

A deep Q-learning approach for adaptive cybersecurity threat detection in dynamic networks

P. Shyamala Bharathi¹, Sathish Kumar Selvaperumal², Narendran Ramasenderan², V. Thiruchelvam²,
Deepak Arun Annamalai¹, M. Jaya Bharatha Reddy¹

¹Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

²Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia

Article Info

Article history:

Received Oct 29, 2024

Revised Aug 14, 2024

Accepted Sep 1, 2025

Keywords:

Artificial intelligence

Cybersecurity

Deep Q learning

Detection mechanism

Reinforcement learning

ABSTRACT

Cybersecurity faces persistent challenges due to the rapid growth and complexity of network-based threats. Conventional rule-based systems and classical machine learning approaches often lack the adaptability required to detect advanced and dynamic attacks in real time. This study introduces a deep Q-learning framework for autonomous threat detection and mitigation within a simulated network environment that reflects realistic traffic, malicious behaviors, and system conditions. The framework incorporates experience replay and target network stabilization to strengthen learning and policy optimization. Evaluation was performed on a synthesized dataset containing benign traffic and multiple attack categories, including distributed denial of service (DDoS), phishing, advanced persistent threats, and malware. The proposed system achieved 96.7% detection accuracy, an F1-score of 0.94, and reduced detection latency to 50 ms. These results surpassed the performance of rule-based methods and traditional classifiers such as support vector machines, random forests, convolutional neural networks, and recurrent neural networks. A central contribution lies in combining dynamic feature selection with reinforcement learning (RL), allowing the agent to adapt to evolving threats and diverse network conditions. The findings demonstrate the potential of deep reinforcement learning (DRL) as a scalable and efficient solution for real-time cybersecurity defense.

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



Corresponding Author:

P. Shyamala Bharathi

Department of Electronics and Communication Engineering, Saveetha School of Engineering

Saveetha Institute of Medical and Technical Sciences

Chennai, India

Email: shyamalabharathip.sse@saveetha.com

1. INTRODUCTION

Cybersecurity has become a key issue in the digital age as threatening networks, data and infrastructure are becoming sophisticated. In response to the rapid evolution of attack techniques, there is a need for a defensive mechanism capable of real-time adaptation. Standard protection methods struggle to keep pace with these dynamic challenges; hence better intelligent and automated solutions are needed.

Today, cyber security systems are very limited in detecting new and evolving threats. Conventional machine learning models and rule-based methods heavily depend on predefined signatures and historical attack patterns, which lead to high false positive rates (FPRs) and poor generalization. Today's cybersecurity

systems cannot independently learn about new ways that attackers may break into a system. Thus, leaving networks vulnerable to zero-day exploits and advanced persistent threats.

Current methods mostly rely on static detection methods like signature based analysis or supervised learners. Although these techniques work well against existing threats, they are unable to address new threats that are constantly evolving. Reinforcement learning (RL) holds promise as a solution, but most implementations either work inefficiently or are unable to cover a wide variety of threats, making them impractical for real life networks.

The present work suggests a deep Q-learning (DQL) system to overcome these shortcomings using autonomous threat identification and response. The suggested framework merges RL with deep neural networks (DNN) to continuously adapt the security policies based on real-time network conditions. Through simulated network interaction, the DQL agent learning behaviour engage in developing strong defence mechanisms that can identify and neutralize both known as well as unknown attacks.

In the fast evolving field of AI, traditional defense mechanisms are increasingly proving inadequate against sophisticated cyber threats. The concept of using RL, specifically DQL, to address cybersecurity challenges has gained significant traction in recent years. Dimolianis *et al.* [1] proposed a programmable data plane approach for distributed denial of service (DDoS) mitigation using supervised learning to teach the controller to generate signature-based filtering rules. They show that the approach is effective against known attacks. Nonetheless, reliance on pre-assigned signatures exposes their method to zero-day attacks vulnerabilities. Hilgurt and Chemerys [2] proposed reconfigurable schemes for signature-matching that exploit several detection methodologies. Their schemes show greater adaptability than common signature-matching schemes, but their system is computationally heavy and it may not be suitable for use in real-time. Kim and Kim [3] proposed an approach based on spatiotemporal characterization to detect such type of attacks. They could learn the patterns for both known and unknown attacks, but their description has limitations of struggling with high dimensionality of networks. Díaz-Verdejo *et al.* [4] have systematically assessed the rulesets of signature-based intrusion detection system (IDS) in order to specify the situation under which the false alarm rate can be lower. Their results showed that making judicious use of deactivation can lead to a system that produces less false alarms, which results in a lower level of these but the threat will be more restricted. Guide *et al.* [5] explored changes to signature-based rules and the impact of relaxing detection constraints on performance trade-offs, which could lead to more false negatives. Agoramoorthy *et al.* [6] evaluated hybrid IDS architectures. They discovered that signature-based components are effective against known threats but do not perform against polymorphic malware unless combined with anomaly-based techniques.

