Improving network security and traffic regulation through deep learning systems

Hind Khalid[1]
[1]College of Political Science, University of Nahrain, Baghdad, Iraq

ABSTRACT : *The network security research community has placed great emphasis on detecting anomalies in network traffic, highlighting the importance of network security. With the exponential expansion of the Internet and its widespread integration into our daily lives, many obstacles must be overcome for network anomaly detection to be effective. Such as the lack of representative datasets, the high cost of errors, and the dynamic nature of network traffic, although the volume of publicly available datasets has recently increased, addressing the issue of how to adapt to the dynamic nature of network traffic and security concerns is often neglected. This study attempts the potential application of a dynamic drawing algorithm known as ANN using existing network traffic statistics. An unsupervised anomaly detection system can adapt to insight transformations in the data and does not require any prior training. We verify the detection performance of the method in a realistic simulation environment by applying hyperparameter tuning, the system shows promising results in detecting anomalies with a comprehensive ability to distinguish between anomalies and normal states outperforming a similar static model.*

KEYWORDS- network security, traffic statistics, anomaly detection, deep learning.

I. INTRODUCTION

We are currently witnessing a major shift towards digitization in various aspects of our lives. The use of devices such as computers, smartphones, and tablets has become increasingly widespread, leading to increased reliance on online services. Social interactions, business operations, and financial transactions are now commonly conducted via computer networks. This trend is more emphasized through recent events such as the wide -ranging adoption of remote work and increased online activities [1, 2].

The network traffic cannot be dealt with as fixed because it is constantly developing over time and vary between different networks. In this thesis, we study the potential use of a dynamic model that can adapt to the changing nature of the network traffic continuously. By doing this, we aim to address the restrictions associated with the current traditional curricula. On the signature, our proposed dynamic model provides the possibility of overcoming these restrictions by providing more accurate and effective anomalies detection capabilities and thus strengthening network safety [3, 4].

The increasing dependence on digital infrastructure emphasizes the decisive importance of network security in both the public and private sectors. Weaknesses or malignant activity can settle the integration of the network and expose the sensitive data of the risks. The discovery of abnormal cases is a specific topic in the field of network safety and is often mentioned in relation to network infiltration systems (NIDS), there are still many current solutions to detect homosexuality in the network. Despite the great research efforts in this field, many variables such as high error rates can be blamed. Unfortunately, people tend to ignore how the dynamic network movement [5, 6].

The ancient Greek word "anomalous," which means uneven or irregular, is where the word "anomaly" originated. It is used frequently in a variety of fields. Other names for anomalies include outliers, aberrations, irregularities, and anomalies. The terms "outliers" and "anomaly" are used Commonly used in machine learning.

Anomaly detection, or outlier detection, entails the process of identifying these deviations from the norm. While these terms are often used interchangeably, outlier detection usually refers to the process of cleaning data by removing outlier samples to improve model performance [7]. Conversely, anomaly detection focuses on Anomalies focus on the anomalies themselves, their characteristics, and their potential causes. Understanding the causes of anomalies can be critical, especially when intervention is required. Anomalies can arise from malicious intent (e.g., credit card fraud, network intrusion) or a critical failure. In system (e.g., predictive maintenance) or medical contexts (e.g., malignant tumors in MRI scanning) and as a result anomaly detection plays a vital role in areas such as fraud detection, fault detection in critical systems, and intrusion detection in cybersecurity [7, 8].

Two important characteristics of anomalies, especially in the field of data mining and machine learning, include [9]:

1. Anomalous cases exhibit distinctive characteristics that distinguish them from normal cases.

2. They are rare occurrences compared to normal cases within the data set.

The definition of what is considered rare varies and there is no fixed threshold. However, the general rule is that extreme samples should not exceed 5% of the data set. By adhering to the definition of anomalies as patterns that deviate from normal behavior, different types of anomalies can be identified at the abstract level [10]. To provide an illustration let's consider examples in the context of network traffic. It is important to note that these examples have been chosen for simplicity and may not necessarily represent anomalies of importance in real-world use cases.

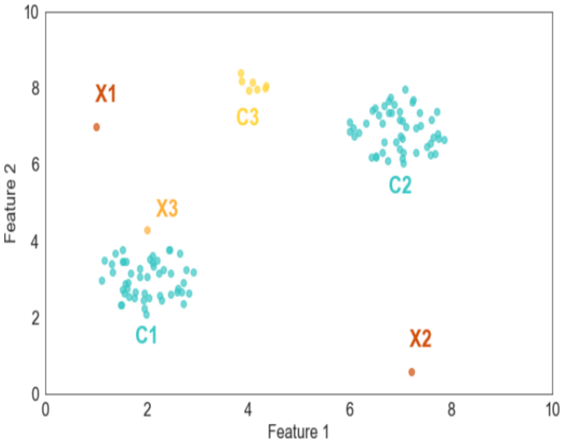

Figure 1. Two-dimensional Anomaly Example
.

