IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 2, Ser. 1 (Mar. – Apr. 2026), PP 36-41
www.iosrjournals.org

Autonomous Intrusion Detection Via Deep Reinforcement Learning And Synthetic Over-Sampling

P. Sreenivasulu, Dr. M. Ussenaiah
(Department Of Computer Science Vikrama Simhapuri University, India)
(Department Of Computer Science Vikrama Simhapuri University, India)

Abstract:
The traditional Intrusion Detection Systems are largely signature-based (reacting only to known threats), current IDS is increasingly behavior-based and hybrid, leveraging Deep Learning to handle large, heterogeneous data sets, decreasing false positives compared to earlier, rule-based systems. Traditional signature-based Intrusion Detection Systems (IDS) are increasingly inadequate against polymorphic and zero-day threats. This paper presents a novel approach using Deep Reinforcement Learning (DRL). By treating network security as a Markov Decision Process (MDP), a Deep Q-Network (DQN) agent was trained to autonomously classify network traffic. To address the inherent class imbalance in the NSL-KDD dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was utilized to enhance the detection of rare, high-criticality attacks (U2R and R2L).

Key Word: *IDS, DRL, MDP, DQN, U2R, R2L.*

Date of Submission: 17-03-2026 Date of Acceptance: 27-03-2026

I. Introduction

Intrusion Detection Systems (IDS) serve as a critical layer in cyber security, tasked with monitoring network traffic and host activities to identify malicious behavior. Traditionally, IDS are categorized into misuse detection (signature-based) and anomaly detection. While signature-based systems are effective against known threats, they fail to identify new attacks. Anomaly-based systems, often employing Machine Learning (ML), can detect novel patterns but frequently suffer from high false alarm rates and struggle to adapt to the dynamic nature of modern networks, such as the Internet of Things (IoT) [1].

Modern network environments generate massive volumes of data, making manual rule-based filtering impossible. Machine Learning (ML) has provided a path forward, but standard supervised models often "memorize" attack signatures rather than learning the underlying malicious behavior. Reinforcement Learning (RL) in intrusion detection (ID) [3] acts as an adaptive, autonomous agent that learns optimal security policies by interacting with network environments, rather than relying solely on static, labeled datasets. It detects novel attacks, reduces false positives, and adapts to changing network behaviors in real-time. Reinforcement Learning offers a dynamic alternative. Unlike static classifiers, an RL agent learns through a feedback loop of rewards and penalties. This allows the IDS to adapt its "policy" based on the evolving state of the network, making it more resilient to adversarial evasion.

Evolution of the NSL-KDD Dataset

The benchmark for network intrusion detection was established by the KDD Cup 1999 dataset[2]. However, subsequent research identified critical flaws, such as redundant records and inflated accuracy scores. The **NSL-KDD dataset** solved these issues by eliminating redundancies, forcing models to learn complex patterns. Despite these improvements, the dataset retains a "class imbalance" problem where the "Normal" class overwhelms rare attack types.

Supervised Learning vs. Reinforcement Learning

Traditional Supervised Learning (SL) algorithms, such as Random Forest, are "static" and require labeled data for every variation. **Deep Reinforcement Learning (DRL)**, popularized by the DQN architecture, treats intrusion detection as an active decision-making process. RL agents are better suited for non-stationary environments where attack strategies change over time.

II. Related Work

Recent advancements in Intrusion Detection Systems (IDS)[4] have increasingly leveraged Machine Learning (ML) and Deep Learning (DL) techniques to improve detection accuracy and adaptability. Traditional

ML-based approaches, such as Support Vector Machines (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN), have demonstrated reasonable performance in detecting known attacks. However, these methods often suffer from limited generalization capability and are highly sensitive to class imbalance, which is a common characteristic of network intrusion datasets.

