Journal of Electrical Engineering and Information Technologies, Vol. **8,** No. 2, pp. 83–92 (2023)
Article 209
Received: October 15, 2023
Accepted: November 29, 2023
In print: ISSN 2545–4250
On line: ISSN 2545–4269
UDC: 004.7.056.57:004.85
https://doi.org/10.51466/JEEIT2382209083s

Original scientific paper

OVERVIEW OF DEEP LEARNING TECHNIQUES FOR NETWORK INTRUSION DETECTION SYSTEMS

Goce Stevanoski[1], Aleksandar Risteski[2], Marko Porjazoski[2]

[1]*Military Academy "General Mihailo Apostolski", Goce Delčev University in Štip, Vasko Karangelevski St., 1000 Skopje, Republic of North Macedonia*
[2]*Faculty of Electrical Engineering and Information Technologies, "Ss. Cyril and Methodius" University in Skopje, Rugjer Bošković bb, P.O.Box 574, 1001 Skopje, Republic of North Macedonia*
goce.stevanoski@ugd.edu.mk

A b s t r a c t: The rapid advances in the new digital world are producing vast amounts of data. This gives more opportunities in business management, but it can also help in implementing new security techniques. Intrusion detection systems (IDS) are enforcing processes for analyzing network data. This study is reviewing the main Deep Learning approaches for intrusion detection in IT network traffic. In the beginning, the study gives an overview of the various IDS types and their usability in the IT network. Then it presents some of the most used Deep Learning techniques proposed by the research community in recent years. By analyzing various papers on the subject, current achievements, and limitations in developing IDS are detected and presented. The study ends by providing the future reach of the newly proposed Deep Learning techniques in monitoring and detecting malicious activities in network traffic.

Key words; intrusion detection systems; machine learning algorithms; deep learning

ПРЕГЛЕД НА ТЕХНИКИ ЗА ДЛАБОКО УЧЕЊЕ НА СИСТЕМИ ЗА ДЕТЕКЦИЈА НА НАПАДИ ВО МРЕЖИ

А п с т р а к т: Брзиот напредок на дигиталното општество предизвикува генерирање голем број податоци. Овие податоци покрај тоа што ги зголемуваат можностите за напредок на бизнисот можат многу да помогнат во справување со безбедносните предизвици во ИТ-инфраструктурата. Улогата на системот за детекција на напади е да врши анализа на мрежниот сообраќај и да ги детектира можните закани. Во трудот се обработени позначајните техники на учењето во длабочина. Разгледани се преку презентирање на нивниот начин на функционирање и како се користат во поновата литература од таа област. Со анализа на релевантни објавени научни трудови на темата се претставени моменталните достигнувања и ограничувања на новите решенија на IDS предложени во последните години користејќи техники на учење во длабочина. На крајот трудот дава преглед на идниот опсег на достигнувањата кои можат да се постигнат со техниките на длабоко учење во делот на мониторинг и детекција на злонамерни активности во мрежен сообраќај.

Клучни зборови: системи за детекција на напади; машинко учење; учење во длабочина

1. INTRODUCTION

Intrusion detection systems (IDs) are a critical segment of every well-established information security posture in an organization. In the security of IT networks, an intrusion is defined as an act of compromising the network resources and servers' infrastructure. An implementation of IDS is providing monitoring, detecting, and intentionally preventing malicious activity on the network to secure the confidentiality, integrity, and availability of the data [1].

The fast-growing evolution of network infrastructure with the introduction of many new services for supporting business continuity is generating enormous quantities of data and information.

This becomes a heavy burden for the IT security personnel and well-positioned IDS can significantly help in mitigating the security risk to organizational data. The implementation of IDS in the network infrastructure can help in detecting and preventing unusual activity in the network traffic.

Traditional IDS are based on technology that is enforcing too many rules on network traffic and this becomes a headache for the IT staff. On the other hand, Machine Learning (ML) technologies are giving an innovative approach to developing and utilizing IDSs in the network infrastructure. Since the beginning of the adoption of ML algorithms in the science community, ML was seen as a new and a potential way ahead for improving the shortcomings of the traditional IDS. The first ML-based IDSs were applying traditional ML algorithms which were able to notable classify the bad traffic from the good network traffic. This led to the introduction of many new ML technologies for IDS implementation.

Nowadays, Deep Learning (DL) techniques are the front runners in the application of ML in IDSs. The innovative approaches are giving remarkable results for pattern recognition and anomaly detection in network traffic. These results provide a very good foundation for enforcing DL techniques in IDS.

?. *Related work*

In the literature, the relevant studies for reviewing the DL approach in building IDS are done for synthesizing the current trends and achievements of the scientific community. In [2] the authors have analyzed the research in IDS using the ML approach until 2019 with a specific interest in the dataset, DL techniques, and metrics. Concluding that soft computing techniques are on the rise and that researchers are using old datasets that can limit the development of ML-based IDSs.