Abnormality detection entails identifying anomalies or abnormal patterns within a given dataset. This process received a great deal of attention and investigation over several decades, which led to the publication of numerous reviews, surveys and books on the subject. Various research areas such as statistics, machine learning, information theory, and others have contributed to addressing the challenges associated with anomaly detection [11, 12].

II. REFRERNCE STUDY

Nearly twenty years ago, network detection has been widely searched in the network of the network, and since the publication of Dorothy Dening paper for the year 1987 entitled 'The Crossing Model' Form, the main field of study on this topic was the infiltration detection systems detection systems (IDS) [13, 14].

In order to discover unusual behavior in computer systems, Denning provided a comprehensive model that takes into account the patterns of use, statistical models, people and things. This approach depends on the idea that the unusual use of the system is necessary for the exploitation of the system weaknesses. IDS systems are still based on this hypothesis. This allows it to identify anomalies without the need to be aware of previously identified weaknesses in the system.

Currently, intrusion detection systems can be classified into two types: misuse detection systems and anomaly detection systems. Anomaly detection systems are particularly useful in identifying new anomalies. In network security research, this area overlaps significantly with outlier detection, and many algorithms have been applied. Outlier detection on network data has also been introduced [15, 16]. These techniques are drawn from different disciplines such as machine learning, statistics, and information theory [17, 18].

In this field, tree-based methods have been widely applied. For example, P. Uyyala presented an end-to-end intrusion detection system (IDS) that uses the Random Forest algorithm to detect misuse and anomalies. They created a service prediction model to create a proximity metric to identify new anomalies. They used supervised learning Random Forest to detect previous attacks [19]. By adding a centroid-based distance metric and using the distance to known anomalies to improve detection, this work [20, 21] proposed modifying the Random Forest for intrusion detection. A hierarchical intrusion detection system was presented using vector machines. Support SVMs and Decision Trees by previous works like [22, 23] in order to train SVMs to classify departures from typical profiles as anomalies Decision trees have been used to identify known attacks and partition normal traffic into smaller subsets. There are several additional published methods that Supervised and unsupervised machine learning techniques such as ANNs, SVM, and k-NN are used, but no clear best practices have been defined for network anomaly detection [24, 25].

However, despite the prevalence of machine learning methods in research (as reported by Hindi and others in a 2018 survey), these methods are rarely implemented in industry settings. Several authors have tried to explain this disparity. Works of [26, 27] conducted a survey of 276 studies and found the majority did not adhere to scientific standards due in part to a lack of scientific rigor in experimental computer science research. A major contributing factor to this was the lack of suitable reference datasets. Public datasets often fail to accurately represent normal traffic and are so large that it cannot be fully evaluated, and as a result researchers resort to data sampling, which distorts the data and hinders its comparability, by challenging the basic assumptions of anomaly detection technology [28, 29]. Gates and Taylor question its feasibility and stress that the possibility that anomalous traffic is rare and that Their characteristics are not different from those of regular traffic [30]. The authors also draw attention to the guidelines that are rarely provided for updating detection models leading to an implicit assumption that network traffic is stationary and ignoring shifts in network traffic composition as well as conceptual drifts [31, 32].

III. METHODOLOGY

Researchers have addressed the dearth of publicly available real-world anomaly datasets in the network by releasing more data in recent years. However, the adoption of these datasets in research has been slow, with many researchers still relying on old DARPA data and its variants [33, 34].

