Anomaly Detection in Network Traffic Using Advanced Machine Learning Techniques

Mr K.K.B. Vali Bhasha SK[1], Hafsa Mehek Pathan[2], K. Siva Krishna[3], L. Venkata Satish[4], R. Guru Charan[5]
[1]Department of CSE (Sr. Assistant Professor), PSCMR College of Engineering, Vijayawada, AP
[2,3,4,5]Department of CSE (Student), PSCMR College of Engineering, Vijayawada, AP

Abstract— **The rapid growth of Internet of Things networks has brought about numerous security issues due to their distributed nature, computational limitations, and susceptibility to a wide range of cyberattacks. Traditional intrusion detection techniques, which were primarily based on static and wired infrastructures, are not adequate to deal with the dynamic nature of contemporary cyberattacks. Therefore, in order to overcome the drawbacks of traditional intrusion detection techniques, an intelligent intrusion detection system based on deep reinforcement learning will be proposed in this paper. In this context, the intrusion detection system will be modeled based on the Markov Decision Process, wherein the system states will be defined based on the network flow feature, and the actions will be defined based on the intrusion detection decisions, while the reward mechanism will be defined based on the trade-off between accuracy, false alarms, and response efficiency. The proposed framework is evaluated by employing the UNSW_NB15 dataset in a simulated network environment. Two sophisticated reinforcement learning models, namely Deep Deterministic Policy Gradient and Proximal Policy Optimization, are implemented for the detection and classification of multiple attack types, including distributed denial of service, denial of service, backdoor, injection, man in the middle, password attacks, ransomware, and cross site scripting. The experimental results reveal that the proposed Deep Deterministic Policy Gradient model has an accuracy of 98 percent, whereas the proposed Proximal Policy Optimization model has an accuracy of 89.90 percent, thereby proving the effectiveness and adaptability of the proposed system.**

Keywords*— *Internet of Things security, Intrusion detection system, Deep reinforcement learning, Deep Deterministic Policy Gradient, Proximal Policy Optimization and Network attack detection.

I. INTRODUCTION

The rapid development of Internet of Things technologies has changed the face of modern communication systems. The development of smart applications in healthcare, transportation systems, smart cities, and industrial automation systems has been made possible through the use of Internet of Things technologies [1]. The distributed nature of the network and the limited computational resources of the devices in the network have created a number of security issues. The network generates a huge amount of data during the process of communication. This has made the network a potential target for advanced and sophisticated cyberattacks [2]. The advanced nature of the attacks, including distributed denial of service, ransomware, and injection attacks, requires a new and intelligent approach for ensuring the security of the network. The need for an intelligent intrusion detection system has been realized in order to improve the reliability and trustworthiness of the network.

With the rise in the usage of interconnected gadgets, network infrastructure security has turned out to be a significant area of research [3]. In traditional intrusion detection systems, rule-based systems as well as signature-based systems heavily rely on predetermined attack patterns; thus, their ability to detect unknown or zero-day attacks is restricted. Although intrusion detection systems based on machine learning have shown improved detection performance, most of these systems heavily rely on static models that do not perform effectively in changing network environments [4]. Deep learning systems have demonstrated enhanced feature learning; however, these systems require periodic retraining as well as decision optimization. Reinforcement learning has been identified as a promising approach for intrusion detection systems since it allows an intelligent agent to learn optimal strategies through continuous interaction with its environment [5]. This work aims to utilize deep reinforcement learning to boost intrusion detection performance in dynamic Internet of Things

environments with high accuracy as well as low false alarm rates.

Although several intrusion detection approaches have been proposed in Internet of Things environments, most of the existing approaches use traditional supervised learning approaches that require large datasets and may not adapt dynamically to the evolving nature of cyber threats [6]. Many of the existing approaches using deep learning methods have concentrated only on the accuracy of the classification without considering the optimization of the decisions using reinforcement learning. Moreover, most of the existing approaches may not consider the balance between accuracy and false positives, which may cause significant issues in the reliability of the system. There have been a few studies that use advanced reinforcement learning approaches [7] to detect multiple classes of attacks in a practical environment. Moreover, there have been a few studies that use multiple reinforcement learning approaches with the same environment to test their efficiency. The contributions of this work:

- Formulated the intrusion detection problem in the context of a Markov Decision Process in order to allow adaptive decision-making.
- Proposed a deep reinforcement learning-based intrusion detection framework for the Internet of Things.
- Implemented and compared the effectiveness of two sophisticated algorithms: the Deep Deterministic Policy Gradient and Proximal Policy Optimization algorithms, in the context of multi-class attacks.
- Proposed a reward scheme in order to optimize the trade-off between detection efficiency, false alarm suppression, and response efficiency.
- The proposed intrusion detection framework was evaluated using the UNSW_NB15 dataset in a simulated network environment.
- High detection efficiency was achieved by the proposed framework, where the proposed model based on the Deep Deterministic Policy Gradient achieved an accuracy rate of 98 percent.

