THE INTEREST OF HYBRIDIZING EXPLAINABLE AI WITH RNN TO RESOLVE DDOS ATTACKS: A COMPREHENSIVE PRACTICAL STUDY

Ahmad Mater Aljohani and Ibrahim Elgendi

Faculty of science and technology, University of Canberra, Canberra, Australia

ABSTRACT

In this paper, we suggest a new method to address attack detection problems in IoT networks, by using a specific emphasis on the user's and network's requirements regarding the application provider and giving a defence against intruders and malicious users. The paper focuses on resolving the problem of detection of attacks in IoT networks using Recurrent Neural Networks and an explainable artificial intelligence technique (XAI) named SHAP. XAI refers to a collection of methods and procedures for explaining how AI models make decisions. Although XAI is useful for understanding the motivations underlying AI models, the information utilized in these discoveries may be a risk. Machine learning models are vulnerable to attacks such as model inversion, model extraction, and membership inference. Such attacks could concentrate on the learning model or on the data used to train and build the model, depending on the involved circumstances and parties. Hence, when the owner of an AI model wishes to grant only black-box access and does not expose the model's parameters and architecture to third parties, solutions like XAI can notably enhance the susceptibility to model extraction attacks.

To guarantee the users regarding possible violations, the proposed RNN-SHAP method is based on a decentralized topology relying on a set of cooperative independent parties including the users of IoT network. Another advantage of the suggested approach is the capitalization on the trust relationships that are part of IoT networks in real life to build trusted mechanisms in the network. Compared to other methods using different metrics, the suggested RNN-SHAP shows its efficiency in resolving the DDoS problem.

KEYWORDS

deep learning (DL), Recurrent neural networks (RNN), SHapley Additive explanation (SHAP), Attack detection, explainable artificial intelligence (XAI), Distributed Denial of Service (DDoS)

1. INTRODUCTION

Daily, millions of users share personal data using different types of devices over IoT networks. It is a known fact nowadays that IoT networks facilitate the daily life of users by expanding their ability to better manage their data and environment. However, various disadvantages are associated with the use of these networks. Among these drawbacks, security issues leading to attacks, distribution of malicious codes, and leakage of critical personal data (locations, date-of-birth, photos...etc.). To resolve these issues, researchers are developing techniques to detect and prevent these potential threats. These specific types of issues give the opportunity to publish high quality research articles tackling the problems of IoT networks using innovant techniques like explainable AI ones.

Any flaws and weaknesses in a system's operating processes, software, or hardware are referred to as network vulnerabilities. These kinds of network vulnerabilities typically involve software or data and are intangible. To prevent authorised users from accessing the network and its services, Denial of Service (DoS) attacks often entail flooding the resources and network bandwidth of the service provider. A denial-of-service (DoS) assault involves taking advantage of wireless connections used by sensors to gather data, wherein the attackers take over various network components.

IoT network analysis reveals how people and heterogeneous devices are connected to one another. Artificial intelligence (AI) has gained popularity over the past ten years, particularly explainable AI (XAI) **[1]**, which is a useful technique for comprehending, elucidating, and interpreting AI conclusions about various events and trends. It is a major technical problem to comprehend, correlate, manage, and make judgments on billions of unique IoT (Internet of Things) little bits of data.

A group of techniques and protocols for elucidating the decision-making processes of AI models are referred to as XAI. Model inversion, model extraction, and membership inference are among the assaults that machine learning (ML) models are susceptible to. Depending on the situation and parties involved, these assaults either targeted the learning model directly or the data that was used to train and develop the model. However, systems such as XAI might significantly increase the vulnerability to model extraction attacks when the owner of an AI model wants to offer only black-box access and does not want to disclose the model's parameters and architecture to other parties.

Machine learning explainable techniques, named XAI, may be used to build threat detection systems for Internet of Things networks. Developing such systems using machine learning presents a challenging issue: producing an effective and realistic training dataset **[2]**. The network data flow should be of the highest quality during an attack as it is the only circumstance that permits attacks to be intercepted. Because of the functioning of several heterogeneous devices in the network, a range of data formats are taken into consideration while constructing the ML model. By solving these challenges, a suitable threat detection system for an IoT environment may be developed. Security needs to be considered right from the start of the design process to ensure that IoT devices are safe from cyberattacks.

Moreover, it may be argued that XAI models are still a developing field that lacks formality and widely recognized principles, despite their rapid progress. Understanding how machine learning algorithms work is essential when analyzing them since they may convert inputs into outputs without the need for manual coding. This leads to a very clear explainability conundrum about the discrimination-related tradeoff, which is the need for basic machine learning models. But if the ML representations are simple, this becomes very biased.

The model's limitations that should hold for the framework to function are significant in different ways depending on the application context. As a result, ML models that meet this requirement for simplicity have larger demands and are more biased; in other words, bias reduces as model complexity increases.

It's also important to keep in mind that efforts to produce results that people will accept and understand are expanding in the field of explainable artificial intelligence research. This is particularly true when it comes to using medical data analysis to provide light on the choices made by machine learning models. Nonetheless, more information about what goes on behind the scenes becomes available when Deep Learning (DL) **[3]** is incorporated into increasingly important societal domains, such as medical diagnosis **[4]** and transport **[5]**. Therefore,

researchers will be able to comprehend a wide range of findings in the fastest and most effective manner feasible with the development of AI techniques with explainability capacity, assisting them in better understanding the solutions to AI outcomes.

It is not necessary for the approach to be explicable because it deals with well-documented challenges that have been addressed by specialists for a long time; nevertheless, this does not mean that all ML models need to be understandable. It makes little sense to explain your bad decisions to an AI model when the repercussions are negligible, like in the case of an AI that can learn to dance. Explainability, however, is essential for machine learning models that directly impact people's lives, such algorithms that decide who gets fired **[6]**.

This is an exemplary case of XAI: A computer will use the data to refine its grasp of bodily components after being trained to recognize various animal species, enabling it to accurately identify, interpret, and describe the animal. Computational AI equipment and programs are the result of science and engineering applied to the design and construction of intelligent technologies. This demonstrates that certain AI systems imitate human thought processes, but it doesn't quantify the impact of this.

Hence, the objective of our study is to design a system and framework for IoT network analysis to detect the attacks and harmful actions in the network. The aim is to detail the general design including the search, storage, capture, and correlations relating the IoT network entities. Various use cases for the data can be presented regarding the analysis of attacks or the detection of malicious content. Besides, the aim is to discuss the issues of managing large IoT devices and datasets and the evolution of the networks considering the risk. Then, suggesting strategies to enhance the analysis of collected data, to predict attacks and to assess risks. Our study incorporates numerous techniques for investigating suspicious events and explaining the roles of users in IoT networks. In addition, this work can incorporate a large-scale strategy for networks investigation. The aim is then to design a system and a framework for a distributed IoT network analysis and to explain the AI taken decisions within the network.

In this context, our main research questions address the issues encountered while designing and implementing an IoT network preserving the of the user’s data and devices. The following questions can be considered for our study:

- What is the potential that explainable AI can give when resolving the problem of attacks detection?
- How to model, design and implement an efficient XAI system for identifying attacks and intrusions in IoT environments?
- How to assess the behavior of the introduced XAI model compared to other ones?
- How the XAI model can successfully identify potential vulnerabilities and threats to correctly assess the degree of ongoing risk and to develop mitigation measures?

