Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Adaptive and Resilient T-PBFT Consensus for Internet of Vehicles Using Multi-Objective Optimization and Reinforcement Learning

**Allagui Maissa**
University of Kairouan
**Harrabi Samira**
samira.harrabi@gmail.com
University of Kairouan

**Research Article**

**Keywords:** Internet of Vehicles, Blockchain, T-PBFT, Multi-Objective Optimization, NSGA-III, Reinforcement Learning, Trust Management

**Posted Date:** November 3rd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11216673/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Adaptive and Resilient T-PBFT Consensus for Internet of Vehicles Using Multi-Objective Optimization and Reinforcement Learning

**Allagui Maissa[1*] and Harrabi Samira[2]**

*Strategies for Modeling and Artificial Intelligence (SMART), Higher Institute of Management, Tunis, University of Tunis, 41, Liberty Street, Urban Area Bouchoucha, 2000 Bardo, Tunis, Tunisia. .*
*[1*]Higher Institute of Computer Science of Kairouan (ISIGK), University of Kairouan, Tunisia..*
*[2]Higher School of Commerce of Tunis (ESCT), University of Manouba, Tunisia..*

Corresponding author(s). E-mail(s): allagui.contact@gmail.com;
*Contributing authors: samira.harrabi@gmail.com;

**Abstract**

The Internet of Vehicles (IoV) enables continuous communication among vehicles and roadside infrastructures, but its highly dynamic conditions, changing network density, resource constraints and exposure to attacks make efficient blockchain consensus challenging. This paper proposes an adaptive T-PBFT consensus framework that combines a trust evaluation mechanism inspired by the principles of EigenTrust, the NSGA-III multi-objective optimization algorithm and Deep Q-Network (DQN) reinforcement learning. The proposed approach jointly considers validator selection and T-PBFT configuration while optimiz-ing consensus latency, communication overhead, computational cost, security risk, validator trust and availability. NSGA-III generates candidate configura-tions representing different trade-offs among these objectives, while the DQN agent selects a suitable configuration according to the current network state. The framework is evaluated through a Python-based simulation under nor-mal, high-density, Sybil-attack, denial-of-service (DoS) and combined stress scenarios. Five independent random seeds are used for each scenario and the results are compared with a fixed T-PBFT configuration and an NSGA-III-only approach. The results show that adaptive approaches substantially improve consensus performance under non-DoS conditions, while the hybrid approach

provides scenario-dependent benefits, particularly under Sybil attacks. How-ever, under severe DoS conditions, all evaluated methods experience complete consensus failure. These findings demonstrate the potential of combining multi-objective optimization and reinforcement learning for adaptive consensus, while highlighting the need for stronger attack-resilience mechanisms.

**Keywords:** Internet of Vehicles, Blockchain, T-PBFT, Multi-Objective Optimization, NSGA-III, Reinforcement Learning, Trust Management

# 1 Introduction

The Internet of Vehicles (IoV) is becoming an important component of intelligent transportation systems, enabling vehicles, roadside units (RSUs) and edge infrastructures to continuously exchange information. This connectivity supports applications such as traffic management, cooperative driving, and road safety. However, IoV networks are highly dynamic: vehicles are constantly moving, network density may change rapidly and communication and computational resources can vary across the network. These characteristics make it challenging to maintain efficient and secure data exchange when consensus decisions must be completed under strict time constraints [1, 2].

Blockchain technology has attracted considerable interest for improving data integrity, traceability and trust among distributed entities. In IoV environments, consortium blockchains are particularly relevant because participation in the consen-sus process can be restricted to. authorized nodes, such as RSUs or edge entities. Trust management can further support this architecture by helping distinguish reli-able participants from potentially malicious nodes [3, 4]. Nevertheless, achieving efficient blockchain consensus in a dynamic vehicular environment remains challenging, particularly when communication and computational overhead must be controlled.

Practical Byzantine Fault Tolerance (PBFT) and its variants are widely considered for permissioned blockchain systems because they provide Byzantine fault tolerance without the computationally intensive mining process used by proof-of-work. However, their communication cost increases with the number of participating validators, which can lead to higher latency and resource consumption in dynamic IoV networks. Several studies have therefore investigated trust-aware validator selection and improved PBFT mechanisms to mitigate these limitations [5–7]. These studies indicate that both validator composition and node trustworthiness can significantly affect consensus performance.

A major challenge is that a consensus configuration suitable under one network condition may become inefficient when the environment changes. For example, a highly trusted validator may experience increased computational load or communication delay, while changes in traffic density may affect the cost of a previously suitable committee configuration. Malicious nodes can further modify the trust and security conditions of the network. Consequently, static validator selection and predefined consensus parameters may be insufficient for highly dynamic IoV environments.

Reinforcement learning (RL) has recently been investigated to introduce greater adaptability into blockchain and IoV decision-making. Existing studies have explored learning-based consensus optimization and adaptive decision-making in blockchain and vehicular environments [8–11]. However, trust mechanisms and RL-based approaches address complementary aspects of the problem: trust mechanisms evalu-ate node reliability, whereas RL learns a decision policy from the observed network state. Moreover, consensus configuration involves several conflicting objectives. Reduc-ing the validator committee size may decrease communication overhead and latency, whereas maintaining a sufficiently reliable committee is important for security and availability. Sybil and DoS/DDoS attacks further increase this complexity by affecting node reliability and network conditions [12, 13].

These challenges motivate the formulation of consensus adaptation as a multi-objective decision problem rather than the optimization of a single performance metric. A multi-objective optimizer can explore trade-offs among latency, communication overhead, computational cost, trust, security, and availability, while a learning-based decision mechanism can select an appropriate configuration according to the current network state.

Based on this motivation, this paper proposes an adaptive and resilient blockchain consensus framework for IoV that combines a trust mechanism inspired by the principles of EigenTrust, NSGA-III and Deep Q-Network (DQN) reinforcement learning within a T-PBFT-based consensus process. The trust mechanism assesses vehicle reliability and supports trust-aware validator selection. NSGA-III jointly optimizes validator selection and T-PBFT configuration according to six objectives: consensus latency, communication overhead, computational cost, security risk, validator trust and validator availability. It generates candidate configurations representing different trade-offs among these objectives, while a DQN agent observes the current IoV state and selects a candidate configuration for the next consensus process.

The proposed framework is evaluated through a Python-based simulation under Normal, High Density, Sybil, DoS, and Combined Stress scenarios. Its behavior is compared with a Fixed T-PBFT configuration and an NSGA-III Only configuration. The evaluation considers consensus success rate, latency, communication overhead, computational cost, validator trust, security risk, and view-change frequency. Rather than assuming that the hybrid approach is uniformly superior, the evaluation characterizes the trade-offs introduced by combining multi-objective optimization with online configuration selection.

The main contributions of this work are as follows:

- An adaptive T-PBFT framework for IoV that integrates trust evaluation, multi-objective optimization and reinforcement learning within a consortium blockchain setting.
- A trust-aware validator selection mechanism in which validator identities and T-PBFT configuration parameters are jointly represented in the NSGA-III optimization process.
- A six-objective optimization formulation considering consensus latency, communication overhead, computational cost, security risk, validator trust and validator availability.

- A DQN-based online configuration selection mechanism that selects among candidate configurations according to the current IoV network state.
- An experimental evaluation under Normal, High Density, Sybil, DoS and Combined Stress scenarios, using comparisons with Fixed T-PBFT and NSGA-III Only configurations.

The remainder of this paper is organized as follows. Section 2 formulates the consensus problem considered in this work. Section 3 reviews related work and discusses the main research gap. Section 4 presents the proposed adaptive T-PBFT framework. Section 5 describes the simulation environment, experimental configurations, attack scenarios, baseline methods and evaluation metrics. Section 6 presents and discusses the experimental results. Finally, Section 7 concludes the paper and outlines future research directions.

# 2 Problem Statement

Achieving efficient and secure consensus in the Internet of Vehicles (IoV) is more challenging than in conventional permissioned blockchain networks. IoV environments are characterized by continuous vehicle mobility, frequent topology changes, variations in traffic density and heterogeneous communication and computational resources. These conditions directly affect the availability, reliability and performance of nodes participating in consensus. Consequently, latency, communication overhead, scalability and fault tolerance remain important challenges for blockchain-based consensus in vehicular environments [1, 2].

PBFT-based mechanisms are attractive for consortium blockchain systems because they provide Byzantine fault tolerance without relying on energy-intensive mining. However, their communication and computational costs increase with the number of participating validators, which can become problematic in dynamic IoV networks. Several studies have therefore investigated improved PBFT mechanisms and trust-aware validator selection. For example, SG-PBFT improves the efficiency and security of PBFT in intelligent IoV environments [6], while other approaches incorporate trust evaluation and dynamic participant management into the consensus process [5, 14]. These studies highlight the importance of validator selection, since the characteristics of participating nodes directly influence consensus performance and reliability.

However, validator selection cannot be considered independently of the changing network state. A vehicle that is suitable for consensus under one condition may experience increased computational load, communication delay, or reduced availability as network conditions evolve. Vehicle mobility may also modify connectivity and committee composition over time. In addition, malicious nodes introduce further security constraints: Sybil attacks can affect the trustworthiness of participating identities, whereas DoS/DDoS attacks can degrade network resources and prevent legitimate validators from completing consensus [4, 12].

These characteristics make fixed validator committees and static $T_{\mathrm{PBFT}}$ parameters unsuitable for rapidly changing IoV conditions. Although reinforcement learning has been investigated for adapting consensus decisions to changing network states [8, 11], consensus configuration remains inherently multi-objective. For example, reducing the

committee size can decrease communication and computational overhead but may affect validator trust and availability. Similarly, increasing the timeout can reduce unnecessary view changes while increasing consensus latency.

The problem addressed in this work is therefore the **joint adaptation of validator selection and $T_{\mathrm{PBFT}}$ configuration** in a dynamic and potentially adversarial IoV environment. At each decision point, the system must determine an appropriate set of vehicles to participate in consensus and select suitable consensus parameters according to the current network conditions. The decision process must simultaneously consider consensus latency, communication overhead, computational cost, validator trust, availability and security risk.

Accordingly, the research problem addressed in this study can be formulated as follows:

> **How can validator selection and $T_{\mathrm{PBFT}}$ parameters be dynamically adapted in a highly dynamic IoV environment to reduce consensus latency and resource con-sumption while maintaining high validator trust and availability and limiting security risks under Sybil and DoS/DDoS attacks?**

This problem is inherently **dynamic and multi-objective**. The suitability of a consensus configuration depends on time-varying factors such as vehicle density, network congestion, communication delay, computational load, validator trust and attack conditions. Therefore, rather than relying on a single static configuration, the objective is to identify feasible configurations representing different trade-offs among the conflicting objectives and to select an appropriate configuration as the IoV network evolves.

# 3 Literature Review

This section reviews recent research on blockchain-based consensus and trust management in the Internet of Vehicles (IoV), with particular emphasis on PBFT-based mechanisms, validator selection, adaptive consensus and reinforcement learning. It summarizes the main approaches proposed to improve consensus efficiency, scalability, trust and security in dynamic vehicular environments. The reviewed studies are then critically examined in terms of their objectives, methodologies, main findings and limitations, with particular attention to their ability to adapt to changing network conditions and adversarial environments. This analysis helps identify the remaining research gaps and provides the basis for the adaptive and multi-objective framework proposed in this work.

## 3.1 Existing Work Review

[6] addressed the high communication overhead and latency of conventional PBFT in dynamic IoV environments. They proposed **SG-PBFT**, a score-grouping PBFT mechanism that selects consensus nodes according to their trust-related scores and periodically replaces low-scoring nodes with higher-scoring candidates. The approach integrates blockchain, PBFT, score-based node selection, cryptographic authentication and a Vehicle–Service Provider–RSU architecture to improve consensus efficiency and

security. Simulation results showed that SG-PBFT significantly reduced transaction delay and communication overhead compared with conventional PBFT and other variants. However, its node selection relies on predefined scoring and replacement rules, limiting its ability to dynamically optimize consensus according to changing network conditions and multiple performance objectives.

[5] proposed C-PBFT to improve the security, efficiency and scalability of blockchain consensus in dynamic IoV environments. The mechanism combines regional and global PBFT with reputation-based committee selection, using the CRMWSL model to identify reliable validators and maintain an honest committee. Randomized selection, digital signatures, and the PBFT view-change mechanism further strengthen the protocol against various attacks. Experimental results demonstrated more accurate malicious-node detection, lower communication overhead and consensus latency, and higher throughput than conventional PBFT and TLPBFT, particularly as the network size increased. However, the approach relies mainly on reputation-based committee selection and predefined consensus mechanisms, limiting its ability to jointly optimize multiple performance objectives and adapt the consensus configuration to rapidly changing IoV conditions.

[14] addressed the high consensus latency, low transaction throughput and communication overhead of blockchain consensus in dynamic IoV environments. They proposed **ESBCP**, which integrates trust assessment, dynamic node partitioning, a Dynamic Unique Node List (DUNL), and an improved **DK-PBFT** consensus mechanism to select reliable nodes and organize consensus according to network conditions. The approach combines trust-aware node selection, dynamic network organization, RSU-based consensus and PBFT to improve the scalability and adaptability of blockchain consensus. Simulation results showed lower consensus latency, more stable throughput and effective malicious-node detection compared with CDBFT. However, the protocol mainly relies on predefined trust and clustering. mechanisms and does not jointly optimize multiple objectives or use a learning-based mechanism to dynamically adapt the consensus configuration to changing IoV conditions.

[15] proposed **AE-PBFT**, an adaptive consensus algorithm designed to address the high latency, limited fault tolerance, and poor adaptability of traditional PBFT in Internet of Vehicles (IoV) environments. Their approach combines a trust management model based on the historical behavior of participating nodes with a dynamic consensus group mechanism. The trust model progressively increases the influence of consistent node behavior, enabling the system to identify and filter potentially malicious nodes more efficiently. In addition, the size of the consensus group is dynamically adjusted according to the overall credit status of the network, allowing the consensus process to adapt to changing environmental conditions. Experimental results show that AE-PBFT reduces consensus latency and communication overhead while achieving higher throughput and better fault tolerance than PBFT, GPBFT, and CBFT. However, the adjustment of the consensus group mainly relies on predefined credit-based rules, which may limit its ability to optimize the consensus configuration under highly dynamic network conditions. Moreover, the approach does not jointly optimize multiple conflicting objectives, such as latency, communication overhead,

computational cost, and security. These limitations highlight the need for more flexible optimization and learning-based mechanisms capable of adapting the consensus process to the continuously changing. conditions of IoV networks.

[8] proposed a lightweight blockchain consensus protocol for resource-constrained IoT networks, combining Proof-of-Communication (PoC) with Multi-Agent Reinforcement Learning (MARL). The proposed approach uses MARL to optimize leader election in unreliable single-hop wireless networks, with the objective of improving both consensus efficiency and fairness. Simulation results showed that the learned policy substantially increased the leader-election success rate and achieved near-optimal consensus efficiency compared with randomized approaches. Despite these results, the evaluation was limited to a single-hop IoT network and did not consider the mobility, dynamic topology, resource heterogeneity, or security challenges inherent to IoV environments. Moreover, the use of reinforcement learning was primarily limited to leader election rather than the dynamic adaptation of the overall consensus configuration.

[10] proposed an AI-based approach for optimizing blockchain consensus by integrating. deep learning and reinforcement learning techniques. Their framework combines a Deep Neural Network (DNN) for extracting relevant features from network conditions and validator behavior with Deep Reinforcement Learning (DRL) for dynamically adapting consensus decisions. This combination enables the system to optimize validator selection and adjust consensus parameters according to variations in transaction load and network conditions. The experimental results demonstrated that the proposed AI-optimized model achieved an average confirmation latency of 320 ms, reduced CPU consumption to 30%, and reached an attack tolerance of 92%, outperforming traditional consensus mechanisms such as PoW, PoS, DPoS, and PBFT. These results highlight the potential of combining deep learning and reinforcement learning to improve blockchain scalability, computational efficiency and resilience against attacks. However, the approach depends on representative training data, while its initial training phase requires substantial computational resources, which may limit its applicability in highly dynamic or resource-constrained environments.

[11] address the challenge of maintaining efficient, secure, and stable blockchain consensus under dynamic network conditions and adversarial scenarios. They propose an AI-driven adaptive consensus framework that integrates **Proximal Policy Optimization (PPO)** with **dynamic graph representations** to adjust valida-tor participation and consensus parameters according to the current network state. The approach combines reinforcement learning, graph-based modeling, malicious-node detection, and Hyperledger Fabric to improve consensus adaptability. Its main con-tribution is the integration of AI directly into the consensus process rather than using it only for passive security detection. The framework was evaluated under **Sybil attacks, network congestion, node failures and critical errors**, using throughput, latency, detection rate, false-positive rate, energy consumption and computational load as evaluation metrics. The results demonstrated improvements in performance, security, stability and energy efficiency. However, the evaluation was mainly conducted in a controlled laboratory environment with synthetic scenarios and further validation on real-world networks and highly constrained edge devices is required.

[16] address the challenges of **trust management, scalability, and security** in dynamic Internet of Vehicles (IoV) environments. They propose **Blockchain-MLTrustNet**, a blockchain-based framework that combines **adaptive graph shard-ing and deep reinforcement learning (DRL)** to dynamically evaluate vehicle trust and improve blockchain scalability. The framework integrates DRL-based trust evaluation with adaptive sharding to reduce transaction latency and cross-shard com-munication overhead while improving the detection of malicious vehicles. Its main contribution is the integration of **trust management and adaptive blockchain scalability** within a unified IoV framework. The authors evaluate the approach using the **VeReMi dataset** and compare it with WGAN-SMOTE, MESMERIC, ICD-DRL, DYNASHARD, and an Ethereum-based approach using metrics such as latency, detection accuracy, trust score, cumulative reward and cross-shard overhead. The results show that the proposed framework achieves **98% detection accuracy** and lower cross-shard overhead than the compared approaches. However, its reliance on adaptive sharding and DRL introduces **computational complexity**, which may limit its applicability in large-scale or resource-constrained IoV environments.

[4] presented a comprehensive survey of security and trust management in the Internet of Vehicles (IoV), with particular emphasis on the challenges posed by malicious attacks, unreliable nodes, and the need for trustworthy communication. The authors reviewed existing security and trust-management approaches and investigated the potential of machine learning (ML) techniques to address these challenges. Their study provides a classification of ML methods and examines existing ML-based solutions for security and trust management in IoV environments. The main contribution of the work is to organize the existing research in this area, identify the major secu-rity requirements and attack types and highlight the role of ML in improving trust and security. Since the study is a survey rather than a newly proposed system, its evaluation is based on the analysis of existing approaches rather than on the imple-mentation and experimental validation of a new method. However, the survey does not introduce or experimentally evaluate a specific adaptive consensus mechanism, leaving room for further research on dynamic and resilient consensus solutions capable of adapting to the changing conditions of IoV networks.

[3] presented a comprehensive survey of blockchain-based trust management approaches for IoV, focusing on security, scalability and adaptability in dynamic vehicular environments. The authors reviewed trust computation, consensus mecha-nisms and blockchain architectures, highlighting a growing trend toward lightweight consensus and AI-based adaptive trust management. Their analysis identified persis-tent trade-offs between security, scalability, and performance and emphasized the need for adaptive mechanisms capable of responding to changing network conditions. In particular, reinforcement learning was identified as a promising approach for dynamic validator selection and resource management.

in terms of security features, computational cost and communication overhead. Although SRPP provides strong protection against several attacks with competitive computational performance, its communication overhead remains higher than some existing schemes and it relies on RSU assistance for pseudonym management.

## 3.2 Critical Discussion and Research Gap

Existing research has made significant progress toward improving blockchain consensus in IoV networks. Several PBFT-based approaches have introduced trust-aware validator selection, reputation mechanisms, dynamic committee organization and adaptive consensus groups to improve efficiency and reliability [5–7, 14, 15]. Other studies have investigated consensus mechanisms specifically designed for the mobility and dynamic characteristics of IoV networks [17, 18]. These studies demonstrate that the composition of the consensus committee and the trustworthiness of participating nodes can significantly affect latency, communication overhead, scalability, and security.

Despite these advances, existing trust-aware and PBFT-based approaches generally rely on predefined scoring, reputation, grouping, or rule-based mechanisms for selecting validators and configuring the consensus process [5, 6, 14, 15]. Such mechanisms can improve node reliability, but their ability to jointly adapt to rapidly changing vehicle density, mobility, communication conditions and computational resources remains limited. In addition, trust and security are often considered as specific components of the consensus process rather than being jointly optimized with performance and resource requirements [3, 4].

Reinforcement learning has provided another direction for adaptive blockchain decision-making. Existing studies have applied RL to consensus optimization, leader or validator selection and adaptive consensus management [8, 10, 11, 16]. These approaches demonstrate the potential of learning-based methods to adapt blockchain decisions according to changing network conditions. However, they generally focus on adapting specific decisions, such as leader selection, validator participation, or consensus parameters, rather than jointly determining **which nodes should participate in consensus and how the consensus process should be configured**. More-over, the considered objectives and network conditions differ across studies and the simultaneous optimization of performance, resource consumption, trust, availability, and security remains insufficiently explored.

Another important issue is the **multi-objective nature of consensus optimization in IoV**. The objectives considered in this context are inherently conflicting. A smaller validator committee can reduce communication and computational overhead, whereas a sufficiently reliable and appropriately sized committee is required to maintain trust, availability and resilience. Similarly, reducing consensus latency may require configurations that increase resource consumption. Therefore, optimizing consensus according to a single performance criterion is insufficient for highly dynamic IoV environments.

The literature consequently reveals a gap in the **joint adaptation of validator selection and consensus configuration as a multi-objective problem**. Exist-ing studies have separately addressed trust-aware validator selection, improved PBFT mechanisms, reinforcement learning, and adaptive blockchain management, but the simultaneous optimization of validator identities and $T_{\mathrm{PBFT}}$ parameters while considering latency, communication overhead, computational cost, trust, security risk and availability remains limited.

**This research addresses this gap by formulating validator selection and $T_{\mathrm{PBFT}}$ configuration as a joint multi-objective optimization problem. The**

**proposed framework combines a trust evaluation mechanism inspired by the principles of EigenTrust, NSGA-III for exploring different trade-offs among six objectives, and DQN-based online selection of candidate con-figurations according to the current IoV state. The framework is further evaluated under normal, high-density, Sybil, DoS and combined stress con-ditions to examine its behavior under both changing network conditions and adversarial scenarios.**

# 4 Proposed Approach

To address the limitations identified in existing IoV blockchain consensus approaches, this work proposes a hybrid adaptive framework that combines trust-aware validator selection, multi-objective optimization and reinforcement learning within a $T_{\mathrm{PBFT}}$ consensus process. The design is motivated by the dynamic and multi-dimensional nature of consensus management in IoV environments, where network density, communication conditions, computational load, trust and security conditions may vary over time [2].

The proposed framework operates at two complementary decision levels. At the first level, NSGA-III periodically explores the solution space and identifies non-dominated combinations of validator committees and $T_{\mathrm{PBFT}}$ parameters. At the second level, a Deep Q-Network (DQN) operates online and selects an appropriate candidate configuration according to the current network state. This separation avoids running the computationally more demanding evolutionary optimization at every decision step while allowing the consensus configuration to adapt to short-term variations in the IoV environment.

The overall workflow consists of four main stages: network-state and trust evaluation, multi-objective optimization, online configuration selection and consensus execution. The resulting consensus performance is then fed back to the adaptation process, forming a closed-loop framework.

## 4.1 System Architecture

The proposed framework consists of four main components: the IoV environment, the trust and network-state evaluation layer, the NSGA-III optimization module and the RL-based online adaptation module. These components operate around a consortium blockchain in which $T_{\mathrm{PBFT}}$ is responsible for executing the consensus process.

The IoV environment represents a dynamic vehicular network composed of vehicles and Roadside Units (RSUs). Vehicles generate and exchange transactions while their communication and computational conditions may change over time. Consequently, the suitability of a vehicle for consensus participation cannot be considered static. Variations in vehicle density, communication latency, congestion, computational load, node behavior and attack conditions can directly affect the composition and performance of the validator committee.

The trust and network-state evaluation layer provides the information required by both optimization and adaptation processes. A simplified trust evaluation mechanism

inspired by the principles of EigenTrust is used to estimate the reliability of participating vehicles from behavioral information and local trust values. In parallel, network indicators such as vehicle density, communication latency, congestion, computational load and attack-related conditions are monitored to characterize the current state of the IoV environment. Trust information is therefore incorporated into consensus management as an additional criterion for evaluating participating vehicles.

Based on the observed information, NSGA-III periodically searches for suitable combinations of validator committees and $T_{\mathrm{PBFT}}$ parameters. Instead of producing a single solution, the optimizer generates a set of non-dominated configurations representing different trade-offs among consensus latency, communication overhead, computational cost, security risk, validator trust and validator availability.

The resulting candidate configurations are provided to the RL adaptation module. At each decision step, the DQN observes the current network state and selects one configuration from the available action library. The selected configuration specifies both the validator committee and the corresponding $T_{\mathrm{PBFT}}$ parameters used for the subsequent consensus process.

After consensus execution, the resulting performance indicators are returned to the adaptation layer. These indicators include consensus latency, communication over-head, computational cost, validator trust, security risk and the consensus outcome. This feedback creates a closed adaptive loop in which the observed network state influences configuration selection, while the resulting consensus performance provides feedback for subsequent decisions. The overall architecture is illustrated in Fig. 1.

## 4.2 Trust and Validator Selection

Validator selection is a central component of the proposed framework because the composition of the consensus committee directly affects the performance and reliability of $T_{\mathrm{PBFT}}$. In a dynamic IoV environment, selecting validators only according to availabil-ity or computational capability may result in committees containing unreliable nodes. Conversely, relying exclusively on trust may produce inefficient configurations in terms of communication and computational requirements. Therefore, validator selection is formulated as part of the multi-objective optimization process.

Previous vehicular blockchain approaches have incorporated reputation, trust, or malicious-node detection into consensus management to improve the reliabil-ity of participating nodes [7]. Building on this principle, the proposed framework explicitly incorporates trust into the validator-selection process while simultaneously considering consensus performance and resource consumption.

Each vehicle $i$ is associated with a trust value $T_i$ representing its estimated reliability. The trust mechanism used in this work is a simplified evaluation mechanism inspired by the principles of EigenTrust rather than a direct implementation of the complete EigenTrust algorithm. It combines observed behavioral information with local trust information, allowing vehicles exhibiting less reliable behavior to receive lower trust scores and consequently become less favorable candidates for consensus participation.

Let $V$ denote the set of available IoV vehicles. A candidate validator committee is defined as




**Fig. 1** Overall architecture of the proposed hybrid adaptive consensus framework for IoV blockchain.

$$C = \{v_1, v_2, \ldots, v_k\}, \qquad v_i \in V, \tag{1}$$

where $k$ represents the number of validators participating in the consensus process. The decision vector used by NSGA-III is defined as

$$x = [p_1, p_2, \ldots, p_N, k, \tau, T], \tag{2}$$

where $p_i$ denotes the optimization priority associated with vehicle $i$, $k$ is the committee size, $\tau$ is the trust threshold, and $T$ represents the $T_{\mathrm{PBFT}}$ timeout.

For a given candidate solution, only vehicles satisfying the trust threshold are considered eligible for validator selection. The eligible vehicles are then ordered according to their optimized priorities and trust values and the first $k$ feasible vehicles are

selected. Thus, the identities of the validators are directly influenced by the optimization process rather than being determined independently after the optimization of the consensus parameters.

The average trust of the selected committee is calculated as

$$T(C) = \frac{1}{k} \sum_{v_i \in C} T_i. \tag{3}$$

A higher value of $T(C)$ indicates a higher average trust level among the selected validators. However, trust is not optimized independently. It is evaluated jointly with latency, communication overhead, computational cost, security risk and val-idator availability. This formulation allows NSGA-III to identify different feasible compromises between validator reliability and consensus efficiency.

## 4.3 Multi-Objective Optimization with NSGA-III

The first decision layer employs NSGA-III to periodically explore the space of possible validator committees and $T_{\mathrm{PBFT}}$ configurations. The use of a multi-objective formula-tion is motivated by the conflicting requirements of blockchain consensus in dynamic IoV environments. For example, increasing the committee size may provide access to a larger set of validators but also increases communication and computational over-head. Conversely, reducing the committee size can decrease consensus overhead while potentially affecting trust and validator availability.

For a candidate solution $x$, the optimization problem is represented by the objective vector

$$F(x) = [f_1(x), f_2(x), f_3(x), f_4(x), f_5(x), f_6(x)]. \tag{4}$$

The six objectives considered in this work are consensus latency, communication overhead, computational cost, security risk, validator trust and validator availability. The first four objectives are minimized, whereas validator trust and availability are maximized.

The objectives are therefore defined conceptually as

$$\min f_1(x) = \mathrm{ConsensusLatency}, \tag{5}$$

$$\min f_2(x) = \mathrm{CommunicationMessages}, \tag{6}$$

$$\min f_3(x) = \mathrm{ComputationalCost}, \tag{7}$$

$$\min f_4(x) = \mathrm{SecurityRisk}, \tag{8}$$

$$\max f_5(x) = \mathrm{TrustLevel}, \tag{9}$$

$$\max f_6(x) = \mathrm{ValidatorAvailability}. \tag{10}$$

Because the NSGA-III implementation uses a minimization formulation, the last two objectives are transformed as

$$f'_5(x) = -f_5(x), \qquad f'_6(x) = -f_6(x). \tag{11}$$

The optimization does not attempt to identify a single configuration that is simultaneously optimal for all objectives. Instead, NSGA-III searches for a set of non-dominated solutions forming an approximation of the Pareto set:

$$P = \{x_1, x_2, \ldots, x_m\}. \tag{12}$$

Each solution represents a different compromise between consensus efficiency, resource consumption, trust, security and validator availability. For example, one configuration may emphasize low latency, whereas another may provide lower communication overhead or a higher level of validator trust.

NSGA-III is executed periodically rather than before every consensus decision. This design limits the computational cost of the evolutionary search while allowing the candidate configuration space to be regenerated when the network conditions change.

## 4.4 RL-Based Online Adaptation

Although NSGA-III provides a set of non-dominated candidate configurations, it does not determine which configuration should be applied at every decision step. