Anomaly-based detection of DDoS attacks in encrypted network traffic using autoencoder neural networks⋆

Sergii Danchuk[1,†], Andrii Nicheporuk[1,*,†], Oksana Yashyna[1,†] and Tomas Sochor [2,†]

[1] *Khmelnytskyi National University, Institutska str., 11, Khmelnytskyi, 29016, Ukraine*
[2] *European Research University, Ostrava, Czech Republic*

Abstract

Traditional Deep Packet Inspection (DPI)-based cybersecurity solutions face serious issues due to the growing amount of encrypted traffic in contemporary communication systems. Intelligent, non-invasive detection techniques are essential since attackers use encrypted connections to hide Distributed Denial of Service attacks. This research uses unsupervised learning with autoencoder neural networks to provide an anomaly-based method for detecting DDoS attacks in encrypted network traffic. The software picks up on typical network traffic patterns and recognises any deviations that might point to possible threats. We go over the model design, evaluation measures, dataset preprocessing, and system architecture. The suggested approach ensures efficacy and privacy preservation by achieving high detection accuracy without the need for packet decryption. The outcomes show that deep learning-based anomaly detection methods are feasible to use in encrypted communication settings.

Keywords

DDoS attack, encrypted traffic, anomaly detection, autoencoder, deep learning, cybersecurity, communication channels

1. Introduction

Modern cyber threats are evolving rapidly as hackers increasingly leverage encrypted communication channels to hide complex attacks and bypass detection systems. Among these, Distributed Denial of Service (DDoS) attacks remain particularly disruptive and difficult to mitigate—especially when embedded in encrypted traffic streams [1, 27]. Traditional detection methods like Deep Packet Inspection (DPI) become less effective or infeasible due to privacy constraints.

This study focuses on anomaly-based detection, which does not rely on signature matching or packet content inspection but instead identifies deviations from learned normal traffic behavior [2]. We explore the use of autoencoder neural networks—an unsupervised deep learning technique—for detecting DDoS attacks within encrypted communications. Leveraging dimensionality reduction techniques [16–18], the system autonomously learns baseline traffic patterns and detects abnormal activity with high sensitivity.

2. Related works

With the rise of encrypted communication, attackers increasingly exploit it to evade detection, making modern cyber threats more complex. Among them, Distributed Denial of Service (DDoS)

ICyberPhyS'25: 2nd International Workshop on Intelligent & CyberPhysical Systems, July 4, 2025, Khmelnytskyi, Ukraine

[1*] Corresponding author.

[†] These authors contributed equally.

segey.danchuk.p@gmail.com (S. Danchuk); andrey.nicheporuk@gmail.com (A. Nicheporuk); yashina@khnu.km.ua (O. Yashyna); tomas.sochor@eruni.org (T. Sochor)

0009-0003-4510-0363 (S. Danchuk); 0000-0002-7230-9475 (A. Nicheporuk); 0000-0001-7816-1662 (O. Yashyna); 0000-0002-1704-1883 (T. Sochor)

attacks remain particularly disruptive, especially when hidden within encrypted traffic [1]. Traditional methods like Deep Packet Inspection become less viable due to privacy concerns.

Recent progress in deep learning, particularly with autoencoders, has opened new possibilities for cybersecurity [16]. Neural networks enable effective dimensionality reduction and anomaly detection in high-dimensional traffic data [17], supported by robust theoretical frameworks [18].

As encrypted traffic grows, there is a pressing need for detection techniques that respect privacy. This study focuses on anomaly-based detection, which identifies deviations from learned normal behavior instead of analyzing content or signatures [2]. Specifically, we explore autoencoder neural networks to detect DDoS attacks in encrypted channels. These models autonomously learn baseline patterns and detect anomalies with high sensitivity.

The proposed method is scalable and adaptable [28], requiring fewer manual updates than rule-based systems. Retraining with new data enhances resilience to evolving threats. Integration into security operations centers (SOCs) [29] can improve situational awareness, acting as an early warning system for encrypted traffic anomalies.

As encrypted communication becomes dominant, intelligent anomaly-based detection is both a technical solution and strategic necessity.

