Class-Focused Evaluation of Deep Learning Techniques for Network Intrusion Detection[*]

Mantas Bacevicius[1,*,†]

[1]*Kaunas Technology University, Faculty of Informatics, Studentu 50, Kaunas, 51368, Lithuania*

Abstract
In an increasingly interconnected world, safeguarding digital systems and networks against cyber threats is of utmost importance. Traditional intrusion detection approaches, relying on rule-based systems or simplistic machine learning models, often struggle to adapt to the evolving threat landscape. Deep Neural Networks (DNNs) offer promising avenues for enhancing Intrusion Detection Systems (IDS) effectiveness, leveraging their hierarchical structure to process complex network traffic data and extract discriminatory features indicative of malicious activity. However, the temporal dynamics inherent in network traffic data pose a unique challenge, prompting exploration into Long Short-Term Memory (LSTM) networks, for their sequential data processing capabilities. This paper investigates the application of deep learning models, including dense neural networks and LSTMs, for classifying network traffic into 28 distinct attack types. By analyzing the architectural design and presenting experimental results on standard benchmark datasets, we demonstrate the practical applicability of our hybrid approach in real-world cybersecurity scenarios, contributing to the advancement of intrusion detection systems through deep learning techniques. Additionally, we explore the challenges posed by class imbalances and dataset characteristics, providing insights into model performance and limitations for various attack types.

Keywords
Intrusion Detection Systems (IDS), LSTM, Dense layers, deep learning, classification.

1. Introduction

With our world ever more connected, protecting digital systems and networks is critical. Cyber threats are constantly getting smarter, so we need better Intrusion Detection Systems (IDS) [] that can quickly spot and stop possible security breaches. Traditional approaches to intrusion detection often rely on rule-based systems [] [] or simple machine learning models [], which may struggle to adapt to the evolving landscape of cyber attacks [] []. In recent years, the emergence of Deep Neural Networks (DNNs) [] [] has offered promising avenues for enhancing the effectiveness of IDS. In the field of cybersecurity, the application of DNN has a great potential to enhance the capabilities of intrusion detection systems. Using the inherent hierarchical structure of DNN, these systems can efficiently process complex network traffic data and extract discriminatory features indicative of malicious activity.

Moreover, the temporal fluctuations ingrained within network traffic data pose a distinctive obstacle in intrusion detection. Therefore, another challenge in intrusion detection is that network traffic data constantly changes over time. DNN architectures frequently encounter difficulties in adeptly capturing temporal dependencies within sequential data, as their proficiency primarily resides in tasks such as image feature extraction, pattern identification, classification, and segmentation. This is where a distinct neural network model, known as a Recurrent Neural Network (RNN), becomes

[*] *IVUS2024: Information Society and University Studies 2024, May 17, Kaunas, Lithuania*
[1,*] Corresponding author
[†] These author contributed equally.
mantas.bacevicius@ktu.edu (M. Bacevicius)
0009-0007-6772-8898 (M. Bacevicius)

relevant []. Unlike DNNs, RNNs excel at processing sequential information and thus could be beneficial for analyzing network traffic. Specifically, Long Short-Term Memory (LSTM) networks [] [ ], a subtype of RNNs, demonstrate remarkable proficiency in learning from sequences and retaining critical information over extended periods. However, directly applying LSTMs to network traffic data presents challenges due to the varying lengths of network data packets and their irregular arrival times.

While classical machine learning methods provide reliable results for classifying network intrusions [], this paper explores the possibility of applying deep learning models given the continuous (numerical) nature of the dataset. Unlike previous works that focused on a limited set of popular attack classes (typically less than 18), our research considers a broader range of 28 different attack types. This study examines the effectiveness of dense neural networks and LSTMs in classifying network traffic into different attack types. We have analyzed the architectural design of this hybrid approach and present experimental results showing the effectiveness of the proposed model on standard benchmark datasets, thus highlighting its practical applicability in real cybersecurity scenarios. Through these studies, we aim to contribute to the continuous advancement of intrusion detection systems using deep learning techniques.

