

A reinforcement learning-based intrusion detection system for internet of things with zero-day resilience

Sowmya Somanath^{1,2}, Usha Banavikal Ajay¹, Sangeetha Kodlipet Nanjundaswamy³

¹Department of Computer Science and Engineering, BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi, India

²School of Computer Science and Engineering, REVA University, Bengaluru, India

³Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jul 9, 2025

Revised Jul 28, 2026

Accepted Aug 12, 2026

Keywords:

Cybersecurity in internet of things

Deep Q-network

Internet of things

Intrusion detection

Reinforcement learning

Zero-day attacks

ABSTRACT

The conventional intrusion detection systems (IDS) are often found to be ineffective when faced with zero-day attacks that are novel compared to previously learned behavior patterns. This paper presents the development of a robust and adaptive IDS for the internet of things (IoT) using a deep Q-network (DQN) agent. A custom environment is designed in which state abstraction is applied to convert high-dimensional feature vectors into discrete state spaces. The proposed agent model is trained to detect normal and malicious traffic using a reward and penalty mechanism to encourage the agent to take correct decisions. The validation of the proposed methodology is carried out using the UNSW-NB15 dataset under both known attack and zero-day attack scenarios. The results showed detection accuracies of 99.56% for known attacks and 99.11% for zero-day attacks.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Sowmya Somanath

Department of Computer Science and Engineering, BMS Institute of Technology and Management

Visvesvaraya Technological University

Doddaballapura Main Rd, Yelahanka, Avalahalli, Bengaluru, Karnataka 560119, India

Email: sowmyasadish@gmail.com

1. INTRODUCTION

The rapid growth of the Internet and digital communication technologies has increased the size and complexity of network infrastructure [1]. Along with this development, the number and sophistication of cyberattacks have increased, including complex and difficult-to-recognize attack methods. Hence, intrusion detection systems (IDS) are crucial for securing networks as they detect security problems or data breaches by finding abnormal behavior [2], [3]. Nevertheless, the adaptability of the conventional rule-based IDS approach is often inadequate in detecting new attacks. The critical challenge associated with the current IDS research is the detection of new and zero-day attacks, which cannot be identified using signature-based methods since there are no signatures of unseen attacks [4], [5]. Zero-day attacks can pass existing IDS systems with disastrous effects, especially in critical domains such as internet of things (IoT), industrial control systems, and cloud computing infrastructure [6]. Furthermore, the inherent class imbalance issues associated with the network traffic data, where higher instances of normal activities exist compared to attacks, which makes even more difficult to create a robust and unbiased IDS [7]. In recent literature work, many computational schemes have been introduced by researchers using machine learning (ML) and deep learning (DL) approaches to detect both known and unknown attacks [8]. In the existing literature, ML classifiers have been widely used to develop lightweight IDSs that are very effective for known attack patterns, but are not very adaptive and efficient when there is complex traffic data and dynamic attacks in the

network. Other issues associated with ML-driven IDSs are that their performance depends on the preprocessing and feature engineering steps, which are manual, time-consuming, and domain expert depended task [9]. To overcome the limitation of ML-driven IDS solutions, researchers used DL models such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) due to their automation in feature extraction and learning both high-level and low-level attributes from the raw traffic data [10]. However, most existing DL-based IDS solutions remain supervised in nature, which are not very adaptive to evolving traffic patterns in the IoT environment.