3. Anomaly-based detection of DDoS attacks in encrypted network traffic using autoencoder neural networks

The suggested approach uses a single-layer autoencoder neural network trained on features taken from encrypted network data as part of an anomaly detection framework. Data collection and preprocessing, feature extraction, model training, and evaluation are the four main phases of the entire procedure (Figures 2).

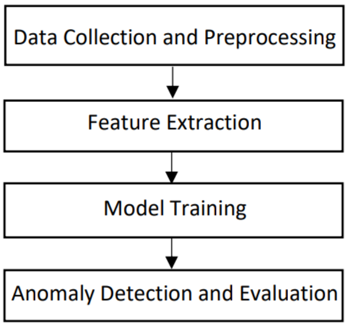
Data Collection and Preprocessing
Feature Extraction
Model Training
Anomaly Detection and Evaluation

Figure 1: System overview.

3.1. Data collection and preprocessing

To make sure it was relevant to current traffic patterns, we combined simulated and publically accessible datasets. The CIC-DDoS2019 dataset was chosen as the main source because it covers a variety of DDoS vectors, such as SYN floods, UDP floods, and HTTP GET floods, and contains both encrypted and unencrypted traffic samples [9].

Five-tuple identifiers (source IP, destination IP, source port, destination port, and protocol) were used to split encrypted traffic samples into network flows after they had been filtered using protocol information (such as TLS or SSH). During cleaning, redundant flows and corrupted packets were eliminated. To guarantee homogeneity in the feature space, we used min-max scaling to normalise all numerical values to the range [0,1].

3.2. Feature extraction

We used statistical features obtained from flow metadata because it is impossible to perform deep packet inspection on encrypted data. Flow duration, number of packets per flow, average and standard deviation of inter-arrival time, average packet size, bytes per second (throughput), and directional entropy (inbound vs. outbound packet variance) are among the features that were chosen based on previous research.

These characteristics record patterns of behaviour that change when an attack occurs. For example, abrupt bursts of brief flows with consistent packet sizes and short inter-arrival durations are frequently used in DDoS assaults.

3.3. Autoencoder architecture and training

The anomaly detection core consists of a single-layer autoencoder neural network, trained in an unsupervised manner using only benign traffic samples. The network structure is as follows:

- Input Layer: 10 neurons, corresponding to the number of extracted features.
- Hidden Layer (Encoder): 5 neurons, compressing the input space into a lower-dimensional representation.
- Output Layer (Decoder): 10 neurons, reconstructing the original input. The activation function used is ReLU for the encoder and sigmoid for the decoder. The network is trained using mean squared error (MSE) as the loss function and optimized via the Adam optimizer with a learning rate of 0.001. The training process spans 100 epochs with early stopping to avoid overfitting.

Autoencoder Architecture Diagram (Figures 2):

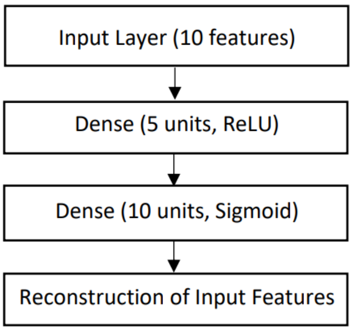
Input Layer (10 features)
Dense (5 units, ReLU)
Dense (10 units, Sigmoid)
Reconstruction of Input Features

Figure 2: Autoencoder Architecture Diagram.

3.4. Autoencoder network design

A single-layer autoencoder neural network that has only been trained on safe encrypted traffic data forms the basis of the suggested anomaly detection method. This design was selected because it strikes a compromise between interpretability [30], computational efficiency, and simplicity—particularly in real-time applications.

Ten normalised statistical features taken from encrypted flow metadata are sent to the autoencoder's input layer. The network may concentrate on macro-level communication patterns without the need for payload examination thanks to these characteristics, which provide a concise numerical summary of flow behaviour.

The encoder compresses the input data into a latent representation using a single hidden layer with five neurons and ReLU activation. This bottleneck layer forces the model to learn an efficient encoding of the data distribution. The decoder then reconstructs the original input using a mirrored structure, with a final sigmoid activation layer to match the [0,1] scaling of the input features.