The contributions of this study focus on proposing an explainable RNN-SHAP model for resolving the DDoS detection problem. In this RNN-SHAP model, the SHAP strategy should explain the decisions taken by the RNN network and enhance its performance in detecting and identifying the DoS attacks. Moreover, this paper will present an implementation which can be executed on different operating systems such as MacOS, Linux or Windows. This implementation is a multithread event-driven application controlled by a set of users who build and keep the P2P and IoT network overlays, achieve cryptography operations and provide facilities in the network like data sharing.

The rest of the paper is organised as follows: section 2 highlights the related state-of-the-art, section 3 details the proposed RNN-SHAP system resolving the DDoS, section 4 illustrates and discusses the results, section 5 concludes the paper.

2. RELATED WORKS

2.1. Literature Review on XAI, and Threats Detection During Attacks

Cybersecurity attributes and threat models, as well as attacks against explainable frameworks in black box scenarios, have not been well explored. In **[7]**, a ML-based intrusion identification system is introduced for IoT networks (named ML-IDS). The aim is to implement supervised ML algorithms to recognize attacks on IoT networks. The research methodology involved a three-stage process. Firstly, the dataset "UNSW-NB15", including a several types of attack and network traffic activities, was used. Then, a reduction module based on "Principal Component Analysis" (PCA) was used. Finally, six ML algorithms were tested on recall, delta, and Mathew correlation coefficient (MCC). According to the authors, the obtained results were promising and proving the efficacy of the proposed ML-IDS.

In **[8]**, AUTOLYCUS, a model of attack extraction is introduced to infer decision boundaries in decision tree models and generate extracted models that behave similarly to the target model. AUTOLYCUS uses XAI to improve the capabilities of cutting-edge algorithms by relaxing the requirements for reconstruction, such as reducing the queries and leveraging current samples as supplementary information. However, the choice of the used datasets (such as Crop and Iris) is not justified.

Another study in **[9]** indicate that Existing attack detection technologies are incapable of detecting DDoS assaults reliably and effectively. To that purpose, the authors present a DDoS detection approach based on explainable artificial intelligence (XAI). This approach detects aberrant network traffic flow characteristics by studying communication at the network layer. Furthermore, it selects the most significant features with influence weight for each anomalous instance and subsequently initiates an appropriate threshold for each feature. As a result, this DDoS attack detection approach sets security rules for application-layer-based, volumetric-based, and transport control protocol (TCP) state-exhaustion-based features depending on each feature threshold value. Because the suggested approach relies on layer three traffic, it can detect DDoS attacks on both IoT and traditional networks. To assess the performance of the proposed system, extensive tests were done using the University of Sannio, Benevento Intrusion Detection System (USB-IDS) dataset, which comprises of several forms of DDoS assaults. Despite the results were clearly presented in this research, the discussion and interpretation of the findings of these results were not achieved. Moreover, the overall system and tests rely only on simulations without any real-world and real-time IoT attack context. The latter give better accuracy to the results of the proposed model.

The study in **[10]** proposes a multi-objective counterfactual explanations (DiCE) model and investigates how to establish explainable artificial intelligence with privacy-preserving. The study helps to improve understanding of the conclusion made through the black-box system; yet, doing so makes the system more visible, leading to a violation of individuals' privacy. However, providing explanations while protecting individuals' privacy is not always easy. As a result, a balance must be struck between the outcome's explainability and the individual's privacy. In this research, the main targeted research questions were: How sensitive to membership inference attack are multi-objective counterfactual explanations? How successful is Differential Privacy in protecting people's privacy, using the counterfactual explanation technique DiCE? How efficient

are the produced counterfactuals once Differential Privacy is applied? How much privacy can be kept while creating multi-objective counterfactual explanations? Experiments demonstrated that as the privacy budget is reduced, the effectiveness of the generated data framework diminishes.

In **[11]**, authors examine the links between model explanations and private data about the training of the model set leaking. They use membership inference attacks to study the privacy hazards of feature-based model explanations: measuring the degree to which predicted models plus their explanations leak data regarding the presence of a datapoint in a model's training set. Over a variety of datasets, the study assesses membership attacks utilizing model explanations based on features. It demonstrates that backpropagation-based explanations can provide important data concerning individual training datapoints. This is since they give statistical information about the model's decision limits for an input, which might indicate its membership. However, the random phenomenon induced by the proposed explainable AI provokes a non-transparency of the model. Besides, designing safe transparency reports is not performed. The latter reports allow more useful (not just an arbitrary noise) and more secure interpretations. This trade-off between security and utility was not considered in the study.

The review in **[12]** performs a detailed investigation on the intersection of XAI and cybersecurity. The research specifically investigates the current research from two different viewpoints: XAI applications to cybersecurity (e.g., detection of intrusions, malware classification) and XAI security (e.g., cyberattacks on XAI pipelines, possible responses). Several security properties discussed in the literature are used by the authors to express the XAI safety. They also explore future research prospects and generate open issues that are either unsolved or inadequately addressed in the literature. Despite the study is interested in uncovered topics before, such as the reusability of XAI techniques, some issues rise in the study regarding the detection of possible threats in the network.

2.2. XAI and IoT Vulnerability to Attacks

The study in **[13]** examines the attack graph association examination of IoT system and the attack graph development methodology, as well as the vulnerability smart early warning technique in the context of the Internet of Things. First, it creates a network safety assessment framework employing the IoT vulnerability analysis of associations using the attack graph technology. The algorithm for creating attack graphs has been enhanced. In the context of the Internet of Things, the node weight metric is used to search for the primary attack path in the attack graph. Utilizing an attack graph, a measuring computation model is employed to achieve a quantitative examination of the security condition of the IoT ecosystem. Then, a dynamic stain propagation approach for detecting early warning vulnerabilities in IoT environments is provided, with an emphasis on the creation of stains and the examination of stains. Based on the IoT as a counterexample, a static detection approach for identifying potential vulnerabilities is suggested. A potential buffer early warning vulnerability is found by the process of identification and context sensitive detection. The driver crawler utilizes function capture to recognize the operation of the stain data and accomplishes automated detection. Experiments show that the proposed model is efficient in detecting vulnerabilities and can unfiltered vulnerability of early detection warnings.

Another study in **[14]** tries to make information about known vulnerabilities easier to understand and more useful for users in the cyber domain. The authors' proposal is in accordance with several dimensions of vulnerability analyses proposed in the Ben Schneiderman’s maxim of "overview first, zoom and filter, details on demand". They offer ideas and use cases that show how supporting these levels is practical and beneficial. The study provides a methodology for

understanding vulnerability clusters, visual encodings for a well-liked vulnerability scoring system that supports prioritization tasks, and an example of the LIME framework for understanding vulnerability descriptions. According to this research, additional cyber analytics, such as network anomaly detection, might benefit from a multi-layer system to interpretability. In these cases, knowing enterprise-wide user behaviours from a population of users could help investigators narrow their focus to specific servers, users, or events.