In another paper, [3], is providing a review of the ML technologies until 2020 by studying the proposed methodology, evaluation metrics, and dataset selection and discussing the strengths and limitations of the proposed solutions. Here, the author highlights various research challenges for the future.

In one of the most recent papers on the topic, the authors of [4] have done an incredibly detailed comparative study on various ML techniques such as artificial neural networks, support vector machines, decision trees, and hybrid classifiers. This study presents some of the datasets used in ML and the performance metrics for evaluating the ML models. Future work is also discussed.

2. INTRUSION DETECTION SYSTEMS

An Intrusion Detection Systems (IDS) makes regular checks on the network traffic for malicious activity on all inbound and outbound packets. If malicious activity is detected IDS can also enforce a security mechanism for preventing damage to the network environment and notify the system administrator about the attack. In comparison to the Firewall, the IDS has better detection of interior attacks, and it provides more reliable strategies for securing the perimeter.

The IDS can be classified into distinct categories depending on the implementation in the operational environment or the detection mechanism. IDS can be implemented as a Host Intrusion Detection System (HIDS) and as a Network Intrusion Detection System (NIDS).

Host intrusion detection system (HIDS) inspects traffic that originates from/to one device in the network. This type is limited in performance and cannot detect what is going on in the other parts of the network environment. Usually detects unusual connections, file changes, and file removal on one system and notifies the user of that system.

For inspecting network traffic, a NIDS is used. NIDS is part of the overall network security infrastructure on the organizational level. Usually is implemented on the main entering network points (gateways, routers) and it inspects the incoming and outgoing network traffic. NIDS checks the traffic for the known attacks' signatures in the data packets. If a match is found, NIDS can prevent damage to the network infrastructure and sends and notifications to the system administrator.

Signature-based detection mechanisms usually mean that the IDS must be regularly updated with the signature of newly discovered attacks. Otherwise, the IDS will be unable to detect and prevent new attacks. So, the inability to detect unknown attacks is one of the main shortcomings of IDS [5] and is a challenge for IDS based on Machine Learning techniques.

Machine learning intrusion detection systems

Machine learning-based implementation of IDS is a process of learning different patterns in data by machines (computer systems) from previously col-

lected data and applying those patterns to newly acquired data. By this, the machines can make predictions about what kind of data is and act on the data or other systems. ML can be categorized based on the learning approach and the functionality of how they work on new data. The three main types of ML techniques categorized by the learning approach are supervised learning, unsupervised learning, and semi-supervised learning.

In supervised learning, the machine learns from labeled data to construct a pattern for future predictions. This learning is used for both classification and regression problems. The following are some of the most common ML algorithms for supervised learning: k-Nearest Neighbor, Decision Tree, Naïve Bayes, Support Vector Machine, Random Forest Algorithm, and Linear Regression Algorithm.

On the other hand, in unsupervised learning, machine learning with unlabeled data detects unknown cases. Some common ML algorithms for unsupervised learning are Hidden Markov Model, K-means Self-Organizing Map.

In semi-supervised learning, the machine learns with both labeled and unlabeled data. Some examples of semi-supervised algorithms are SVM, Gaussian Fields Approach, and Spectral Graph Transducer.

Furthermore, ML algorithms can be classified as Shallow Learning and Deep Learning. ML techniques with few layers are known as Shallow Learning (SL) and they are better for less complex tasks. The newly raised technique which uses more layers of a neural network is named Deep Learning (DL). DL is used for complex tasks and on a larger dataset.

ML techniques are noted as one of the best approaches for the effective development of IDS by providing, a high positive alarm rate, low false alarm rate, and improved detection rate [6]. These types of IDSs are using learning-based systems that can detect classes of attacks by comparing normal (benign) and bad (attack) traffic behavior. In our work we discuss the latest DL techniques for intrusion detection in IT networks.

3. DEEP LEARNING TECHNIQUES FOR IDS

Neural Networks have brought a new approach to building ML models. The DL techniques have shown that traditional ML techniques can be replaced by more efficient and more accurate ML models. In the following part, we are presenting some of the most used DL techniques for building ML models for IDS.