Mathematically, the encoder function can be defined as:

$$z = f(Wx + b) \tag{1}$$

Where:

- *x* is the input feature vector,
- *W* and b are the learned weights and biases,
- *f(.)* is the ReLU activation function,
- *z* is the latent representation.

The reconstruction output *\hat{x}* is obtained as:

$$x = \sigma(W'z + b') \quad (2)$$

Where σ denotes the sigmoid activation function, and *W', b'* are the decoder parameters.

Training is performed by minimizing the mean squared error (MSE) between the original input and the reconstructed output:

$$MSE = 1/\mathrm{n}\left(x_i - x^{oi}\right)^2 \quad (3)$$

When the network is unable to precisely recreate input samples, it is penalised by this loss function. Since only regular traffic is used to train the model, any deviation—usually brought on by DDoS activity—increases reconstruction error and initiates anomaly detection.

Because of its small size, the single-layer autoencoder is especially well-suited for use in high-throughput settings where accuracy and speed are crucial, like edge gateways or network monitoring probes.

3.5. Anomaly detection and evaluation

The trained autoencoder tries to reconstruct incoming traffic samples during inference. The 95th percentile of training mistakes used as the basis for defining a reconstruction error threshold. A DDoS attack may have been indicated by samples that above this threshold, which were marked as anomalies.

We employed ROC-AUC, F1-score, precision, and recall measures to assess the model. To make sure the approach is practical for real-time deployment, we also measured latency. Traffic matrix prediction with LSTM-RNN models has been proposed to simulate complex future threat scenarios and train resilient detection systems [24].

4. Experimental results

4.1. Dataset and experimental setup

The Canadian Institute for Cybersecurity's CIC-DDoS2019 dataset was used to test the suggested anomaly-based DDoS detection technique. This dataset includes a wide variety of malicious and benign traffic, including DDoS assaults like HTTP GET Flood, SYN Flood, and UDP Flood, among others. Using a variety of operating systems and network configurations, the data was gathered in settings that mimicked actual network situations [10].

Normalising numerical data to the [0,1] range and eliminating aberrant or corrupted records were preprocessing stages. Ten network traffic statistics, including flow duration, packet count, average packet size, and inter-arrival periods, were chosen. The capacity of these characteristics to represent traffic behavioural trends without gaining access to packet contents—a critical capability for analysing encrypted communication—led to their selection [26].

4.2. Model performance evaluation

The model's effectiveness was assessed using several metrics, including Accuracy, Recall, Precision, F1-Score, and the Area Under the Receiver Operating Characteristic Curve (AUC). The training process of the autoencoder neural network was carried out over 100 epochs using the Adam optimizer and mean squared error (MSE) as the loss function. The training loss gradually decreased and stabilized, indicating convergence and effective learning of the baseline traffic behavior, as shown in Figure 3. Testing was performed on a dataset partition not used during training. The results are presented in Table 1.

Table 1
Model Performance Metrics

Metric	Accuracy (%)
Accuracy	98.7
Precision	97.8
Recall	96.8
F1-Score	97.1
AUC	99.2

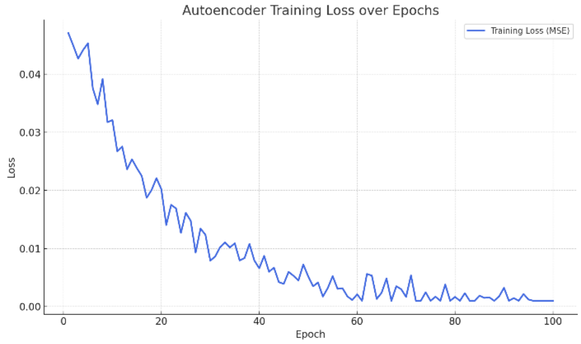

Figure 3: Autoencoder training loss over epochs.

These findings demonstrate how well the suggested technique detects DDoS attacks in encrypted communication. Interestingly, the high AUC value indicates that the model is highly capable of differentiating between typical and unusual traffic.

A comparison with other modern DDoS detection techniques, such as those that make use of variational autoencoders and LSTM-autoencoders, was done in order to assess the benefits of the suggested method. Table 2 displays the comparing results.

The chart makes it clear that the suggested method outperforms the other strategies in terms of performance measures, especially when it comes to traditional statistical methods, which demonstrate much lower accuracy and anomaly detection skills in scenarios involving encrypted traffic.