2.3. XAI for IoT Networks Attack Detection in Smart Cities

Due to the engineers' acknowledged fear of the often-opaque inner workings of AI, the use of AI techniques for tracking and monitoring remains limited. The "black box" character of AI models should be clarified to the engineers to increase trust in AI, leading to what is known as "explainable artificial intelligence" (XAI).

On the other hand, decision makers are becoming more concerned about the difficulty of explaining AI-based solutions, which is seen as a key barrier to public acceptance and confidence in these types of AI-based solutions. With the advent of XAI, people's lives have been improved, and the idea of smart cities has been envisioned. Smart cities would utilize deliberate activities, improved user perceptions, and decisive decision-making procedures.

In **[15]**, the use of XAI in cybersecurity is thoroughly reviewed in the current paper. System, network, and software protection from various threats is made possible by cybersecurity. The potential for using XAI to anticipate such attacks is enormous. This study offers a succinct introduction of attack types and cybersecurity. The use of conventional AI techniques and the difficulties connected with them is then examined, which pave the way for the use of XAI in a variety of applications. Additionally, the XAI implementations of numerous academic projects and business are shown.

Authors in **[14]** said there is an opportunity with commercial XAI frameworks such Local LIME, which assigns sensitivity in an ML local prediction instance to specific input features, there is a chance to increase interpretability in ML-based cyber analytics. These technologies have mostly been tested in computer vision-related fields where there are spatially connected features and concrete classification problems. Because of the competitive character of attacks and the proof they leave in disparate data streams, these scenarios vary dramatically from many other cyber applications. Authors examine portions of cyber data—specifically, vulnerability descriptions—where previous methods have been effective on related data types. They go over their findings about defining vulnerability classes, but they neglect the issues of local ML explanations in the field of cyber security.

Another investigation in **[16]** offered a thorough analysis of XAI in intelligent connected cars (ICVs) for the purpose of intrusion detection and mitigation. Because of the vulnerabilities present in linked devices, ICVs/IoV is an expanded use of the IoT in smart transport systems (ITSs). However, because most detection systems use artificial intelligence in a "black box" fashion, transparency remains a challenge, necessitating the development of explainable AI. A thorough overview of the current XAI frameworks, and their utilization to ICV security are covered in this study. To promote rule based XAI's acceptance in crucial areas like the automotive industry, XAI developers must also address the problem of bias caused by this technology. Moreover, an interesting research topic is to recognize the requirements for dependability, the minimal needed computing complexity, and the inclusion of user-friendly XAI modules.

Since XAI needs to comprehend why a system issues a security warning to properly react to the occurrence in a manner that the computer might not (for example, instructing network users to refrain from performing specific behaviours). A summary of these requirements in the context of software security and malware classification, emphasizing that the absence of accessibility in deep neural networks poses major obstacles in building confidence on the AI model or on successfully troubleshooting and categorizing errors.

The work in **[17]** has presented a first step in adapting XAI methodologies in smart monitoring by drawing on developments in existing ML solutions for it. The incomprehensibility of many of the ML algorithms used in smart monitoring has been examined, and XAI techniques have been put out as a solution. Depending on their goal and the target audience of the human beings they are watching, ML algorithms may need varying degrees of explanations for smart monitoring. The study concludes that for promoting the development of smart monitoring and smart cities, a thorough examination of the explainability and degrees of explanation for ML algorithms is necessary.

The study in **[18]** provides a thorough analysis of recent and upcoming advances in XAI technology for smart cities. It also emphasizes the technical, industrial, and sociological developments that spur the development of XAI for smart cities. It provides a detailed explanation of what is essential to implementing XAI solutions for smart cities. The idea of XAI for smart cities, numerous XAI application cases, problems, applications, potential substitute solutions, and present and future research advancements are also covered in the paper. Detailed descriptions of research initiatives and activities, such as efforts to standardize the development of XAI for smart cities, are provided.

In the framework of a project financed by the European Commission **[19]**, an Explainable AI solution utilizing DL and semantic web techniques is suggested to create a hybrid classifier for the application of smart cities flood monitoring. In this mixed model, the DL component determines the existence and degree of object coverage while the categorization is done using semantic rules that were carefully developed in close contact with experts. The experimental findings, which were presented with a practical application, revealed that this hybrid technique to image classification performs on average 11% better (F-Measure) than DL-only classifiers.

The main goal of the review in **[20]** is to examine how smart cities have evolved in relation to AI, ML, and deep reinforcement learning. The methods mentioned above are effectively utilized to build the best possible policies for a variety of complicated issues pertaining to smart cities. In this review, authors provide detailed information on the ways that the earlier techniques have been applied to ITSs (intelligent transportation systems), cyber security, energy-efficient use of 5G networks and smart grids, communications services, and healthcare platforms in smart cities.

The IoT-based intelligent devices would have access to private information about network infrastructure, creating significant privacy and security concerns. Smart systems based on the IoT are quite vulnerable to botnet assaults. Examples of such assaults include the distributed denial of service (DDoS) attacks caused by the Mirai and BASHLITE viruses, which are often used in smart cities and are launched by hacked surveillance systems. These DDoS attacks on IoT gadgets and networks put the burgeoning idea of sustainable smart cities in even greater danger.

In this regard, the study in **[21]** introduces the IoTBoT-IDS architecture, to defend IoT-based intelligent infrastructure against botnet attacks, as a revolutionary probabilistic learning-based botnet detection system. IoTBoT-IDS uses statistical learning-based approaches such as the Beta Mixture Model (BMM) to represent the typical behavior of IoT networks. Any departure from the

expected pattern of behavior is recognized as an abnormal occurrence. Three comparisons set of data produced from actual IoT networks were utilized to assess IoTBoT-IDS.

These previously mentioned recent research on machine learning interpretability methodologies indicate that more improvement in this area is still possible. These studies underscore the potential benefits and enhancements that XAI approaches may provide to the existing machine learning processes. However, they also draw attention to the limitations and deficiencies inherent in these techniques.

3. RNN-SHAP FOR DETECTION OF DDOS ATTACKS

The most reachable category of XAI is explainable data. However, given the massive amounts of data that could be utilized for training an AI algorithm, "attainable" is not as simple as it sounds. An extreme example is the GPT-3, or natural language algorithm. Although the model can mimic human speech, it additionally accepted a lot of harmful internet content during training.

According to Google, an "AI system is best understood by its basic training data and process, along with the resulting AI model." This knowledge necessitates the capacity to link a trained AI model to the same dataset that was used to train it, as well as the ability to thoroughly study that data, even if it has been years since a version of it was created.

Paying careful consideration to the information used to train a model is one of the simplest methods to improve its explainability. During the design phase, teams must decide where the data to train an algorithm will come from, whether that data was obtained legally and ethically (assuming it exists), if the information includes bias, and what can be done to mitigate that bias. This is a large task that should not be ignored; 67% of organizations use over twenty data sources for AI.

It is also critical to thoroughly filter data that is or must be unrelated to the conclusion. A loan confirmation process may make judgments based in significant part on an applicant's zip code. The simplest approach to verify that an algorithm's output isn't dependent on an irrelevant feature, such as a zip code, which is frequently used as a proxy for race, is to avoid involving such information in the set used for training or on the input data **[22]**.

