The ISC Int'l Journal of Information Security
ISeCure

January 2026, Volume 18, Number 1 (pp. 49–60)
http://www.isecure-journal.org

A Novel Reinforcement Learning-based Congestion Control Algorithm for DDoS-Induced Adversarial Conditions in Blockchain and Distributed Networks

Ehsan Abedini [1], Amir Jalaly Bidgoly [1,*], and Mohsen Nickray [1]
[1] *Department of Computer Engineering, University of Qom, Qom, Iran.*

ARTICLE INFO.

Article history:
Received: April 19, 2025
Revised: October 13, 2025
Accepted: November 17, 2025
Published Online: November 19, 2025

Keywords:
Reinforcement Learning, Congestion Control, DDoS, Blockchain Security, Distributed Networks, Adversarial Conditions.

Type: Research Article

doi: 10.22042/isecure.2025.515662.1221

ABSTRACT

Distributed Denial-of-Service (DDoS) attacks are among the most critical security threats to distributed network infrastructures, including blockchain systems. These attacks degrade performance, cause congestion, and disrupt service delivery or transaction processing. Traditional mitigation techniques have undergone extensive development. However, they often fail to intelligently detect and manage traffic patterns and struggle to adapt to dynamic conditions in decentralized environments. The method dynamically adjusts the congestion window (CWND) according to traditional TCP principles. It uses network signals such as delay and packet loss to guide its adjustment process. What distinguishes our approach is that the RL-agent interprets persistent or abnormal congestion patterns as potential indicators of adversarial high-load conditions (e.g., DDoS-induced congestion) and adapts CWND adjustments more intelligently to reduce their effects. Leveraging the Q-learning algorithm, the proposed approach adapts dynamically to fluctuating traffic and conditions. Its learning capability enables continuous monitoring of behavior and timely responsiveness to anomalies, including sustained congestion patterns often associated with adversarial traffic surges. Simulation results were obtained across several DDoS scenarios and compared with conventional CC algorithms. The proposed method achieved significant improvements in key performance indicators. These include reduced latency, better bandwidth utilization, improved stability, lower packet loss, and higher throughput. The proposed Q-learning-based CC operates at the peer-to-peer layer, regulating flow among blockchain nodes. It is independent of consensus mechanisms while indirectly improving consensus efficiency by reducing message delays and packet loss. This method provides a scalable and intelligent solution for CC under adversarial conditions. It enhances robustness and efficiency in both distributed systems and blockchain-based networks.

1 Introduction

In recent years, distributed networks—including blockchain [1] systems—have emerged as a cornerstone of modern computing infrastructures, enabling a wide range of applications in finance, Internet of Things (IoT) [2], data management, and decentralized services [3]. Among these, blockchain technology has gained considerable attention due to its features such as transparency [4], tamper-resistance, and de-

* Corresponding author.
Email addresses: e.abedini@stu.qom.ac.ir, jalaly@qom.ac.ir, m.nickray@qom.ac.ir
ISSN: 2008-2045

centralized trust mechanisms. These characteristics have promoted its adoption in multiple domains. Examples include cryptocurrencies, smart contracts, and supply chain systems [5].

However, as both blockchain and other distributed applications continue to grow, so do the associated security challenges [6]. One of the most serious threats to such networks is the Distributed Denial-of-Service (DDoS) attack [7]. In this attack, illegitimate traffic floods the network, consuming bandwidth and resources. As a result, legitimate users and services are denied access [8].

Conventional DDoS mitigation techniques—such as firewalls, Content Delivery Networks (CDNs), and rule-based detection algorithms—have demonstrated limited scalability and adaptability to evolving threats in dynamic environments [9]. These limitations highlight the need for intelligent, context-aware mechanisms that can autonomously analyze traffic conditions and respond proactively.

This paper introduces a reinforcement learning–based (RL-based) congestion control (CC) algorithm. The proposed algorithm employs the Q-learning technique to optimize its control process. This method does not estimate the Congestion Window (CWND) but dynamically adjusts it in line with traditional TCP principles based on congestion signals such as delay and packet loss. The key differentiator of our approach lies in the RL-agent's ability to construe persistent or abnormal congestion as potential indicators of adversarial high-load conditions (e.g., DDoS-induced congestion), thereby adapting its CWND adjustments more judiciously to alleviate detrimental effects in both blockchain and general distributed network environments. Thus, the method should be regarded as a reinforcement learning-based congestion control mechanism that operates under adversarial conditions, rather than as direct malicious traffic filtering. By learning from real-time network feedback and adapting to dynamically changing conditions, the proposed method operates through three key stages:

(1) **Network State Analysis and Identification:** Monitoring key parameters such as packet loss rate, bandwidth, and delay.
(2) **Adaptive Action Execution:** Adjusting transmission rates through Congestion Window (CWND) Sizing.
(3) **Reward Metric Definition:** Evaluating system performance based on reduced delay and packet loss, along with improved bandwidth efficiency.

The proposed approach has been evaluated in a simulated environment and compared against traditional CC algorithms. The experimental results confirm its strong capability to strengthen congestion management under conditions that may arise or be exacerbated by DDoS attacks, thereby improving the performance and resilience of distributed network infrastructures, including blockchain systems.

The remainder of this paper is organized as follows: Section 2 reviews the related work on DDoS mitigation and reinforcement learning-based congestion control methods. Section 3 presents the proposed methodology, detailing the RL-based congestion control model, environment setup, agent actions, and reward function. Section 4 describes the performance metrics used to evaluate the system. Section 5 discusses the experimental setup, simulation results, and comparative analysis against existing techniques. Section 6 addresses the implementation limitations and outlines potential directions for future work. Finally, Section 7 concludes the paper by summarizing the key contributions and findings.