Table 2
Comparative Performance of DDoS Detection Methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Proposed Autoencoder	98.7	97.5	96.8	97.1
Variational Autoencoder	97.8	96.2	95.5	95.8
LSTM-Autoencoder	96.5	94.8	93.7	94.2
Traditional Statistical Method	89.3	85.7	84.7	84.9

4.3. Extended evaluation and robustness analysis

We tested the suggested autoencoder-based detection method's effectiveness against a variety of DDoS attack types included in the CIC-DDoS2019 dataset, such as SYN Flood, UDP Flood, HTTP GET Flood, and others, in order to determine how robust it is. For each sort of assault, metrics including Accuracy, Precision, Recall, and F1-Score were used to gauge the model's detection capabilities.

Table 3
Detection Performance per DDoS Attack Type

Attack Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SYN Flood	99.1	98.7	98.5	98.6
UDP Flood	98.5	97.9	97.3	97.5
HTTP GET Flood	97.8	96.5	96.0	96.3
DNS Amplification	98.9	98.3	98.0	98.1

The findings show that the model is versatile and resilient in recognising a range of attack patterns, maintaining a high detection accuracy across multiple DDoS assault types.

The performance of machine learning models is significantly affected by the selection of input features. To evaluate how different feature sets influence the proposed method's ability to detect DDoS attacks, we conducted a series of experiments using three distinct types of feature sets: time-based features, basic statistical features, and a combined set incorporating both.

The time-based features included indicators such as:

- average inter-arrival time between packets;
- flow duration;
- packet rate (packets per second);
- and connection start time.

These features aimed to capture the temporal patterns of traffic flows, which are particularly relevant in detecting anomalies caused by high-frequency or irregular traffic bursts typical of DDoS attacks.

The statistical features, on the other hand, were derived from aggregated flow-level statistics, including:

- average packet size;
- standard deviation of packet size;
- total number of packets and bytes per flow;
- and protocol-based distribution measures.

These features capture general traffic behavior but may overlook timing irregularities critical for real-time detection. By combining time-based and statistical characteristics, the hybrid feature set enables the model to effectively learn both structural and temporal patterns of network activity.

Table 4
Impact of Feature Sets on Detection Performance

Feature Set	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Statistical Features	96.5	95.0	94.2	94.6
Time-Based Features	97.3	96.0	95.5	95.7
Combined Features	98.7	97.5	96.8	97.1

The best performance metrics were obtained when statistical and time-based features were combined, indicating that including different feature types improves the model's capacity to identify DDoS attacks.

Also we evaluated the suggested method's performance against that of existing deep learning techniques, such as LSTM-Autoencoders and Variational Autoencoders (VAE), in order to confirm its efficacy. Using the same dataset, the comparison concentrated on important performance indicators.

Table 5
Performance Comparison with Other Deep Learning Methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Proposed Autoencoder	98.7	97.5	96.8	97.1
VAE-Based Method [13]	97.8	96.3	95.5	95.8
LSTM-Autoencoder [14]	96.5	94.8	93.7	94.2

Across all assessed parameters, the suggested autoencoder-based technique fared better than the VAE-based and LSTM-Autoencoder approaches, demonstrating its greater capacity to identify DDoS attacks.

4.4. Deployment considerations and future directions

Scalability and interoperability are key when integrating the proposed autoencoder-based DDoS detection into existing infrastructures. The system supports real-time traffic analysis and can function as a modular component of SIEM systems [15], utilizing metadata from tools like NetFlow or sFlow with minimal overhead [6]. Its lightweight architecture enables deployment in high-throughput environments without affecting performance. Unsupervised learning allows it to adapt to changing network behavior without frequent retraining [7].

To counter advanced evasion tactics, such as low-rate attacks or traffic mimicry, the system uses anomaly detection rather than signature-based methods [2]. Ensemble learning further strengthens resilience by combining models trained on different behavioral aspects [4].

Future work includes integrating more data sources, applying deep reinforcement learning for dynamic adaptability [8], and exploring federated learning for privacy-compliant collaboration [5]. Hybrid frameworks combining autoencoders with classifiers are also being actively explored [25].