3.1. RNN

RNN stands for Recurrent Neural Network **[23]**, which is a type of artificial neural network that can process sequences of data, such as speech, text, or time series1. RNNs have an internal memory that allows them to remember past inputs and use them to make predictions or generate outputs23. RNNs are used in many applications such as natural language processing, speech recognition, machine translation, and image captioning.

Recurrent neural networks (RNNs) can be used for attack detection in various fields such as cybersecurity, intrusion detection, and network security. The advantage of using RNNs is that they can process sequential data where each timestamp depends on the previous one. This is crucial in detecting cyber-attacks as hackers may execute their attacks over a set of steps.

One example of using RNNs for attack detection is in detecting network intrusion attempts. The RNN can be trained on a set of normal network traffic data and then used to identify anomalous traffic patterns that potentially signify an attack. The RNN can detect these patterns based on the sequence of incoming traffic and flag them as potential attacks.

Another use case of RNNs in attack detection is in detecting phishing attacks. The RNN can be trained on a set of legitimate emails and then be used to detect any potential phishing emails received. The RNN can detect the pattern of language used in the email and flag it as a potential phishing attempt.

Overall, RNNs are an effective tool for attack detection as they can analyse sequential data and detect any anomalous patterns. Through training on a set of normal data, RNNs can identify any deviations from this pattern and flag them as potential attacks.

Recurrent Neural Networks (RNN) can be used for detecting DDoS attacks. However, they are not typically used for launching DDoS attacks. RNNs are a type of machine learning algorithm that is commonly used for time-series prediction and sequential pattern recognition. In the context of DDoS detection, RNNs can be trained on network traffic data to identify patterns of traffic that are indicative of a DDoS attack. The trained RNN can then be used to classify new traffic data in real-time and trigger an alert or mitigation action if a DDoS attack is detected.

To detect DDoS attacks using RNN, the network traffic data is collected and pre-processed. The RNN model is then trained on the pre-processed data to learn the normal traffic pattern of the network. Once the RNN model is trained, it can be used to predict whether the incoming traffic is normal or malicious.

During the detection process, the RNN model compares the incoming traffic with the learned normal traffic pattern. If the incoming traffic deviates from the normal traffic pattern, the RNN model identifies it as a potential DDoS attack and sends an alert to the network administrator.

The advantage of using RNN for DDoS detection is that it can analyse the traffic data in real-time and detect DDoS attacks quickly, minimizing the damage caused by the attack. Additionally, RNN can adapt to changes in the traffic pattern and update its detection parameters accordingly.

3.2. SHAP

SHAP (SHapley Additive exPlanations) **[24] [25]** is a unified framework for interpreting predictions of machine learning models. It is a game theoretic approach that assigns feature importance values to each feature in a prediction. This method is based on the concept of Shapley values, which is a solution concept in cooperative game theory that assigns a fair allocation of rewards to players in a coalition game. In the context of machine learning, these rewards are the contributions of each feature to the final prediction. SHAP values provide a more accurate and reliable interpretation of machine learning models compared to other feature importance methods.

3.3. RNN-SHAP

To explain a recurrent neural network (RNN) using SHAP **[26]**, we can apply the following general process:

1. Train the RNN on your data.
2. Use SHAP to compute the SHAP values for each input feature and for each output of the RNN.
3. Visualize the SHAP values to understand how much each feature contributed to the model's output at each time step.

In the case of an RNN, there are some specific considerations to keep in mind, as the model has an internal hidden state that is used to propagate information through time.
One way to approach this is to unroll the RNN over time, so that we have a separate set of SHAP values for each time step. This lets us see how the contribution of each feature changes as the hidden state (and therefore, the input to the model) evolves over time.

It's also important to ensure that the SHAP values are being computed in a way that considers the sequential nature of the input data. One approach is to use the "masking" technique, which involves setting the SHAP values for any inputs that are padded with zeros to zero as well. This helps to ensure that the SHAP values are only reflecting the information that's relevant to the model's output.

Overall, using SHAP to explain an RNN involves some additional complexity due to the model's recurrent nature, but it can still provide valuable insights into how the model is making its predictions over time.

3.4. RNN-SHAP for DDoS

A distributed denial-of-service (DDoS) attack is a malicious attempt to disrupt normal traffic of a targeted server, service, or network by overwhelming the target or its surrounding infrastructure with a flood of Internet traffic. It has caused great harm to the security of the network environment.

DDoS attack is a variants of DoS attack in which attacker uses authorized user IP address to attack on a particular victim. There are several types of DDoS attacks, namely, SYN-flood, ACK-flood, UDP-flood, Connection DDoS, DNS Reflect, ICMP flood and so on. The main aim of the attackers is to jam the resources to deny services to the recipients. Attackers can use several strategies to achieve this goal, one of which is by flooding the network with bogus requests. DDoS attack is distributed in the way that the attacker is using multiple computers to launch the denial-of-service attack.

DDoS attacks pose an immense threat to the Internet, and many defence mechanisms have been proposed to combat the problem.

Different methods were used to resolve DDoS, such as **ANN [27]**, **RNN [28]**, **LSTM [29]**, **transformer-based [30]**, **CNN [31]** or using **hybrid methods** such as **RNN-ELM [32]**, **CNN-LSTM [33]**, **RNN-LSTM [34]**, **Fuzzy Set Based Neural Network [35]**. Using SHAP to explain the predictions of an RNN-based DDoS detection model can help us understand which features are most important in determining if an incoming traffic flow is malicious or not.

To implement RNN-SHAP for DDoS, we can follow these steps:

- Pre-process the data: We need to pre-process the dataset to extract relevant features and transform them into numeric format that the RNN model can work with.
- Train the model: Train an RNN model with the pre-processed dataset.
- Compute SHAP values: Use the SHAP implementation, such as `shap.Explainer`, to compute the SHAP values for each feature and for each output of the RNN.
- Visualize the SHAP values: Visualize the SHAP values to see which features are most important for predicting whether a traffic flow is malicious or not. This can help us identify patterns and behaviours that are indicative of a DDoS attack, such as a sudden increase in traffic from a specific IP address.

- Refine the model: Based on the insights gained from the SHAP values, refine the RNN model to improve its performance in detecting DDoS attacks.

To sum up, few studies concern the optimization of RNN by explaining it using SHAP, such as **[26]**, **[36]**. Other studies focus on the use of RNN for DDoS, such as **[28]**, **[37]**, **[38]**, **[39]**. Our main idea of contribution is then to apply SHAP to explain RNN used to resolve the DDoS problem. By using RNN-SHAP for DDoS, we can gain insights into the factors that influence the model's predictions and improve its effectiveness in detecting these types of cyber-attacks.

4. NUMERICAL TESTS AND FINDINGS

This section evaluates the suggested SHAP-RNN's performance using UNSW_NB15 **[40]** datasets and contrasts it with alternative paradigms including RNN, RL, and SVM. Data-driven solutions are what we employ as a data. Our data came from a well-known dataset that was found in the literature, which is available in UNSW_NB15. The UNSW-NB15 dataset was introduced by the Australian Centre for Cyber Security (ACCS) and gathers network packet data produced by a combination of modern synthetic attack behaviors and real-world, modern normal activities. This data was generated at the Cyber Range Lab of the Australian Centre for Cyber Security using the IXIA PerfectStorm program. We collected 100 GB of the raw data as Pcap files using the Tcpdump utility. There are nine different kinds of attacks in it: backdoors, exploits, DoS, worms, shellcode, reconnaissance, generic, and analysis.