Restricted Boltzmann machine

Restricted Boltzmann machine (RBM) can find patterns in the data by reconstructing the input. Hinton in [7] introduced an approach that created the restricted Boltzmann machine. RBM, depicted in Figure 1, is a shallow two-layer network with visible and hidden layers. Each node in the visible layer relates to each node in the hidden layer. RBM is considered restricted because no two nodes in the same layer share a connection. The nodes are conditionally independent of each other in the same layer. In the forward pass, RBM takes the input and translates that into the set of numbers that encodes the input. In the backward pass, RBM takes the output set of numbers and sends them back to the visible layer to reconstruct the input. A well-trained network will be able to make translations with a remarkably high level of accuracy. In RBM weights and biases have a particularly important role. They help the RBM to determine the relationships among the import features and help RBM to decide which features are most important when they detect patterns. The training process consists of many forward and backward passes that are helping the RBM to reconstruct the input data. In the beginning, every input is combined with every individual weight and one overall bias. The result may or may not activate the hidden neurons. The same process is done in the next step with the backward process. In the backward pass, the input is combined with every individual weight and the overall bias. In the visible layer, the returned result is compared to the initial input to determine the quality of the results.

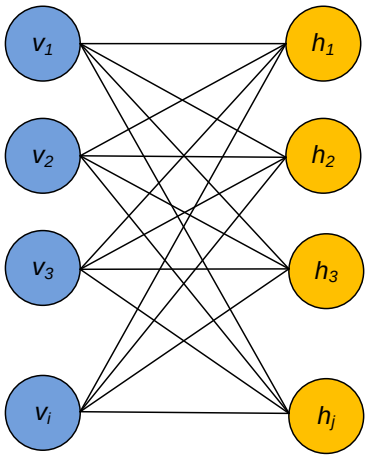

Fig. 1. Architecture of restricted Boltzmann machine (RBM)

This process is repeated until the input and the reconstructed result are as close as possible. Here learning means assigning a probability to every possible pair of visible/hidden vectors. The estimation of updating values is presented with Eqs. 1, 2, and 3 [8].

$$\Delta w_{ij} = \epsilon \left(v_i^{(0)} h_j^{(0)} - v_i^{(T)} h_j^{(t)} \right) \quad (1)$$

$$\Delta a_i = \epsilon \left(v_i^{(0)} - v_i^{(T)} \right) \quad (2)$$

$$\Delta b_j = \epsilon \left(h_j^{(0)} - h_j^{(t)} \right) \quad (3)$$

where:

v_i – is state of the visible unit i,

h_j – is state of the hidden unit j,

a_i – is bias of visible unit i,

b_j – is bias of hidden unit j,

w_{ij} – weight between visible unit and the hidden unit,

Δw_{ij} – is updating value for w_{ij},

Δa_i – is updating value for a_i,

Δb_j – is updating value for b_j,

ϵ – is the learning rate,

T – times of probabilistic distribution,

t – ????.

The application of RBM for network intrusion detection does not have many applications proposed by the scientific community. The authors in [9] are proposing an intrusion detection method with RBM as one phase in the methodology. In this paper, the authors are proposing the application of RBM for feature extraction since RBM has proven performance for unsupervised feature extraction which could be efficient to learn user behaviors from raw traffic data. Furthermore, the proposed method is utilizing the Feed-Forward Neural Network, automated FFNN, Random Forest (RF), and Support Vector Machine (SVM) ML techniques for classifying the network traffic in the observed domain. Their work shows that the proposed methodology can detect DDoS attacks efficiently and accurately.

Convolutional neural networks

Convolutional neural networks (CNNs) were introduced to overcome the problem of too many weights in neural networks. Used mostly in image recognition, CNN doesn't need data preprocessing, it can process raw data and extract the needed features. CNN is using spares connections and weight sharing which lowers the need for computational power and helps in the time complexity of the training. The data is transitioning over two layers, the convolutional layer, and the pooling layer of CNN. The process of training, depicted in Figure 2, is repeated for several convolutional – pooling iterations. In the convolutional layer the data is convoluted with fix sized kernels which sample the data to clusters – convoluted representation of the data and the kernel. Further, the pooling layer reduces the dimensionality of the convoluted data clusters. This reduction can be enforced with distinct functions but the most established are max-pooling and average pooling. With max-pooling the maximum value of the cluster is taken for the next iteration, whereas average pooling the average value of the processed cluster is taken for the next iteration.

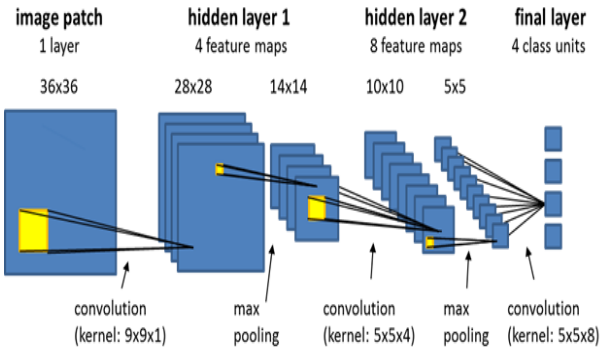

Fig. 2. Overview of the training process for CNN

As CNN is used for image recognition its use in IDS is challenging. There is much research done with CNN for IDS but developing a CNN model for IDS that will be reliable and effective is still a problem. The authors in [10] have notated this and proposed a method that will use evolutionary convolutional neural network (ECNN) for IDS. They are suggesting the use of multi-objective immune algorithm to optimize the accuracy and weights of CNNs. Their model has shown higher detection accuracy when compared to some state-of-the-art algorithms. Furthermore, the authors of [11] have suggested the use of IDS which can identify attacks by using model based on CNN. Here they are removing the redundant and irrelevant features in the network traffic data using various dimensionality reduction methods. Later the features from the dimensionality reduction data are extracted using CNN and the traffic vector had been converted to