2. Dataset

Developed at the University of New Brunswick by the Canadian Institute for Cybersecurity (CIC), both CIC-IDS2017 [] and the CSE-CIC-IDS2018 [] [] datasets are significant aggregations of network traffic data, vital for evaluating the performance and reliability of Intrusion Detection Systems (IDS) and related security technologies. These datasets are widely used for artificial intelligence model development for cybersecurity because of their 1) size – they are extensive, containing a substantial amount of network traffic data, encompassing both benign and malicious samples (number of samples); 2) range of attacks – datasets encompass a broad spectrum of cyber threats and irregularities, including Denial of Service (DoS), Distributed Denial of Service (DDoS), infiltration attacks and diverse intrusion attempts; 3) authentic scenarios – constructed to mirror real-world network traffic situations, these datasets are invaluable for assessing the effectiveness of intrusion detection systems in practical environments; 4) data labelling – each instance of network traffic in the datasets is categorized, indicating whether it represents normal (benign) activity or malicious behavior. This labelling aids in employing supervised learning techniques to construct and assess intrusion detection models; 5) catholic features – a variety of characteristics are extracted from network traffic data in this dataset, such as packet attributes, protocol details and traffic patterns. These attributes function as input parameters for machine learning models aimed at classifying network traffic as either benign or malicious.

This study combined two network intrusion detection datasets, CIC-IDS2017 and CSE-CIC-IDS2018. To improve data consistency, we merged similar malicious categories and removed features deemed uninformative. Additionally, we replaced missing values (NaN) with zeros. This process resulted in a new dataset containing 28 distinct network traffic classes and a total of 19,063,687 entries. It must be noted, that both CIC-IDS2017 and CSE-CIC-IDS2018 datasets are highly unbalanced as it is shown in Figure 1.

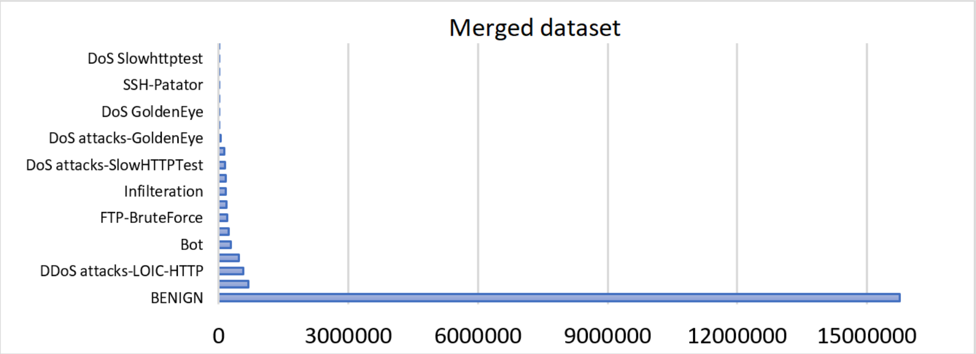


Figure 1: Merged Dataset providing Top10 categories of networks attacks

It has been observed that deep learning approach is applied to both the binary network intrusion problem and the multi-class [] [] []. While the goal might be multi-class classification, many intrusion detection tasks are simplified by grouping similar attack types into a smaller number of categories. This typically results in 5 to 14 classes, depending on the specific needs and chosen dataset [] []. This is because, the network intrusion datasets typically have a large portion of normal traffic (Benign) and a significantly smaller portion of malicious traffic categorized into various attack types ("SQL Injection", "Heartbleed", "FTP-Patator", etc.). This creates a class imbalance, where the majority class (the Benign class) is prioritised when training the model.

Long Short-Term Memory (LSTM) networks are popular choices for intrusion detection systems due to their ability to learn from sequential data. LSTM is typically used for anomaly detection or for binary tasks by classifying attacks into benign and malicious [].

However, their accuracy depends on the specific dataset and the model architecture itself. For example, the Kddcup99 dataset allows for a highest accuracy of 0.98 for the ACC metric [].

3. Related works

The CIC-IDS2017 dataset is a well-known benchmark dataset for Intrusion Detection Systems (IDS) as it contains network traffic data collected from real-world environments with different types of attacks and normal traffic. In order to improve the classification performance of this dataset, there are a number of studies with detailed analysis including feature selection, class grouping, data cleaning and processing. The classifiers to use for this task also vary widely, from classical RF or MLP (multi-layer perceptron) to deep learning architectures. The results of certain multi-class classification studies are presented in Table 1 showing the F1-score and the number of classification classes.