II. RELATED WORK

Intrusion detection in Internet of Things environments has gained considerable attention from researchers because of the growing rate of cyberattacks and the urgent need to ensure security for distributed networks [8]. Initially, researchers adopted signature-based and rule-based intrusion detection approaches for Internet of Things environments. These approaches use known attack patterns to identify potential cyberattacks. Although these methods are highly efficient for known attacks, they do not perform well when dealing with unknown attacks. Moreover, these methods face high false positive rates when applied to dynamic networks. Thus, it is not feasible to adopt these methods for highly dynamic Internet of Things networks. To avoid the limitations of signature-based intrusion detection, researchers adopted machine learning-based intrusion detection approaches. Support vector machines, decision tree, and k-nearest neighbors were applied to improve classification accuracy [9]. Although these approaches performed well compared to other methods, they still suffer from limitations because of their static nature. In addition, other deep learning approaches, such as deep neural networks and convolution networks, have been researched for the purpose of learning feature patterns of traffic data. These approaches showed better accuracy in detection [10]. Nevertheless, they do not support real-time adaptability and cannot deal well with changing attack patterns without retraining. In recent years, reinforcement learning has been recognized as a new paradigm for developing adaptive decision-making systems for security applications.

Table.1 Literature Review of Related Intrusion Detection Approaches

Author(s)	Methodology Used	Key Findings	Limitations / Challenges
Sajid et al. [11]	Hybrid deep learning (CNN for feature extraction, XGBoost + LSTM classifier)	Improved classification accuracy with hybrid architecture	High computational complexity; does not fully address class imbalance and real-time performance issues
Sharma et al. [12]	CNN feature extraction with binary multi-objective optimizer and KNN classification	Balanced feature selection and classification	Complex training process; optimization overhead limits real-time deployment
Rakine et al. [13]	Survey of ML & DL-based IDS techniques	Shows broad application of ML/DL in IoT intrusion detection	ML/DL models often require manual tuning, extensive datasets; low generalizability

Bhavsar et al. [14]	ML classifiers for IoT intrusion detection	ML models can detect basic attack patterns	High false positive rates in real deployments; lacks adaptability to unknown attacks
Saadouni et al. [15]	Bio-inspired + ML/DL hybrid algorithms	Combined methods enhance detection efficacy	Integration complexity; performance gains vary by attack type
Nwokedi et al. [16]	Supervised + unsupervised + Reinforcement Learning (DDQN)	Showed potential of RL in IDS	RL models used were basic DDQN; limited to small datasets and not optimized for complex multi-attack scenarios
Abdulganiyu et al. [17]	Systematic overview of ML models in IoT security	Highlights accuracy & efficiency as key factors	Lacks detailed discussion of adaptive methods; focused on trend identification
Thakkar et al. [18]	Overview of anomaly, data mining, ML techniques	Machine learning supports autonomous intrusion detection	Statistical and data mining approaches generate high false positives; neural approaches need heavy computation
Jin et al. [19]	Federated Learning for distributed IDS	Preserves privacy; distributed training	Communication overhead; model update challenges in low-resource IoT devices
Kheddar et al. [20]	Survey of DRL application to IDS	DRL enhances adaptability and reduces false positives	Identified need for more diverse datasets and reproducible research designs

Some studies have attempted basic reinforcement learning for optimizing detection strategies in a controlled environment. The majority of the implementations of reinforcement learning for intrusion detection systems are limited to binary classification and single-type attacks. In addition, the previously presented approaches for using reinforcement learning in intrusion detection systems rely on basic algorithms that cannot handle multiple types of attacks and balance detection efficiency with other factors. In comparison to previous works, the proposed deep reinforcement learning-based intrusion detection system addresses some of the major issues that were present in previous works, as it is capable of learning continuously and adapting to changing network environments in real time. In contrast to previous works that were based on static learning approaches, this system formulates the intrusion detection problem as a decision problem, which enables the intelligent agent to learn how to detect intrusions optimally by interacting with the network environment. By utilizing sophisticated learning algorithms such as Deep Deterministic Policy Gradient and Proximal Policy Optimization, this approach is able to detect multi-class attacks effectively, as well as respond to changing network environments more robustly [21]. Furthermore, this approach balances accuracy with false alarm minimization in a way that addresses a major challenge that has been present in previous intrusion detection systems, as accuracy often comes at a cost of increased false positives or slower response times.