2 Materials and Methods

DDoS attacks represent a critical threat to decentralized blockchain systems. They can disrupt network functionality by saturating bandwidth and exhausting processing resources. Over the years, extensive research has been conducted to counter these threats [10]. This section presents a review of prior studies on conventional DDoS mitigation techniques.

2.1 DDoS Mitigation Approaches

With the growing complexity and volume of DDoS attacks, traditional mitigation methods have proven insufficient in dynamic and decentralized environments such as blockchain networks. In recent years, researchers have introduced intelligent and hybrid solutions that leverage blockchain, Machine Learning (ML), and Software-Defined Networking (SDN) to enhance detection, scalability, and responsiveness against DDoS threats.

Ilyas *et al.* [11] proposed a blockchain-integrated deep neural network model based on a Poaching–Raptor optimization algorithm. This model, embedded in smart contracts, achieved over 95% accuracy in detecting and filtering malicious traffic, demonstrating the feasibility of combining deep learning with decentralized infrastructures. In another study, Jmal *et al.* [12] proposed a hybrid security architecture for IoT systems. This architecture integrates SDN, blockchain, and artificial neural networks (ANNs) to improve adaptability and security. This design enabled centralized control with decentralized verification, effectively identifying and

mitigating DDoS traffic in real time.

Abdullah and Hussein [13] proposed a novel blockchain mechanism that reverses TCP request flows to detect DDoS attacks. Their model achieved more than 99% detection accuracy and reduced the time to identify attackers to under 0.75 seconds. Similarly, Li *et al.* [14] developed an ensemble ML framework for adaptive DDoS detection in blockchain environments, combining classifiers to improve detection accuracy and reduce false positives.

From the perspective of SDN-based defense, Kavitha and Ramalakshmi [15] presented a machine learning–driven detection system that combines Logistic Regression (LR), K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), and Random Forest (RF) algorithms. Their model achieved over 99.99% accuracy in identifying DDoS traffic in SDN–IoT networks. Vanlalruata *et al.* [16] implemented a Deep Neural Network (DNN)–based method for real-time DDoS detection in SDN infrastructures, reporting more than 99.9% accuracy across several standard datasets.

From a broader perspective, Shah *et al.* [17] and Sharyar *et al.* [18] conducted comprehensive surveys covering blockchain-enabled DDoS mitigation. These studies analyzed decentralized strategies, smart contracts, Fuzzy Neural Networks (FNNs), and SDN integration. They identified limitations in scalability and coordination. Additionally, they outlined promising future directions, particularly reinforcement learning–based models. A comparative overview of these techniques is provided in Table 1.

2.2 CWND Adjustment and Adaptive Control

Dynamic and adaptive adjustment of the CWND is essential for effective congestion control. This is especially important in environments where traffic patterns change rapidly. Such variations can be caused by abnormal events, including DDoS-induced loads. Standard TCP variants (e.g., Reno, NewReno, CUBIC, BBR) rely on heuristic AIMD mechanisms [19]. While effective in stable networks, these methods struggle in highly dynamic conditions [20]. Recent research has thus focused on machine learning and deep learning to improve CWND control.

Majid *et al.* [21] introduced a Weighted Ensemble Deep Reinforcement Learning (WEDRL) framework combining DQN, PPO, DDPG, and TD3 for congestion control in TCP/IP networks. Their model outperformed both single DRL agents and traditional TCP schemes, achieving approximately 4% higher throughput and 10.5% lower delay.

Kumar *et al.* [22] proposed an adaptive contention window design for wireless networks using deep Q-learning. Their Q-network agent dynamically adjusts the Minimum Contention Window (DQL-CWA). Through this adaptation, it achieves near-optimal performance under varying network conditions.

Molia [23] developed TCP with Reinforcement Learning–based Adaptive Congestion Control (TCP-RLACC). This method integrates reinforcement learning to dynamically select the CWND growth function—linear, polynomial, or exponential—based on wireless network conditions. Implemented in ns 3, TCP RLACC improved throughput and packet delivery ratio compared to TCP Westwood+.

Xing and Shahzad [24] introduced AppSpec-RL, a deep reinforcement learning (DRL) framework for TCP. It allows different congestion objectives—such as minimizing latency or maximizing throughput—to be optimized dynamically. AppSpec-RL operates in a client-server architecture to maintain scalability and achieve superior performance across multiple metrics. Table 2 summarizes the key characteristics and outcomes of these methods.

2.3 Positioning of the Present Research

Despite recent advancements in machine learning–based DDoS effects mitigation, most existing approaches still rely on pre-trained models or static classification rules. These methods cannot make dynamic decisions in real time. This limitation becomes critical in decentralized and rapidly changing network environments, such as blockchain systems. Moreover, limited attention has been given to CC as a fundamental mechanism for alleviating the adverse effects of DDoS-induced traffic surges. Many studies focus solely on traffic classification or intrusion detection, without addressing the core issue of congestion caused by high-volume malicious traffic.

Similarly, in the area of CWND adaptive adjustment and control, RL—particularly Q-learning and deep RL—has demonstrated promising results in dynamic network settings. These approaches outperform traditional heuristic methods by learning optimal congestion window strategies from real-time feedback. However, they have been predominantly applied to single-path TCP scenarios or wireless access networks. Their applicability to decentralized blockchain networks, especially under adversarial conditions like DDoS attacks, remains underexplored.

Furthermore, many existing techniques struggle to scale efficiently. They often depend on manual configuration, which limits their effectiveness against complex and constantly evolving attack patterns. Impor-

Table 1. Comparison of recent DDoS mitigation approaches.