In recent state-of-the-art works, a wide range of solutions using advanced DL and hybrid models were explored towards developing intelligent and robust IDS for traditional wireless networks and IoT ecosystems [11]. The work carried out by Bhavani and Mangla [12] suggested a regularization technique to reduce the impact of class imbalance on detection performance of the supervised models. The outcome showed improved accuracy and F1-scores for detecting different attacks. Jamalipour and Murali [13] introduced a lightweight IDS model using a support vector machine (SVM) classifier, whose feature selection process is emphasized to refine IDS performance. Zhao *et al.* [14] presented a hybrid learning model that used a generative adversarial network (GAN) DL model to make training data more dynamic to address class imbalance issues and used a k-nearest neighbor (KNN) classifier to perform intrusion detection. Novaes *et al.* [15] implemented a long short-term memory (LSTM) network, an RNN model, where principal component analysis (PCA) is used for data dimensionality reduction and similar work is also seen by Imrana *et al.* [16], where they used a bidirectional long short-term memory (BiLSTM) model for stronger IDS resilience. Halbouni *et al.* [17] presented a hybrid model based on the joint application of LSTM-CNN to learn both spatial features and temporal patterns in network traffic. Chen *et al.* [18] focus on using an unsupervised learning approach where k-means clustering is used with ant-lion optimization to effectively isolate anomalous behavior. Siahmarzkooh and Alimardani [19] used hierarchical clustering with whale optimization to build a scalable IDS. Oluwasanmi *et al.* [20] built an anomaly detection model by using autoencoders, whereas stacked sparse autoencoders with enhanced gaussian mixture models are suggested by Zhang *et al.* [21]. Another contribution by Bhavani and Mangla [22] is the development of an improved anomaly-based IDS model where autoencoders with clustering techniques and supervised classifiers were tested for multi-class intrusion detection. Tharewal *et al.* [23] adopted a reinforcement learning (RL) technique to develop an adaptive IDS that follows a Markov decision process (MDP) where the agent is trained based on environmental feedback. Kumar *et al.* [24] also used an RL approach where deep Q-learning is used in IDS to mitigate denial of service (DoS) attacks. Suratkar *et al.* [25] used a Q-learning-based honeypot system for dynamic attack analysis. Hossain [26] integrated a deep encoder network with Q-networks for real-time detection and classification of cyber threats. Badr [27] presented a dueling deep Q-network (DQN) framework for processing complex traffic data and a robust IDS agent model for efficient response against dynamic threats. Despite many research contributions, significant open challenges persist:

- i) Most existing works on IDS using ML and DL still lack adaptability in detecting unknown attacks.
- ii) The existing IDS using RL algorithms are subjected to weak reward mechanisms and agent training to counter high-risk misclassifications.
- iii) There is also inadequate handling of the class imbalance issue in feature extraction, which biases model learning towards majority classes.
- iv) Limited exploration and integration of state abstraction mechanisms in an RL-driven solution that can reduce the complexity of high-dimensional data.

In order to address identified research gaps, the proposed research work presents the development of an adaptive network anomaly detection framework using a deep RL algorithm. The proposed framework employs a state abstraction mechanism using a clustering scheme with an efficient reward mechanism to prioritize critical learning and perform correct actions for anomaly detection by the agent. The core contributions are as follows: i) zero-day detection setup, ii) development of a DQN-driven agent with a reward function that strongly penalizes false negatives, improving attack detection reliability, iii) comprehensive data preprocessing including class balancing, data scaling, feature selection, and encoding, and iv) evaluation on both known and zero-day attack samples. The novelty of this work is the integration of deep Q-learning into a complete IDS pipeline that is both zero-day-aware and reinforcement-driven.

2. METHOD

The proposed system architecture is data-driven and consists of a number of processes, including data preprocessing, environment modelling, agent training, and agent evaluation. The data preprocessing stage includes standardization, feature selection, and class balancing using oversampling (synthetic minority over-sampling technique or SMOTE), undersampling, and hybrid methods. This process results in a cleaned and balanced dataset with retaining relevant features that are used as input to the model, which increases

learning efficiency and reduces the likelihood of overfitting. The preprocessed data is used to generate an RL environment, where high-dimensional input features are mapped into discrete states using MiniBatchKMeans clustering. The environment is defined by a custom object (NetworkTrafficEnvironment) that simulates traffic flow and produces rewards based on the correctness of classifying traffic actions as normal or attack instances. A custom object also defines a reward function that imposes a higher penalty on false negatives, which is significant for anomaly detection scenarios. In the DQN agent training module, agents are trained using a DQN architecture composed of multilayer perceptron (MLP) layers. The agent uses an experience replay buffer that allows it to learn from past state-action transitions and uses a target network to provide training stability. The agent interacts with the environment for a limited number of episodes, during which it improves its policy through Q-learning and ϵ -greedy decision making. Figure 1 illustrates the system model and implementation workflow.