The Argus and Bro-IDS tools are used to construct the dataset, and twelve distinct methods were used to generate a total of 49 features with class labels. This dataset is frequently used by researchers to create and assess machine learning models for the identification of network-based threats and to enhance cybersecurity safeguards. Several measures, including precision, F1 value, and recall (to evaluate the feature/class performance of the DDoS traffic), accuracy (to evaluate the classifier performance), AUC, and ROC, are used to evaluate the proposed RNN-SHAP on UNSW_NB15.

Fig. 1 displays the number of entries in UNSW-NB15 pertaining to DDoS and other forms of attacks, derived from the dataset UNSW-NB15.

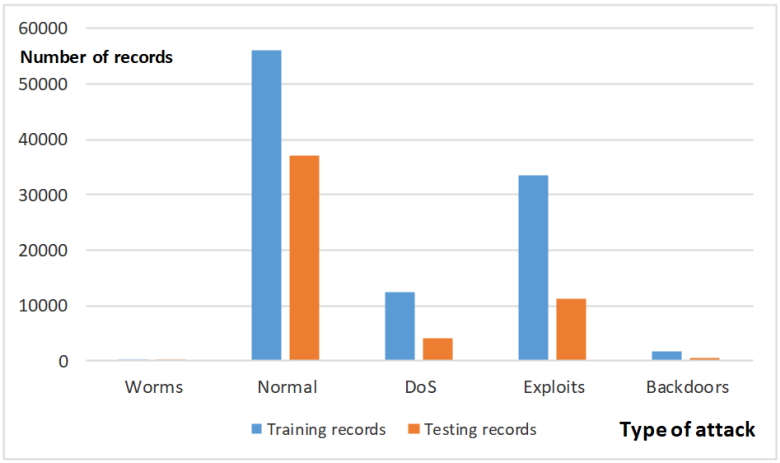

Fig. 1 UNSW-NB15 categories of records

An additional parameter used to evaluate the effectiveness of algorithms is the training time. Table 1 indicates that, even with RNN-SHAP's previously demonstrated performance, its training duration is not superior to that of other paradigms.

Table 1 Training time for the various methods in use

Model	RNN **[23]**	RNN-SHAP **[26]**	GRU-RNN **[41]**	SVM **[42]**	Random forest (RF) **[43]**
Training time (s)	3291	4832	3628	4872	2983
Testing time (s)	1417	2192	1376	2193	1248

The rate of accurate predictions made from all record data is shown by a different statistic called AUC. AUC of 1 indicates that the records are entirely accurate. Fig. 2 demonstrates that RNN-SHAP has the greatest prediction rate since it has the highest AUC values.

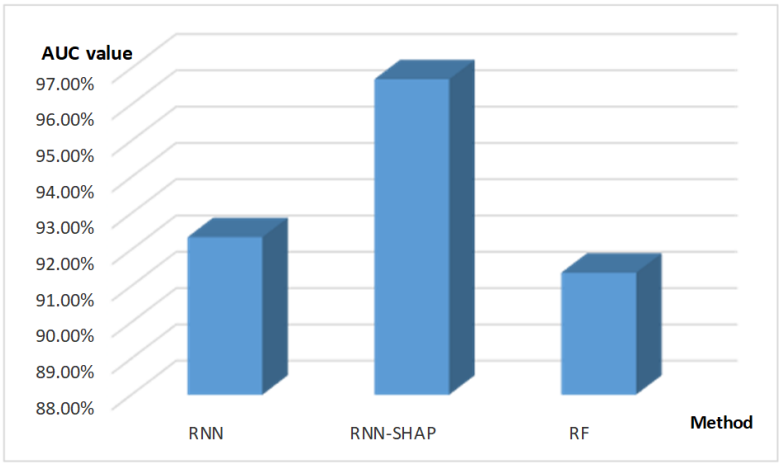

Fig. 2 RNN, RNN-SHAP, and RF AUC values

Another test that uses the ROC curve to assess how algorithms behave is called a ROC analysis. The outcomes are shown in Table 2. show the following: The ROC of RF shows that 95% of TP values are obtained for a forecast with 15% error. The RNN's ROC shows that 80% of TP values are obtained for a prediction with 20% error. The RNN-SHAP's ROC shows that 95% of TP values are obtained for a prediction with 10% error.

Table 2 ROC curve for the algorithms in use

	0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
Random forest [43]	0.2	0.75	**0.94**	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
RNN [23]	0.3	0.5	**0.8**	0.88	0.9	0.9	0.96	0.96	0.96	0.96	0.96
RNN-SHAP [26]	0.2	0.6	**0.95**	0.95	0.96	0.96	0.97	0.97	0.98	0.98	0.98

The formula A = Σ(TN+TP) / Σ(FP+FN+TN+TP) (xx) is the basis for the accuracy metric. The acronyms in A stand for False Positive (FP), and False Negative (FN), True Negative (TN), and True Positive (TP). Table 3 displays the accuracy rates for 100 epochs.

Table 3 Analysing the RNN-SHAP accuracy statistics

	0	**20**	**40**	**60**	**80**	**100**
Testing phase	83.51	95.69	96.32	96.89	97.76	98.98
Training phase	77.73	94.34	95.28	96.26	97.66	98.91

F1 = (2TP) / (2TP+FP+FN) (xxx) is the formula for the F1 score. The formula for precision is P=TN/(TN+FP) (xxx). The recall is R=TP/(TP+FN) (xxx). Fig. 3 displays the confusion matrix of testing instances for the RNN-SHAP, which demonstrates a high accuracy of predictions.

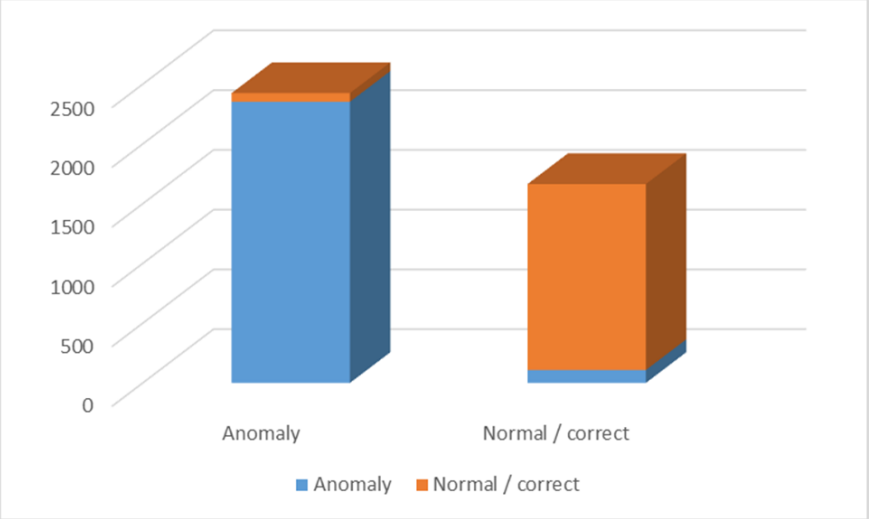

Fig. 3 Testing instance confusion matrix created using RNN-SHAP

Figures 4 and 5 include a complete list of metrics values for every algorithm that has been examined.