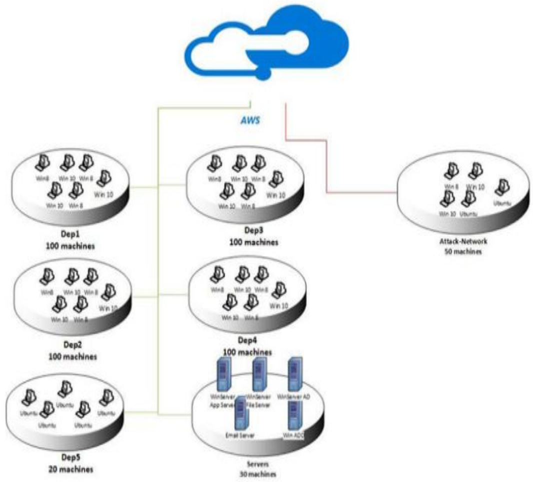

Figure 2.CSE-CIC-IDS2018 Testbed

The limited use of these datasets can partly be attributed to the lack of a clear benchmark winner as different datasets cover different scenarios. For example, the CIC-IDS2017 dataset [35, 36] includes diverse attacks but presents challenges in creating a comprehensive test set due to the attacks being distributed across different days. The large volume of data and different research goals lead to modifications to data sets. Researchers often use only subsets of the data to validate them or apply different data pre-processing techniques. These modifications make it difficult to replicate an accurate data set, which in turn hinders the validation of scientific claims [22].

Typically, most anomaly detection algorithms develop a profile of normal behavior and recognize anomalies as data points that differ from this specific profile. It is important to keep in mind that these algorithms are often not built with anomaly detection in mind and instead they can be mainly used for tasks that involve clustering or classification [10]. Additionally, many methods have also been introduced, which rely on density or distance metrics to detect anomalies, which can put significant computational pressure on the system especially in high-dimensional environments [37].

The basic principle of this approach is that anomalies are easier to isolate from the rest of the data than regular occurrences if specific parameters are met (for example, anomalies must be rare within the data and have distinct properties) and isolation in this sense refers to keeping the abnormal instance away from other instances.

Binary trees are a particularly useful data structure for this task, but the exact method used for isolation may vary. Binary trees are a useful tool to illustrate the idea of effectively identifying anomalies through isolation. This data structure is used through a number of techniques such as robust random forest and isolation trees. Spanning and isolation trees [38] are well known for their rule in finding anomalies. Although the basic structure of the binary tree is the same for both systems, the construction and scoring techniques are different. We consider an isolation tree. Any node N in the tree can be an internal node with a test and two child nodes. Exactly (Nl , Nr) or an external (leaf) node without children The feature q and the split value p form a test and whether the sample passes Nl or Nr is determined by the condition q < p.

Definition 1 provides a basic understanding of an isolation tree given a set of data points and qmax Using the test q < p an internal node is created and the data points are partitioned appropriately, the two subsets Xl = and Memory of a tree Given the number of instances (n) that are part of the tree, a tree has n leaves and n!-1 internal nodes when each of its instances is isolated for a total of 2n!-1 nodes in the tree.

Anomalies should be found closer to the root of the tree than they would be in typical cases Once the tree structure has been formed, data points can be classified according to the degree of anomaly they exhibit using path length as the primary metric [39].

The Gaussian distribution, also known as the normal distribution, is a probability distribution widely used in statistics and data analysis that is characterized by a bell-shaped curve and is defined by two parameters: the mean (μ) and the standard deviation (σ). The Gaussian circulation is mainly helpful in noticing anomalies for it provides. A way to model the expected behavior or an information normality set. Through fitting a Gaussian model to the data, we can evaluate the mean and standard variation that describe a circulation [38].

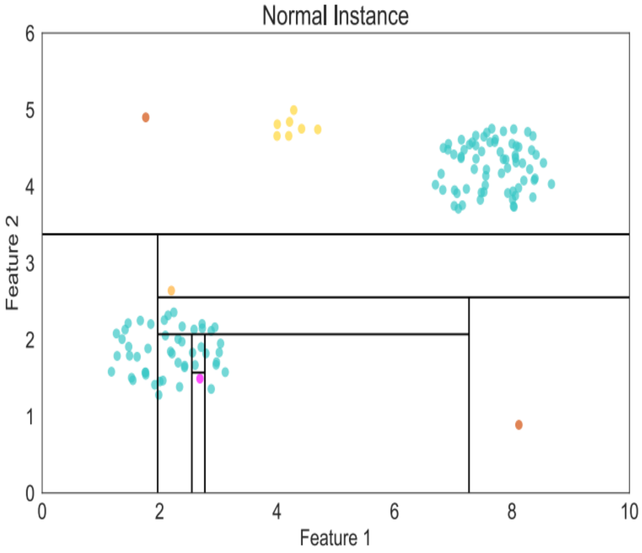

Figure 3. Binary Tree Examples Isolation of a Normal Instance

To make abnormality detection by means of a Gaussian circulation, we trace these steps**:**

1 .Fit Gaussian model**:**

Compute the mean (mean) and normal perversion of an agreed data set. The mean signifies the central slope or medium the data value. The normal deviation measures the spread or the data scuttle about the mean. These parameters describe the Gaussian distribution of the data set**.**