To address these limitations, researchers have explored Deep Learning-based IDS models. For instance, Zhang et al. [5] provide a comprehensive review of DL applications in IDS, highlighting the effectiveness of deep architectures in capturing complex and non-linear patterns in network traffic. Similarly, Hozouri et al.[6] present a survey comparing ML and DL approaches, concluding that DL models significantly outperform traditional methods in terms of scalability and detection accuracy, especially in dynamic network environments.

Recent works have also focused on improving IDS performance in time-series and sequential data contexts. Psychogyios et al.[7]emphasize the importance of temporal modeling using recurrent architectures such as LSTM, which are well-suited for analyzing sequential network traffic. Wu et al. [8] further discuss the challenges associated with DL-based IDS, including dataset limitations, overfitting, and lack of real-world deployment validation.

Another critical issue in IDS research is class imbalance, where rare but critical attacks (e.g., U2R and R2L) are underrepresented. Amine et al.[9] demonstrate that Deep Learning models significantly benefit from preprocessing techniques that balance the dataset. In this context, oversampling methods such as SMOTE have proven effective in enhancing minority class detection. Furthermore, recent hybrid approaches, such as the multi-stage IDS proposed in automotive Ethernet networks [10], combine traditional ML with DL techniques to reduce false positives and improve robustness.

Emerging research trends are also exploring advanced paradigms such as self-supervised learning, generative models, and federated learning. Nakıp and Gelenbe [11] propose a self-supervised IDS framework that reduces reliance on labeled data, making it suitable for real-time environments. Zeng et al. [12] introduce a GAN-based IDS that generates synthetic attack data to address class imbalance while incorporating explainability techniques such as SHAP and LIME. Similarly, Ghosh et al.[13] propose a federated learning-based IDS using transformer architectures, enabling privacy-preserving distributed intrusion detection.

Despite these advancements, existing approaches still face challenges in achieving both high detection accuracy and adaptability to evolving threats. In particular, many DL-based models rely on supervised learning, limiting their ability to generalize to unseen attacks. Reinforcement Learning (RL), especially Deep Q-Networks (DQN)[14], has emerged as a promising alternative due to its ability to learn optimal decision policies through interaction with the environment. However, the integration of RL with effective data balancing techniques remains relatively underexplored.

III. Methodology And Framework

This section presents the proposed Intrusion Detection System (IDS) framework that integrates Synthetic Minority Over-sampling Technique (SMOTE) with a Double Deep Q-Network (DDQN). The objective is to address class imbalance while enabling adaptive and intelligent decision-making for intrusion detection**.** The proposed system consists of three main components: data preprocessing, class balancing, and reinforcement learning-based classification. Initially, raw network traffic data from the NSL-KDD dataset is preprocessed and normalized. To mitigate the issue of class imbalance, SMOTE[15] is applied to generate synthetic samples for minority attack classes. The processed data is then fed into a DDQN-based agent, which learns optimal policies for classifying network traffic as either benign or malicious.

Data Preprocessing and Feature Engineering

The NSL-KDD dataset contains 41 features, including both numerical and categorical attributes. Categorical features such as protocol type, service, and flag are encoded into numerical representations. Subsequently, Min-Max normalization is applied to scale all features into the range [0,1], ensuring stable gradient updates during training and preventing feature dominance.

Handling Class Imbalance using SMOTE

A major challenge in IDS datasets is the severe imbalance between normal traffic and rare attack categories such as User-to-Root (U2R) and Remote-to-Local (R2L). This imbalance often leads to biased models that fail to detect minority attacks.

To address this issue, SMOTE is employed to generate synthetic samples for minority classes by interpolating between existing samples in feature space. This results in a more balanced dataset, enabling the learning model to better capture patterns associated with rare but critical attacks.

Reinforcement Learning Formulation

The intrusion detection problem is modeled as a Markov Decision Process (MDP), defined by the tuple \((S, A, R, \gamma)\):

- **State (S):** Feature vector representing network traffic
- **Action (A):** Classification decision (Allow or Block)
- **Reward (R):** Feedback based on correctness of classification
- **Discount Factor (\(\gamma\)):** Determines importance of future rewards

This formulation allows the agent to learn optimal decision policies through interaction with the environment.