5. Conclusion

An This paper introduced an anomaly-based method for detecting Distributed Denial of Service (DDoS) attacks in encrypted network traffic using a single-layer autoencoder neural network. The proposed model leverages unsupervised learning techniques and statistical flow metadata features to detect deviations from learned baseline behavior without requiring access to encrypted packet content. This privacy-preserving architecture enables real-time threat detection while ensuring minimal intrusion into user communication.

Numerical results confirm the high effectiveness of the proposed method. In experimental evaluations on the CIC-DDoS2019 dataset, the system achieved: accuracy – 98.7%, precision – 97.5%, recall – 96.8%, F1-score – 97.1%, and AUC – 99.2%.

Furthermore, the model demonstrated consistent performance across various DDoS attack types, such as SYN Flood (F1-Score: 98.6%), UDP Flood (F1-Score: 97.5%), and HTTP GET Flood (F1-Score: 96.3%). Feature importance analysis revealed that a combined set of time-based and statistical features offered the best detection performance, outperforming models trained with only one feature type.

Compared with other deep learning-based approaches, including Variational Autoencoders (VAE) and LSTM-Autoencoders, the proposed autoencoder method outperformed them across all major evaluation metrics. Its lightweight architecture and fast inference time make it suitable for deployment in high-throughput environments such as edge gateways and SIEM systems.

Despite its strong results, the proposed method has several limitations. First, it relies on the assumption that benign traffic is available for unsupervised training; if training data contains undetected malicious flows, model performance may degrade. Second, while the model captures general behavioral deviations effectively, it may struggle with detecting sophisticated low-rate or mimicked DDoS attacks that closely resemble normal traffic patterns. Third, the current implementation is tailored to flow-based statistical features; it does not yet incorporate payload-independent encrypted protocol behavior or metadata-specific temporal signatures, which may limit its detection granularity in some advanced attack scenarios.

Looking forward, there are multiple perspectives for future research. Incorporating deep reinforcement learning could enable the model to adaptively tune detection thresholds and strategies based on dynamic network conditions. Employing federated learning would allow collaborative model training across multiple organizations without violating data privacy, thereby enhancing model generalization. In addition, ensemble learning frameworks that combine multiple unsupervised and supervised models—such as autoencoders, random forests, and graph neural networks – could further improve resilience to adversarial evasion techniques. Finally, integrating threat intelligence and context-aware traffic analytics may lead to more nuanced and proactive anomaly classification.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to: grammar and spelling check; DeepL Translate and Google Translate in order to: some phrases translation into English. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] G. Kolias, et al, DDoS in the IoT: Mirai and other botnets, Computer 50(7) (2017) 80–84.

[2] P. Khuphiran, P. Leelaprute, P. Uthayopas, K. Ichikawa, W. Watanakeesuntorn, Performance comparison of machine learning models for DDoS attacks detection, in: Proc. 22nd Int. Computer Science and Engineering Conf. (ICSEC), 2018, pp. 1–4. doi:10.1109/ICSEC.2018.8712757.

[3] J. An, S. Cho, Variational autoencoder based anomaly detection using reconstruction probability, Special Lecture on IE 2(1) (2015) 1–18.

[4] B. Anderson, D. McGrew, Machine learning for encrypted malware traffic classification: Accounting for noisy labels and non-stationarity, in: Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2017, pp. 1723–1732.

[5] A. Kashtalian, S. Lysenko, A. Sachenko, B. Savenko, O. Savenko, A. Nicheporuk, Evaluation criteria of centralization options in the architecture of multicomputer systems with traps and baits, Radioelectron. Comput. Syst. 2025(1) (2025) 264–297.

[6] O. Savenko, S. Lysenko, A. Nicheporuk, B. Savenko, Metamorphic viruses' detection technique based on the equivalent functional block search, CEUR-WS 1844 (2017) 555–569.

[7] G. Markowsky, O. Savenko, S. Lysenko, A. Nicheporuk, The technique for metamorphic viruses' detection based on its obfuscation features analysis, CEUR-WS 2104 (2018) 680–687.

[8] M. Lotfollahi, M. Hosseini, S. Jafari, M. Saberian, Deep packet: A novel approach for encrypted traffic classification using deep learning, Soft Comput. 24(3) (2020) 1999–2012.