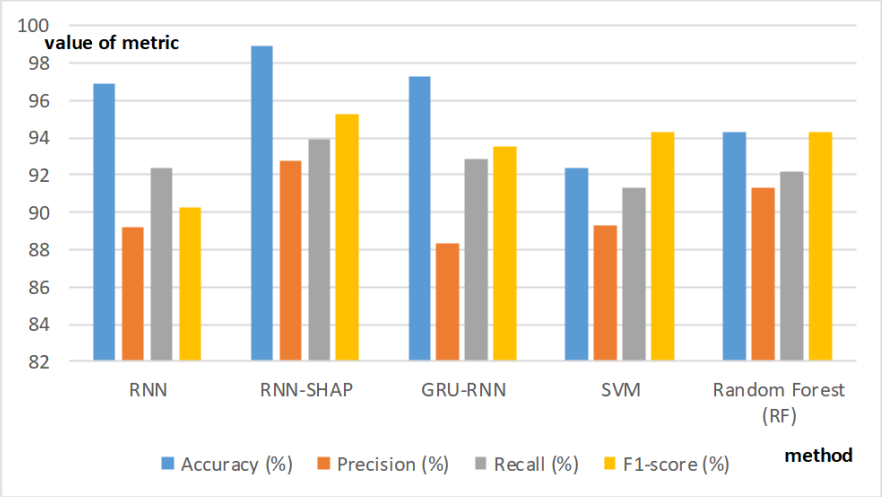

Fig 4 Training phase metrics

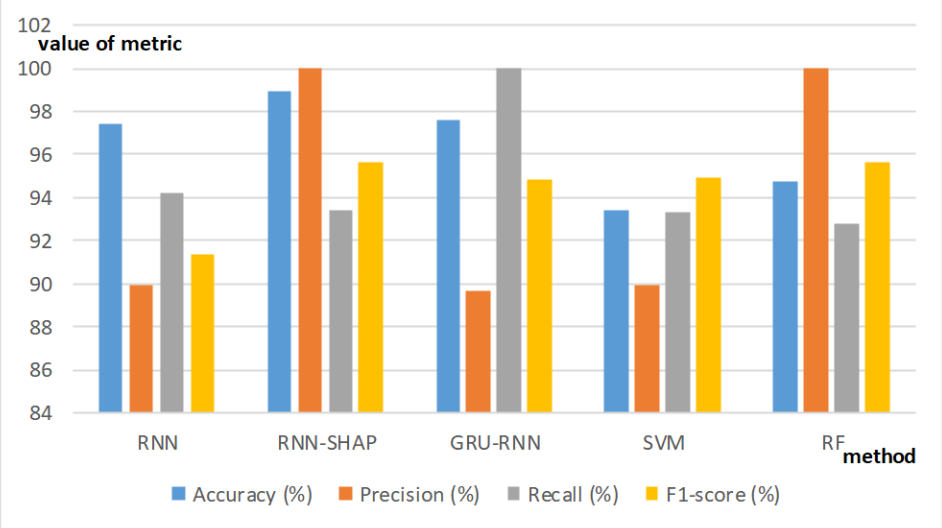

Fig. 5 Testing phase metrics

From the related works and our numerical tests, we conclude that machine learning algorithms, especially XAI, can be used to build attack detection systems for IoT networks. The challenging issue in building such systems using ML is in generating a realistic and high-quality training dataset. A good quality of data flow in network should exist during the process of attack because interception is possible only with continuous flow of data. Different heterogeneous devices are to be operated in the network so diverse data are considered in building the ML model. By overcoming these challenges, a suitable attack detection system can be created for IoT environment. Security features must be considered early in the design process to ensure that IoT devices are protected from cyber-threats.

Moreover, it can be concluded that despite its rapid growth, explainable artificial intelligence is still not a mature and well-established field, often suffering from a lack of formality and not well agreed upon definitions. Consequently, although a great number of machine learning

interpretability techniques and studies have been developed, they rarely form a substantial part of machine learning workflows and pipelines.

Besides, the volume of studies on machine learning interpretability methods over the past years demonstrated the room for improvement that exists. These studies showcasing the benefits and enhancements that XAI methods can bring to existing machine learning workflows, and exposed their flaws, weaknesses, and how much they lack performance-aside.

According to the investigated state-of-the-art, it is our belief that explainable artificial intelligence still has unexplored aspects and a lot of potential to unlock in the coming years in relation to IoT attacks detection.

It is crucial to keep in mind that to interpret ML algorithms and comprehend what an ML algorithm is, there has been a paradigm shift from 'traditional programming' where it was necessary to explicitly pass all heuristics to a new idea where, rather than saving each operation the model achieves, simply provide numerous instances, and let ML understand what the best selections to take.

Because ML algorithms may change feeds into outcomes without being directly coded, it is important to understand how they work when evaluating ML algorithms.

As a result, there is a highly explicit dilemma about explainability in respect to the compromise linked to discrimination, namely, the requirement for Simple models for ML algorithms for them to be comprehensible.

However, this is quite biased if the representations are straightforward.

The significance of the model's reductions and the hypotheses that must be established for the framework to function varies depending on the context in which it is used. Additionally, simpler ML models under this criterion are more biased models with greater requirements; in other words, the bias gets smaller as model complexity increases.

It's also important to note that efforts to produce results that users will accept and understand are expanding research into explainable artificial intelligence, particularly when it comes to the investigation of medical data to explain the choices taken by ML models.
However, it is reasonable that more information might be obtained about what goes on behind the scenes as DL is incorporated into more important spheres of the community, such as diagnosis in medicine.

Therefore, a wide range of findings will be given to people to grasp, in the most effective and fastest manner, how to address issues and better comprehend the replies of these results through the invention of AI techniques having this capacity (explainable).

Since dealing with well-documented issues that have been handled by professionals for a lengthy period, it is not required for the method to be explainable, this does not imply that all ML models must be comprehensible. When the consequences of your poor choices are minimal, such as in the case of an AI that can learn to dance, it doesn't even make sense to explain them to an AI model. However, explainability is crucial for ML models that directly affect people's life, such as algorithms that determine who will be fired **[6]**.

Speaking of healthcare that use cutting-edge digital intelligence, in certain situations in which an AI-based healthcare assistance system suggests surgical treatment rather than radiation therapy, indicating a suggestion that can be counterintuitive. Recent AI systems may recommend and indicate decisions, but they cannot explain the reasoning behind those decisions.

A computer that has been trained to recognize animals will learn about various forms of body parts, assembling this knowledge (acquired from the data) to accurately determine, understand, and define the animal. AI is associated with the science and engineering of creating and building smart machines (computational equipment and programs).

This methodology establishes and demonstrates that some AI systems simulate human thought processes, although it fails to demonstrate the impact.