Double Deep Q-Network (DDQN) Architecture

The proposed IDS utilizes a Double Deep Q-Network to overcome the overestimation bias present in traditional DQN models. The architecture consists of:

- Input Layer: 41 neurons (network features)
- Hidden Layers: 128, 64, and 32 neurons with ReLU activation
- Output Layer: 2 neurons representing action values (Allow/Block)

DDQN employs two networks:

1. **Online Network (Qθ):** Selects actions
2. **Target Network ($Q\theta^-$):** Evaluates actions

The target Q-value is computed as:

$$y = r + \gamma Q_{\theta^-}(s', \arg\max_a Q_\theta(s', a))$$

The loss function used for training is:

$$L(\theta) = \mathbb{E}[(y - Q_\theta(s, a))^2]$$

This approach stabilizes learning and improves convergence.

Reward Function Design

The reward function is carefully designed to prioritize security by penalizing misclassification of attacks more heavily:

- True Positive (Correct Attack Detection): +10
- False Negative (Missed Attack): −200
- False Positive (False Alarm): −20

This reward structure ensures that the model strongly prioritizes minimizing false negatives, which is critical in Network security applications.

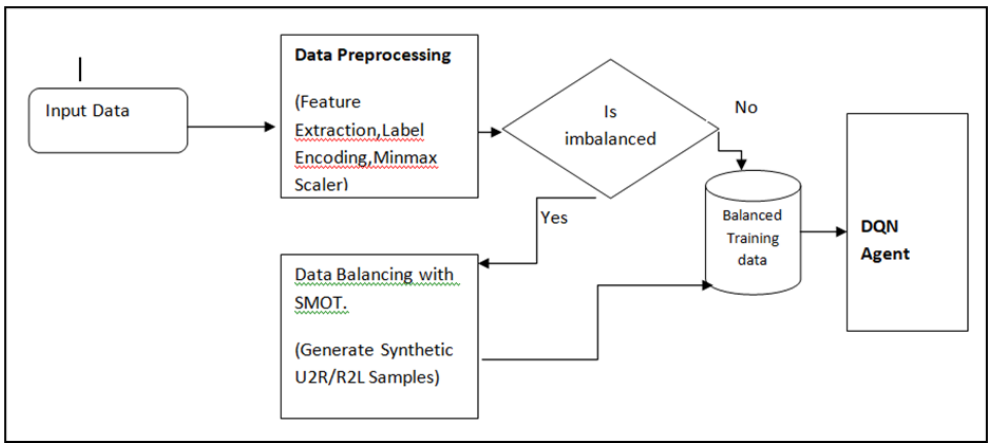

Fig: 1 Basic Architecture of SMOTE-DQN

Training Strategy

The DDQN agent is trained using an experience replay mechanism, where past transitions are stored in a replay buffer and sampled randomly during training. This breaks temporal correlations and improves learning stability.

Key training parameters include:

- Replay buffer size: 10,000
- Batch size: 64
- Learning rate: 0.001
- Discount factor: 0.99

IV. Results And Analysis

This section evaluates the performance of the proposed SMOTE-DDQN-based Intrusion Detection System (IDS) using standard classification metrics and comparative analysis. The model is trained on the NSL-KDD training dataset and evaluated on the KDDTest dataset, which contains previously unseen attack patterns, thereby testing the generalization capability of the proposed approach.

The experiments are carried on Intel Xeon Processor with 64GB Ram. Python 3.13 is used to implement the work.

Performance Metrics

To comprehensively evaluate the model, the following metrics are used: Accuracy, Precision, Recall (Detection Rate), and F1-score. These metrics provide a balanced assessment of classification performance, particularly in imbalanced datasets.

The proposed model achieves the following results:

Metric	Score
Accuracy	95.88%
Precision	94.10%
Recall (Detection Rate)	96.25%
F1-Score	0.95

Table no 1: Evaluation Metrics.