2 .Verify data points:

Estimate the distance from the mean and use the studied mean and standard perversion to mark the abnormality disclosure sill, for all new information point. Common methods include sighted data points that are standard deviations certain number far from the mean as abnormality. You can regulate this verge founded on the specific supplies and sensibility wanted for the system disclosure of your anomalies**.**

3 .Identify anomalies**:**

Put mark it as an abnormality, If the distance between the data point and the mean overrides a limit of fixed. Abnormality are information points that distort safely the predictable behavior represented by a Gaussian distribution. Gaussian allocation is efficient in revealing anomalies in different fields like analysis, data of economic, monitor product quality, detection of service outage, identification of financial fraud, and other uses that depend on anomaly revealing and analysis of statistical. through fitting data to a Gaussian model and contrasting points of new data to the expected allocation, you can reveal anomalous styles or outliers that may indicate anomalies or uncommon behavior.

IV. RESULTS AND DISCUSSION

With the creation of a data set for the monitoring system with the following parameters:

- The total observation time is 3600 seconds.
- The interval between readings is 1 second.
- The average data rate is 100 kilograms per second.
- The data rate difference is 20 kbps.

It counts the number of readings based on the total time and interval. It then creates arrays of time and data rates using a normal distribution around the average data rate with the specified variance. An outlier is added at index 50 with a data rate of 500 KB/s.

To implement an anomaly detection algorithm to detect abnormal behavior in server computers according to deep learning and Gaussian distribution, with the features of measuring throughput (MB/s) and response time (milliseconds) of the response. For each server, our goal is to identify anomalies within that dataset. We assume that the majority of examples are “normal” (non-anomalous) cases of normally operating servers, but there may be some anomalies as well.

We will start by visualizing the data set in 2D format to get a better understanding of what the algorithm is doing. This visualization will involve fitting a Gaussian distribution to the dataset and identifying values with very low probabilities, which can be considered anomalies. After successfully visualizing the 2D dataset, we will proceed to apply the anomaly detection algorithm to a larger, multi-dimensional dataset.

The first method is to analyze the data and detect anomalous readings using the Mahalanobis distance [40] and after creating a data set to interpret the data over a specific period of time. Therefore the `normrnd` function is used to control a value that follows a Gaussian distribution (a specified mean and standard deviation) to generate data where anomalous data is added at a specific position in the set, where the value of the data rate at that position is set to a different and unusual value representing the anomaly The data set is plotted Temporal data adjustment to show the beginning in rates over time (Figure.4) The data is analyzed using the Gaussian taxi, where the formal calculation and the elastic deviation of the data rates are calculated according to the choice of Gauss to distribute the data, and using the Mahalanobis distance, the distance between each reading and the rest of the readings in the group is determined. The Mahalanobis distance and the classification of the differences between a particular reading and the rest of the readings are used to evaluate the statistical opinion poll and the anomaly threshold is used which is calculated using the chi-square distribution of the anomaly readings. The leadership of anomalous readings is determined whether the Mahala Nobis distance of a particular reading exceeds the researcher's testimony.

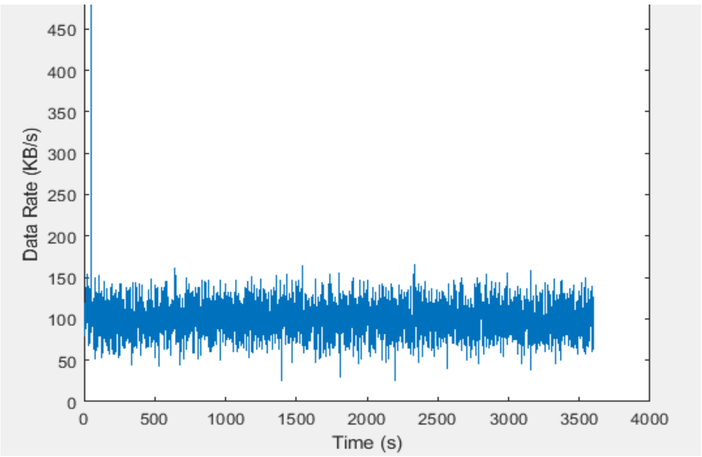

Figure 4. Data traffic in the network

The main benefit of using Mahala Nobis distance is that it takes into account the structure of the variation of the data. In multi -dimensional data, we do not only have average data but also distributing and varied. Using Mahala Nobis distance, we can improve the accuracy of the distance between a specific point and the statistical distribution of data and by using the CT scan, we can determine the relationship between different variables in data and determine their distribution. With the help of Mahala Nobis, we can calculate the distance between individual points in data and distribute data more accurately in terms of contrast. The Mahala Nobis distance is used to identify unusual or abnormal readings in the dataset and the

Mahala Nobis distance is used to classify models and determine how similar they are. The Mahala Nobis distance can be used in the analysis of key components (PCA) to determine the relationship between different variables in the data and extract the main variables from them, as shown in Figure 5.