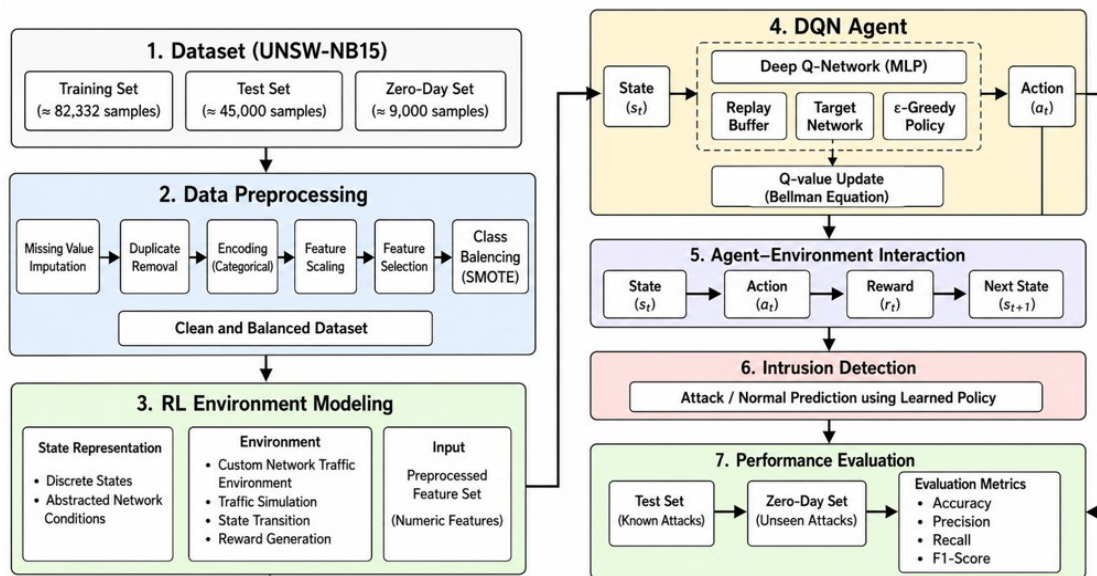


Figure 1. Design architecture and workflow of the proposed predictive system

2.1. Dataset description

The proposed study utilizes the UNSW-NB15 dataset, which is an extensive dataset of network traffic, widely used in the research area of network threat detection and forensic data analysis. The dataset is from the Australian Centre for Cyber Security (ACCS) and aims to enhance and expand beyond older datasets such as KDD99, providing more diverse and realistic representations of network activities that are both normal and malicious. Table 1 presents the dataset description, which includes approximately 2.5 million records that encompass real legitimate traffic along with synthetically generated attacks. This ensures diversity in the dataset involving normal and malicious scenarios. The dataset has 41 features that detail network behavior, such as flow features (source/destination IPs, ports, and protocols), time-based features (packet inter-arrival times), and content-based features (such as hypertext transfer protocol (HTTP) methods and domain name system (DNS) queries), among others, to facilitate the analysis of network traffic behavior and to effectively identify anomalies. The dataset classifies attacks into nine distinct categories, such as fuzzers, DoS, exploits, reconnaissance, worms, and many other attack types, which enable experiments on multi-class classification.

Table 1. Dataset description

Dataset split	Sample count (approximately)	Description
Training set	82,332	Filtered data with known attack types and normal samples. Used to train the DQN agent.
Testing set	45,000	Evaluation set comprising known attack types and normal traffic. Measures detection performance on seen attacks.
Zero-day set	9,000	Contains unseen (zero-day) attack types along with normal samples. Used to assess the generalization capability of the model.

2.2. Dataset preprocessing

An extensive preprocessing operation is carried out to make a quality dataset before training the RL agent model. The preprocessing aids in feature consistency, relevance, and balance, which are crucial aspects for an accurate learning process in the case of high-dimensional network traffic data. The dataset is first analyzed to detect missing values, which are then treated by imputing them with the mean of the values of the corresponding feature. This process reduces the chances of any outliers being introduced or distorting the data while maintaining the proportions of the entire dataset. Then, duplicate samples are evaluated and removed to eliminate redundancy, which is also important to identify previously unseen patterns that help avoid model overfitting. In the next step, all categorical features present in the dataset are encoded using label encoding, where all unique string labels are assigned an integer value. The next step carried out is the feature scaling process, where the StandardScaler method is utilized to normalize the traffic features to obtain zero-centered values and unit variance to make the training operation of the RL agent more stabilized and minimize the influence of features having comparatively larger numerical values. Afterwards, a correlation filtering operation is performed where features with a Pearson correlation coefficient greater than 0.9 are removed to reduce multicollinearity and redundant information. This process also helps in lowering the computational complexity of the model. The final stage of the preprocessing module involves feature selection, where the top 20th features are retained using the SelectKBest method based on the analysis of variance (ANOVA) F1-score. This feature set is then used for model training, as they show the strongest relationship with the target class, which enables the model to focus on important traffic patterns and also reduces the training time.