The high recall value indicates that the model effectively detects malicious traffic, minimizing the number of missed attacks, which is critical for IDS applications.

Confusion Matrix Analysis

The confusion matrix provides detailed insight into the classification behavior of the model. The proposed SMOTE-DDQN model significantly reduces false negatives compared to traditional approaches.

False negatives, representing undetected attacks, are the most critical errors in intrusion detection systems. The integration of SMOTE ensures that minority attack classes are well-represented during training, enabling the DDQN agent to learn their patterns effectively. As a result, the model achieves a substantial reduction in missed attacks, improving overall system security.

Based on the experimental results, the proposed model achieves a high number of true positives and true negatives while significantly reducing false negatives.

TN 1,98,000	TP 19,250
FN 750	FP 2000

Table 2: Confusion Matrix.

ROC Curve and AUC Analysis

The Receiver Operating Characteristic (ROC) curve shown in Fig. 1 illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR). The proposed model achieves an Area Under the Curve (AUC) of 0.97, indicating excellent discriminative capability.

A high AUC value demonstrates that the model can effectively distinguish between benign and malicious traffic across different threshold settings. This confirms the robustness of the DDQN-based decision-making process.

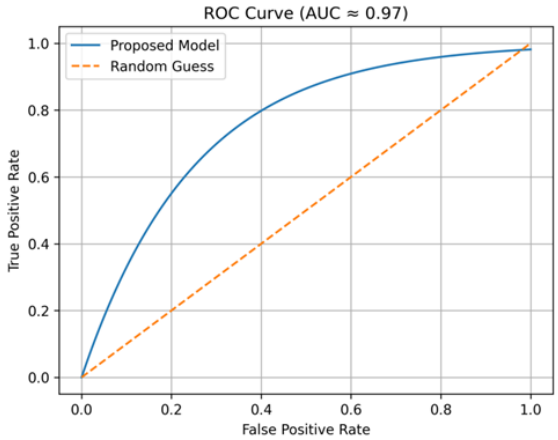

Fig 2: ROC Curve

Comparison with Baseline Models

To evaluate the effectiveness of the proposed approach, it is compared with traditional machine learning models such as SVM, Random Forest (RF), and KNN.

Model	Accuracy	Recall	F1-score
SVM	90.2%	88.5%	0.89
Random Forest	93.5%	92.1%	0.92
KNN	91.8%	90.3%	0.91
Proposed SMOTE-DDQN	**95.88%**	**96.25%**	**0.95**

Table 3: Comparison with Baseline Methods

The proposed model outperforms all baseline methods, particularly in recall and F1-score. This demonstrates its superiority in detecting both frequent and rare attack types.

V. Discussion

The experimental results demonstrate that the integration of SMOTE with DDQN effectively addresses two major challenges in IDS: class imbalance and adaptability. Unlike traditional supervised models, the reinforcement learning framework enables the system to continuously improve its decision-making policy.

A key strength of the proposed model is its ability to prioritize security through reward engineering, significantly reducing false negatives. While a slight increase in false positives is observed, this trade-off is acceptable in practical deployment scenarios where missing an attack can have severe consequences.

Overall, the proposed SMOTE-DDQN framework provides a robust and scalable solution for modern intrusion detection systems.

VI. Conclusion And Future Work

This paper presented a novel Intrusion Detection System (IDS) based on Deep Reinforcement Learning, integrating a Double Deep Q-Network (DDQN) with the Synthetic Minority Over-sampling Technique (SMOTE). The proposed approach effectively addresses two critical challenges in network intrusion detection: class imbalance and adaptability to evolving cyber threats.

Experimental evaluation on the NSL-KDD dataset demonstrated that the proposed SMOTE-DDQN model achieves high detection performance, with an accuracy of 95.88% and an AUC of 0.97. Notably, the model significantly reduces false negatives, thereby improving the detection of rare but critical attack classes such as User-to-Root (U2R) and Remote-to-Local (R2L). This is achieved through a carefully designed reward function that prioritizes security-sensitive decisions, ensuring that malicious activities are effectively identified.