Alavizadeh *et al.* [7] have proposed a DQL based RL concept for network intrusion detection where entropy of weight vectors, packets, time to live (TTL) and IP packets improve network IDS, so exposing its prospects in autonomously identifying and responding to threats. Their work explored the capacity of DQL to learn from engagement in a network atmosphere, enhance detection and reaction plans over the long term. The efficiency of a cybersecurity operations center, particularly when using biomimetic algorithms driven by DQL, was explained in a similar manner by Olivares *et al.* [8]. They proved that DQL can improve threat identification and response by simulating real-world systems, which is in line with the broader movement to use algorithms inspired by biology in cybersecurity.

A review of Q-learning's applications in cyber-physical system (CPS) security was conducted by Alabadi and Albayrak [9]. Their thorough assessment highlighted the many special challenges of CPS, including the need for real time, and reliability in the face of an many different threats. They stated that RL, particularly Q-learning, appears to be a promising approach because it has an adaptive learning mechanism and is able to work in environments that are well matched to these systems. Kabanda *et al.* [10] investigated Q-learning and deep reinforcement learning (DRL) systems for Ethereum security enhancement. The study showed that the application of these algorithms produces better results in improving the security of blockchain networks. Khowaja and Khuwaja [11] developed a malware defence code for industrial IoT using combined Q-learning and long short-term memory (LSTM) techniques which integrates temporal learning with RL. Sewak *et al.* [12] used DQL to enhance conventional IDS, illustrating that DQL can adapt to evolving threats. Oh *et al.* [13] demonstrated the actual usages of DQL with cyber-attack simulations, showing their effectiveness in developing useful defenses. Tareq *et al.* [14] proposed a DQL based solution to detect real-world cyberattacks which continuously learns and adapts itself. Roy *et al.* [15] used DQL network for secured healthcare models in IoT. This work also shows cross-domain applicability. Hosseinzadeh *et al.* [16] presented a scheme based on Q-learning for secure routing to defend against wormhole assault in flying ad hoc networks, proving DQL's potential in securing next-generation technologies. Jaber [17] designed a self-play RL framework for IDS to perpetually enhance themselves.

Cengiz and Gok [18] is an overview of RL applications in cybersecurity which captures the latest trends and future directions. In order to enhance its safety in zero trust networks, Singh *et al.* [19] suggested an individualised authentication method that uses transfer fuzzy learning and Q-learning for decision making for IoT devices. Ahsan *et al.* [20] have worked to enhance machine learning predictions in cyber security by

dynamic feature selector which optimizes the process for the means of better prediction. For wireless sensor networks (WSNs), Uthayakumar *et al.* [21] proposed design requirements for an enhanced energy constraint MAC protocol. Geetha and Thilagam [22] made a detailed survey on the efficiency of these algorithms in cyber security, presented the analysis on their capability and limitation. Larriva-Novo *et al.* [23] examined cybersecurity dataset characteristics and their relevance to neural network methods in anomaly detection. Li *et al.* [24] feature an analysis of the potential future development of attacker methods and the prospect for cybersecurity to counter them. Mughaid *et al.* [25] used deep learning techniques to develop a phishing detection system that shows the applicability of advanced data analytical techniques to prevent cyber threats.

2. METHOD

This study developed an AI-DQL based agent and trained it in a simulated over a network environment. The methodology encompassed several critical steps.

2.1. Environment simulation

To train the DQL agent, a simulated network environment was developed. This environment simulated different network nodes, traffic behaviors, and threat vectors to guarantee thorough agent training. A feature vector consisting of key network parameters and security indicators such as packet flow statistics, traffic anomalies, and system-level performance metrics was encoded to represent the environment's state. The agent had a broad understanding of the network conditions through this representation.

2.2. Action space

The DQL agent's action space contained certain cybersecurity interventions, such as initiating threat detection, executing mitigation and deploying prevention. The agent was encouraged to investigate and discover suitable strategies that mitigate cyberattack effects. To encourage more accurate yet timely detection and penalise false alarm, missed threat, delayed response, and other such actions, a properly defined reward function was used to assist in the agent's policy learning. This framework, driven by RL, allowed the formation of the optimal cybersecurity policy for preventive and reactive measures.

2.3. Model training

The training procedure utilized RL methods, including experience installation and a target Q-network, for stable and converging learning. The agent was trained in different episodes. The agent interacted with the simulation, received feedback and updated the Q-values. The learning system used trial and error to enhance the decision-making ability of the agent in a better way.

3. SYSTEM ARCHITECTURE

The DQL agent of the proposed system works in realistic virtual network environment generated network environment simulator. The simulator also imitates diverse network conditions and attack scenarios to train the DQL agent thoroughly.