Table 1. Multi-classification accuracy results for CIC-IDS2017 dataset.

Method	Classifier	F-1 Score, %	Classes
Bulavas et al. []	ADA	99.9	17
	QDA	94.4	
	MLP	98.0	
Sharafaldin et al. []	QDA	92.0	17
Zhong et al. []	DMTR	99.3	15
Liu et al. []	CNN-MLP	88.16	5
Belarbi et al. []	DBN	86.62	6
Jiang et al. []	LSTM-RNN	87.38	4

The CSE-CIC-IDS-2018 dataset comprises 28 categories of network traffic, with each category representing distinct network activities. However, during classification exercises, these categories are typically consolidated into 7 main groups (such as "Benign," "DDoS," "DoS," "Brute Force," "Bot," "Infiltration," and "Web"), occasionally 15, and the highest number of observed categories for classification stands at 17 (see Table 2).

Table 2. Multi-classification accuracy results for CSE-CIC-IDS2018 dataset.

Method	Classifier	F-1 Score, %	Classes
Bulavas et al. []	ADA	99.9	17
	QDA	59.7	
	MLP	95.8	
Karatas et al. []	LDA	99.0	5
Liu et al. []	DNN	97.0	7
Gamage et al. []	DNN	97.8	5
Kunang et al. []	DNN-AE	95.1	15

Mezina et al. []	TCN+LSTM	97.7	14
Al-Fawa’reh et al. []	DNN	98.7	15

4. Methodology

To address overfitting in the highly unbalanced merged network traffic dataset, we employed several deep learning models. First, a Deep Autoencoder-Deep Neural Network (DAE-DNN) was used to potentially reduce dimensionality and extract relevant features. Second, individual deep learning models were built for each malicious traffic category to improve class-specific accuracy. Finally, a Long Short-Term Memory (LSTM) model was constructed to evaluate the dataset's suitability for time series analysis.

For time distributed LSTM, model was constructed containing 4 LSTM layers with 195✸125✸125✸195 layout with RepeatVector layer in the middle, as it is shown in Figure 2. Model was trained with 30 epochs. For training dataset was split into 1280 batches.

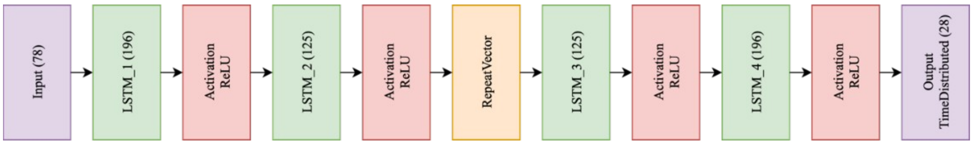


Figure 2: Time distributed LSTM model architecture

To examine the DAE-DNN architecture fitness to the dataset, couple of these models with varying numbers of DAE layers and DNN layers, were constructed. First DAE-DNN model consisted of 3 dense layers 80✸60✸80, performing the DAE function, with ReLU activation in each layer. Then, a single dense layer was added, to perform DNN function consisting of 80 neurons, again with ReLU activation function. Lastly, another dense layer was appended for output, this time with sigmoid activation function. The architectural components of this model are provided in Figure 3.

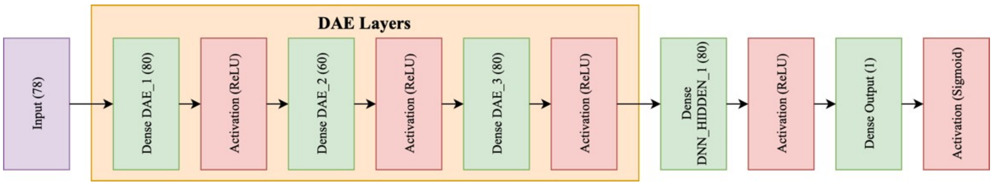


Figure 3: DAE-DNN model architecture

Then, the architecture of DAE-DNN was modified, so that number of DAE function performing layers was changed from 3 to 9 with 80✸70✸40✸30✸25✸30✸40✸70✸80 layout. Number of DNN layers was also changed from 1 to 3, with 80 neurons each. Activation functions stayed the same – ReLU for DAE and DNN layers, sigmoid for the output. The architectural components of the extended DAE-DNN model are provided in Figure 4.

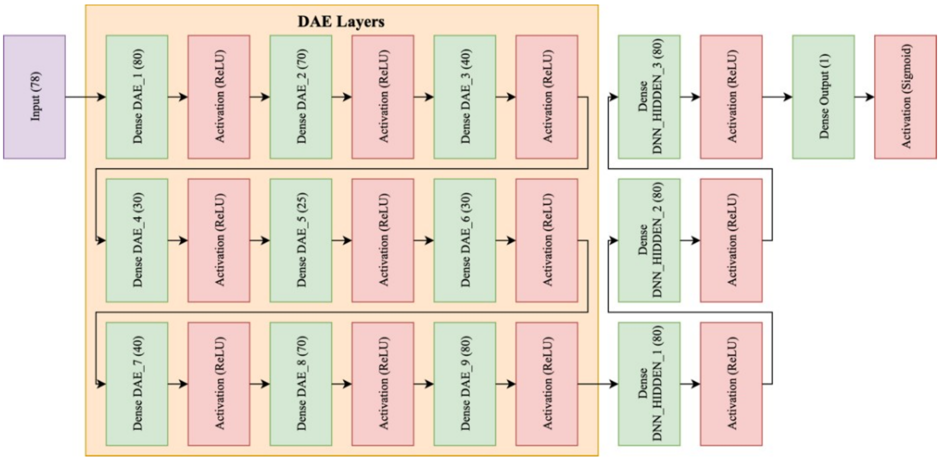
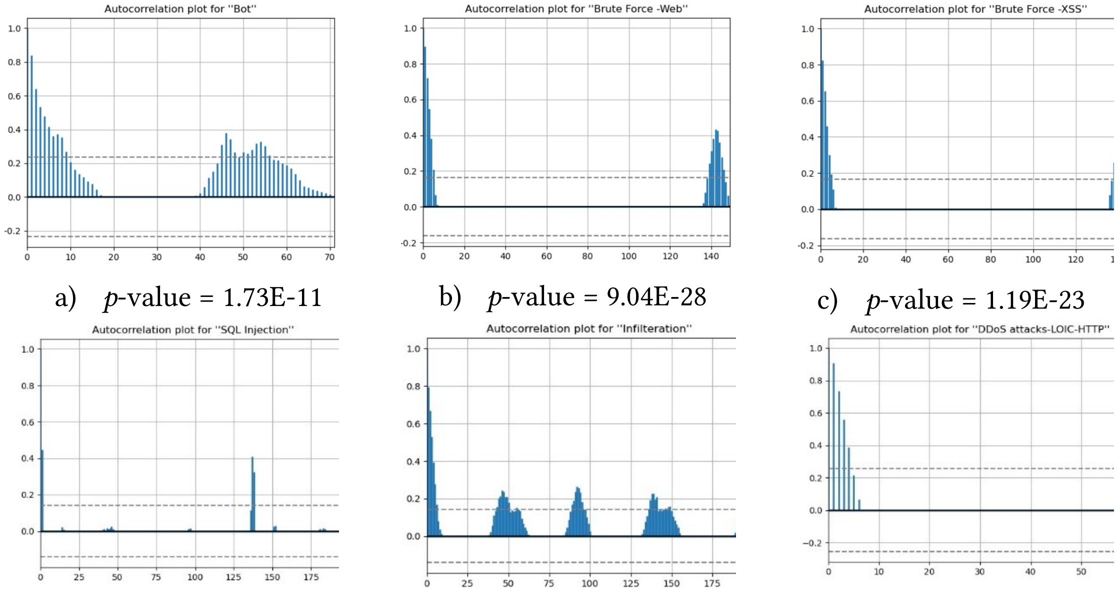


Figure 4: DAE-DNN extended model architecture

Last approach was to make a separate DNN model for each of the malicious class in the dataset. Dataset for each model contained both benign and class-specific malicious data. The model architecture for each class was the same non-expanded DAE-DNN model, reviewed in previous section, provided in the Figure 3.

5. Results

For time series testing, the dataset was split into 10-minute intervals, based on the timestamp field. Gaps in the timeline were filled with previous dataset entry, while data entries in intervals that contained more than one entry where removed, leaving only the first recorded malicious entry, if it was registered during the 10 minute period. If not, only the first recorded benign entry was left. Resulting dataset was tested with Ljung box test for autocorrelation. As it is visible from Figure 5, significant autocorrelation (p-value < 0.05, and 99% confidence band identify statistically significant autocorrelation values) can be observed only in minority of classes – “Brute Force -Web”, “Bot”, “Brute Force -XSS”, “SQL Injection”, “Infilteration” and “DDoS attacks-LOIC-HTTP”.



a) *p*-value = 1.73E-11

b) *p*-value = 9.04E-28

c) *p*-value = 1.19E-23

d) *p*-value = 5.38E-10

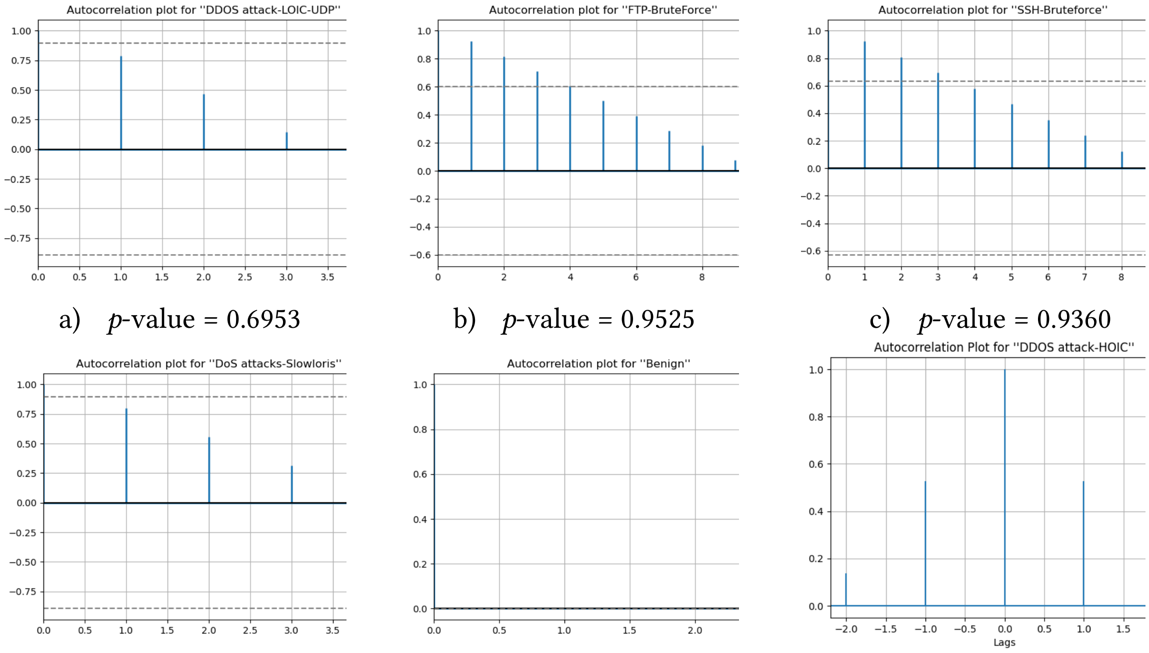
e) *p*-value = 4.92E-28

f) *p*-value = 3.79E-13

Figure 5: Ljung box test results for dataset class with significant p-value

The absence of autocorrelation in other classes (including "Benign") suggests that the data points within those classes are independent of past observations (see Figure 6). p-values of more than 0.9

have been observed for such classes as “FTP-BruteForce”, “SSH-Bruteforce”, “DDOS-Attack-HOIC” and “DoS attacks Slowloris”.



a) *p*-value = 0.6953

b) *p*-value = 0.9525

c) *p*-value = 0.9360

d) *p*-value = 0.9839

e) *p*-value = 0

f) *p*-value = 0.9962

Figure 6: Ljung box test results for the attacks with non-significant p-value

The results of the DAE-DNN and extended DAE-DNN model classification (weighted average F1-score) are presented in Table 3. It is notable that the extended model generally performs better, achieving a score of 0.9361 for F1-score and 0.9508 for recall value.

Table 3. Weighted average F1-score for Classification results including 28 classes.

	DAE-DNN	Extended DAE-DNN
Precision	0.9424	0.9410
Recall	0.9266	0.9508
F1-score	0.9178	0.9361

The split DAE-DNN classification models achieved the best results comparing to other two models for 12 classes (see Figure 7). Results for the remaining classes are not included due to either zero cases during testing or an F1-score of 0 in the classification results for those classes. DNN split models failed to classify “Brute Force - Web”, “Heartbleed”, “Infiltration”, “Web Attack Brute Force”, “Web Attack Sql Injection” and “Web Attack XSS” due to low count of data entries in those classes. We can assume that the improved architecture may be an appropriate solution for classifying attacks in this way, by creating a class-by-class model.

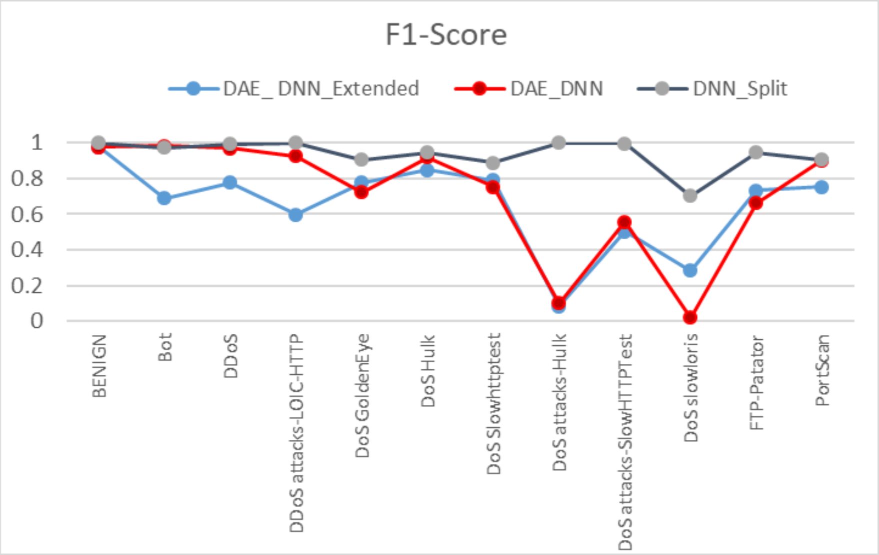


Figure 7: DAE-DNN models comparison F-1 score results for 12 selected classes.

While the extended DAE-DNN model demonstrated superior classification results overall, a closer examination of the F1-Score values for individual attack classifiers revealed that more classes performed better with the simple DAE-DNN model. However, the extended model's significant superiority was evident in specific classes, like “SSH-Patator”, “Portscan”, “Bot”, etc.

6. Conclusions

This study underscores the urgent need for advanced Intrusion Detection Systems (IDS) in the face of escalating cyber threats. While traditional methods struggle to keep pace, Deep Neural Networks (DNNs) with addition of Deep Autoencoders (DAE) show promise in efficiently processing complex network traffic data to pinpoint malicious activity. Our experimentation, involving DAE-DNN, separate DNN models for each class and LSTM models, demonstrates tangible progress in classifying network traffic into 28 distinct attack types. However, challenges such as class imbalances persist, impacting detection accuracy. For instance, our findings reveal a 66% to 99% F-1 score range across various attack types, with certain classes posing persistent challenges due to limited data availability without using oversampling methods. These results underscore the necessity for ongoing refinement and optimization in IDS methodologies to ensure robust cybersecurity defenses in today's interconnected world.

7. References

[1] 1. M. Ozkan-Okay, R. Samet, O. Aslan, and D. Gupta, “A Comprehensive Systematic Literature Review on Intrusion Detection Systems,” *IEEE Access*, vol. 9, pp. 157727–157760, 2021, doi: 10.1109/ACCESS.2021.3129336

[2] S. Jin, J.-G. Chung, and Y. Xu, “Signature-Based Intrusion Detection System (IDS) for In-Vehicle CAN Bus Network,” in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2021, pp. 1–5. doi: 10.1109/ISCAS51556.2021.9401087. Available: https://ieeexplore.ieee.org/abstract/document/9401087. [Accessed: Feb. 27, 2024]

[3] F. Erlacher and F. Dressler, “FIXIDS: A high-speed signature-based flow intrusion detection system,” in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, Apr. 2018, pp. 1–8. doi: 10.1109/NOMS.2018.8406247. Available: https://ieeexplore.ieee.org/abstract/document/8406247. [Accessed: Feb. 27, 2024]

[4] D. Preuveneers, V. Rimmer, I. Tsingenopoulos, J. Spooren, W. Joosen, and E. Ilie-Zudor, “Chained Anomaly Detection Models for Federated Learning: An Intrusion Detection Case Study,” *Appl. Sci.*, vol. 8, no. 12, Art. no. 12, Dec. 2018, doi: 10.3390/app8122663

[5] Y. Li and Q. Liu, “A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments,” *Energy Rep.*, vol. 7, Sep. 2021, doi: 10.1016/j.egyr.2021.08.126

[6] Z. Yang *et al.*, “A systematic literature review of methods and datasets for anomaly-based network intrusion detection,” *Comput. Secur.*, vol. 116, p. 102675, May 2022, doi: 10.1016/j.cose.2022.102675

[7] M. Maithem and G. A. Al-sultany, “Network intrusion detection system using deep neural networks,” *J. Phys. Conf. Ser.*, vol. 1804, no. 1, p. 012138, Feb. 2021, doi: 10.1088/1742-6596/1804/1/012138

[8] M. Ramaiah, V. Chandrasekaran, V. Ravi, and N. Kumar, “An intrusion detection system using optimized deep neural network architecture,” *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 4, p. e4221, 2021, doi: 10.1002/ett.4221

[9] “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore.ieee.org/abstract/document/8066291. [Accessed: Feb. 27, 2024]

[10] “LSTM-Based Intrusion Detection System for In-Vehicle Can Bus Communications | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore.ieee.org/abstract/document/9216166. [Accessed: Feb. 27, 2024]

[11] Y. Imrana, Y. Xiang, L. Ali, and Z. Abdul-Rauf, “A bidirectional LSTM deep learning approach for intrusion detection,” *Expert Syst. Appl.*, vol. 185, p. 115524, Dec. 2021, doi: 10.1016/j.eswa.2021.115524

[12] M. Bacevicius and A. Paulauskaite-Taraseviciene, “Machine Learning Algorithms for Raw and Unbalanced Intrusion Detection Data in a Multi-Class Classification Problem,” *Appl. Sci.*, vol. 13, no. 12, Art. no. 12, Jan. 2023, doi: 10.3390/app13127328

[13] A. Rosay, E. Cheval, F. Carlier, and P. Leroux, “Network Intrusion Detection: A Comprehensive Analysis of CIC-IDS2017:,” in *Proceedings of the 8th International Conference on Information Systems Security and Privacy*, Online Streaming, SCITEPRESS - Science and Technology Publications, 2022, pp. 25–36. doi: 10.5220/0010774000003120. Available: https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0010774000003120. [Accessed: Feb. 27, 2024]

[14] J. L. Leevy and T. M. Khoshgoftaar, “A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 Big Data,” *J. Big Data*, vol. 7, no. 1, p. 104, Nov. 2020, doi: 10.1186/s40537-020-00382-x

[15] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization:,” in *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications, 2018, pp. 108–116. doi: 10.5220/0006639801080116. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006639801080116. [Accessed: Feb. 27, 2024]

[16] Y. N. Kunang, S. Nurmaini, D. Stiawan, and B. Y. Suprapto, “Attack classification of an intrusion detection system using deep learning and hyperparameter optimization,” *J. Inf. Secur. Appl.*, vol. 58, p. 102804, May 2021, doi: 10.1016/j.jisa.2021.102804

[17] S. N. Mighan and M. Kahani, “A novel scalable intrusion detection system based on deep learning,” *Int. J. Inf. Secur.*, vol. 20, no. 3, pp. 387–403, Jun. 2021, doi: 10.1007/s10207-020-00508-5

[18] Z. Wang, Y. Liu, D. He, and S. Chan, “Intrusion detection methods based on integrated deep learning model,” *Comput. Secur.*, vol. 103, p. 102177, Apr. 2021, doi: 10.1016/j.cose.2021.102177

[19] R. I. Farhan, A. T. Maolood, and N. F. Hassan, “Optimized Deep Learning with Binary PSO for Intrusion Detection on CSE-CIC-IDS2018 Dataset,” *J. Al-Qadisiyah Comput. Sci. Math.*, vol. 12, no. 3, Nov. 2020, doi: 10.29304/jqcm.2020.12.3.706. Available: https://jqcsm.qu.edu.iq/index.php/journalcm/article/view/706. [Accessed: Mar. 17, 2024]

[20] Y.-C. Wang, Y.-C. Houng, H.-X. Chen, and S.-M. Tseng, “Network Anomaly Intrusion Detection Based on Deep Learning Approach,” *Sensors*, vol. 23, no. 4, Art. no. 4, Jan. 2023, doi: 10.3390/s23042171

[21] “CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore.ieee.org/abstract/document/9889698. [Accessed: Mar. 17, 2024]

[22] R. Z. of Science *et al.*, “Recurrent Neural Networks & Deep Neural Networks Based on Intrusion Detection System,” *Open Access Libr. J.*, vol. 07, no. 03, Art. no. 03, 2020, doi: 10.4236/oalib.1106151

[23] V. Bulavas, V. Marcinkevičius, and J. Rumiński, “Study of Multi-Class Classification Algorithms’ Performance on Highly Imbalanced Network Intrusion Datasets,” *Informatica*, vol. 32, no. 3, pp. 441–475, Jan. 2021, doi: 10.15388/21-INFOR457

[24] M. Zhong, M. Lin, and Z. He, “Dynamic multi-scale topological representation for enhancing network intrusion detection,” *Comput. Secur.*, vol. 135, p. 103516, Dec. 2023, doi: 10.1016/j.cose.2023.103516

[25] W. Liu *et al.*, “Intrusion Detection for Maritime Transportation Systems With Batch Federated Aggregation,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2503–2514, Feb. 2023, doi: 10.1109/TITS.2022.3181436

[26] O. Belarbi, A. Khan, P. Carnelli, and T. Spyridopoulos, “An Intrusion Detection System Based on Deep Belief Networks,” in *Science of Cyber Security*, C. Su, K. Sakurai, and F. Liu, Eds., in Lecture Notes in Computer Science, vol. 13580. Cham: Springer International Publishing, 2022, pp. 377–392. doi: 10.1007/978-3-031-17551-0_25. Available: https://link.springer.com/10.1007/978-3-031-17551-0_25. [Accessed: Apr. 18, 2024]

[27] “Deep Learning Based Multi-Channel Intelligent Attack Detection for Data Security | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore-ieee-org.ezproxy.ktu.edu/abstract/document/8259310. [Accessed: Apr. 18, 2024]

[28] G. Karatas, O. Demir, and O. K. Sahingoz, “Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset,” *IEEE Access*, vol. 8, pp. 32150–32162, 2020, doi: 10.1109/ACCESS.2020.2973219

[29] “Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore-ieee-org.ezproxy.ktu.edu/abstract/document/9311173. [Accessed: Apr. 18, 2024]

[30] S. Gamage and J. Samarabandu, “Deep learning methods in network intrusion detection: A survey and an objective comparison,” *J. Netw. Comput. Appl.*, vol. 169, p. 102767, Nov. 2020, doi: 10.1016/j.jnca.2020.102767

[31] “Network Anomaly Detection With Temporal Convolutional Network and U-Net Model | IEEE Journals & Magazine | IEEE Xplore.” Available: https://ieeexplore-ieee-org.ezproxy.ktu.edu/abstract/document/9583228. [Accessed: Apr. 18, 2024]

[32] “Cyber threat intelligence using PCA-DNN model to detect abnormal network behavior - ScienceDirect.” Available: https://www-sciencedirect-com.ezproxy.ktu.edu/science/article/pii/S1110866521000785#b0155. [Accessed: Apr. 18, 2024]