image so to reduce the computational cost. The authors proved that they method have not only theoretical value but also and the practical value by experimenting with the method on standard KDD-CUP99 dataset.

The authors at [12] have depict the usage of Binary Grey Wolf Optimization for detection of optimal features from the data. They have proposed a CNN approach named TreeNets. This method is identifying the attacks and segregates them into binary outcomes. With their work they exhibit three variants of TreeNets and made a comparison with a known state-of-the-art ML and DL models. The experiments have shown respectable results in detecting suspicious activity.

Overall, the scientific community is working on CNN application for IDS, but the work is not providing reliable and effective results. CNN can significantly improve the accuracy of classification but the convergence speed and ability for generalization of CNN is a problem that should be addressed in the future.

Recurrent neural networks

Recurrent Neural Network (RNN) is a type of Deep Learning network which is used for processing sequential data in layers of neurons. RNN is mostly used for supervised classification learning and incorporates input, hidden and output layers. Much of the work for data processing is done in the hidden layer. The architecture of RNN is shown in Figure 3. This network forms connections between the nodes in a directed structure and encompasses a feedback loop for applying the output values in the input layer. In building bigger models RNN is suffering from a vanishing/exploding gradient problem and the solution can be found in implementing Long-Short Term Memory (LSTM) and Gated Recurrent Units (GRU) models. These models possess an internal cell for storing input values which helps in processing directly corelated data from the dataset. The best results are achieved in applications where the output is predicted by analyzing the previous values of the data.

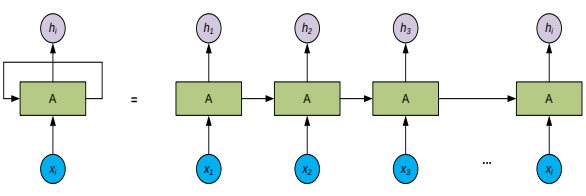

Fig. 3. Architecture of RNN

In [12] the authors are suggesting a hybrid DL model for IDS by using RNNs. The method is based on LSTM type of RNN and is proposed improved long-time memory tree model. This approach should have a secondary detection for solving problems with high false negative rate in RNN based IDS. For proving their work, they have made experiments with NSL-KDD data set. In another method, the authors of [13] have address the intrusion detection with combining two DL techniques CNN and RNN. Their work is done in two parts. The first is analyzing the payloads with CNN classification and the second part is detecting an attack with RNN classification. They are suggesting that the proposed model is learning the features without feature engineering, supports end-to-end detection and shows better results compared to some state-of-the-art methods. Similar approach is proposed by authors of [14] suggesting a hybrid model with CNN and LSTM-RNN networks. They are presenting that the model has obtained a high accuracy as 94.4% with previous hyper parameter tuning done by using deep learning architecture.

The ability of RNNs to store input values and use them later for making data correlation is a significant advantage in comparison to other NNs when are sequential data is processed. Combination of RNNs with various ML techniques in hybrid models has been seen as a right approach in application of RNNs for IDS.

AutoEncoder neural network

AutoEncoder (AE) is an unsupervised DL technique. It is used for building artificial neural networks that can learn data representations on the input of the network, to reconstruct as an output. AE is learning the features of a set of data, the purpose of it is for dimensionality reduction. The network, depicted in Figure 4, is symmetrically constructed and the numbers of neurons and layers on the input are the same as on the output. In the middle between the input and output layer, there is the smallest layer of all which is called the bottleneck layer.

The main reason for constructing this type of layer is to encode the input data from the input layer into the bottleneck layer and later decode the data on the output layer. Because of this functionality, the input layer is known as the encoder and the output layer is known as the decoder.

This process of reconstructing the input data on the output of the AE comes with some loss of level of data accuracy.