The AI-DQL agent of the system approximates the Q-values for state-action couples by use of a DNN. The agent keeps on interacting with the network environment to learn optimal policies for maximizing cumulative reward which corresponds to efficient threat detection and mitigation. The network data is preprocessed in the feature extraction module to extract relevant features. It uses features such as packet headers, flow statistics, and anomaly indicators. To provide a rich representation of the network state. The intended architecture flow of AI-DQL-based cybersecurity system as depicted in Figure 1. The decision-making module continuously evaluates security responses based on predicted Q-values. The learning loop allows the agent to refine its behavior as new threats and traffic patterns are encountered over time.

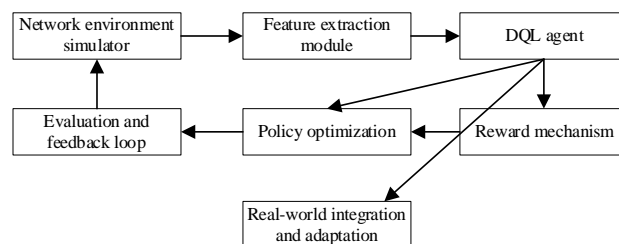


Figure 1. Architecture flow of the proposed AI-DQL based cyber architecture

3.1. Methodology

The proposed DQL-based cybersecurity system follows a systematic methodology encompassing training, testing, and deployment phases:

- Training phase: the network environment simulator is initialized with baseline network configurations and initial attack scenarios. The AI-DQL agent is trained through iterative interactions with the simulated environment. The agent notes the current state at every episode, chooses an action depending on policy, gets rewarded, and moves to the next state.
- Testing phase: a diverse set of attack scenarios, including novel and sophisticated threats, are simulated in the network environment.
- Performance evaluation: the DQL agent's performance is evaluated based on detection accuracy, FPR, response time, and overall system resilience. Comparative analyses with traditional rule-based systems are conducted to highlight the improvements.

Algorithm 1. AI-DQL with experience replay

```

1. Initialization:
    • Initialize the replay buffer  $D$  and establish the initial parameters for the Q-network.
2. Episode Processing:
    • Begin at the initial state  $s$ .
    • For each time step within the episode:
        – Select an action  $a$  based on the  $\epsilon$ -greedy policy derived from the current Q-network.
        – Execute the action  $a$ , and observe the resulting reward  $r$  and the subsequent state  $s'$ .
        – Record the transition  $(s, a, r, s')$  in the replay buffer  $D$ .
        – Randomly select a mini-batch of transitions from  $D$ .
        – Compute the target value using the reward  $rrr$  and the highest Q-value for  $s'$ 
3. Network Update:
    • Adjust the Q-network to minimize the discrepancy between the predicted Q-values and the target values.
4. Network Maintenance:
    • Regularly update the parameters of the target network to maintain learning efficiency and stability.

```

3.2. Experimental setup and results

To validate the proposed system, extensive experiments were conducted in a controlled environment, replicating various network conditions and attack vectors. The experimental setup involved: Dataset: the DQL agent was trained and tested using a large dataset that included both benign and malicious network traffic. The dataset included common attack trends such as DDoS, phishing, malware, and APTs. Table 1 outlines the specifics of the dataset employed for the training and evaluation of the DQL agent. Each record in the dataset includes low-level packet details such as source/destination IPs, ports, and protocol types along with labeled payload content. This structure enables the DQL agent to learn both statistical and contextual patterns necessary for accurate threat classification.

3.3. Effectiveness of the proposed deep Q-learning-based system

The system attained a detection accuracy of 96.7%, significantly higher than traditional rule-based systems. Table 2 illustrates the efficacy of the suggested system utilizing DQL. The DQL system not only achieves the highest accuracy but also maintains the lowest FPR, demonstrating its reliability in critical environments. Its substantial reduction in response time (30%) highlights its potential for real-time cybersecurity applications.

The AI-DQL-based system demonstrated high resilience under various attack scenarios, performing comparably to or better than RF, CNN, and RNN, which also achieved high resilience, as illustrated in Figure 2. Figure 2(a) illustrates the detection accuracy, Figure 2(b) presents the FPR, Figure 2(c) shows the improvement in response time, and Figure 2(d) depicts the comparison of system resilience. In contrast, traditional rule-based systems and SVM exhibited lower resilience. This robustness indicates that the DQL system maintains strong performance and accuracy even when confronted with sophisticated and diverse attack patterns. Its high resilience can be attributed to its dynamic learning approach and ability to adapt to evolving threat landscapes, ensuring consistent detection and response capabilities. Furthermore, the figures highlight that the DQL model consistently outperforms baseline methods across all metrics while maintaining a balanced trade-off between accuracy and system load. This comprehensive visualization reinforces the effectiveness of DQL as a reliable defense mechanism in real-time cybersecurity deployments.