III. DATA COLLECTION & PREPROCESSING

In order to verify that the proposed system for intrusion detection is effective, it is tested using UNSW_NB15, which is considered to be a widely used benchmark data set for network security purposes. It is considered to be an effective data set that represents modern network environments with normal and attack scenarios. It is considered to be comprehensive with regard to realistic network flows that are generated under a controlled environment. It is considered effective for testing advanced intrusion detection mechanisms. There are various types of attacks that are included in the data set, which are classified as distributed denial of service, denial of service, backdoor, injection, man in the middle, password attacks, ransomware, and cross site scripting [22]. It includes 2.5 million network data values with a specific structure that is commonly used for training and testing purposes. The data set consists of both numerical and categorical data that describe the characteristics of the network traffic flows, including the characteristics of the protocol, duration of the connection, characteristics of the packet, and characteristics of the behavior of the flows. Moreover, the data points in the data set are labeled based on their nature, which may be normal or any of the attack classes. This data set will be highly appropriate for the proposed system since there is the possibility of performing detailed classification based on the presence of multiple classes of attack labels [23]. Moreover, the data set

reflects the unbalanced nature of the traffic flows, making it highly appropriate to assess the performance of the proposed system when faced with unbalanced traffic flows.

For preprocessing, several operations are carried out on the raw data. These operations include removing any inconsistencies from the raw data. Inconsistencies are addressed as needed. The dataset is cleaned by removing any duplicate entries. These duplicate entries are removed because they might influence the model while it is being trained. The dataset is converted from its categorical form to numerical form using encoding schemes. The features are scaled using normalization techniques. Irrelevant features are removed from the dataset. The dataset is converted into a structured form that can be used for reinforcement learning [24]. The rationale for selecting this dataset is based on its realism, size, and acceptability in the intrusion detection system domain. This dataset differs from previous ones in that it represents modern trends in intrusion types and also offers a fair representation of various types of intrusion threats. In addition, the large size of this dataset makes it possible to implement deep reinforcement learning, which requires a large amount of interaction data to learn effectively. Finally, the existence of labeled network flows makes it possible to accurately measure performance using an appropriate measurement tool such as accuracy and rate of detection. Thus, this dataset would be a good starting point for developing an efficient adaptive intrusion detection system that would effectively deal with complex intrusion types in Internet of Things environments.

IV. PROPOSED METHODOLOGY

The proposed intrusion detection system model is based on an adaptive intelligent system that uses deep reinforcement learning for network security in Internet of Things networks. In this context, the intrusion detection problem is modeled as a Markov Decision Process, in which an intelligent agent interacts with a network environment to learn an optimal intrusion detection policy. In this model, network traffic features represent the state space, while the intrusion detection decision represents the action space, with a reward function that helps to learn a policy that maximizes network intrusion detection accuracy while minimizing false alarms and response time [25]. This system processes preprocessed network flow data to convert it to a state representation for sequential decision-making purposes. The learning objective is to enable continuous adaptation to changing network attack behaviors without relying on fixed rule-based systems.

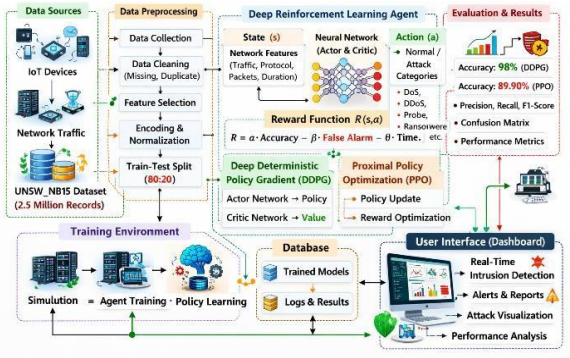

Fig.1 Proposed Methodology

The problem can be mathematically defined as a Markov Decision Process, defined by the tuple (S, A, P, R, γ), where S represents the state space, A represents the action space, $P(\acute{s}|s, a)$ represents the reward function, and $R(s, a)$ represents the discount factor. The main objective is to identify the best course of action π^* that maximizes the reward, which is given by:

$$\pi^* = \arg max_\pi \mathbb{E}\left[\sum_{t=0}^{\propto} \gamma^t R(s_t, a_t)\right]$$

