems", March 2020.
5. CHUANLONG YIN , YUEFEI ZHU, JINLONG FEI, AND XINZHENG HE, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks", October 12, 2017-IEEE Access.
6.Furqan Alama, Rashid Mehmoodb, Iyad Katiba, Aiiad Albeshria, "Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT)", ELSEVIER-Procedia Computer Science 98 (2016) 437 – 442.
7. Afsana Anjum, Ayasha Siddiqua, Shaista Sabeer, Sunanda Kondapalli, Chamandeep Kaur and Khwaja Mohd Rafi, "ANALYSIS OF SECURITY THREATS, ATTACKS IN THE INTERNET OF THINGS", International Journal of Mechanical Engineering-Vol. 6 No. 3 December, 2021.
8.Abderrahmane Boudi, Ivan Farris, Miloud Bagaa and Tarik," Lightweight Virtualization based security framework for Network Edge", IEEE Conference on Standards for Communications and Networking (CSCN) 2018.
9. Peter Bull, Ron Austin, Evgenii Popov, Mak Sharma & Richard Watson, "Flow Based Security for IoT Devices using an SDN Gateway", IEEE 4th International Conference on Future Internet of Things and Cloud 2016.
10.Janice Canedo, Anthony Skjellum, "Using Machine Learning to Secure IoT Systems", Auburn Cyber Research Center, 14th Annual Conference on Privacy, IEEE, Security and Trust (PST) 2016.
11.Muhammad Naveed Aman, Member, IEEE, Uzair Javaidy, Student Member, IEEE,and Biplab Sikdary, Senior Member, IEEE, " Security Function Virtualization for IoT Applications in 6G Networks", Member, article in IEEE Communications Standards Magazine, July 2021
12.Elike Hodo, Xavier Bellekens, Andrew Hamilton, Pierre-Louis Dubouilh, Ephraim Iorkyase, Christos Tachtatzis and Robert Atkinson, "Threat analysis of IoT networks Using Artificial Neural Network Intrusion Detection System", IEEE 2016.
13. KarthikKumar Vaigandla1, Nilofar Azmi, RadhaKrishna Karne, "Investigation on Intrusion Detection Systems (IDSs) in IoT", International Journal of Emerging Trends in Engineering, Research Volume 10. No.3, March 2022.
14. Praveen Kumar Kollu1, R. Satya Prasad, "Intrusion Detection System Using Recurrent Neural Networks and Attention Mechanism", International Journal of Emerging Trends in Engineering Research, August 2019.
15.Ali Haider Shamsan, Arman Rasool Faridi, "A Novel SDNFV IoT Architecture Leveraging Softwarization Technology Services to Alleviate IoT Network Resource Restrictions", Seventh Sense Research Group, February 2022.
16.Fabio Martinellia, Fiammetta Marullib, Francesco Mercaldoa, "Evaluating Convolutional Neural Network for Effective Mobile Malware Detection", Procedia Computer Science 112 (2017) 2372–2381 , International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2017, 6-8 September 2017, Marseille, France.
17.Bimal Kumar Mishra, Dinesh Kumar Saini, "SEIRS epidemic model with delay for transmission of malicious objects in computer network", Elsevier, Applied Mathematics and Computation, August 2013.

18.Nour Moustafa, IEEE student Member, Jill Slay, “UNSW-NB15: A Comprehensive Data set for Network Intrusion Detection systems”, November 2015.

19.Mohammad Nauman, Tamleek Ali Tanveer, Sohail Khan, Toqeer Ali Syed, “Deep neural architectures for large scale android malware analysis”, Springer Science+Business Media New York 2017

20.Robin Nix,Jian Zhang, “Classification of Android Apps and Malware Using Deep Neural Networks”, IEEE, Electronic ISSN: 2161-4407,International Joint Conference on Neural Networks (IJCNN), ,2017.

21.Mike Ojo, Davide Adami and Stefano Giordano, “A SDN-IoT Architecture with NFV Implementation”, 2016 IEEE Globecom Workshops (GC Wkshps) ISBN:978-1-5090-2483-4.

22.Eslam Amer , Ivan Zelinka, Shaker El-Sappagh, ”A Multi-Perspective malware detection approach through behavioral fusion of API call sequence”, Volume 110, November 2021, 102449, Elsevier.

23.Nickolaos Koroniotisa, Nour Moustafaa, , Elena Sitnikovaa, Benjamin Turnbulla, “Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset”, Volume 100, November 2019, Pages 779-796, Elsevier.

24.Milan Vojnovic´, Member, IEEE, and Ayalvadi J. Ganesh, “On the Race of Worms, Alerts, and Patches”, IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 16, NO. 5, Page(s): 1066 – 1079, OCTOBER 2008.