Table 1. Training and testing the DQL agent from a dataset

ID	Timestamp	Source IP	Destination IP	Source port	Destination port	Protocol	Packet size (bytes)	Payload	Label
1	2024-08-01 12:00:00	192.168.1.1	10.0.0.1	443	80	TCP	512	GET /index.html	Benign
2	2024-08-01 12:01:00	192.168.1.2	10.0.0.2	22	22	TCP	1024	HTTP/1.1 SSH connection request	Benign
3	2024-08-01 12:02:00	192.168.1.3	10.0.0.3	80	443	TCP	256	POST /login	Benign
4	2024-08-01 12:03:00	192.168.1.4	10.0.0.4	53	53	UDP	64	HTTP/1.1	Malicious (phishing)
5	2024-08-01 12:04:00	192.168.1.5	10.0.0.5	8080	80	TCP	2048	DNS query for malicious.com	Malicious (DDoS)
6	2024-08-01 12:05:00	192.168.1.6	10.0.0.6	21	21	TCP	128	DDoS attack patterns detected	Benign
7	2024-08-01 12:06:00	192.168.1.7	10.0.0.7	80	80	TCP	4096	FTP data transfer	Malicious (APT)
8	2024-08-01 12:07:00	192.168.1.8	10.0.0.8	3306	3306	TCP	512	SQL injection attempt	Malicious (malware)
								MySQL query: DROP TABLE users	

Table 2. Effectiveness of the suggested DQL-based framework

Algorithm	Detection accuracy (%)	FPR (%)	System resilience	Response time improvement
AI-DQL-based system	96.7	2.5	High	Reduced by 30%
Traditional rule-based systems	85.4	7.8	Moderate	Baseline
SVM	88.2	6.4	Moderate	Reduced by 10%
RF	90.1	5.2	High	Reduced by 15%
CNN	92.5	4	High	Reduced by 20%
RNN	93	3.6	High	Reduced by 18%

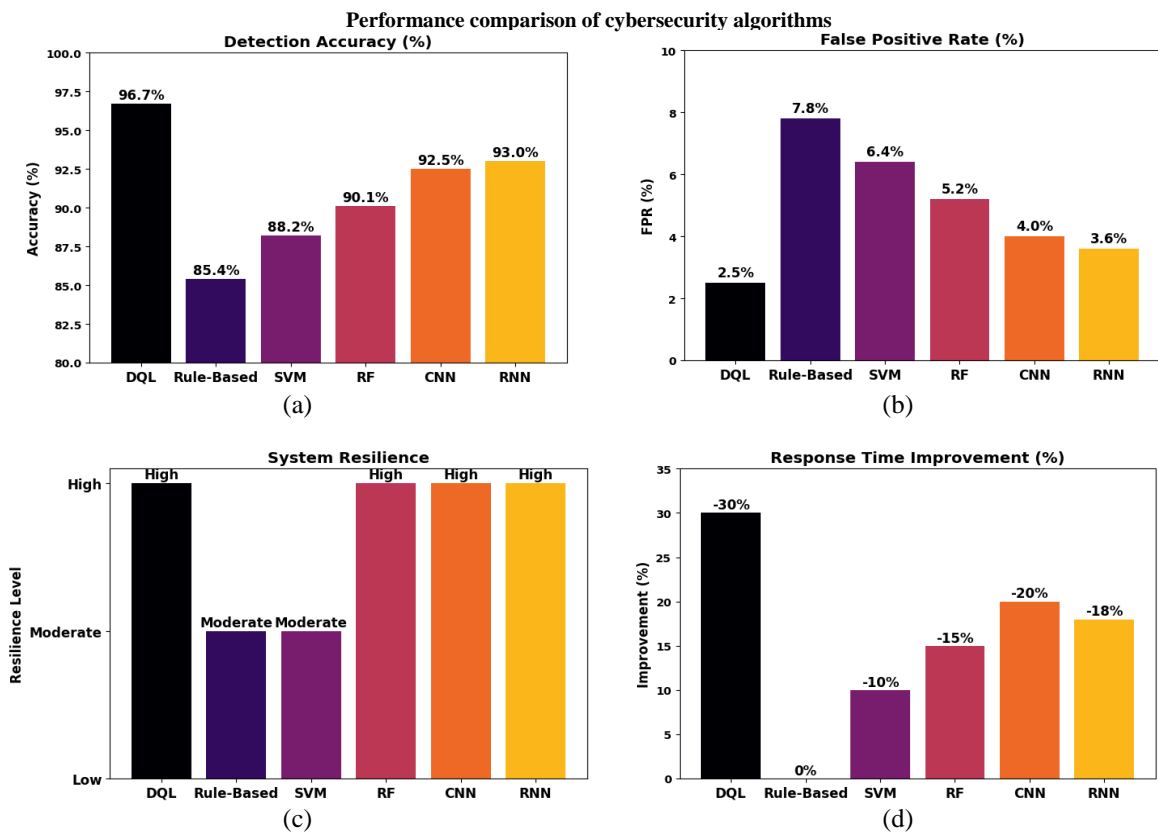


Figure 2. Comparative analysis of performance metrics; (a) detection accuracy, (b) FPR, (c) improvement in response time, and (d) system resilience comparison