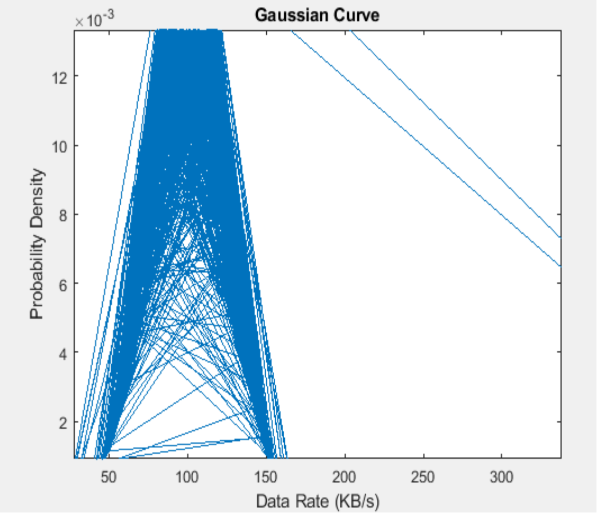

Figure 5. Regulate network traffic via Gaussian anomaly detection

The second method used is to use a deep learning system to improve network security and traffic control. An artificial neural network is used to regulate the data rate passing through the network. The network is trained using the collected data which contains the required target rates. Structured data rates are set as inputs to the neural network which must match the actual data rates (Inputs) passing into the network and after the network is trained, it is run to pass the actual data. The results extracted from the network (Outputs) represent the data rates produced that must be proportional to the required regulated rates. These extracted results and regulated rates are represented on the time axis to obtain a graph showing improved data rate regulation. By using a neural network, significant improvement in network security and traffic control can be achieved. A neural network can learn patterns and changes in traffic and determine appropriate rates for data passing through the network. This helps reduce security risks such as hacking attacks and network congestion because, as we can see in the figure (6), it takes a shorter time in the detection process with the smoothing feature and the ability to predict the anomalies that occur. Therefore, it is a more practical method in the field of regulating data traffic and protection. Using deep learning systems gives better results. Traditional models such as Gaussian distribution help improve network security and control traffic. Deep learning allows neural networks to effectively detect and analyze complex patterns, which helps in achieving better performance and improving adaptability to changes in the network.

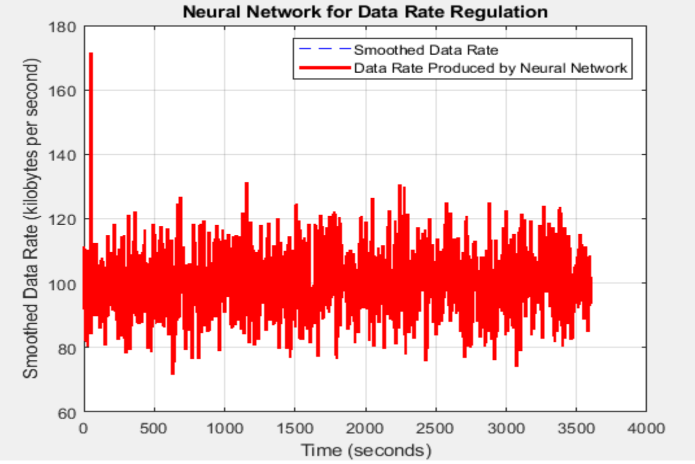

Figure 6. Regulating data traffic in the network via ANN anomaly detection

The network training and testing process, as can be seen in Figure 7, used to improve network security and control traffic through deep learning systems, a set of data is used that contains the required structured data rates (Target Rates) and the actual data (Inputs) passing through the network. This data is set as input and output to the neural network. During the training process, the network parameters are modified so that the results extracted from the network (Outputs) match the required regularization rates. After completing the training process, a set of new data (not used in the training process) is used to test the performance of the network. This data is passed as input to the network and the extracted results (Outputs) are extracted. The accuracy of the network is measured by comparing the results extracted with the actual organized rates and the accuracy of the network is calculated by comparing the results extracted from the network with the actual organized rates. Various accuracy measures such as average absolute deviation or medium Spring deviation are used to assess the network performance. The goal is to achieve a high accuracy in network predictions so that the actual data and extracted results are as identical as possible.