25.Liang Xiao, Xiaoyue Wan, Xiaozhen Lu, Yanyong Zhang, and Di Wu,” IoT Security Techniques Based on Machine Learning”, IEEE Signal Processing Magazine , Page(s): 41 – 49,Volume: 35, Issue: 5, September 2018.

26.Sarp Koksal, Yaser Dalveren, Bamoye Maiga, Ali Kara,” Distributed denial-of-service attack mitigation in network functions virtualization-based 5G networks using management and orchestration”, 25 March,2021 John Wiley & Sons Ltd.

27.Misun Ahn, SeungGwan Lee and Sungwon Lee, “Virtualized Network Function Orchestration System and Experimental Network Based QR Recognition for a 5G Mobile Access Network”, Symmetry 2017.

28.Parag Verma, Ankur Dumka, Rajesh Singh, Alaknanda Ashok, Anita Gehlot, Praveen Kumar Malik, Gurjot Singh Gaba and Mustapha Hedabou, “A Novel Intrusion Detection Approach Using Machine Learning Ensemble for IoT Environments”, Applied Sciences (2076-3417), 2021, Vol 11, Issue 21, p10268, Appl. Sci. 2021.

29.Daniyal Alghazzawi, Omaimah Bamasag, Hayat Ullah and Muhammad Zubair Asghar, “Efficient Detection of DDoS Attacks Using a Hybrid Deep Learning Model with Improved Feature Selection”, Appl. Sci. 2021,Volume 11, Issue 24,Year 2021.

30.Bo Cao, Chenghai Li, Yafei Song , Yueyi Qin and Chen Chen, “Network Intrusion Detection Model Based on CNN and GRU”, Applied Sciences 12(9):4184,2021.

31.Khalid Albulayhi, Qasem Abu Al-Haija , Suliman A. Alsuhibany , Ananth A. Jillepalli , Mohammad Ashrafuzzaman and Frederick T. Sheldon,” IoT Intrusion Detection Using Machine Learning with a Novel High Performing Feature Selection Method”, Appl. Sci. *2022,* 12*(10), 5015*.

32. CHUANLONG YIN , YUEFEI ZHU, JINLONG FEI, AND XINZHENG HE, “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks”, IEEE Access (Volume: 5), Page(s): 21954 – 21961, Electronic ISSN: 2169-3536, 12 October 2017.

33.Shahid Latif, Zil e Huma , Sajjad Shaukat Jamal , Fawad Ahmed, Jawad Ahmad, Adnan Zahid , Kia Dashtipour , Muhammad Umar Aftab , Muhammad Ahmad, Qammer Hussain Abbasi , “Intrusion Detection Framework for the Internet of Things Using a Dense Random Neural Network ”, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 18, NO. 9, SEPTEMBER 2022.

34.Jiong Zhang, Mohammad Zulkernine, and Anwar Haque,” Jiong Zhang, Mohammad Zulkernine, and Anwar Haque”, IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS PART C: APPLICATIONS AND REVIEWS, VOL. 38, NO. 5, SEPTEMBER 2008.

35.Zhida Li, Ana Laura Gonzalez Rios and Ljiljana Trajkovi, “Machine Learning for Detecting Anomalies and Intrusions in Communication Networks “,IEEE Journal on Selected Areas in Communications, Volume: 39, Issue: 7, July 2021.

36.Ning Lu, Dan Li, Wenbo Shi , Pandi Vijayakumar , Francesco Piccialli , Victor Chang, “An efficient combined deep neural network based malware detection framework in 5G environment”, Volume 189, 22 April 2021, 107932.

37. Hajar Moudoud, Lyes Khoukhi, and Soumaya Cherkaoui, “Prediction and Detection of FDIA and DDoS Attacks in 5G Enabled IoT”, IEEE Network, Volume: 35, Issue: 2, March/April 2021.

38.Zhongru Ren,Haomin Wu,Qian Ning,Iftikhar Hussain,BingcaiChen, “End-to-end malware detection for android IoT devices using deep learning”, Volume 101, 15 April 2020, 102098,Elsevier .

39. Syed Rizvi , RJ Orr , Austin Coxa, Prithvee Ashokkumar a, Mohammad R. Rizv, “Identifying the attack surface for IoT network”, Journal- Internet of Things (Netherlands), Volume-9, Elsevier 2020.

40.Danish Vasana,b, Mamoun Alazabc, Sobia Wassand, Hamad Naeeme, Babak Safaeif , Qin Zhenga, “IMCFN: Image-based malware classification using fine-tuned convolutional neural network architecture”, Volume 171, 22 April 2020, 107138.

41. Anli Yan, Zhenxiang Chen, Haibo Zhang , Lizhi Peng , Qiben Yan, Muhammad Umair Hassan, Chuan Zhao, Bo Yang, “Effective detection of mobile malware behavior based on explainable deep neural network”, Neurocomputing 453(2),ELSEVIER-2020.