2.3. Agent and environment interaction

The proposed RL framework is formulated as an MDP represented by the tuple $(S_t, A_t, P_{sa}, R_w, \gamma)$, where S_t is the environment state, which mimics network characteristics such as traffic data, packet-level features, user requests, system commands, and protocol level characteristics. The action space A_t includes the decisions performed by the agent, such as intrusion detection and alert generation. The state transition probability P_{sa} defines the probability of moving from the current state to the next state after executing action a . For each executed action, the agent gets a feedback signal (reward) R_w from the environment. This feedback feature is incorporated to encourage the agent model to improvise its strategy or action policy towards getting accumulated reward. In the case when the agent performs an inappropriate action, it gets negative rewards as a penalty applied by the environment. The variable γ denotes the discount factor that ranges in $[0,1]$ to determine the current value of future rewards, and it optimizes the agent's action policy towards getting all future reward values. The action selection policy denoted as $\pi(a | s)$ defines the action chosen for each observed state s , $\pi: S \rightarrow A$. The framework also employs a state-value function $V^\pi(s)$, which estimates the expected cumulative reward signal from the given state s , numerically defined as (1) and (2).

$$V^\pi(s) = E[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t=s] \quad (1)$$

$$Q(s, a) = E[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t=s, A_t=a] \quad (2)$$

Where $V^\pi(s)$ is the long-term reward value from the current state (s), and the action-value function $Q(s, a)$ in (2) computes the reward expected after executing a specific action in the same state according to the current policy π . Both $V^\pi(s)$ and $Q(s, a)$ functions ensure that the agent model always adopts an optimal policy to achieve a balance between immediate response to intrusions by taking correct action and maintaining it over successive training episodes.

2.4. Deep Q-networks

The proposed agent is developed using the DQN algorithm, where two neural networks are maintained during training, namely the current network (Q_{cur}) and the target network (Q_{tar}). The current network predicts the Q-values associated with the observed state-action pairs, whereas the target network provides stable target values for updating the current network. The target network parameters are periodically synchronised with the current network, which improves the stability of the learning process and prevents oscillations during training. The agent selects its actions according to the ϵ -greedy strategy, which ensures the agent explores the environment by selecting a random action; it exploits the learned policy by selecting the action corresponding to the maximum predicted Q-value. This decision process is defined as (3).

$$\begin{cases} \text{random_action} & \text{with probability } \epsilon \\ \arg \max_a Q_{current}(S_t, a_t) & \text{with probability } (1 - \epsilon) \end{cases} \quad (3)$$

3. RESULTS AND DISCUSSION

The proposed RL-driven IDS was designed and developed using the Python programming language running in the anaconda distribution installed on Windows 11 with 16 GB RAM and GPU acceleration provided by NVIDIA hardware. Figure 2 shows the performance analysis of the DQN-based IDS on a test dataset consisting of known attack types. The results of the confusion matrix show that the system has a very strong classification ability, as the model correctly classified 55,729 instances as normal activities and 1,118,834 instances as attacks. The model only misclassified 271 normal samples as attacks (false positives) and only 507 attack samples as normal traffic (false negatives).

It can be seen that the classification accuracy is 99.76%, and the attack detection rate (recall rate) is as high as 99.86%. The following analysis of zero-day attacks is shown in Figure 3, which contains the corresponding confusion matrix. The results show that the model correctly identified 55,729 normal cases and 28,195 attack cases. Among them, 271 normal cases were misclassified as attack cases (false positives) by the model, and 480 zero-day attack cases were misclassified as normal cases (false negatives) by the model.