Ref	Environment	Core Technique	ML/DL	Key Benefits
[11]	Blockchain	Poaching Raptor Deep Neural Network	✓	>95% accuracy, smart-contract enforcement
[12]	IoT + SDN	SDN + Blockchain + ANN	✓	Decentralized trust, agile anomaly detection
[13]	Blockchain	Blockchain + Reversed TCP requests	–	>99% detection, fast attacker identification
[14]	Blockchain	Ensemble ML models	✓	Adaptive, low false positives
[15]	IoT + SDN	LR/KNN/MLP/RF classifiers	✓	∼99.99% detection
[16]	SDN	Deep Neural Network	✓	>99.9% accuracy, real-time mitigation
[17]	Blockchain + IoT	Systematic survey	–	Comprehensive taxonomy
[18]	Blockchain	Mixed methods: FNN, smart contracts, SDN	✓	Fuzzy NN detection, immutable logging

Table 2. Comparison of intelligent CWND iontrol techniques.

Ref	Context	Technique	CWND Task	Key Results
[21]	TCP/IP networks	Ensemble DRL (DQN/PPO/etc.)	CWND control + AQM	+4% throughput, −10.5% delay vs. Random Early Detection (RED)
[22]	Wireless networks	Deep Q-learning	MCW adjustment	Near optimal performance in MCW tuning
[23]	Wireless mesh networks	Reinforcement learning	CWND growth function selection	Throughput ↑, PDR ↑ vs. Westwood+
[24]	TCP client-server networks	DRL (AppSpec-RL framework)	Multi-objective CWND adaptation	Outperforms specialized CC schemes

tantly, a critical yet under-addressed aspect of DDoS attacks is their ability to induce severe network congestion, significantly degrading performance. Current solutions rarely incorporate adaptive congestion control mechanisms that respond to real-time variations in traffic load and network conditions.

In summary, a clear research gap exists in developing intelligent, scalable, and decentralized methods that combine resilience to adversarial (DDoS-induced) effects with real-time congestion management. Our research aims to fill this gap by introducing a Q-learning–based congestion control strategy specifically designed for blockchain networks. The proposed method dynamically adjusts transmission behavior based on real-time network state observations. It should therefore be regarded as a reinforcement learning–based congestion control mechanism that integrates a resilience to adversarial conditions perspective (DDoS-induced congestion), rather than as direct malicious traffic filtering.

2.4 Integration of the Proposed RL-based Congestion Control with Blockchain Architectures

The proposed RL-based congestion control algorithm is primarily designed to function at the network (peer-to-peer) layer of blockchain systems. This layer is responsible for disseminating blocks, transactions, and state updates among nodes. By employing Q-learning, our algorithm dynamically regulates the congestion window and adapts the transmission rate. This helps reduce packet loss and ensures efficient bandwidth utilization, even under adversarial high-load conditions such as DDoS-induced congestion. In terms of the consensus process, the algorithm is consensus-agnostic—it does not directly intervene in consensus protocols such as PoW, PoS, or PBFT. However, by improving the stability and timeliness of message delivery across the network, it indirectly supports faster block propagation and more reliable consensus finalization.

3 Proposed Methodology

The primary objective of this research is to design and implement an innovative RL-based CC method to mitigate congestion under DDoS-induced adversarial conditions within blockchain networks. This approach utilizes intelligent congestion control principles to regulate network traffic and reduce the adverse effects of such attacks. The following sections provide a detailed explanation of the methodology's operation and its novel contributions.

3.1 Environment and Its Features

In this model, the blockchain network is conceptualized as the learning environment. The environment consists of multiple nodes operating in a distributed network. Network traffic includes both legitimate packets and malicious traffic generated by DDoS attacks. A fundamental characteristic of this environment is its decentralized architecture, meaning that no central node is responsible for decision-making. Therefore, the proposed method requires a decentralized mechanism that can operate autonomously at each node. At the same time, it should collectively enhance overall network performance.

The environment is described using the following key parameters:

(1) **CWND:** Indicates the transmission capacity of each node.
(2) **Packet Loss Rate:** Indicates network congestion or potential attacks.
(3) **Delay Rate:** Represents congestion levels across network paths.
(4) **Throughput:** Quantifies overall network efficiency.
(5) **Variability:** Captures short-term fluctuations in key performance parameters (e.g., delay or throughput) to represent dynamic changes in network conditions. It is computed as the normalized standard deviation over a moving observation window, providing the RL agent with awareness of temporal instability or volatility in the environment.

These parameters serve as input states to the reinforcement learning algorithm and form the foundation for its decision-making process.

3.2 Agent Actions

The learning agent in this model is capable of performing two principal actions:

(1) **Increase the Congestion Window Size:** Executed under low traffic conditions to improve throughput.
(2) **Decrease the Congestion Window Size:** Triggered upon detection of congestion or attack conditions to alleviate network load.

These actions are strategically formulated to detect and manage malicious traffic while maintaining efficient utilization of network resources.

3.3 Reward Function Design

One of the most innovative aspects of this research is the design of the reward function, which aims to optimize network performance while minimizing the impact of attacks. The reward function is formulated as Equation 1:

$$R = \eta \cdot T - \beta \cdot D - \lambda \cdot P - \mu \cdot V \tag{1}$$

where:

- T: Throughput
- D: Delay
- P: Packet Loss Rate
- V: Variability of congestion signals over time (temporal consistency penalty)
- η, β, λ, and μ: Adjustable weights that specify the relative importance of each metric.