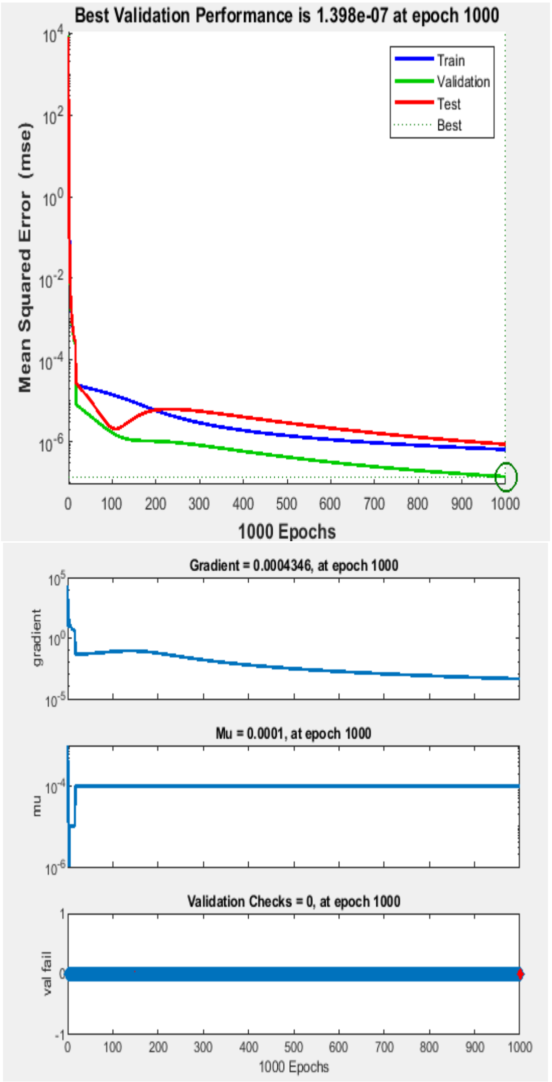

Figure 7. The network training and testing process

V. CONCLUSION

It highlights the importance of the network safety and the importance of discovering abnormal cases of traffic on the network. The challenges facing the detection of abnormal cases of traffic on the network are highlighted, such as the lack of representative data groups, the high cost of errors, and the dynamic nature of the network traffic. The study provides a dynamic model known as 'deep learning' that uses the current network traffic statistics in which the illegal error discovery system can adapt to data transformations in data and does not require any prior training. The performance of the method of discovering impurities was verified in a realistic representative environment by controlling parameters and showing promising results in discovering impurities, with the total ability to distinguish between impurities and normal cases and outperform a similar fixed model. The suggested impurities discovery form has been tested in realistic simulation. The adjustments were applied to improve the performance of the model, with promising results in detecting impurities. The program user model enables the data modification and can perform better than the fixed form for itself.

A study decided that the use of a dynamic drawing algorithm to detect impurities in network movement can be effective and contribute to improving network safety. It is better than fixed systems such as the gossip. The proposed form is characterized by the ability to adapt to the transformations of data and does not require pre - training, making it an option. Promising to apply them in real environments.

Continuing developments in the field of discovering abnormal cases in future traffic will continue through deep learning techniques. By improving network quality, diversity of data groups, and developing analysis and learning technologies, networks can be improved to discover abnormal cases in network movement.

REFERENCES

[1.] Sahut JM, Lissillour R. The adoption of remote work platforms after the Covid-19 lockdown: New approach, new evidence. Journal of Business Research. 2023;154:113345.https://doi.org/10.1016/j.jbusres.2022.113345

[2.] Oluka A, Kader A. Adoption of remote work: implications for tax practitioners. Technology audit and production reserves. 2023;3(4):17-24.https://doi.org/10.15587/2706-5448.2023.284026

[3.] Kumar A, Behera RP, Kumar A, Narasimhan S, editors. Strengthening Network Security in Safety-Critical I&C Systems of Nuclear Reactors: Design and Implementation of a Robust Data Diode. 2023 IEEE 20th India Council International Conference (INDICON); 2023 14-17 Dec.