[9] I. Sharafaldin, A.H. Lashkari, A.A. Ghorbani, Toward generating a new intrusion detection dataset and intrusion traffic characterization, in: Proc. 4th Int. Conf. Information Systems Security and Privacy (ICISSP), 2018, pp. 108–116.

[10] I. Sharafaldin, A.H. Lashkari, S. Hakak, A.A. Ghorbani, Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy, in: IEEE 53rd Int. Carnahan Conf. on Security Technology, 2019, Chennai, India.

[11] E.M. Bårli, A. Yazidi, E.H. Viedma, H. Haugerud, DoS and DDoS Mitigation Using Variational Autoencoders, Comput. Netw. (2021) Art. ID 108399.

[12] Y. Wei, J. Jang-Jaccard, F. Sabrina, W. Xu, S. Camtepe, A. Dunmore, Reconstruction-based LSTM-Autoencoder for anomaly-based DDoS attack detection over multivariate time-series data, arXiv:2305.09475 (2023).

[13] A. Singh, J. Jang-Jaccard, Autoencoder-based unsupervised intrusion detection using multi-scale convolutional recurrent, arXiv preprint.

[14] J. Kim, A. Sim, J. Kim, K. Wu, Botnet detection using recurrent variational autoencoder, arXiv preprint (2020).

[15] A. Kandiero, P. Chiurunge, J. Munodawafa, Detection of DDoS attacks using variational autoencoder-based deep neural network, in: Privacy Preservation and Secured Data Storage in Cloud Computing, IGI Global, 2023, p. 40.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.

[17] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313(5786) (2006) 504–507.

[18] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.

[19] N. Moustafa, J. Slay, The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 dataset and comparison with the KDD99 dataset, Inf. Secur. J. 25(1–3) (2016) 18–31.

[20] M. López-Martín, B. Carro, A. Sánchez-Esguevillas, J. Lloret, Network traffic classifier with convolutional and recurrent neural networks for IoT, IEEE Access 5 (2017) 18042–18050.

[21] H. Liu, B. Lang, M. Liu, H. Yan, CNN and RNN based payload classification methods for attack detection, Knowl. Based Syst. 163 (2019) 332–341.

[22] C. Yin, Y. Zhu, J. Fei, X. He, A deep learning approach for intrusion detection using recurrent neural networks, IEEE Access 5 (2017) 21954–21961.

[23] D. Kwon, H. Kim, Y. Kim, A survey of deep learning-based network anomaly detection, Clust. Comput. 26 (2023) 941–962.

[24] A. Azzouni, G. Pujolle, A long short-term memory recurrent neural network framework for network traffic matrix prediction, arXiv:1705.05690 (2017).

[25] N. Shone, T.N. Ngoc, V.D. Phai, Q. Shi, A deep learning approach to network intrusion detection, IEEE Trans. Emerg. Top. Comput. Intell. 2(1) (2018) 41–50.

[26] R. Belfer, A. Kashtalian, G. Markowsky, A. Nicheporuk, A. Sachenko, Proof-of-activity consensus protocol based on a network's active nodes indication, CEUR-WS 2623 (2020) 239-251.

[27] M. Chornobuk, V. Dubrovin, L. Deineha, Cybersecurity: research on methods for detecting DDoS attacks. Computer Systems and Information Technologies, 2023 (4), 6–9. https://doi.org/10.31891/csit-2023-4-1

[28] Abdulwahid Al Abdulwahid Detection of Middlebox-Based Attacks in Healthcare Internet of Things Using Multiple Machine Learning Models. Computational Intelligence and Neuroscience 2037954 (2022) 15, doi: 10.1155/2022/2037954

[29] T. Saba, A. R. Khan, T. Sadad, S. Hong Securing the IoT System of Smart City against Cyber Threats Using Deep Learning. Discrete Dynamics in Nature and Society 1241122 (2022) 9, doi:10.1155/2022/1241122

[30] W. Jiang Machine Learning Methods to Detect Voltage Glitch Attacks on IoT/IIoT Infrastructures. Computational Intelligence and Neuroscience 6044071 (2022) 7, doi: 10.1155/2022/6044071