The additional term V captures short-term fluctuations in congestion indicators (e.g., sudden spikes in delay or packet loss). Low-rate or pulsing DDoS attacks, such as Shrew attacks [25], often manifest as intermittent bursts that periodically reduce the congestion window of legitimate flows without sustaining continuous congestion. By incorporating V, the reward function penalizes inconsistent or highly variable congestion patterns less severely than sustained congestion. This design allows the RL-agent to distinguish between transient anomalies and persistent congestion, thereby improving robustness against low-rate adversarial conditions while maintaining sensitivity to genuine congestion events.

This formulation enables the agent to prioritize decisions that maximize throughput, minimize delay, and reduce packet loss.

3.4 RL-Based Congestion Control Under DDoS Conditions

One of the main motivations behind the proposed RL-based congestion control algorithm is to handle severe congestion that occurs during DDoS attacks. These attacks generate large volumes of malicious traffic. This excessive load overwhelms network resources and increases both packet loss and delay.

The proposed approach addresses this problem by dynamically adjusting the CWND size in real time. It relies on congestion indicators such as packet loss and latency to guide its adjustments. Persistent or abnormal increases in these parameters are interpreted as potential symptoms of DDoS-induced congestion. The agent interprets an increase in packet loss or delay as a potential sign of a DDoS event. When this occurs, it immediately reduces the CWND, lowering the node's transmission rate. This response helps to alleviate overall network pressure and mitigate the impact of such anomalies on throughput and stability. Through continuous feedback from the environment, the RL agent improves its responsiveness to abnormal congestion patterns and enhances the overall robustness of congestion control under high-load conditions, while operating without differentiating between legitimate and malicious flows.

Unlike traditional defense mechanisms that rely on explicit packet filtering or static thresholds, this model enables each node to autonomously and intelligently react to emerging congestion patterns. Because congestion is a key symptom of DDoS activity, this strategy does not directly filter malicious packets. Instead, it manages network behavior under adversarial, DDoS-like conditions at the transport layer.

Therefore, the CWND adjustment by the agent

should not be viewed merely as a generic congestion response. Instead, it represents a reinforcement learning–driven congestion control mechanism that incorporates a DDoS-aware perspective, focusing on congestion effects rather than traffic classification. Consequently, the proposed method provides a practical bridge between congestion management and resilience against DDoS-induced congestion surges in distributed blockchain networks.

4 Implementation of the RL-based Algorithm

This study employs the Q-learning algorithm [26] to derive optimal policies. Q-learning is a lightweight and effective reinforcement learning technique that does not require complex modeling and learns incrementally. The algorithm proceeds through the following stages:

(1) **Initialization:** The Q-Table is initialized with zeros.
(2) **Interaction Generation:** At each iteration, the agent selects an action based on the ε-greedy policy.
(3) **Reward Calculation:** The chosen action is executed within the environment, resulting in a new state and an associated reward, which are returned to the agent.
(4) **Q-Table Update:** The Q-Table is updated according to the following rule:

$$Q(s,a) \leftarrow Q(s,a) + \eta \left[R + \lambda \cdot \max_{a'} Q(s',a') - Q(s,a) \right] \tag{2}$$

where:

- η: Learning rate
- λ: Discount factor
- Q(s', a'): The Q-value for state s and action a
- $\max_{a'}$ Q(s', a'): The maximum Q-value in the new state for all possible actions
- s: Current state
- a: Selected action
- R: Received reward
- s': New state
- a': New action

This iterative process continues until the agent learns an optimal decision-making policy for different network conditions. The detailed algorithm for adjusting the next CWND value is presented in Algorithm 1.

Algorithm 1 RL-based CC (cwnd Adjustment)

Input:
  s_t: Current state (e.g., {cwnd, delay, throughput, packet_loss_rate, variability})
  t: Current time step
  η, λ: Learning rate and discount factor
  ϵ: Exploration rate

Output:
  Updated congestion window $cwnd(t)$

Initialize:
  $Q(s,a) = 0$ for all s and a.
  Define reward $r_t(s,a)$ based on throughput, delay, packet loss, and variability.

At $t = 0$:
  a. Observe initial state s_0.
  b. Initialize $cwnd(0)$ with a default or predefined value.

At each time step t:
  a. Observe current state s_t (e.g., network metrics).
  b. Select action a_t using ϵ-greedy:
    With probability ϵ, select a random action (exploration).
    Otherwise, select $a_t = \arg\max_a Q(s_t, a)$ (exploitation).
  c. Execute action a_t:
    Adjust $cwnd$:
    Increase, decrease, or maintain window size based on a_t
  d. Observe new state s_{t+1} and reward $r_t(s_t, a_t)$.
    Update Q-function:
      Update Q-Values:
      $Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \eta\left[r_t + \lambda \cdot \max_{a'} Q(s_{t+1}, a'_t) - Q(s_t,a_t)\right]$
    Adjust $cwnd(t)$:
    Use $cwnd(t) = f(Q(s,a))$, where f directly computes adjustments.
    Example: Simple mapping from $\arg\max Q(s_t,a)$ to $cwnd(t)$

Repeat Steps 3–5 for each time step in the simulation.
Return $cwnd(t)$: Output updated congestion window size at each step.

4.1 Advantages of the Proposed Method

The proposed method offers several notable advantages, as outlined below:

- **Adaptability to Changing Network Conditions:** By leveraging reinforcement learning, the method dynamically adapts to fluctuations in network conditions, enhancing its responsiveness to changing environments.
- **Improved Network Efficiency:** Through in-

telligent and real-time adjustment of the congestion window size, the proposed method increases throughput. As a result, it optimizes overall network performance and resource utilization.
- **Distributed Mechanisms:** The approach functions in a decentralized manner, eliminating single points of failure and strengthening the robustness of the entire system.
- **Dynamic Congestion Management under Adversarial Conditions:** Instead of detecting and filtering malicious traffic, this model proactively strengthens congestion control. This reinforcement improves network resilience against sustained high-load, DDoS-like conditions.