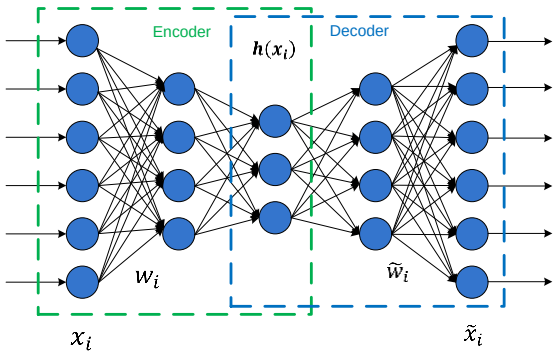

Fig. 4. Architecture of AutoEncoder neural network

If the AE is well-trained the loss is minimal. Eq. 4 represents the cost function of the AE.

$$J_{A,E} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{2}\|\hat{x}_i - x_i\|^2\right), \qquad (4)$$

where,

$J_{A,E}$ – is the value of the cost for input data,
n – is number of input samples,
x_i – is the value of the input data,
$\hat{x}_i$ – is the value of the output data.

The AE training is done in two stages in the first stage a Contrastive Divergence (CD) is used between neighboring pairs of input layers. In the second part with back backpropagation, the whole network is for finetuning. The process of reconstructing the data on the output is useful in many cases. A well-trained model can enforce noise reduction on the data and emphasize the meaningful feature of the input data, variational autoencoder can be trained to generate new data and by application on pattern recognition problems, the AE can detect the anomalies in sequential data.

In IDS AE is used for detecting anomalies in network traffic by segregating the usual traffic from the unusual traffic. This application of AE is constantly evolving, and the results are improving. In [15] the authors proposed a deep learning classification method in the process of data preprocessing for feature extraction. This approach has led them to the improved classification of performance and detection speed. The authors in [16] have proposed a similar method by combining DL and SL techniques. They presented an effective stacked contractive autoencoder (SCAE) for feature extraction from the raw network traffic. Further, they have used the SVM classification algorithm for improving the detection performance on two different evaluation datasets KDD Cup 99 and NSL-KDD. In the next paper [17] the authors are bringing together the AE and the Improved Genetic Algorithm BP (IGA-BP) in one proposed method. AE is used for the elimination of redundant information and for reducing data dimensionality. The IGA-BP network model solved the problems with slow detection rate and getting easy into local optimality in the BP network. Their findings from the experiments have shown that the proposed method has a significant effect on classification accuracy, false positives, and detection rate.

Overall, the work with autoencoders in the recent period is focused on improving the feature extraction in the preprocessing phase of the model building and combining the results with different ML techniques.

Generative adversarial networks

Generative Adversarial Networks (GANs) are a class of artificial intelligence algorithms used in unsupervised machine learning. GANs are implemented by a system of two neural networks contesting with each other in a zero-sum game framework. From a wide perspective GANs consist of two parts: the generator (G) and the discriminator (D), depicted in Figure 5.

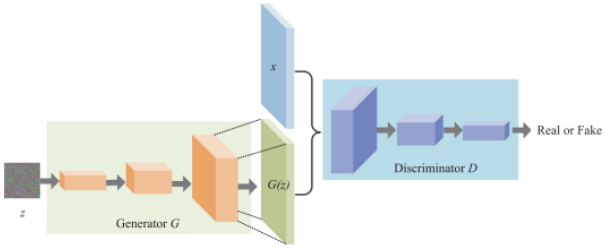

Fig. 5. The illustration of generative adversarial network

The generator creates data that is as realistic as possible, while the discriminator evaluates the data, trying to distinguish between real and generated data.

The primary goal of GANs is to generate high quality, realistic data, which makes them powerful tools for image generation, video generation, time-series data generation and more. In the system architecture generator and discriminator are typically deep neural networks. The architecture is designed such that both networks improve their performance in response to the other, creating a dynamic learning environment. During training, the generator's goal is to produce data that is indistinguishable from real data, thereby 'fooling' the discriminator. The discriminator learns to become better at distinguishing

real data from the fake data created by the generator. This adversarial process continues until the generator produces data so close to real that the discriminator cannot reliably distinguish fake from real. Eq. 5 represents objective function of GANs that can be transformed into a minimax game,

$$\min\max V(D,G) = E_{x\sim p_{data(x)}}[\log D(x)] - \\ - E_{z\sim p_{z(z)}}[log(1-D(G(z)))], \quad (5)$$

where x indicates the real data distribution from $p_{data(x)}$, E indicates the expectation, z indicates the vector from the random noise distribution $p_z(z)$, $G(z)$ and $D(x)$ indicates the samples generated by the generator, and the probability that D discriminates x as real data, respectively.

$D\big(G(z)\big)$ indicates the probability that D determines the data generated by G. For the generator G, to fraud D, the discriminator probability $D\big(G(z)\big)$ needs to be maximized, so $\log\Big(1-D\big(G(z)\big)\Big)$ will be minimized. For the discriminator D, a cross entropy function is used to distinguish between $G(z)$ and x, and D wants $V(D,G)$ to be maximized. In practice, G will be fixed firstly, and the parameters of discriminator D are updated to maximize the accuracy of D. And then, D is fixed to optimize G. When G and D have a sufficient capacity, the model will converge, and these two parts will reach the Nash equilibrium. At this time, $p_{data(x)} = pg(x)$, and the discriminator cannot determine the differences between these two distributions.