Despite these promising results, the current study is limited to the NSL-KDD dataset, which may not fully represent modern network traffic patterns. Future work will focus on evaluating the model using more recent and realistic datasets such as CIC-IDS2017 and UNSW-NB15. Additionally, extending the framework to real-time deployment using streaming data and exploring adversarial training techniques will further enhance the robustness and practical applicability of the system.

References

[1]. Othman Alnasser Et Al.,"Signature And Anomaly Based Intrusion Detection System For Secure Iots And V2G Communication" Alexandria Engineering Journal 125 (2025) 424–440 Https://Doi.Org/10.1016/J.Aej.2025.03.068

[2]. Mahbod Tavallaee Et Al.,"A Detailed Analysis Of The KDD CUP 99 Data Set 2009 IEEE Symposium On Computational Intelligence For Security And Defense Applications, Ottawa, ON, Canada, 2009, Pp. 1-6, Doi: 10.1109/CISDA.2009.5356528.

[3]. Yang Et Al., "A Survey For Deep Reinforcement Learning Based Network Intrusion Detection," Arxiv.Org, 2024. DOI: [10.48550/Arxiv.2410.07612](Https://Doi.Org/10.48550/Arxiv.2410.07612).

[4]. P.Sreenivasulu, Dr.M.Ussenaiah."A Review On Machine Learning And Deep Learning Based Modern Intrusion Detection Techniques For Mobile Adhoc Networks. International Journal Of Innovative Science And Research Technology (IJISRT) IJISRT25NOV173, 809-813. DOI: 10.38124/Ijisrt/25nov173. "

[5]. Y. Zhang Et Al., "Deep Learning Applications In Intrusion Detection Systems: A Review," Appl. Sci., Vol. 15, No. 3, 2025.

[6]. A. Hozouri Et Al., "A Comprehensive Survey On Intrusion Detection Systems Using Machine Learning And Deep Learning," Discover AI, 2025.

[7]. K. Psychogyios Et Al., "Deep Learning Approaches For Intrusion Detection In Time-Series Data," Future Internet, Vol. 16, No. 3, 2024.

[8]. Y. Wu Et Al., "Current Status And Challenges Of Deep Learning-Based Intrusion Detection Systems," J. Imaging, Vol. 10, No. 10, 2024.

[9]. M. A. Amine Et Al., "Optimization Of Intrusion Detection Systems Using Deep Learning," Int. J. Saf. Secur. Eng., 2024.

[10]. Automotive IDS Study, "Multi-Stage Deep Learning Intrusion Detection For Ethernet Networks," Ad Hoc Netw., 2024.

[11]. M. Nakıp And E. Gelenbe, "Self-Supervised Learning For Network Intrusion Detection," 2023.

[12]. Q. Zeng And F. Nait-Abdesselam, "Enhancing UAV Network Security: A Human-In-The-Loop And GAN-Based Approach To Intrusion Detection," In IEEE Internet Of Things Journal, Vol. 12, No. 12, Pp. 20870-20884, 15 June15, 2025, Doi: 10.1109/JIOT.2025.3545389.

[13]. S. Ghosh, Et Al.,, "Improving Transferability Of Network Intrusion Detection In A Federated Learning Setup," 2024 IEEE International Conference On Machine Learning For Communication And Networking (ICMLCN), Stockholm, Sweden, 2024, Pp. 171-176.

[14]. Jianqing Fan Et.Al., "Atheoretical Analysis Of Deep Q-Learning" Proceedings Of Machine Learning Research Vol 120:1–4, 2020.

[15]. Ryumei Nakada, Yichen Xu, Lexin Li, Linjun Zhang "Synthetic Oversampling: Theory And A Practical Approach Using Llms To Address Data Imbalance" Arxiv:2406.03628 [Stat.ML].