5 Results and Discussion

This section presents the results of simulations evaluating the performance of the proposed method. The main objective of these simulations is to compare the effectiveness of the proposed approach against various existing DDoS mitigation techniques within blockchain networks. To this end, a series of experiments was conducted under low, medium, and high attack intensities.

These experiments assess key performance metrics, including throughput, packet loss rate, network delay, bandwidth utilization, and network stability, providing a comprehensive evaluation of the RL-based congestion control algorithm's effectiveness.

5.1 Experiment Configuration

In this section, we describe the simulation environment and experimental settings used to evaluate the performance of the proposed RL-based congestion control algorithm.

The blockchain network was simulated using a distributed topology consisting of 10 legitimate nodes and two attacker nodes. To simulate a realistic network environment, a heterogeneous traffic mix was generated. It consisted of 80% TCP flows and 20% UDP flows. Legitimate traffic was composed of transaction requests and block propagation messages.

DDoS attacks were introduced at different intensity levels—low, medium, and high. They were triggered at varying time intervals to reflect diverse attack dynamics. The attacker nodes continuously generated high volumes of malicious traffic, including connection requests and dummy transactions, causing artificial congestion and bandwidth exhaustion.

To ensure clarity and realism, the DDoS intensity levels were quantitatively defined as follows:

- **Low Attack:** 5,000 packets/sec for 10 seconds
- **Medium Attack:** 10,000 packets/sec for 20 seconds
- **High Attack:** 20,000 packets/sec for 30 seconds

These attacks combined TCP- and UDP-based flooding patterns. They were executed at randomized time intervals to simulate typical bursty and unpredictable DDoS behavior, including SYN floods and slow-rate attacks.

To simulate a realistic adversarial environment, traffic characteristics such as burstiness and variability were introduced to mimic common DDoS patterns, including flooding and slow-rate attacks.

The performance of the proposed method was compared with both classical and recent congestion control techniques. The classical methods include five baseline algorithms: RED, TCP Reno, NewReno, Tahoe, and standard TCP. RED [27] is a proactive congestion control mechanism that randomly drops packets before buffer overflow. TCP and its variants—Reno, NewReno, and Tahoe—represent classic transport protocols with predefined congestion response strategies such as fast retransmit and recovery [28].

In addition to these, we included three recent RL-based approaches from the literature: (i) WEDRL [21], (ii) DQL-CWA [22], and (iii) AppSpec-RL for TCP Congestion Control [24]. These approaches were selected for their relevance to dynamic congestion control. Their recent contributions to the field are further discussed in Section 2.2.

These baseline methods were selected not only for their role as classical congestion control mechanisms but also because they are commonly targeted or exploited by attackers in DDoS scenarios. As such, evaluating the proposed method against them helps to assess both its congestion resilience and security relevance.

The simulation was implemented entirely in Python. Each scenario was run over 100 iterations to ensure statistical robustness. Simulation parameters were selected based on real-world distributed network characteristics:

- **Default bandwidth:** 100 Mbps
- **Round-Trip Time (RTT):** randomized between 50–150 ms
- **Queue capacity:** 500 packets

Performance metrics included throughput, delay, packet loss rate, bandwidth utilization, and stability under attack. The simulation measured performance across the entire network (not individual flows), en-

suring that the evaluation reflects overall network behavior under attack conditions.

5.2 Results of Experiments

This section presents the simulation results for each performance criterion evaluated.

5.2.1 Throughput

Throughput is a critical performance metric in blockchain networks. The simulation results indicate that the proposed method significantly enhances throughput under DDoS attack conditions.

Figure 1 illustrates the throughput performance under various DDoS attack intensities across different congestion control algorithms.

Compared with other methods, the RL-based algorithm adaptively detects congestion and optimizes the congestion window size. This adaptive behavior leads to a significant increase in data transmission rates.

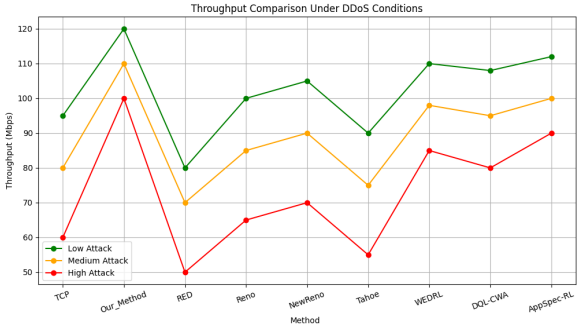

Figure 1. Comparison of throughput across blockchain networks with different CC methods under different DDoS attacks.

5.2.2 Packet Loss Rate

A major consequence of DDoS attacks is increased packet loss. As illustrated in Figure 2, the proposed method demonstrates a significant reduction in packet loss rate under attack conditions. Compared with traditional approaches, the RL-based method consistently maintains a lower packet loss rate, highlighting its effectiveness in managing congestion and preserving data integrity during DDoS attacks.