While primarily known for data generation in various fields including art generation, photo-realistic image synthesis, style transfer, and more, GANs can also be applied in anomaly detection for intrusion detection, where they learn to generate 'normal' data, and deviations from this can be flagged as anomalies. GANs are actively researched, and many researchers have shown the potential of GANs for intrusion detection.

AnoGAN was one of the first GAN that was proposed for anomaly detection in image data [18], and latter was applied to GAN-based intrusion detection in [19]. In AnoGAN, the generator receives information about a real instance and, if this was anomalous, it would learn to produce a fake benign instance using a technique called feature matching. A major disadvantage is that AnoGAN required a time-consuming backpropagation process to define the mapping for anomaly scoring and reconstruction of the data [20].

The project Efficient GAN-Based Anomaly Detection (EGBAD) improved the performance of AnoGAN by using a BiGAN design [19]. The EGBAD and AnoGAN were evaluated on the KDD-99 network intrusion dataset [21]. The influence of BiGAN encoder helped in eliminating much of the computation necessary for the scoring and reconstruction that originally influenced AnoGAN's training performance. This helped in increasing binary classification performance. For evaluation the authors used F1-scoring, which gives a clearer estimation of the performance then using raw accuracy. Results show that AnoGAN achieved an F1-score of 78.52%, EGBAD achieved 93.72% [18].

In IDSGAN, the generator was used for creating malicious traffic and the discriminator received responses from a simulated intrusion-detection system as input. In the research the model is tested on NSL-KDD dataset [21]. The IDSGAN was implemented based on a GAN using the performance metric of Wasserstein loss [22], this produced an "authenticity" score instead of a probability that an instance is part of the real dataset. With this the generator receives more specific information about how to adjust its weights and with that can create more stable model [23]. Here the intrusion-detection system was simulated with one of the following machine-learning algorithms: a Support-Vector Machine, Logistic Regression, Naive Bayes, k-Nearest Neighbors, Linear Programming, and Random Forest.

One proposal that has notably results was the Generative Adversarial Network Intrusion-Detection System (GIDS). The authors in this proposal are implementing a raw Controller Area Network (CAN) traffic data [24]. The CAN protocol it resembles industrial control system protocols and is used in automotive vehicles. The researchers tested two discriminators. The first one was trained on real traffic data, both malicious and benign. The second discriminator was trained on both fake and real data generated from the generator. Here the generator took a combination of normal and noise traffic data as input. For evaluation of the proposal the authors used accuracy, which is the proportion of correct predictions over total tested instances. In this research the first discriminator has detected over 99% of the malicious data that it was tested on, on the other hand, the second discriminator reached 98% accuracy on the attacks that were not known for the first discriminator. The proposed model shows that with combination of the two discriminators a 100% accuracy can be achieved.

Continuous advancements in GAN technology have led to more efficient and effective models. For instance, some research has focused on stabilizing the training process, which is traditionally challenging due to the adversarial nature of the networks.

In summary, GANs represent a significant breakthrough in the field of generative models, with their unique adversarial approach to learning, enabling a wide range of applications from realistic image generation to complex anomaly detection systems.

4. CHALLENGES AND OPPORTUNITIES IN FUTURE DL BASED IDS

Deep learning for intrusion detection systems (IDS) is a rapidly evolving field, with significant potential for advancements and innovations. We can highlight several areas that are interesting for future research and development. These areas not only focus on enhancing the accuracy and efficiency of intrusion detection but also address broader challenges such as adaptability, scalability, and integration with emerging technologies. Some of the future challenges include:

1. Federated learning for IDS [25]

This approach is working on developing federated learning approaches for IDS that enable collaborative learning across multiple devices or networks without sharing raw data, enhancing privacy and data security. This can lead to development of robust models that can learn from diverse network environments and adapt to local conditions without compromising sensitive data.

2. Explainable AI (XAI) in IDS [26]

Integrating XAI methods to make deep learning-based IDS more transparent and interpretable, helping security analysts understand and trust the decisions made by the system. They have the potential to develop IDS that provide actionable insights and explanations for detected threats, improving decision-making and response strategies.

3. Anomaly detection with advanced GANs [27]

In this approach advanced Generative Adversarial Networks (GANs) can be used for more sophisticated anomaly detection, especially for identifying novel or sophisticated attacks. In the future works GANs can simulate complex attack scenarios for better training of IDS and more effective detection of previously unseen attack types.

4. Integration with blockchain technology [28]

In this approach a combination of deep learning-based IDS with blockchain technology is used for ensuring secure and trustable network transactions. This can potentially lead to development of IDS solutions that ensure data integrity and traceability, enhancing accountability in security operations.