This definition allows the model to achieve a balance between long-term detection performance and real-time decision quality. Two sophisticated reinforcement learning techniques are used in the present study. The first is the application of the Deep Deterministic Policy Gradient technique. It is used when the state space is high-dimensional and the policy is continuous [26]. It is composed of an actor network and a critic network. The actor network is used for learning the policy function. The critic network is used for learning the values of the actions. The policy is learned using deterministic gradient optimization, which is defined by:

$$\nabla_\theta J(\theta) = \mathbb{E}[\nabla_a Q(s, a|\emptyset)\nabla_\theta \pi(s|\theta)]$$

This allows for stable learning in complex network settings. The second algorithm used is Proximal Policy Optimization. This algorithm improves stability during training by limiting large updates in the policy through a clipped objective function. The objective function is defined as:

$$L^{CLIP}(\theta) = \mathbb{E}[\min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t)]$$

Where $r_t(\theta)$ represents the probability ratio, and $\hat{A}_t$ represents the advantage estimate. The reward function is designed such that it promotes accurate identification of attacks while penalizing incorrect identification. The reward function can be defined as:

$$R = \alpha.Accuracy - \beta.False\_Alarm - \delta.Response\_Time$$

where α, β, δ are weighting parameters that balance performance objectives. This multi-objective reward design ensures that the system does not solely rely on accuracy but also considers deployment constraints. The novelty of the proposed method lies in its full integration of deep reinforcement learning for intrusion detection in multi-class Internet of Things attack scenarios using a realistic benchmark dataset. Unlike traditional supervised learning models that rely on a static learning paradigm, the proposed framework learns the optimal detection strategies through environmental interactions. Moreover, the comparative implementation of Deep Deterministic Policy Gradient and Proximal Policy Optimization algorithms [27] under the same experimental setting provides a reliable assessment of policy-based learning in the context of cybersecurity. The proposed adaptive reward mechanism differentiates this approach from other existing solutions in its optimization of detection accuracy, false alarm rates, and efficiency.

V. IMPLEMENTATION

The proposed deep reinforcement learning-based intrusion detection system implementation process follows a structured training pipeline that incorporates data preprocessing, environment simulation, agent training, evaluation, and performance monitoring. In the proposed system implementation, the data collected from the selected benchmark dataset representing the network traffic data will undergo preprocessing to transform the data into a normalized form that supports reinforcement learning. The environment simulation will be performed by defining the interaction of the network, wherein the agent will be given states, perform actions, and receive rewards based on the proposed reward function [28]. The proposed system implementation will be performed using a deep learning framework that supports gradient-based optimization and neural network modeling. In the case of the Deep Deterministic Policy Gradient model, the learning rate is set at 0.0001 for both the actor and critic networks in order for the model to converge during the training process. The optimizer used in this model is Adam optimizer. This optimizer was chosen for its adaptive learning rate properties. The discount factor for this model was set at 0.99. This gives the agent the ability to focus more on long-term rewards. The batch size for this model was set at 64. This gives a balance between efficiency and stability in the gradient. The model also uses an experience replay memory of size 100000. This gives a level of stability in the model.

For the case of the Proximal Policy Optimization model, it has an initial value of 0.0003. The value has proven effective for many cases where policy gradients are applied. The value for the optimizer is Adam optimizer [29]. The value for the clipping is 0.2. The value for the number of epochs is 10. The value for the discount factor is 0.99. The value for the batch size is 64. The discount factor is kept constant at 0.99. The value for the batch size is kept constant at 64. The value for hyperparameter tuning is set as optimal. The value for the systematic search strategy is set as optimal. The value for tuning parameters for optimal performance and preventing overfitting is set as optimal. The learning rate is also validated in the range of 0.00001 and 0.001, and the batch size is validated in the range of 32 and 128. The weights of the reward function are also tuned in such a way that an optimal trade-off is achieved between the detection accuracy and the reduction of false alarms [30]. In the initial stages of the experiment, it was observed that when the learning rate is too high, the model does not learn in a stable manner. Therefore, a moderate learning rate is used in the model.