2023.https://doi.org/10.1109/INDICON59947.2023.10440821

[4.] Shetty S, Thrisha MS, Vandana HM, Rizawan NS, editors. Intelligent Network Traffic Control with AI and Machine Learning. 2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks (CICN); 2024 22-23 Dec. 2024; Indore, India https://doi.org/10.1109/CICN63059.2024.10847397

[5.] Chimento M, Farine DR. The contribution of movement to social network structure and spreading dynamics under simple and complex transmission. Philosophical Transactions of the Royal Society B: Biological Sciences. 2024;379(1912):20220524.https://doi.org/10.1098/rstb.2022.0524

[6.] Pan Y, Liu X, Yao F, Zhang L, Li W, Wang P. Identification of dynamic networks community by fusing deep learning and evolutionary clustering. Scientific Reports. 2024;14(1):23741.https://doi.org/10.1038/s41598-024-74361-0

[7.] Jiang W, Han H, He M, Gu W. Machine Learning-based Multi-Class Traffic Management for Smart Grid Communication Network. Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing; Cancun, Quintana Roo, Mexico: Association for Computing Machinery; 2023. p. 694–9.https://doi.org/10.1145/3594739.3612909

[8.] Hwang FS, Confrey T, Reidy C, Picovici D, Callaghan D, Culliton D, et al. Review of battery thermal management systems in electric vehicles. Renewable and Sustainable Energy Reviews. 2024;192:114171.https://doi.org/10.1016/j.rser.2023.114171

[9.] Paheding S, Saleem A, Siddiqui MFH, Rawashdeh N, Essa A, Reyes AA. Advancing horizons in remote sensing: a comprehensive survey of deep learning models and applications in image classification and beyond. Neural Computing and Applications. 2024;36(27):16727-67.https://doi.org/10.1007/s00521-024-10165-7

[10.] Olteanu M, Rossi F, Yger F. Meta-survey on outlier and anomaly detection. Neurocomputing. 2023;555:126634.https://doi.org/10.1016/j.neucom.2023.126634

[11.] Afandizadeh S, Abdolahi S, Mirzahossein H. Deep Learning Algorithms for Traffic Forecasting: A Comprehensive Review and Comparison with Classical Ones. Journal of Advanced Transportation. 2024;2024(1):9981657.https://doi.org/10.1155/2024/9981657

[12.] Zou X, Chung E, Ye H, Zhang H. Deep Learning for Traffic Prediction and Trend Deviation Identification: A Case Study in Hong Kong. Data Science for Transportation. 2024;6(3):27.https://doi.org/10.1007/s42421-024-00112-2

[13.] Naghib A, Gharehchopogh FS, Zamanifar A. A comprehensive and systematic literature review on intrusion detection systems in the internet of medical things: current status, challenges, and opportunities. Artificial Intelligence Review. 2025;58(4):114.https://doi.org/10.1007/s10462-024-11101-w

[14.] Jangra R, Kajal A, editors. A Review of Deep Learning based Intrusion Detection Systems. 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS); 2023 3-4 Nov. 2023; Greater Noida, India https://doi.org/10.1109/ICCCIS60361.2023.10425595

[15.] Panjei E, Gruenwald L. Discovering outlying attributes of outliers in data streams. Data & Knowledge Engineering. 2024;154:102349.https://doi.org/10.1016/j.datak.2024.102349

[16.] Dash CSK, Behera AK, Dehuri S, Ghosh A. An outliers detection and elimination framework in classification task of data mining. Decision Analytics Journal. 2023;6:100164.https://doi.org/10.1016/j.dajour.2023.100164

[17.] Mienye ID, Swart TG. A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications. Information. 2024;15(12):755.https://doi.org/10.3390/info15120755

[18.] Faiz FA, Wibirama S, Nurlatifa H, Setiawan NA, editors. A Systematic Review of Deep Learning Approaches to Visual Saliency Prediction on Webpage Images. 2024 7th International Conference on Informatics and Computational Sciences (ICICoS); 2024 17-18 July 2024.https://doi.org/10.1109/ICICoS62600.2024.10636890

[19.] Saranya R, Paneerselvam K, Prateeksha Y, Raghunath AP, Ridish R, editors. Sign Language Recognition Using Convolutional Neural Network. 2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG); 2023 8-9 Dec. 2023; Indore, India https://doi.org/10.1109/ICTBIG59752.2023.10456160

[20.] Gayathri D, Ramar A, Karpagam S, Sudharsanan J, Rishwan S, Sudalaimani PK. Sign Language Recognition Using Convolutional Neural Network. International Journal of Intelligent Systems and Applications in Engineering. 2024;12(17):329 - 37

[21.] Shankara NV, Sneha VM, Padmavathi S, editors. Continuous Sign Language Recognition using Convolutional Neural Network. 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE); 2024; Vellore, India.https://doi.org/10.1109/ic-ETITE58242.2024.10493715