XAI allows for AI-oriented systems to explain the environment in which intelligent devices function by developing underpinning explanatory AI-oriented models that make it feasible to describe actual procedures. Finally, without creating a cause-and-effect connection, it is inconsistent to develop AI technology into a heavenly force that people will crave. On the other hand, it is impractical to disregard the digital knowledge that this gives society. In essence, it is vital to develop interpretable and adaptable AI models that allow collaboration with professionals that have expertise and academic standing in many fields.

Overall, with the ability to adequately clarify their logic, artificial thinking, technique, and to express an appropriate understanding of how the models operate, XAI will therefore be crucial for future operators to comprehend, handle, and ultimately believe the next wave of AI-oriented machines.

5. CONCLUSION

The application of explainable AI techniques to the problem of DDoS threat identification in Internet of Things networks is the aim of this research.

To create ML models that are particularly well-suited for online threat detection, it is imperative to close the knowledge gap between ML techniques and IoT application security. In this regard, the use of explainable AI combined with neural networks is a relevant research field.

This study suggests the use of SHAP as a XAI technique, to explain the RNN decisions when resolving the DDoS problem. The study determines how explainable AI might be used efficiently on networking and security applications.

The results and findings are valuable to the networking community, users of distributed systems, and any type of network service involving several stakeholders sharing documents, data, and resources.

As a future research direction, large-scale attack detection datasets need to be generated to build efficient deep learning IoT threat detection systems.

REFERENCES

[1] Marwa Keshk, Nickolaos Koroniotis, Nam Pham, Nour Moustafa, Benjamin Turnbull, Albert Y. Zomaya. An explainable deep learning-enabled intrusion detection framework in IoT networks, Information Sciences, Vol 639, 2023, https://doi.org/10.1016/j.ins.2023.119000.

[2] A. A. Alashhab, M. S. M. Zahid, M. Abdullahi and M. S. Rahman, "Real-Time Detection of Low-Rate DDoS Attacks in SDN-Based Networks Using Online Machine Learning Model," 2023 7th Cyber Security in Networking Conference (CSNet), Montreal, QC, Canada, 2023, pp. 95-101, http://doi.org/10.1109/CSNet59123.2023.10339791.

[3] G. Karatas, O. Demir and O. Koray Sahingoz, "Deep Learning in Intrusion Detection Systems," 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), Ankara, Turkey, 2018, pp. 113-116, http://doi.org/10.1109/IBIGDELFT.2018.8625278.

[4] Laddha S, Mnasri S, Alghamdi M, Kumar V, Kaur M, Alrashidi M, Almuhaimeed A, Alshehri A, Alrowaily MA, Alkhazi I. COVID-19 Diagnosis and Classification Using Radiological Imaging and Deep Learning Techniques: A Comparative Study. Diagnostics. 2022; 12(8):1880. https://doi.org/10.3390/diagnostics12081880.

[5] Mnasri, S., K. Zidi, and K. Ghedira. "A heuristic approach based on the multi-agents negotiation for the resolution of the DDBAP." The 4th international conference on metaheuristics and nature inspired computing, Sousse, Tunisia. 2012.

[6] Mohamed Lahby, Utku Kose, Akash Kumar Bhoi. Explainable Artificial Intelligence for Smart Cities. 1st Edition, November 2021, CRC Press, https://doi.org/10.1201/9781003172772.

[7] Saheed Yakub K., Abiodun Aremu Idris, Misra Sanjay, Monica Kristiansen Holone, Ricardo Colomo-Palacios. A machine learning-based intrusion detection for detecting internet of things network attacks, Alexandria Engineering Journal,61(12),2022, pp. 9395-9409, https://doi.org/10.1016/j.aej.2022.02.063.

[8] Abdullah Caglar Oksuz, Anisa Halimi, Erman Ayday. AUTOLYCUS: Exploiting Explainable AI (XAI) for Model Extraction Attacks against Decision Tree Models. 2023. The Network and Distributed System Security Symposium (NDSS) 2023. https://doi.org/10.48550/arXiv.2302.02162.

[9] Kalutharage, C.S.; Liu, X.; Chrysoulas, C.; Pitropakis, N.; Papadopoulos, P. Explainable AI-Based DDOS Attack Identification Method for IoT Networks. Computers 2023, 12, 32. https://doi.org/10.3390/computers12020032.

[10] Nelson D.J. M.Sc. Thesis: Privacy-Preserving Counterfactual Explanations To Help Humans Contest AI Based Decisions. 2023. University of Twente. Faculty of Electrical Engineering, Mathematics and Computer Science.

[11] Reza Shokri, Martin Strobel, and Yair Zick. 2021. On the Privacy Risks of Model Explanations. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21), May 19–21, 2021, Virtual Event, USA. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3461702.3462533.

[12] Charmet, F., Tanuwidjaja, H.C., Ayoubi, S. et al. Explainable artificial intelligence for cybersecurity: a literature survey. Ann. Telecommun. 77, 789–812 (2022). https://doi.org/10.1007/s12243-022-00926-7.

[13] Wong L. J., Headley W. C., Michaels A. J. Specific Emitter Identification Using Convolutional Neural Network-Based IQ Imbalance Estimators, IEEE Access, vol. 7, pp. 33544-33555, 2019, http://doi.org/10.1109/ACCESS.2019.2903444.

[14] Alperin K. B., Wollaber A. B., Gomez S. R. Improving Interpretability for Cyber Vulnerability Assessment Using Focus and Context Visualizations, 2020 IEEE Symposium on Visualization for Cyber Security (VizSec), Salt Lake City, UT, USA, 2020, pp. 30-39, http://doi.org/10.1109/VizSec51108.2020.00011.

[15] Srivastava, Gautam et al. "XAI for Cybersecurity: State of the Art, Challenges, Open Issues and Future Directions." ArXiv abs/2206.03585 (2022).

[16] Nwakanma, C.I.;Ahakonye, L.A.C.; Njoku, J.N.; Odirichukwu, J.C.; Okolie, S.A.; Uzondu, C.; Ndubuisi Nweke, C.C.; Kim, D.-S. Explainable Artificial Intelligence (XAI) for Intrusion Detection and Mitigation in Intelligent Connected Vehicles: A Review. Appl. Sci. 2023, 13, 1252. https://doi.org/10.3390/app13031252.

[17] Luckey, D., Fritz, H., Legatiuk, D., Dragos, K., Smarsly, K. (2021). Artificial Intelligence Techniques for Smart City Applications. In: Toledo Santos, E., Scheer, S. (eds) Proceedings of the 18th International Conference on Computing in Civil and Building Engineering. ICCCBE 2020. Lecture Notes in Civil Engineering, vol 98. Springer, Cham. https://doi.org/10.1007/978-3-030-51295-8_1.

[18] Javed, A.R.; Ahmed,W.; Pandya, S.; Maddikunta, P.K.R.; Alazab, M.; Gadekallu, T.R. A Survey of Explainable Artificial Intelligence for Smart Cities. Electronics 2023, 12, 1020. https://doi.org/10.3390/electronics12041020.

[19] Dhavalkumar Thakker, Bhupesh Kumar Mishra, Amr Abdullatif, Suvodeep Mazumdar and Sydney Simpson. Explainable Artificial Intelligence for Developing Smart Cities Solutions. Smart Cities 2020, 3, pp. 1353–1382; http://doi.org/10.3390/smartcities3040065.