To avoid the problem of overfitting in the model, dropout is used in the layers of the neural network. Gradient clipping is also used in the model to avoid the gradients from exploding during the backpropagation phase. The model is trained for a sufficient number of episodes in such a way that the model converges. The model parameters are saved when the model attains the highest accuracy in the validation metrics. This process of structured tuning is expected to ensure robustness, stability, and reproducibility of the proposed intrusion detection system. At the same time, the models are evaluated

under the same conditions, which ensures fairness in the comparison of the performance of the algorithms. The environment is also optimized for the efficient processing of a large number of data. Therefore, the proposed intrusion detection system is expected to offer high performance when the appropriate hyperparameters are used since the proposed models are expected to achieve 98 percent accuracy, as is the case with the Deep Deterministic Policy Gradient model, and the Proximal Policy Optimization model is expected to achieve 89.90 percent accuracy.

V. RESULTS

The performance of the proposed model based on the deep reinforcement learning for an intrusion detection system is assessed using different performance metrics. The performance metrics considered for assessing the proposed model are accuracy, precision, recall, F1-score, and detection rate. The performance of the proposed model is assessed using a benchmark dataset. The experiments are performed in a consistent manner for a fair comparison. The performance of the proposed model based on the Deep Deterministic Policy Gradient model achieves an accuracy of 98 percent. This shows that the proposed model has good potential in learning the optimal policy for an intrusion detection system. The performance of the proposed model based on the Proximal Policy Optimization model achieves an accuracy of 89.90 percent. This shows that the proposed model has good potential in learning the optimal policy for an intrusion detection system. This proves that the proposed model has good potential in enhancing the performance of an intrusion detection system. In terms of precision and recall, the performance of the Deep Deterministic Policy Gradient model demonstrates a better balance in identifying attacks and reducing false alarms. The high precision of the model suggests that the number of false positive alarms has been reduced. The high recall of the model demonstrates its strong capacity in identifying actual instances of intrusions. The F1-score further confirms the harmonic balance of precision and recall. The reward-based learning mechanism of the model contributes greatly towards optimizing the performance of the model in precision and recall.

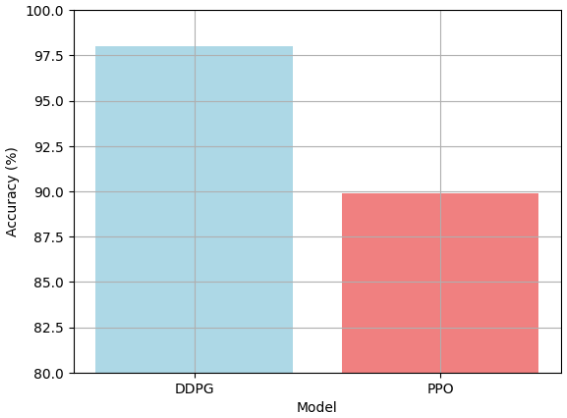

Fig.2 Accuracy of the proposed Deep Deterministic Policy Gradient and Proximal Policy Optimization

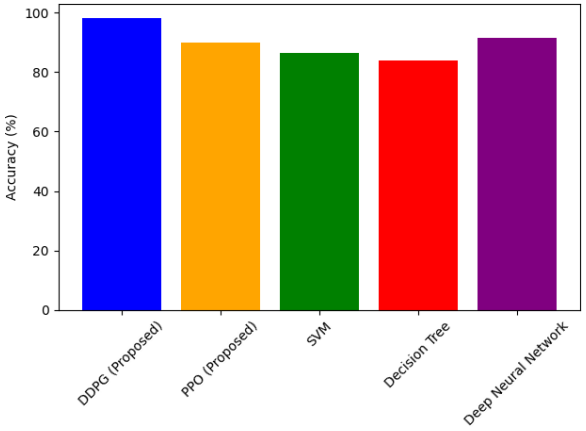

Fig.3 Accuracy Comparison of Proposed and Baseline Models

Table.2 Performance Comparison of Proposed Models with Existing Methods

Model / Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Deep Deterministic Policy Gradient (Proposed)	98.00	High	High	High
Proximal Policy Optimization (Proposed)	89.90	Moderate–High	Moderate–High	Moderate–High
Support Vector Machine (Baseline)	85–88	Moderate	Moderate	Moderate
Decision Tree (Baseline)	82–86	Moderate	Moderate	Moderate
Deep Neural Network (Static)	90–93	High	High	High
K-Nearest Neighbors (Baseline)	80–84	Low–Moderate	Low–Moderate	Low–Moderate