[22.] Selmy HA, Mohamed HK, Medhat W. Big data analytics deep learning techniques and applications: A survey. Information Systems. 2024;120:102318.https://doi.org/10.1016/j.is.2023.102318

[23.] Jahani H, Jain R, Ivanov D. Data science and big data analytics: a systematic review of methodologies used in the supply chain and logistics research. Annals of Operations Research. 2023.https://doi.org/10.1007/s10479-023-05390-7

[24.] Fesli U, Ozdemir MB. Electric Vehicles: A Comprehensive Review of Technologies, Integration, Adoption, and Optimization. IEEE Access. 2024;12:140908-31.https://doi.org/10.1109/ACCESS.2024.3469054

[25.] Madaram VG, Biswas PK, Sain C, Thanikanti SB, Balachandran PK. Advancement of electric vehicle technologies, classification of charging methodologies, and optimization strategies for sustainable development - A comprehensive review. Heliyon. 2024;10(20):e39299.https://doi.org/10.1016/j.heliyon.2024.e39299

[26.] Ahmed U, Nazir M, Sarwar A, Ali T, Aggoune E-HM, Shahzad T, et al. Signature-based intrusion detection using machine learning and deep learning approaches empowered with fuzzy clustering. Scientific Reports. 2025;15(1):1726.https://doi.org/10.1038/s41598-025-85866-7

[27.] Hore S, Ghadermazi J, Shah A, Bastian ND. A sequential deep learning framework for a robust and resilient network intrusion detection system. Computers & Security. 2024;144:103928.https://doi.org/10.1016/j.cose.2024.103928

[28.] Landauer M, Onder S, Skopik F, Wurzenberger M. Deep learning for anomaly detection in log data: A survey. Machine Learning with Applications. 2023;12:100470.https://doi.org/10.1016/j.mlwa.2023.100470

[29.] Himeur Y, Ghanem K, Alsalemi A, Bensaali F, Amira A. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Applied Energy. 2021;287:116601.https://doi.org/10.1016/j.apenergy.2021.116601

[30.] Wu Y, Sicard B, Gadsden SA. Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. Expert Systems with

Applications. 2024;255:124678.https://doi.org/https://doi.org/10.1016/j.eswa.2024.124678

[31.] Alam MA, Sajib M, Rahman F, Ether S, Hanson M, Sayeed A, et al. Implications of Big Data Analytics, AI, Machine Learning, and Deep Learning in the Health Care System of Bangladesh: Scoping Review. Journal of medical Internet research. 2024;26:e54710.https://doi.org/10.2196/54710

[32.] Dhaygude AD, Varma RA, Yerpude P, Swarnkar SK, Jindal RK, Rabbi F, editors. Deep Learning Approaches for Feature Extraction in Big Data Analytics. 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON); 2023; Gautam Buddha Nagar, India IEEE.https://doi.org/10.1109/UPCON59197.2023.10434607

[33.] Lippmann RP, Fried DJ, Graf I, Haines JW, Kendall KR, McClung D, et al., editors. Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation. Proceedings DARPA Information Survivability Conference and Exposition DISCEX'00; 2000 25-27 Jan. 2000.https://doi.org/10.1109/DISCEX.2000.821506

[34.] Thomas C, Sharma VP, Balakrishnan N. Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security - Usefulness of DARPA dataset for intrusion detection system evaluation: SPIE; 2008.

[35.] Chen X. CICIDS2017 and UNBSW-NB15. IEEE Dataport; 2023.

[36.] Boukhamla A, Gaviro JC. CICIDS2017 dataset: performance improvements and validation as a robust intrusion detection system testbed. Int J Inf Comput Secur. 2021;16(1–2):20–32.https://doi.org/10.1504/ijics.2021.117392

[37.] Souto Arias LA, Oosterlee CW, Cirillo P. AIDA: Analytic isolation and distance-based anomaly detection algorithm. Pattern Recognition. 2023;141:109607.https://doi.org/10.1016/j.patcog.2023.109607

[38.] 38. Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y, et al. MONAI: An open-source framework for deep learning in healthcare. ArXiv. 2022;abs/2211.02701.https://doi.org/10.48550/arXiv.2211.02701

[39.] Uyyala P. Sign language recognition using convolutional neural networks. Journal of Interdisiplinary Cycle Research. 2022;14(1):1198-207

[40.] Brereton RG. The Mahalanobis distance and its relationship to principal component scores. Journal of Chemometrics. 2015;29(3):143-5.https://doi.org/10.1002/cem.2692