5.2.3 Delay

Delay is another important metric that reflects the impact of DDoS attacks on network performance. As shown in Figure 3, the proposed method effectively reduces network delay under various attack conditions. Traditional methods such as NewReno and RED show limited responsiveness under severe DDoS conditions. As a result, they cause increased latency and slower recovery. In contrast, RL-based

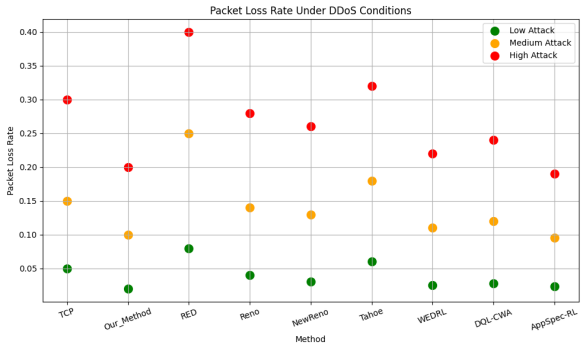

Figure 2. Comparison of packet loss rates in blockchain networks with different methods under DDoS attacks.

approaches show greater adaptability to congestion. They achieve more stable operation and lower delay performance. Among all compared methods, the proposed Q-learning-based strategy consistently achieves the lowest latency across different DDoS intensities, confirming its effectiveness in maintaining network responsiveness under attack.

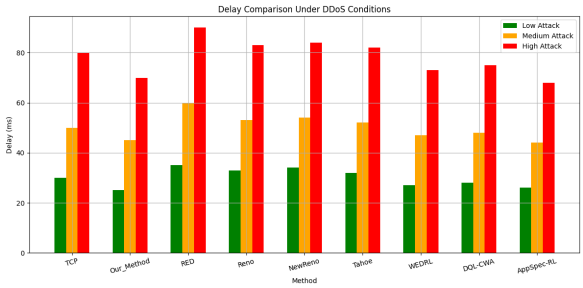

Figure 3. Comparison of delay across blockchain networks with different CC methods under different DDoS attacks.

5.2.4 Bandwidth Utilization

Optimizing bandwidth utilization is another key performance indicator. Under DDoS conditions, the proposed method demonstrates efficient use of available network bandwidth.

The RL-based algorithm intelligently allocates bandwidth by adapting to varying traffic patterns, thereby ensuring consistent performance even under heavy attack loads. Figure 4 presents the bandwidth utilization across different DDoS intensities using both classical and modern RL-based congestion control algorithms.

As shown, our proposed method consistently achieves the highest utilization rates across all attack levels, highlighting its ability to adaptively manage bandwidth resources in real time and outperform both legacy and contemporary approaches.

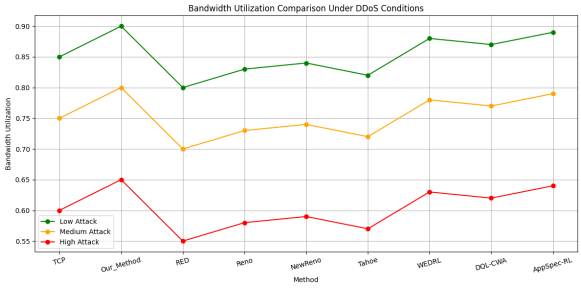

Figure 4. Comparison of bandwidth utilization across blockchain networks with different CC methods under different DDoS attacks.

5.2.5 Network Stability

Network stability refers to the network's ability to maintain consistent performance under high traffic volumes and attack scenarios. The RL-based algorithm enhances stability by dynamically adjusting the congestion window and detecting DDoS attacks in real time.

These additional comparisons enable us to benchmark our proposed method not only against traditional protocols but also against modern intelligent congestion control techniques.

As shown in Figure 5, the proposed RL-based method consistently achieves the highest network stability across low, medium, and high attack intensities. These results confirm that the proposed approach adapts effectively to DDoS-induced fluctuations while maintaining reliable performance in decentralized environments.

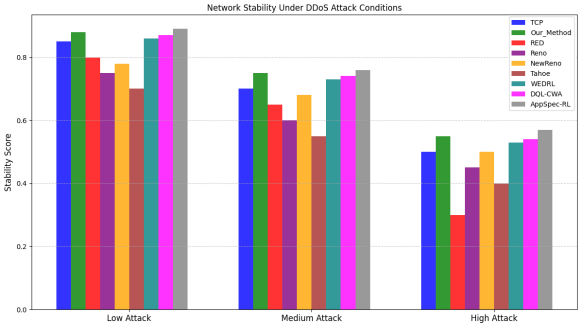

Figure 5. Comparison of network stability across blockchain networks with different CC methods under different DDoS attacks.

5.3 Comparative Analysis

For a comprehensive comparative analysis, the performance of the proposed RL-based method was evaluated against both classical and recent reinforcement learning-based congestion control techniques across multiple key metrics. Specifically, the benchmark methods include TCP, RED, Reno, NewReno, Tahoe, and three recent RL-based approaches: WEDRL, DQL-CWA, and AppSpec-RL.

As shown in Figure 6, the proposed method outperforms all compared techniques across throughput, delay, packet loss rate, bandwidth utilization, and network stability.

Among the RL-based baselines, our method demonstrates superior adaptability and robustness under DDoS conditions, validating its practical advantage in securing decentralized networks.

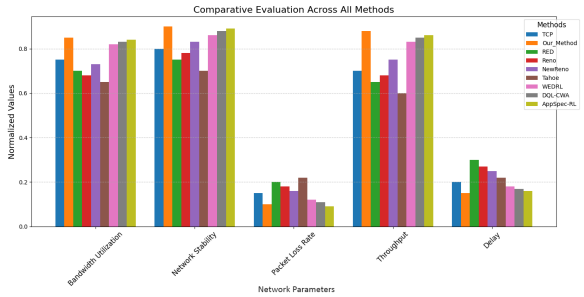

Figure 6. Comparison of all network parameters in different methods.

These findings highlight the effectiveness of the algorithm in managing congestion and mitigating DDoS impacts in distributed and blockchain environments. They also emphasize its relevance for real-time and adaptive security strategies.