[20] Zaib Ullah, Fadi Al-Turjman, Leonardo Mostarda, Roberto Gagliardi. Applications of Artificial Intelligence and Machine learning in smart cities, Computer Communications, 154, 2020, pp. 313-323, https://doi.org/10.1016/j.comcom.2020.02.069.

[21] Ashraf Javed , Keshk Marwa, Moustafa Nour, Abdel-Basset Mohamed, Khurshid Hasnat, D. Bakhshi Asim , R. Mostafa Reham. IoTBoT-IDS: A novel statistical learning-enabled botnet detection framework for protecting networks of smart cities, Sustainable Cities and Society, 72, 2021, https://doi.org/10.1016/j.scs.2021.103041.

[22] Wojciech Samek, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, Klaus-Robert Müller. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Springer Cham, Lecture Notes in Computer Science, 2019, ISSN 0302-9743, https://doi.org/10.1007/978-3-030-28954-6.

[23] Gouhara K., Watanabe T. and Uchikawa Y. "Learning process of recurrent neural networks," [Proceedings] 1991 IEEE International Joint Conference on Neural Networks, Singapore, 1991, pp. 746-751 vol.1, http://doi.org/10.1109/IJCNN.1991.170489.

[24] SHAP: Accessed: https://shap.readthedocs.io/en/latest/ ; available : February 2024

[25] SHAP: Accessed: https://github.com/shap/shap ; available : Mars 2024

[26] RNN-SHAP: Accessed: https://github.com/shap/shap/issues/213 ; available : Mars 2024

[27] Shah, A., Rathod, D., Dave, D. (2021). DDoS Attack Detection Using Artificial Neural Network. In: Chaubey, N., Parikh, S., Amin, K. (eds) Computing Science, Communication and Security. COMS2 2021. Communications in Computer and Information Science, vol 1416. Springer, Cham. https://doi.org/10.1007/978-3-030-76776-1_4.

[28] Li Y., Shi H. and Fan M. "DDoS Attack Traffic Identification Using Recurrent Neural Network," 2021 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), Colombo, Sri Lanka, 2021, pp. 1-6, http://doi.org/10.1109/SLAAI-ICAI54477.2021.9664685.

[29] Gaur V. and Kumar R. "DDoSLSTM: Detection of Distributed Denial of Service Attacks on IoT Devices using LSTM Model," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 01-07, http://doi.org/10.1109/IC3IOT53935.2022.9767889.

[30] Ahmed S. W., Kientz F. and Kashef R. "A Modified Transformer Neural Network (MTNN) for Robust Intrusion Detection in IoT Networks," 2023 International Telecommunications Conference (ITC-Egypt), Alexandria, Egypt, 2023, pp. 663-668, http://doi.org/10.1109/ITC-Egypt58155.2023.10206134.

[31] Amma, N.G.B., Selvakumar, S. Optimization of vector convolutional deep neural network using binary real cumulative incarnation for detection of distributed denial of service attacks. Neural Comput & Applic 34, 2869–2882 (2022). https://doi.org/10.1007/s00521-021-06565-8.

[32] Hariprasad S., Deepa T. and Bharathiraja N. Detection of DDoS Attack in IoT Networks Using Sample Selected RNN-ELM. Intelligent Automation & Soft Computing http://doi.org/10.32604/iasc.2022.022856.

[33] Dora, V.R.S., Lakshmi, V.N. Optimal feature selection with CNN-feature learning for DDoS attack detection using meta-heuristic-based LSTM. Int J Intell Robot Appl 6, 323–349 (2022). https://doi.org/10.1007/s41315-022-00224-4.

[34] Kona, Siva Sarat (2020) Detection of DDoS attacks using RNN-LSTM and Hybrid model ensemble. Master's thesis, Dublin, National College of Ireland.

[35] Wang, Y., Gu, D., Wen, M., Xu, J., Li, H. (2010). Denial of Service Detection with Hybrid Fuzzy Set Based Feed Forward Neural Network. In: Zhang, L., Lu, BL., Kwok, J. (eds) Advances in Neural Networks - ISNN 2010. ISNN 2010. Lecture Notes in Computer Science, vol 6064. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-13318-3_71.

[36] Truong Pham N., Dzung Nguyen S., Song Thuy Nguyen V., Hong Pham B. N., Minh Dang D. N. (2023) Speech emotion recognition using overlapping sliding window and Shapley additive explainable deep neural network, Journal of Information and Telecommunication, 7:3, 317-335, http://doi.org/10.1080/24751839.2023.2187278.

[37] Chen C. -Y., Chen L. -A., Cai Y. -Z. and Tsai M. -H., "RNN-based DDoS Detection in IoT Scenario," 2020 International Computer Symposium (ICS), Tainan, Taiwan, 2020, pp. 448-453, http://doi.org/10.1109/ICS51289.2020.00094.

[38] Nadeem M. W., Goh H. G., Aun Y. and Ponnusamy V. "A Recurrent Neural Network based Method for Low-Rate DDoS Attack Detection in SDN," 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS), IPOH, Malaysia, 2022, pp. 13-18, http://doi.org/10.1109/AiDAS56890.2022.9918802.

[39] Karthika R. A., Sriramya P. and Rohini A. "Detection and Classification of DDoS Attacks in Cloud Data Using Hybrid LSTM and RNN for Feature Selection," 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), Kollam, India, 2023, pp. 1491-1495, http://doi.org/10.1109/ICCPCT58313.2023.10244979.

[40] Accessed: UNSW_NB15 dataset: https://www.kaggle.com/datasets/mrwellsdavid/unsw-nb15 ; available: May 2023.

[41] Tang, T.A., McLernon, D., Mhamdi, L., Zaidi, S.A.R., Ghogho, M. (2019). Intrusion Detection in SDN-Based Networks: Deep Recurrent Neural Network Approach. In: Alazab, M., Tang, M. (eds) Deep Learning Applications for Cyber Security. Advanced Sciences and Technologies for Security Applications. Springer, Cham. https://doi.org/10.1007/978-3-030-13057-2_8.

[42] Abusitta, A., Bellaiche, M. & Dagenais, M. An SVM-based framework for detecting DoS attacks in virtualized clouds under changing environment. J Cloud Comp 7, 9 (2018). https://doi.org/10.1186/s13677-018-0109-4.

[43] Awotunde, J.B., Ayo, F.E., Panigrahi, R. et al. A Multi-level Random Forest Model-Based Intrusion Detection Using Fuzzy Inference System for Internet of Things Networks. Int J Comput Intell Syst 16, 31 (2023). https://doi.org/10.1007/s44196-023-00205-w.

AUTHORS

Ahmad Mater Aljohani was a teacher at the University of Tabuk (Computer Sciences of Tabuk, Applied College, Tabuk, Saudi Arabia). He achieved his Master degree from the University of Bedfordshire, UK and he is currently a PhD student at the University of Canberra, Australia. His research interest focuses on security of IoT networks, attacks detection, trust, privacy and DDoS attacks on IoT devices.

Ibrahim Elgendi holds a Ph.D. in Information Technology from University of Canberra, Australia. He is currently a lecturer in Networking and Cybersecurity. His research focuses on Mobile and Wireless Networks, Internet-of- Things, Machine Learning, and Cyber-Physical-Security