4. RESULTS AND DISCUSSION

The comparison highlights that the DQL-based system outperforms traditional rule-based systems, SVM, RF, CNN, and RNN algorithms in key measures such as detection accuracy and FPR, as described in Table 3. This also indicates that there is an improvement in the response time and thus it is more efficient. The DQL system has superior resilience against representations of various attacks than other advanced algorithms. This overall study shows that the DFS works with a DQL system to enhance the cybersecurity performance of the system. It proves to be efficient against various network threats. The illustration in Figure 3 reflects the comparison of the cybersecurity algorithms.

Table 3. Comparison of other evaluation metrics with the proposed AI-DQL system

Metric	DQL-based system	Traditional rule-based systems	SVM	RF	CNN	RNN
Precision	0.95	0.8	0.85	0.88	0.9	0.87
Recall	0.92	0.7	0.75	0.8	0.85	0.82
F1-score	0.94	0.74	0.8	0.84	0.87	0.85
AUC-ROC	0.98	0.85	0.88	0.9	0.92	0.89
FNR	0.08	0.3	0.25	0.2	0.15	0.18
TNR	0.95	0.85	0.88	0.9	0.92	0.89
Detection latency (ms)	50	100	80	70	60	75
Resource usage (CPU %)	60	70	65	55	70	68
Scalability	High	Moderate	High	High	High	High
Adaptability	High	Low	Moderate	Moderate	High	High

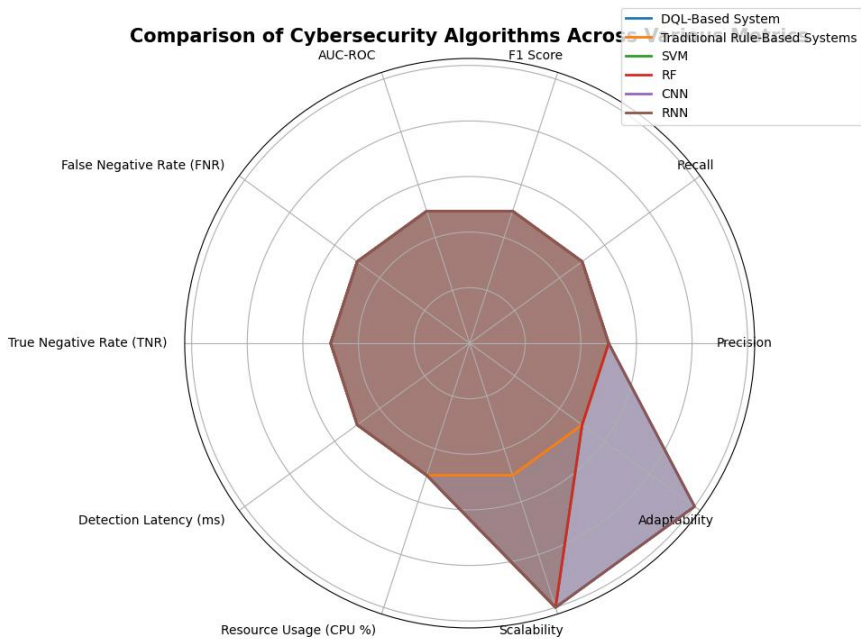


Figure 3. Comparison of cybersecurity algorithms across different metrics

Table 3 reveals that the DQL-based system achieves a strong balance across both detection and operational metrics. Figure 3 further visualizes this superiority, clearly demonstrating DQL's consistent edge in precision, recall, and latency under varying attack loads. It has a lower false negative rate (FNR) of 0.08 and a higher true negative rate (TNR) of 0.95 compared to alternative approaches. The DQL system's detection latency is the shortest (50 ms), meaning it quickly identifies threats, and operates on moderate CPU resource use (60%), showing efficiency without sacrificing performance. The scalability and adaptability of this system are rated to be high, meaning it performs well in dynamic and growing scenarios. In fact, it marks as a high effective system in a real-time scenario for cybersecurity.