6 Future Works

Future research could focus on enhancing the underlying reinforcement learning models, adapting the solution to real-world distributed and blockchain environments, and expanding its applicability to other forms of cyberattacks. In addition, addressing computational efficiency and scalability challenges in large-scale distributed and blockchain networks remains a critical area for development. Integrating the proposed techniques with existing blockchain protocols may lead to more robust, scalable, and secure systems for congestion management and resilience against DDoS-induced traffic surges.

Moreover, the framework can be extended through the integration of smart contracts in future designs. A dedicated monitoring contract could be deployed to record abnormal network behavior, such as transaction floods indicative of high-load or adversarial conditions, and to trigger adaptive responses via the RL-based module. This approach provides a decentralized and tamper-resistant mechanism for automated congestion management. It focuses on enhancing network resilience rather than directly detecting or filtering attacks.

7 Conclusion

This study introduces a reinforcement learning-based congestion control algorithm to manage DDoS-induced congestion rather than directly filtering malicious traffic in blockchain and distributed networks. Simulation results show that the proposed method adapts dynamically to congestion signals. It responds effectively to both normal load fluctuations and attack-driven surges. As a result, it improves network performance across key metrics such as throughput, delay, packet loss rate, and stability.

When compared with existing techniques, the RL-based approach offers superior flexibility and efficiency, particularly under dynamic network conditions. Its comparative evaluation includes both traditional TCP-based protocols and three recent RL-based approaches, showing consistent improvements across all evaluation criteria. Its adaptive mechanism enables real-time adjustments to shifting traffic patterns and supports resilience against congestion that may result from adversarial high-load conditions (e.g., DDoS-induced congestion). However, further research is necessary to assess the method's performance in real-world scenarios, optimize computational resource usage, and address evolving attack strategies. This work lays the groundwork for future advancements to strengthen the security and resilience of blockchain and general-purpose distributed infrastructures. These findings are expected to support ongoing efforts to enhance network robustness against congestion and traffic overloads such as those triggered by DDoS attacks.

By employing the Q-learning algorithm, the model continuously analyzes network state parameters such as throughput, delay, and packet loss, and dynamically adjusts the CWND size in line with TCP principles based on congestion indicators.

Unlike traditional methods that passively respond to congestion, the proposed approach interprets persistent or abnormal congestion patterns as potential indicators of DDoS activity. It reacts by adjusting transmission rates more intelligently to alleviate their effects.

Simulation results confirm that the proposed method significantly improves throughput and bandwidth utilization while reducing delay and packet loss, particularly under adverse conditions exacerbated by attack traffic.

Our method provides a dual contribution. First, it functions as a congestion control mechanism that operates effectively under adversarial (DDoS-induced) conditions. Second, it offers a scalable and intelligent solution that strengthens the resilience of distributed and blockchain networks.

Acknowledgment

The authors would like to thank the editorial team and the anonymous reviewers for their valuable comments.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest to disclose.

References

[1] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] S. Singh *et al.*, Blockchain Security Attacks, Challenges, and Solutions for the Future Distributed IoT Network, IEEE Access, vol. 9, pp. 13938–13959, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3051602

[3] C. D. Morar and D. E. Popescu, A Survey of Blockchain Applicability, Challenges, and Key Threats, *Computers*, vol. 13, no. 9, p. 223, Sept. 2024. [Online]. Available: https://doi.org/10.3390/computers13090223

[4] J. Liu and J. Wu, A Comprehensive Survey on Blockchain Technology and Its Applications, Highlights in Science, Engineering and Technology, vol. 85, pp. 128–138, Mar. 2024. [Online]. Available: https://doi.org/10.54097/r0ggyr24

[5] L. Wang *et al.*, Security and Privacy Issues in Blockchain and Its Applications, IET Blockchain, vol. 3, no. 4, pp. 169–171, Dec. 2023. [Online]. Available: https://doi.org/10.1049/blc2.12051

[6] X. Li *et al.*, Blockchain Security Threats and Collaborative Defense: A Literature Review, Computers, Materials & Continua, vol. 76, no. 3, pp. 2597–2629, 2023. [Online]. Available: https://doi.org/10.32604/cmc.2023.040596

[7] R. Chaganti *et al.*, A Survey on Blockchain Solutions in DDoS Attacks Mitigation: Techniques, Open Challenges and Future Directions, Computer Communications, vol. 197, pp. 96–112, Jan. 2023. [Online]. Available: https://doi.org/10.1016/j.comcom.2022.10.026

[8] M. Conti *et al.*, A Survey on Security and Privacy Issues of Bitcoin, IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 3416–3452, 2018. [Online]. Available: https://doi.org/10.1109/COMST.2018.2842460

[9] R. Chaganti *et al.*, A Comprehensive Review of Denial of Service Attacks in Blockchain Ecosystem and Open Challenges, IEEE Access, vol. 10, pp. 96538–96555, 2022. [Online]. Available: https://doi.org/10.1109/ACCESS.2022.3205019