5. Handling encrypted traffic [29]

The main course of action for future work is in developing methods to detect malicious activities in encrypted traffic without decryption, respecting privacy while ensuring security. This can lead to potential novel approaches or models capable of analyzing encrypted data patterns to identify potential threats.

Future research in deep learning for intrusion detection systems is likely to focus on enhancing adaptability, explainability, real-time processing, and integration with other emerging technologies like IoT, Blockchain etc. These advancements aim not only to improve detection accuracy but also to address broader challenges in these technologies.

4. CONCLUSION

In this study, we made an overview of some of the most researched DL approaches applied in IDS. DL techniques are becoming more and more sophisticated and are taking over the field of ML for IDS from the traditional ML techniques. This is done by developing new DL applications or integrating the DL models with other techniques in various ensemble methods.

DL techniques, particularly those involving neural networks, are highly effective in detecting complex patterns and anomalies in data. They can identify subtle, non-linear relationships that traditional methods might not detect. Additionally, they can learn from new data continuously, allowing them to adapt to evolving cyber threats more effectively than traditional systems, which often rely on predefined rules and signatures. Applying these models it can help in automatically identify and extract relevant features from data, reducing the time and expertise required for system setup and maintenance.

Besides the positive aspects of DL for IDS there are shortcomings of these techniques that need to be addressed. The significant computational resources and power that these techniques require

may not be feasible for smaller organizations or in environments with limited infrastructure. Additionally, due to their complexity, DL models are prone to overfitting, especially if not trained with sufficiently large and diverse datasets. Transparency and interpretability can be an issue too, since the DL models, particularly deep neural networks, are often seen as 'black boxes' due to their complex internal workings, making it difficult to understand the rationale behind specific detections or predictions.

The work has shown that the future of DL for IDS also brings many challenges and opportunities for development of new applications in various emerging technologies. This shows that DL for IDS is a very broad point for discussion in the field of cybersecurity and can contribute to the efforts for enhancing that field.

REFERENCES

[1] McHugh, J. (Aug. 2001): Intrusion and intrusion detection, *Int J Inf Secur*, Vol. **1**, no. 1, pp. 14–35. DOI: 10.1007/s102070100001

[2] Kok, S., Abdullah, A., Zaman, N., Supramaniam, M. (Nov. 2019): A review of intrusion detection system using machine learning approach, *International Journal of Engineering Research and Technology*, Vol. **12**, pp. 8–15.

[3] Ahmad, Z., Shahid Khan, Wai Shiang, A., C., Abdullah, J., Ahmad, F. (Jan. 2021): Network intrusion detection system: A systematic study of machine learning and deep learning approaches, *Transactions on Emerging Telecommunications Technologies*, Vol. **32**, no. 1. DOI: 10.1002/ett.4150

[4] Prethija, G., Katiravan, J. (2022): *Machine Learning and Deep Learning Approaches for Intrusion Detection: A Comparative Study*, Springer, pp. 75–95. DOI: 10.1007/978-981-16-5529-6_7

[5] Patcha, A., Park, J.-M. (Aug. 2007): An overview of anomaly detection techniques: Existing solutions and latest technological trends, *Computer Networks*, Vol. **51**, no. 12, pp. 3448–3470. DOI: 10.1016/j.comnet.2007.02.001

[6] Dayal, N., Maity, P., Srivastava, S., Khondoker, R. (Dec. 2016): Research trends in security and DDoS in SDN, *Security and Communication Networks*, Vol. **9**, no. 18, pp. 6386–6411. DOI: 10.1002/sec.1759

[7] Hinton, G. E. (Aug. 2002): Training products of experts by minimizing contrastive divergence, *Neural Comput*, Vol. **14**, no. 8, pp. 1771–1800. DOI:10.1162/089976602760128018

[8] Ogawa, S., Mori, H. (2019): A Gaussian-Gaussian-Restricted-Boltzmann-Machine-based Deep Neural Network Technique for Photovoltaic System Generation Forecasting, *IFAC-PapersOnLine*, Vol. **52**, no. 4, pp. 87–92. DOI: 10.1016/j.ifacol.2019.08.160

[9] Elsaeidy, A., Munasinghe, K. S., Sharma, D., Jamalipour, A. (Jun. 2019): Intrusion detection in smart cities using Restricted Boltzmann Machines, *Journal of Network and Computer Applications*, Vol. **135**, pp. 76–83. DOI: 10.1016/j.jnca.2019.02.026

[10] Chen, Y., Chen, S., Xuan, M., Lin, Q., Wei, W. (May 2021): Evolutionary convolutional neural network: An application to intrusion detection, In: *2021 13th International Conference on Advanced Computational Intelligence (ICACI)*, IEEE, pp. 245–252. DOI: 10.1109/ICACI52617.2021.9435859