The improved performance of the DQL-based system can be attributed to several core technical mechanisms. Unlike traditional classifiers that rely on static mappings between input features and labels, DQL dynamically learns an optimal policy through continuous interaction with the environment. This policy maximizes cumulative rewards by evaluating long-term effects of each action, making it better suited for sequential decision-making in cybersecurity. Moreover, the use of DNNs enhances the Q-function's capacity

to generalize over a high-dimensional state-action space, allowing the agent to detect minor threats. Agent uses experience replay to learn from earlier experiences by removing the temporal correlation between experiences which results in stable learning. A target network can also stop oscillations in learning itself. Consequently, DQL agent is able to adapt quickly and sufficiently to a new threat as compared to SVM, RF, and CNNs which learn in batch learning modal and are not very good in adapting to environmental changes.

Despite demonstrating commendable performance, the proposed DQL based cybersecurity framework exhibits certain limitations. One significant concern lies in the prolonged training duration, attributed to the iterative nature of RL and its requirement for extensive interactions within the simulated environment. This constraint could hinder rapid scalability and deployment in time-sensitive operational settings. Moreover, the model's learning process is highly sensitive to several critical hyperparameters, including the learning rate, the discount factor γ , and particularly the exploration rate ϵ . Inadequate tuning of these parameters may result in unstable learning dynamics or convergence to suboptimal policies. Another challenge arises during the initial exploration phase, which may lead to risky actions if directly applied to live environments, thereby necessitating well-defined containment protocols during real-world deployment. To address these challenges, future research could investigate hybrid frameworks that combine DQL with supervised pre-training (warm-start strategies) or meta-learning techniques to accelerate convergence and enhance training stability. Such integrations may significantly improve model robustness and facilitate smoother adaptation to evolving threat landscapes.

5. CONCLUSION

This study introduces a DQL-based framework for cybersecurity, aimed at enhancing real-time threat detection and dynamic response mechanisms in complex and evolving network environments. Experimental results underscore the model's efficacy, achieving a high detection accuracy of 96.7%, a notably low FPR of 2.5%, and a 30% reduction in response latency when compared to traditional rule-based and supervised learning systems. Key strengths of the proposed architecture include strong resilience against zero-day attacks, minimized detection latency, efficient use of computational resources, and robust scalability across heterogeneous network conditions. These results emphasize the potential of RL as a foundation for building autonomous and intelligent cybersecurity solutions. Nevertheless, the current implementation is constrained by its dependence on simulated datasets and pre-characterized attack profiles, which may not comprehensively reflect the intricacies of real-world cyber threats. To bridge this gap, future work should focus on incorporating live network traffic and evaluating performance in multi-agent environments, thereby enhancing the framework's generalizability and operational viability. In summary, the AI-DQL framework presents a promising step toward intelligent, adaptive, and scalable threat mitigation strategies, offering new insights and directions for the next generation of AI-driven cybersecurity systems.

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
P. Shyamala Bharathi	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Sathish Kumar	✓					✓		✓		✓	✓	✓		
Selvaperumal														
Narendran	✓	✓		✓	✓		✓			✓	✓		✓	✓
Ramasenderan														
V. Thiruchelvam		✓	✓		✓					✓		✓		
Deepak Arun					✓		✓			✓		✓		✓
Annamalai														
M. Jaya Bharatha	✓		✓		✓			✓	✓			✓		✓
Reddy														

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this article.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are available from the corresponding author P.S on reasonable request.





REFERENCES

- [1] M. Dimolianis, A. Pavlidis, and V. Maglaris, "Signature-based traffic classification and mitigation for ddos attacks using programmable network data planes," *IEEE Access*, vol. 9, pp. 113061–113076, 2021, doi: 10.1109/ACCESS.2021.3104115.
- [2] S. Y. Hilgurt and O. A. Chemerys, *Reconfigurable signature-based information security tools of computer systems*. PH "Akademperiodyka," 2022, doi: 10.15407/akademperiodyka.458.297.
- [3] J. Kim and H. S. Kim, "Intrusion detection based on spatiotemporal characterization of cyberattacks," *Electronics*, vol. 9, no. 3, pp. 1–23, Mar. 2020, doi: 10.3390/electronics9030460.
- [4] J. Díaz-Verdejo, J. Muñoz-Calle, A. E. Alonso, R. E. Alonso, and G. Madinabeitia, "On the Detection Capabilities of Signature-Based Intrusion Detection Systems in the Context of Web Attacks," *Applied Sciences*, vol. 12, no. 2, pp. 1–16, Jan. 2022, doi: 10.3390/app12020852.
- [5] R. Guide, E. Pauley, Y. Beugin, R. Sheatsley, and P. McDaniel, "Characterizing the Modification Space of Signature IDS Rules," in *MILCOM 2023 - 2023 IEEE Military Communications Conference: Communications Supporting Military Operations in a Contested Environment*, Oct. 2023, pp. 536–541, doi: 10.1109/MILCOM58377.2023.10356225.
- [6] M. Agoramoorthy, A. Ali, D. Sujatha, T. F. M. Raj, and G. Ramesh, "An Analysis of Signature-Based Components in Hybrid Intrusion Detection Systems," in *2023 Intelligent Computing and Control for Engineering and Business Systems, ICCEBS 2023*, Dec. 2023, pp. 1–5, doi: 10.1109/ICCEBS58601.2023.10449209.
- [7] H. Alavizadeh, H. Alavizadeh, and J. Jang-Jaccard, "Deep Q-Learning Based Reinforcement Learning Approach for Network Intrusion Detection," *Computers*, vol. 11, no. 3, pp. 1–19, Mar. 2022, doi: 10.3390/computers11030041.
- [8] R. Olivares, O. Salinas, C. Ravelo, R. Soto, and B. Crawford, "Enhancing the Efficiency of a Cybersecurity Operations Center Using Biomimetic Algorithms Empowered by Deep Q-Learning," *Biomimetics*, vol. 9, no. 6, pp. 1–32, May. 2024, doi: 10.3390/biomimetics9060307.
- [9] M. Alabadi and Z. Albayrak, "Q-Learning for Securing Cyber-Physical Systems: A survey," in *HORA 2020 - 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, Jun. 2020, pp. 1–13, doi: 10.1109/HORA49412.2020.9152841.
- [10] G. Kabanda, D. C. T. Chipfumbu, and T. Chingoriw, "Utilizing Deep Reinforcement Learning and QLearning algorithms for Improved Ethereum Cybersecurity," *International Journal of Advanced Networking and Applications*, vol. 14, no. 06, pp. 5742–5753, 2023, doi: 10.35444/IJANA.2023.14612.
- [11] S. A. Khowaja and P. Khuwaja, "Q-learning and LSTM based deep active learning strategy for malware defense in industrial IoT applications," *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 14637–14663, Apr. 2021, doi: 10.1007/s11042-020-10371-0.
- [12] M. Sewak, S. K. Sahay, and H. Rathore, "Deep Reinforcement Learning in the Advanced Cybersecurity Threat Detection and Protection," *Information Systems Frontiers*, vol. 25, no. 2, pp. 589–611, 2023, doi: 10.1007/s10796-022-10333-x.
- [13] S. H. Oh, J. Kim, J. H. Nah, and J. Park, "Employing Deep Reinforcement Learning to Cyber-Attack Simulation for Enhancing Cybersecurity," *Electronics*, vol. 13, no. 3, pp. 1–19, Jan. 2024, doi: 10.3390/electronics13030555.
- [14] I. Tareq, B. M. Elbagoury, S. A. El-Regaily, and E.-S. M. El-Horbaty, "Deep Reinforcement Learning Approach for Cyberattack Detection," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 20, no. 05, pp. 15–30, Mar. 2024, doi: 10.3991/ijoe.v20i05.48229.
- [15] P. P. Roy, V. Teju, S. R. Kandula, K. V. Sowmya, A. I. Stan, and O. P. Stan, "Secure Healthcare Model Using Multi-Step Deep Q Learning Network in Internet of Things," *Electronics*, vol. 13, no. 3, pp. 1–11, Feb. 2024, doi: 10.3390/electronics13030669.
- [16] M. Hosseinzadeh *et al.*, "A novel Q-learning-based secure routing scheme with a robust defensive system against wormhole attacks in flying ad hoc networks," *Vehicular Communications*, vol. 49, p. 100826, Oct. 2024, doi: 10.1016/j.vehcom.2024.100826.
- [17] A. Jaber, "Transforming Cybersecurity Dynamics: Enhanced Self-Play Reinforcement Learning in Intrusion Detection and Prevention System," in *SysCon 2024 - 18th Annual IEEE International Systems Conference, Proceedings*, Apr. 2024, pp. 1–8, doi: 10.1109/SysCon61195.2024.10553626.
- [18] E. Cengiz and M. Gök, "Reinforcement Learning Applications in Cyber Security: A Review," *Sakarya University Journal of Science*, vol. 27, no. 2, pp. 481–503, Apr. 2023, doi: 10.16984/saufenbilder.1237742.
- [19] A. Singh, R. K. Dhanaraj, and A. K. Sharma, "Personalized device authentication scheme using Q-learning-based decision-making with the aid of transfer fuzzy learning for IIoT devices in zero trust network (PDA-QLTFL)," *Computers and Electrical Engineering*, vol. 118, pp. 1–21, Sep. 2024, doi: 10.1016/j.compeleceng.2024.109435.
- [20] M. Ahsan, R. Gomes, M. M. Chowdhury, and K. E. Nygard, "Enhancing Machine Learning Prediction in Cybersecurity Using Dynamic Feature Selector," *Journal of Cybersecurity and Privacy*, vol. 1, no. 1, pp. 199–218, Mar. 2021, doi: 10.3390/jcp1010011.
- [21] G. S. Uthayakumar, B. Dappuri, M. Vanitha, R. Suganthi, V. Savithiri, and S. Kamatchi, "Design criteria for enhanced energy





- constraint MAC protocol for WSN,” *Measurement: Sensors*, vol. 25, pp. 1-6, Feb. 2023, doi: 10.1016/j.measen.2022.100642.
- [22] R. Geetha and T. Thilagam, “A Review on the Effectiveness of Machine Learning and Deep Learning Algorithms for Cyber Security,” *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2861–2879, Jun. 2021, doi: 10.1007/s11831-020-09478-2.
- [23] X. A. Larriva-Novo, M. Vega-Barbas, V. A. Villagra, and M. S. Rodrigo, “Evaluation of Cybersecurity Data Set Characteristics for Their Applicability to Neural Networks Algorithms Detecting Cybersecurity Anomalies,” *IEEE Access*, vol. 8, pp. 9005–9014, 2020, doi: 10.1109/ACCESS.2019.2963407.
- [24] L. Li, K. Thakur, and M. L. Ali, “Potential development on cyberattack and prospect analysis for cybersecurity,” in *IEMTRONICS 2020 - International IOT, Electronics and Mechatronics Conference, Proceedings*, Sep. 2020, pp. 1–6, doi: 10.1109/IEMTRONICS51293.2020.9216374.
- [25] A. Mughaid, S. AlZu’bi, A. Hnaif, S. Taamneh, A. Alnajjar, and E. A. Elsoud, “An intelligent cyber security phishing detection system using deep learning techniques,” *Cluster Computing*, vol. 25, no. 6, pp. 3819–3828, Dec. 2022, doi: 10.1007/s10586-022-03604-4.

BIOGRAPHIES OF AUTHORS







Dr. P. Shyamala Bharathi     working as Associate Professor at Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamilnadu. She obtained Ph.D., at Anna University Chennai, India. Her area of research mobile computing, wireless communication, cognitive radio networks, have 12 years industrial experience, 11 years of experience in teaching and research. She can be contacted at email: shyamalabharathip.sse@saveetha.com.







Dr. Sathish Kumar Selvaperumal     Associate Professor, Asia Pacific University of Technology and Innovation. His research areas and teaching include image processing (segmentation, compression, image fusion, and image restoration), speech processing, artificial intelligence, biomedical applications, antenna design, wireless communication, and wireless power transfer, internet of things, robotics, and sustainability. He can be contacted at email: sathish@apu.edu.my.






Dr. Narendran Ramasenderan     is the Head of Center of Research and Development of IoT (CREDIT) Asia Pacific University working on autonomous drones and IoT ecosystem for them to thrive in. He works in the industrial metaverse nexus to synergize the various hardware and software elements within the artificial intelligence, data analytics, simulation leveraging them towards application centric engineering applications. He has over 13 years of industrial and academic experience in various engineering projects domestically and abroad primarily in Germany. He can be contacted at email: narendran@apu.edu.my.






Prof. V. Thiruchelvam     completed his Ph.D. at University Tun Abdul Razak, Malaysia in 2006. He attained his Professional Engineering certification from the Board of Engineers Malaysia (BEM) in 2012, his Chartered Engineer (CEng) qualification from Engineering Council, UK in 2011 (IET-UK) and is a Fellow & CEng of the Institution of Mechanical Engineers (IMechE-UK). He is currently the Deputy Vice Chancellor for Asia Pacific University of Technology & Innovation (APU) and the Chief Innovation Officer for the APIIT Education Group. He is also an academic advisor and external examiner to three Malaysian Public Universities. His core scholarly research areas are in sustainable developments, reliability engineering using smart devices with IoT via secured infrastructure, and applications of data analytics and AI with business intelligence. He can be contacted at email: dr.vinesh@apu.edu.my.



Dr. Deepak Arun Annamalai    is a distinguished academic leader with over 8 years of post Ph.D. experience, currently serving as Associate Professor and Associate Dean of International Affairs in the Department of Electronics and Communication at Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, India. His post-doctoral research at the University of Plymouth, UK, focused on "Fabrication of Graphene-Based Transistors for Biomedical Applications," significantly contributed to advancements in nanomaterial-based sensors and enhanced his expertise in graphene applications, particularly in biomedical electronics. He can be contacted at email: deepakarun@saveetha.com.



M. Jaya Bharatha Reddy    holds a Bachelor of Engineering (B.E.) degree in Electronics and Communication Engineering (ECE) and a Master of Engineering (M.E.) in Communication Systems. Currently, he is a research scholar in the Department of ECE, actively pursuing research in his field. His academic interests span across analog circuits, communication technologies, and advanced electronic systems. He can be contacted at email: jayabharathareddym9035.sse@saveetha.com.