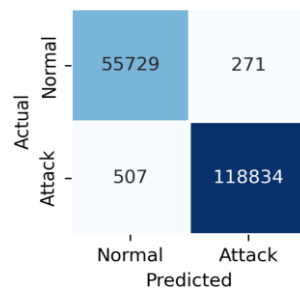


Figure 2. Analysis of confusion matrix for test dataset

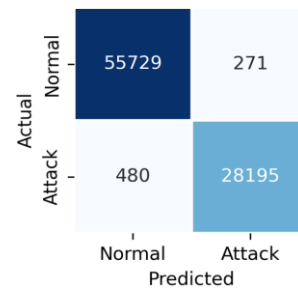


Figure 3. Analysis of confusion matrix for zero-day attacks

The next Figure 4 shows the comparative analysis of accuracy and F1-score on the standard test dataset and the zero-day dataset. It can be seen that the model achieved a high accuracy of 99.56% and an F1-score of 99.56% on the test dataset, showing that it has strong generalization ability for known attack patterns. In the zero-day dataset with unknown attack scenarios, the model achieved an accuracy of 99.11% and an F1-score of 98.71%, which demonstrates its ability to identify new and previously unseen threats. The results confirm the effectiveness of the proposed RL-based agent detection method under both traditional and adversarial network environments. The agent learning performance is shown in Figure 5 in terms of average reward vs. epsilon for 15 training epochs. It can be seen that initially there is a continuous upward trend for reward, which stabilizes after epoch 5, and the epsilon value remains constant throughout the training process, which indicates a good balance between the exploration and exploitation rate of the model. This shows that the agent learned the optimal policy through its interaction with the environment, where the proposed reward mechanism prevents the agent from becoming overconfident too early and allows sufficient exploration to refine its policy.

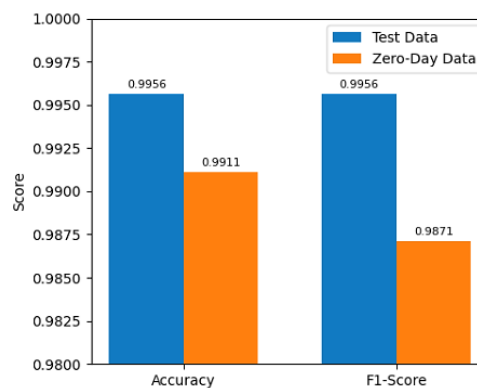


Figure 4. Comparative analysis for both known and zero-day attacks

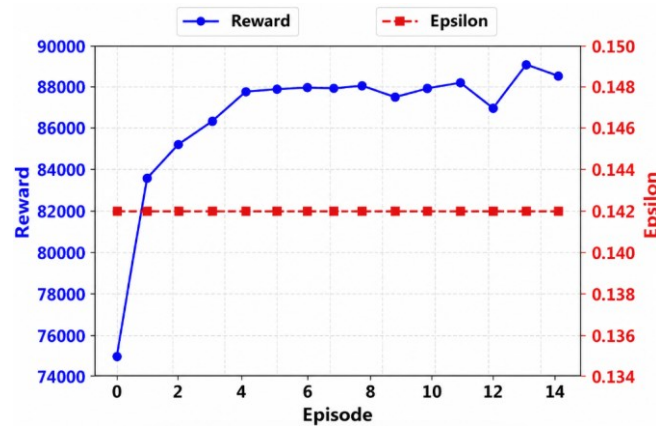


Figure 5. Reward and epsilon analysis during training episodes

Figure 6 shows the receiver operating characteristic (ROC) curves obtained on the test dataset and the zero-day dataset. The model achieved an area under the curve (AUC) of 0.9955 on the test dataset and 0.9892 on the zero-day dataset, which indicates good discrimination between normal and attack traffic. The ROC curve for the test dataset is closer to the upper-left corner, thereby reflecting strong performance on known attacks. A slight reduction in the AUC is observed for the zero-day dataset; however, the ROC curve is still closer to the upper-left corner, indicating that the model has maintained good detection performance for previously unseen attacks.

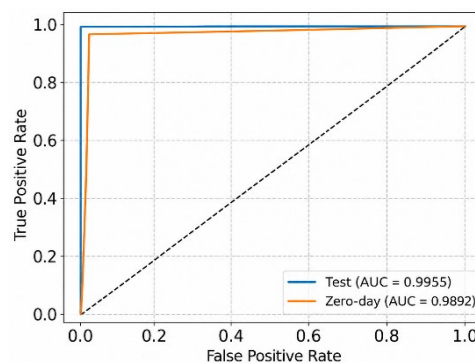


Figure 6. ROC performance on the test and zero-day datasets

Figure 7 exhibits the detection rate (recall) and F1-score at different episodes, i.e., training stages. The analysis reveals that the recall remained high from the initial stage of training, thereby indicating that the agent consistently detects most intrusion samples effectively. Also, the F1-score increases during the early training stages and gradually stabilizes with smaller variations. Further improvement is noted after stage 9, indicating that the model reduces misclassifications as training progresses, which demonstrates that the RL-based approach maintains stable and accurate classification of both benign and malicious network traffic. Table 2 compares the proposed DQN-based IDS with existing methods in terms of accuracy and F1-score. The existing methods such as memory-augmented deep auto-encoder (MemAE) [28], fast persistent contrastive divergence-based deep belief network (FPCD-DBN) [29], and Gaussian mixture model (GMM) [30] reported accuracy values below 90%, whereas the hybrid method Stack ML [31] and LSTM-based IDS [32] improved the detection performance, with accuracy of 96.24% and 90%, respectively. The proposed DQN-based system achieved an accuracy of 99.58% and an F1-score of 99.56%, which are better than the existing works. The performance improvement is due to the fact that the DQN learning strategy, together with the customized reward environment, guided the agent towards better classification decisions. In addition, experience replay, target networks, and state-space clustering further contributed to achieving stable learning and reliable intrusion detection for both known and zero-day attacks.