[10] P. Kamboj *et al.*, Detection Techniques of DDoS Attacks: A Survey, in Proc. IEEE UPCON, 2017, pp. 675–679. [Online]. Available: https://doi.org/10.1109/UPCON.2017.8251130
[11] B. Ilyas *et al.*, Prevention of DDoS Attacks Using an Optimized Deep Learning Approach in Blockchain Technology, Trans. Emerging Telecommunications Technologies, vol. 34, no. 4, p. e4729, Apr. 2023. [Online]. Available: https://doi.org/10.1002/ett.4729
[12] R. Jmal *et al.*, Distributed Blockchain-SDN Secure IoT System Based on ANN to Mitigate DDoS Attacks, Applied Sciences, vol. 13, no. 8, p. 4953, Apr. 2023. [Online]. Available: https://doi.org/10.3390/app13084953
[13] A. A. Abdullah and S. A. Hussein, Detection and Mitigation Distribution Denial of Service Attack Based on Blockchain Concept, Ingénierie Des Systèmes d'Information, vol. 29, no. 3, pp. 1043–1049, Jun. 2024. [Online]. Available: https://doi.org/10.18280/isi.290322
[14] X. Li *et al.*, An Adaptive DDoS Detection and Classification Method in Blockchain Using an Integrated Multi-Models, Computers, Materials & Continua, vol. 77, no. 3, pp. 3265–3288, 2023. [Online]. Available: https://doi.org/10.32604/cmc.2023.045588
[15] Kavitha and Ramalakshmi, Machine Learning-Based DDOS Attack Detection and Mitigation in SDNs for IoT Environments, Journal of the Franklin Institute, vol. 361, no. 17, p. 107197, Nov. 2024. [Online]. Available: https://doi.org/10.1016/j.jfranklin.2024.107197
[16] V. Hnamte *et al.*, DDoS Attack Detection and Mitigation Using Deep Neural Network in SDN Environment, Computers & Security, vol. 138, p. 103661, Mar. 2024. [Online]. Available: https://doi.org/10.1016/j.cose.2023.103661
[17] Z. Shah *et al.*, Blockchain Based Solutions to Mitigate Distributed Denial of Service (DDoS) Attacks in the Internet of Things (IoT): A Survey, Sensors, vol. 22, no. 3, p. 1094, Jan. 2022. [Online]. Available: https://doi.org/10.3390/s22031094
[18] S. Wani *et al.*, Distributed Denial of Service (DDoS) Mitigation Using Blockchain—A Comprehensive Insight, Symmetry, vol. 13, no. 2, p. 227, Jan. 2021. [Online]. Available: https://doi.org/10.3390/sym13020227
[19] V. Jacobson, Congestion Avoidance and Control, ACM SIGCOMM Computer Communication Review, vol. 18, no. 4, pp. 314–329, Aug. 1988. [Online]. Available: https://doi.org/10.1145/52325.52356
[20] S. Ha *et al.*, CUBIC: A New TCP-Friendly High-Speed TCP Variant, ACM SIGOPS Operating Systems Review, vol. 42, no. 5, pp. 64–74, Jul. 2008. [Online]. Available: https://doi.org/10.1145/1400097.1400105
[21] M. H. Ali and S. Öztürk, Efficient Congestion Control in Communications Using Novel Weighted Ensemble Deep Reinforcement Learning, Computers and Electrical Engineering, vol. 110, p. 108811, Sep. 2023. [Online]. Available: https://doi.org/10.1016/j.compeleceng.2023.108811
[22] A. Kumar *et al.*, Adaptive Contention Window Design Using Deep Q-Learning, arXiv, 2020. [Online]. Available: https://doi.org/10.48550/ARXIV.2011.09418
[23] H. K. Molia, Reinforcement Learning Based Adaptive Congestion Control for TCP over Wireless Networks, Int. J. of Computer Networks and Applications, vol. 11, no. 5, pp. 607–616, Oct. 2024. [Online]. Available: https://doi.org/10.22247/ijcna/2024/39
[24] J. Xing and M. Shahzad, A Reinforcement Learning Framework for Application-Specific TCP Congestion-Control, arXiv, 2025. [Online]. Available: https://doi.org/10.48550/ARXIV.2505.07042
[25] N. Gogoli *et al.*, Shrew DDoS Attack Detection Based on Statistical Analysis, The ISC International Journal of Information Security, no. Online First, Jun. 2024. [Online]. Available: https://doi.org/10.22042/isecure.2024.420803.1032
[26] B. Jang *et al.*, Q-Learning Algorithms: A Comprehensive Classification and Applications, IEEE Access, vol. 7, pp. 133653–133667, 2019. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2941229
[27] S. Floyd and V. Jacobson, Random Early Detection Gateways for Congestion Avoidance, IEEE/ACM Trans. on Networking, vol. 1, no. 4, pp. 397–413, Aug. 1993. [Online]. Available: https://doi.org/10.1109/90.251892
[28] K. Fall and S. Floyd, Simulation-Based Comparisons of Tahoe, Reno and SACK TCP, ACM SIGCOMM Computer Communication Review, vol. 26, no. 3, pp. 5–21, Jul. 1996. [Online]. Available: https://doi.org/10.1145/235160.235162

Ehsan Abedini is currently pursuing a Ph.D. Candidate in the Department of Computer Engineering at the University of Qom, Qom, Iran. He received his M.Sc. degree in information technology engineering in 2016. His recent research interests include computer networks and computer security.

Amir Jalaly Bidgoly received his M.Sc. degree in Software Engineering from the Iran University of Science and Technology (IUST) in 2009 and his Ph.D. degree in Software Engineering from the University of Isfahan, Iran, in 2015. He is currently an Associate Professor with the Department of Computer Engineering at the University of Qom, Iran. His research interests include computer security and machine learning.

Mohsen Nickray received his B.Sc. and M.Sc. degrees in Computer Engineering from the Iran University of Science and Technology and the University of Tehran in 2004 and 2007, respectively. He obtained his Ph.D. degree in Computer Architecture from the University of Tehran in 2012. Currently, he is an assistant professor in the Department of Computer Engineering at the University of Qom, Iran. His recent research interests include resource management and task scheduling in Cloud and Fog computing.