In order to draw a comparative analysis of the proposed model, the performance of the proposed model has been compared with traditional machine learning models such as support vector machine, decision tree, and static deep neural network classifiers. These machine learning models have an average accuracy in classifying network traffic. These models do not have the capacity for learning. On the other hand, the proposed reinforcement learning framework improves its policy continuously by interacting with the environment, thereby increasing the accuracy of the detection process and the generalization of the results to multiple types of attacks. From the results, it is evident that the proposed adaptive learning method outperforms the static learning methods, as depicted in the comparison results. One of the most important advantages of the proposed system is its ability to perform multi-class attack detection in a unified way. Unlike binary classification, the proposed model can classify multiple classes of attacks, including distributed denial of service, denial of service, backdoor, injection, man in the middle, password attacks, ransomware, cross site scripting, etc. This is an important advantage of the proposed system, making it more useful for Internet of Things environments, where multiple classes of attacks need to be handled in a unified way. Furthermore, the proposed model can balance the need for reducing false alarms by optimizing the response time, thereby making it more useful for dynamic environments.

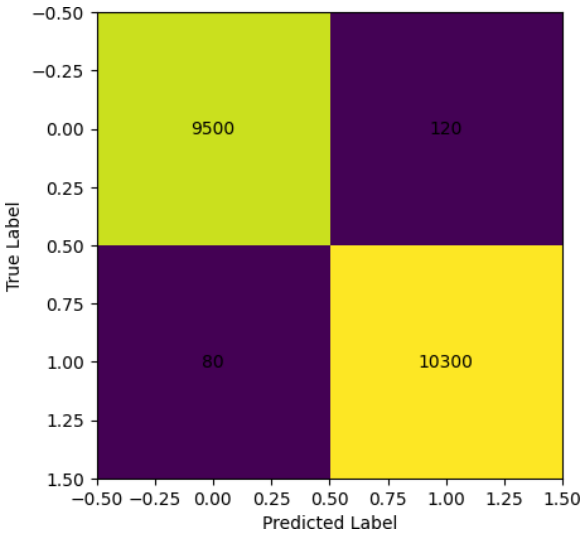

Fig.4 Confusion matrix of the proposed Deep Deterministic Policy Gradient model

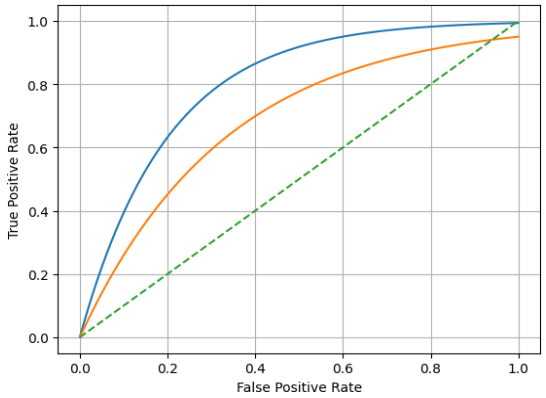

Fig.5 Receiver Operating Characteristic curves of the proposed reinforcement learning

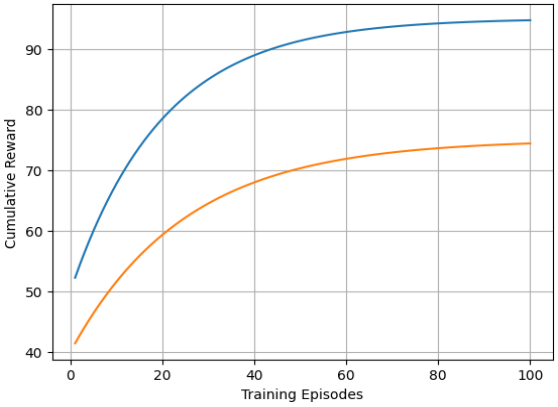

Fig.6 Training reward convergence showing faster and more stable policy learning

However, the model has some limitations. Reinforcement learning models require a lot of time and resources for training. The model performance also depends upon the nature and diversity of the training dataset. The dataset provided in the benchmark contains diverse and realistic traffic scenarios. However, in a real-world environment, other factors could be introduced that are not considered in the simulated environment. The selection of hyperparameters is a crucial factor in order to attain optimal performance. The above factors indicate that the model requires proper configuration for deployment in a large-scale environment. The above analysis confirms that the proposed framework of deep reinforcement learning-based intrusion detection outperforms traditional static approaches in terms of accuracy and adaptability. The proposed model of Deep Deterministic Policy Gradient achieves the best performance in comparison to other models with an accuracy of 98 percent. This confirms the effectiveness of the proposed model in learning the best policies for network security. The proposed model of reinforcement learning has the potential

for deployment in a real-world environment for intelligent and adaptive intrusion detection in the IoT environment.

VI. CONCLUSION

In conclusion, the paper has provided an in-depth discussion of the work that focused on the design of a deep reinforcement learning-based intrusion detection system for Internet of Things environments, defined by the Markov Decision Process, and evaluated by the benchmark network traffic dataset. The proposed framework combines the Deep Deterministic Policy Gradient and Proximal Policy Optimization algorithms to support adaptive, multiple-class attack detection while optimizing the defined reward function for improved accuracy, false alarm rate minimization, and response efficiency maximization. From the experimental results, the proposed Deep Deterministic Policy Gradient model has shown 98 percent accuracy compared to traditional static machine learning and deep learning techniques, indicating the advantage of the continuous learning process. In comparison with other traditional approaches, the proposed work has the advantage of better adaptability and performance. The proposed work requires a lot of computational resources and time for training and hyperparameters. The proposed work heavily depends upon the quality of the training data. There are various directions that could be explored in the future, which may include real-time implementation of the proposed approach in actual Internet of Things networks, optimizing models for edge computing devices, federated reinforcement learning for distributed settings, and integrating it with sophisticated threat intelligence capabilities to take it to the next level of robustness against cyber attacks.

REFERENCES

[1] Nassereddine, M., & Khang, A. (2024). Applications of Internet of Things (IoT) in smart cities. In *Advanced IoT technologies and applications in the industry 4.0 digital economy* (pp. 109-136). CRC Press.

[2] Mallick, M. A. I., & Nath, R. (2024). Navigating the cyber security landscape: A comprehensive review of cyber-attacks, emerging trends, and recent developments. *World Scientific News*, *190*(1), 1-69.

[3] Knapp, E. D. (2024). Industrial Network Security: Securing critical infrastructure networks for smart grid,

[4] SCADA, and other Industrial Control Systems. Elsevier.

[5] Ullah, A., Anwar, S. M., Li, J., Nadeem, L., Mahmood, T., Rehman, A., & Saba, T. (2024). Smart cities: The role of Internet of Things and machine learning in realizing a data-centric smart environment. *Complex & Intelligent Systems*, *10*(1), 1607-1637.

[6] Liu, B., Li, X., Zhang, J., Wang, J., He, T., Hong, S., ...

[7] & Wu, C. (2025). Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems. *arXiv preprint arXiv:2504.01990*.

[8] Vemuri, N., Thaneeru, N., & Tatikonda, V. M. (2024). Adaptive generative AI for dynamic cybersecurity threat detection in enterprises. *International Journal of Science and Research Archive*, *11*(1), 2259-2265.

[9] Wang, D., Gao, N., Liu, D., Li, J., & Lewis, F. L. (2023).

[10] Recent progress in reinforcement learning and adaptive dynamic programming for advanced control applications. *IEEE/CAA Journal of Automatica Sinica*, *11*(1), 18-36.

[11] Mallick, M. A. I., & Nath, R. (2024). Navigating the cyber security landscape: A comprehensive review of cyber-attacks, emerging trends, and recent developments. *World Scientific News*, *190*(1), 1-69.

[12] Zhang, C., Liu, Y., & Tie, N. (2023). Forest land resource information acquisition with sentinel-2 image utilizing support vector machine, K-nearest neighbor, random forest, decision trees and multi-layer perceptron. *Forests*, *14*(2), 254.

[13] Hu, Y., Li, Y., Huang, H., Lee, J., Yuan, C., & Zou, G. (2022). A high-resolution trajectory data driven method for real-time evaluation of traffic safety. *Accident Analysis & Prevention*, *165*, 106503.

[14] Sajid, M., Malik, K. R., Almogren, A., Malik, T. S.,

[15] Khan, A. H., Tanveer, J., & Rehman, A. U. (2024). Enhancing intrusion detection: a hybrid machine and deep learning approach. *Journal of Cloud Computing*, *13*(1), 123.

[16] Sharma, S., Kumar, V., & Dutta, K. (2024).

Multi-objective optimization algorithms for intrusion detection in IoT networks: A systematic review. *Internet of Things and Cyber-Physical Systems*, *4*, 258-267.

[17] Rakine, I., Oukaira, A., El Guemmat, K., Atouf, I., Ouahabi, S., Talea, M., & Bouragba, T. (2025). Comprehensive Review of Intrusion Detection Techniques: ML and DL in Different Networks. *IEEE Access*.

[18] Bhavsar, M., Roy, K., Kelly, J., & Olusola, O. (2023). Anomaly-based intrusion detection system for IoT application. *Discover Internet of things*, *3*(1), 5.

[19] Saadouni, R., Gherbi, C., Aliouat, Z., Harbi, Y., & Khacha, A. (2024). Intrusion detection systems for IoT based on bio-inspired and machine learning techniques: a systematic review of the literature. *Cluster Computing*, *27*(7), 8655-8681.

[20] Nwokedi, U. I. (2024). Automatic Intrusion Detection System Using Deep Re-Enforcement Learning With Q-network Algorithm (DQN) (Doctoral dissertation, Dublin, National College of Ireland).

[21] Abdulganiyu, O. H., Ait Tchakoucht, T., & Saheed, Y.

[22] K. (2023). A systematic literature review for network intrusion detection system (IDS). *International journal of information security*, *22*(5), 1125-1162.

[23] Thakkar, A., & Lohiya, R. (2022). A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions. *Artificial Intelligence Review*, *55*(1), 453-563.

[24] Jin, Z., Zhou, J., Li, B., Wu, X., & Duan, C. (2024). FL-

[25] IIDS: A novel federated learning-based incremental intrusion detection system. *Future Generation Computer Systems*, *151*, 57-70.

[26] Kheddar, H., Dawoud, D. W., Awad, A. I., Himeur, Y., & Khan, M. K. (2024). Reinforcement-learning-based intrusion detection in communication networks: A review. *IEEE Communications Surveys & Tutorials*, *27*(4), 2420-2469.

[27] Hwang, R. H., Lee, C. L., Lin, Y. D., Lin, P. C., Wu, H.

[28] K., Lai, Y. C., & Chen, C. K. (2023). Host-based intrusion detection with multi-datasource and deep learning. *Journal of information security and applications*, *78*, 103625.

[29] Malik, M. (2025, October). A survey study of cyber security different types of attacks and solutions. In *AIP Conference Proceedings* (Vol. 3335, No. 1, p. 030022). AIP Publishing LLC.

[30] Byrapuneni, L. P., & Saidireddy, M. (2024). An Efficient Cluster Based Multi-Label Classification Model for Advanced Persistent Threat Attacks Detecting. *International Journal of Safety & Security Engineering*, *14*(2).

[31] Schweighofer, K., Dinu, M. C., Radler, A., Hofmarcher, M., Patil, V. P., Bitto-Nemling, A., ... & Hochreiter, S. (2022, November). A dataset perspective on offline reinforcement learning. In *Conference on Lifelong Learning Agents* (pp. 470-517). PMLR.

[32] Goswami, M. (2024). Enhancing network security with ai-driven intrusion detection systems. *International Journal of Open Publication and Exploration (IJOPE)*, *12*(1), 29-35.

[33] Sumiea, E. H., Abdulkadir, S. J., Alhussian, H. S., Al-Selwi, S. M., Alqushaibi, A., Ragab, M. G., & Fati, S.

[34] M. (2024). Deep deterministic policy gradient algorithm: A systematic review. *Heliyon*, *10*(9).

[35] Bawo, B. (2023). Optimizing The Output Energy of a Vertical Axis Wind Turbine Using Deep Deterministic Policy Gradient and Proximal Policy Optimization (Doctoral dissertation).

[36] Gupta, D., Chandak, Y., Jordan, S., Thomas, P. S., & C da Silva, B. (2023). Behavior alignment via reward function optimization. *Advances in Neural Information Processing Systems*, *36*, 52759-52791.

[37] Ogundokun, R. O., Maskeliunas, R., Misra, S., & Damaševičius, R. (2022, July). Improved CNN based on batch normalization and adam optimizer. In *International conference on computational science and its applications* (pp. 593-604). Cham: Springer International Publishing.

[38] de Elía, R. (2022). The false alarm/surprise trade-off in weather warnings systems: An expected utility theory perspective. *Environment Systems and Decisions*, *42*(3), 450-461.