[11] Xiao, Y., Xing, C., Zhang, T., Zhao, Z. (2019): *An intrusion detection model based on feature reduction and convolutional neural networks*, IEEE Access, Vol. **7**, pp. 42210–42219. DOI: 10.1109/ACCESS.2019.2904620

[12] Mishra, S., Dwivedula, R., Kshirsagar, V., Hota, C. (Jan. 2021): Robust detection of network intrusion using tree-based convolutional neural networks, In: *8th ACM IKDD CODS and 26th COMAD*, New York, NY, USA: ACM, pp. 233–237. DOI: 10.1145/3430984.3431036

[13] Wang, S., Xia, C., Wang, T. (May 2019): A novel intrusion detector based on deep learning hybrid methods, In: *2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, IEEE, pp. 300–305. DOI: 10.1109/BigDataSecurity-HPSC-IDS.2019.00062

[14] Liu, H., Lang, B., Liu, M., H. Yan, M. (Jan. 2019): CNN and RNN based payload classification methods for attack detection, *Knowl Based Syst*, Vol. **163**, pp. 332–341. DOI: 10.1016/j.knosys.2018.08.036

[15] Akarsh, S., Simran, K., Poornachandran, P., Menon, V. K., Soman, K. P. (Mar. 2019): Deep learning framework and visualization for malware classification, In: *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, IEEE, pp. 1059–1063. DOI: 10.1109/ICACCS.2019.8728471

[16] Lee, J., Pak, J., Lee, M. (Oct. 2020): Network Intrusion detection system using feature extraction based on deep sparse autoencoder, In: *International Conference on Information and Communication Technology Convergence (ICTC)*, IEEE, pp. 1282–1287. DOI: 10.1109/ICTC49870.2020.9289253

[17] Deng, H., Yang, T. (Jul. 2021): Network intrusion detection based on sparse autoencoder and IGA-BP network, *Wirel Commun Mob Comput*, Vol. **2021**, pp. 1–11. DOI: 10.1155/2021/9510858.

[18] Di Mattia, F., Galeone, P., De Simoni, M., Ghelfi, E. (Jun. 2019): A Survey on GANs for Anomaly Detection.

[19] Zenati, H., Foo, C. S., Lecouat, B., Manek, G., Chandrasekhar, V. R. (Feb. 2018): *Efficient GAN-Based Anomaly Detection*.

[20] Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., Langs, G. (March 2017): *Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery*.

[21] Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A. A. (July 2009): A detailed analysis of the KDD CUP 99 data set, In: *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, IEEE, pp. 1–6. DOI: 10.1109/CISDA.2009.5356528

[22] Martin Arjovsky, S. C. L. B. (2017): Wasserstein generative adversarial networks, In: *ICML'17: Proceedings of the 34th International Conference on Machine Learning*, pp. 214–223.

[23] Lucic, M., Kurach, K., Michalski, M., Gelly, S., Bousquet, O. (Nov. 2017): *Are GANs Created Equal?* A Large-Scale Study.

[24] Seo, E., Song, H. M., Kim, H. K. (Aug. 2018): GIDS: GAN based intrusion detection system for In-Vehicle Network, In: *16th Annual Conference on Privacy, Security and Trust (PST)*, IEEE, pp. 1–6. DOI: 10.1109/PST.2018.8514157

[25] Agrawal, S. et al. (Nov. 2022): Federated Learning for intrusion detection system: Concepts, challenges and future directions, *Comput Commun*, Vol. **195**, pp. 346–361. DOI: 10.1016/j.comcom.2022.09.012

[26] Patil, S. et al. (Sep. 2022): Explainable artificial intelligene for intrusion detection system, *Electronics (Basel)*, Vol. **11**, no. 19, p. 3079. DOI: 10.3390/electronics11193079

[27] Jain, S., Seth, G., Paruthi, A., Soni, U., Kumar, G. (Apr. 2022): Synthetic data augmentation for surface defect detection and classification using deep learning, *J Intell Manuf*, Vol. **33**, no. 4, pp. 1007–1020. DOI: 10.1007/s10845-020-01710-x

[28] Mathew, S. S., Hayawi, K., Dawit, N. A., Taleb, I., Trabelsi, Z. (Dec. 2022): Integration of blockchain and collaborative intrusion detection for secure data transactions in industrial IoT: a survey, *Cluster Comput*, Vol. **25**, no. 6, pp. 4129–4149. DOI: 10.1007/s10586-022-03645-9

[29] Bakhshi, T., Ghita, B. (Sep. 2021): Anomaly detection in encrypted internet traffic using hybrid deep learning, *Security and Communication Networks*, Vol. **2021**, pp. 1–16. DOI: 10.1155/2021/5363750