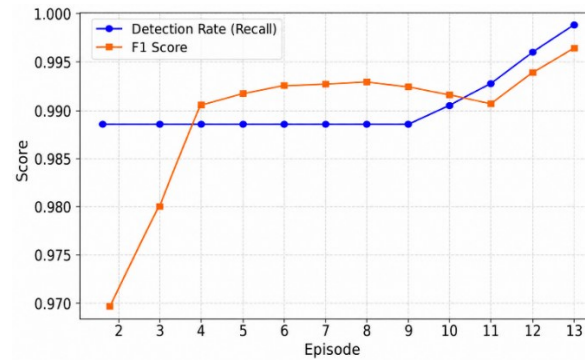


Figure 7. Analysis of detection rate

Table 2. Comparative analysis

Method	Accuracy (%)	F1-score (%)
MemAE [28]	85.30	85.26
FPCD-DBN [29]	86.62	88.28
GMM [30]	89.00	89.00
Stack ML [31]	96.24	-
LSTM [32]	90.00	89.00
Proposed (DQN)	99.58	99.56

4. CONCLUSION

This paper presented a DQN-based IoT security learning framework to achieve adaptive defense strategies from network traffic data. The proposed scheme utilized the UNSW-NB15 dataset, which contains diverse attack categories and realistic traffic patterns. The data pre-processing process consisted of data normalization, feature extraction, and class balancing. The proposed DQN agent interacted with the RL environment, where it learned to classify network traffic as normal or malicious. A reward function was designed as a feedback mechanism that imposes a higher penalty on false negative decisions. The proposed system was evaluated under both standard and zero-day test scenarios. The experimental results showed that the framework maintained high detection accuracy and a low false negative rate for previously unseen attacks. Future work will focus on improving the proposed methodology to handle severe class imbalance in multi-class intrusion scenarios. The framework will be extended towards integrating federated learning to achieve privacy-aware distributed intrusion detection without sharing raw network traffic data. In addition, collaborative learning will be emphasized for distributed edge nodes to support large-scale IoT environments. An adaptive reward will be designed with an online policy updating mechanism that will further improve scalability and detection performance against evolving attack patterns.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sowmya Somanath	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Usha Banavikal Ajay		✓				✓		✓	✓	✓	✓	✓		
Sangeetha Kodlipet	✓		✓		✓	✓	✓	✓	✓	✓	✓			
Nanjundaswamy														

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nvestigation

R : **R**esources

D : **D**ata Curation

O : Writing - **O**riginal Draft

E : Writing - Review & **E**ditng

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [SS], upon reasonable request.

REFERENCES

- [1] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019, doi: 10.1109/ACCESS.2019.2924045.
- [2] N. Neshenko, E. B. -Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2702–2733, 2019, doi: 10.1109/COMST.2019.2910750.
- [3] R. Ahmad and I. Alsmadi, "Machine learning approaches to IoT security: a systematic literature review," *Internet of Things*, vol. 14, Jun. 2021, doi: 10.1016/j.iot.2021.100365.
- [4] E. Schiller, A. Aidoo, J. Fuhrer, J. Stahl, M. Ziörjen, and B. Stiller, "Landscape of IoT security," *Computer Science Review*, vol. 44, May 2022, doi: 10.1016/j.cosrev.2022.100467.
- [5] K. Albulayhi, A. A. Smadi, F. T. Sheldon, and R. K. Abercrombie, "IoT intrusion detection taxonomy, reference architecture, and analyses," *Sensors*, vol. 21, no. 19, Sep. 2021, doi: 10.3390/s21196432.
- [6] K. Albulayhi, Q. A. Al-Hajja, S. A. Alsuhbany, A. A. Jillepalli, M. Ashrafuzzaman, and F. T. Sheldon, "IoT intrusion detection using machine learning with a novel high-performing feature selection method," *Applied Sciences*, vol. 12, no. 10, May 2022, doi: 10.3390/app12105015.
- [7] Y. Yang *et al.*, "A CE-GAN based approach to address data imbalance in network intrusion detection systems," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, doi: 10.1038/s41598-025-90815-5.
- [8] A. Thakkar and R. Lohiya, "A review on machine learning and deep learning perspectives of IDS for IoT: recent updates, security issues, and challenges," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 3211–3243, Jun. 2021, doi: 10.1007/s11831-020-09496-0.
- [9] J. Jose and D. V. Jose, "Deep learning algorithms for intrusion detection systems in internet of things using CIC-IDS 2017 dataset," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, pp. 1134–1141, Feb. 2023, doi: 10.11591/ijece.v13i1.pp1134-1141.
- [10] A. Fatani, A. Dahou, M. A. A. Al-qaness, S. Lu, and M. A. Abd Elaziz, "Advanced feature extraction and selection approach using deep learning and Aquila optimizer for IoT intrusion detection system," *Sensors*, vol. 22, no. 1, Dec. 2021, doi: 10.3390/s22010140.
- [11] O. Yousuf and R. N. Mir, "DDoS attack detection in internet of things using recurrent neural network," *Computers and Electrical Engineering*, vol. 101, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108034.
- [12] A. D. Bhavani and N. Mangla, "A novel approach of intrusion detection system for IoT against modern attacks using deep learning," in *Software Engineering Methods in Systems and Network Systems*, Cham, Switzerland: Springer, 2024, pp. 172–182, doi: 10.1007/978-3-031-53549-9_18.
- [13] A. Jamalipour and S. Murali, "A taxonomy of machine-learning-based intrusion detection systems for the internet of things: a survey," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9444–9466, Jun. 2022, doi: 10.1109/IJOT.2021.3126811.
- [14] X. Zhao, S. Guan, Y. Xue, and H. Pan, "HS-CGK: a hybrid sampling method for imbalance data based on conditional tabular generative adversarial network and K-nearest neighbor algorithm," *Computing and Informatics*, vol. 43, no. 1, pp. 213–239, 2024, doi: 10.31577/cai_2024_1_213.
- [15] M. P. Novaes, L. F. Carvalho, J. Lloret, and M. L. Proenca, "Long short-term memory and fuzzy logic for anomaly detection and mitigation in software-defined network environment," *IEEE Access*, vol. 8, pp. 83765–83781, 2020, doi: 10.1109/ACCESS.2020.2992044.
- [16] Y. Imrana, Y. Xiang, L. Ali, and Z. A. -Rauf, "A bidirectional LSTM deep learning approach for intrusion detection," *Expert Systems with Applications*, vol. 185, Dec. 2021, doi: 10.1016/j.eswa.2021.115524.
- [17] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "CNN-LSTM: hybrid deep neural network for network intrusion detection system," *IEEE Access*, vol. 10, pp. 99837–99849, 2022, doi: 10.1109/ACCESS.2022.3206425.
- [18] J. Chen, X. Qi, L. Chen, F. Chen, and G. Cheng, "Quantum-inspired ant lion optimized hybrid k-means for cluster analysis and intrusion detection," *Knowledge-Based Systems*, vol. 203, Sep. 2020, doi: 10.1016/j.knosys.2020.106167.
- [19] A. T. Siahmarzkooh and M. Alimardani, "A novel anomaly-based intrusion detection system using whale optimization algorithm WOA-based intrusion detection system," *International Journal of Web Research*, vol. 4, no. 2, pp. 8–15, 2021, doi: 10.22133/ijwr.2022.305467.1106.
- [20] A. Oluwasanmi, M. U. Aftab, E. Baagyer, Z. Qin, M. Ahmad, and M. Mazzara, "Attention autoencoder for generative latent representational learning in anomaly detection," *Sensors*, vol. 22, no. 1, Dec. 2021, doi: 10.3390/s22010123.
- [21] T. Zhang, W. Chen, Y. Liu, and L. Wu, "An intrusion detection method based on stacked sparse autoencoder and improved gaussian mixture model," *Computers & Security*, vol. 128, May 2023, doi: 10.1016/j.cose.2023.103144.
- [22] A. D. Bhavani and N. Mangla, "A novel network intrusion detection system based on semi-supervised approach for IoT," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 4, 2023, doi: 10.14569/IJACSA.2023.0140424.
- [23] S. Tharewal, M. W. Ashfaq, S. S. Banu, P. Uma, S. M. Hassen, and M. Shabaz, "Intrusion detection system for industrial internet of things based on deep reinforcement learning," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, Jan. 2022, doi: 10.1155/2022/9023719.
- [24] A. Kumar, S. Dutta, and P. Pranav, "FQBDDA: fuzzy Q-learning based DDoS attack detection algorithm for cloud computing environment," *International Journal of Information Technology*, vol. 16, no. 2, pp. 891–900, Feb. 2024, doi: 10.1007/s41870-023-01509-y.
- [25] S. Suratkar *et al.*, "An adaptive honeypot using Q-learning with severity analyzer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 10, pp. 4865–4876, Oct. 2022, doi: 10.1007/s12652-021-03229-2.

- [26] M. A. Hossain, "Deep Q-learning intrusion detection system (DQ-IDS): A novel reinforcement learning approach for adaptive and self-learning cybersecurity," *ICT Express*, vol. 11, no. 5, pp. 875–880, Oct. 2025, doi: 10.1016/j.ict.2025.05.007.
- [27] Y. Badr, "Enabling intrusion detection systems with dueling double deep Q -learning," *Digital Transformation and Society*, vol. 1, no. 1, pp. 115–141, Aug. 2022, doi: 10.1108/DTS-05-2022-0016.
- [28] B. Min, J. Yoo, S. Kim, D. Shin, and D. Shin, "Network anomaly detection using memory augmented deep autoencoder," *IEEE Access*, vol. 9, pp. 104695–104706, 2021, doi: 10.1109/ACCESS.2021.3100087.
- [29] J. Jeong, S. Park, J. Lim, J. Kang, D. Shin, and D. Shin, "A study on network anomaly detection using fast persistent contrastive divergence," *Symmetry*, vol. 16, no. 9, Sep. 2024, doi: 10.3390/sym16091220.
- [30] E. Anthi, L. Williams, M. Slowinska, G. Theodorakopoulos, and P. Burnap, "A supervised intrusion detection system for smart home IoT devices," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9042–9053, Oct. 2019, doi: 10.1109/JIOT.2019.2926365.
- [31] M. H. Kabir, M. S. Rajib, A. S. M. T. Rahman, M. M. Rahman, and S. K. Dey, "Network intrusion detection using UNSW NB15 dataset: Stacking machine learning based approach," in *2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEES)*, Feb. 2022, pp. 1–6, doi: 10.1109/ICAEES54957.2022.9836404.
- [32] A. Chehri, I. Fofana, and X. Yang, "Security risk modeling in smart grid critical infrastructures in the era of big data and artificial intelligence," *Sustainability*, vol. 13, no. 6, Mar. 2021, doi: 10.3390/su13063196.

BIOGRAPHIES OF AUTHORS



Sowmya Somanath     received the B.E. degree in Computer Science and Engineering from Visvesvaraya Technological University in 2008; M.Tech. degree in Computer Science and Engineering from Visvesvaraya Technological University in 2015; and currently pursuing Ph.D. under Visvesvaraya Technological University in security in IoT using machine learning domain. Currently, she is an assistant professor at the Department of Computer Science and Engineering, REVA University. She is a member of International association of Engineers (IAENG). Her research interests include internet of things, network security, and machine learning. She can be contacted at email: sowmyasadish@gmail.com.



Dr. Usha Banavikal Ajay     is a dedicated motivational individual committed to maximizing learning opportunities in diverse academic settings using consistent and organized practices. Energetic and ambitious professional with 18 years of experience in teaching. She has completed her research in the area of information security. She obtained her Ph.D. from Visvesvaraya Technological University, Belagavi, Karnataka, India in the year 2016. She has published more than 35 research papers in reputed international journal and conferences where some of them are Scopus indexed and also has good impact factors. Currently, she is working as professor in the Department of Information Science and Engineering, BMS Institute of Technology and Management, Bengaluru. Her research interests include IoT, information security, cyber security, data mining, artificial intelligence, and machine learning. She can be contacted at email: ushaajay1@gmail.com.



Dr. Sangeetha Kodlipet Nanjundaswamy     presently working as assistant professor in JSS Academy of Technical Education, Bengaluru. She completed her bachelor's degree, master degree, and doctor of philosophy in Electronics and Communication Engineering from Visveswaraya Technological University, Belagavi. Her area of specialization includes digital communication, image processing, information security, cryptography, and python programming. She has 13 years of teaching experience so far. She is an enthusiastic, committed individual who uses systematic, consistent methods to maximize learning opportunities in a variety of academic environment. She was selected as a reviewer for reputed international conference. She published papers in reputed international journal and conferences. She can be contacted at email: sangeethakn@jssateb.ac.in.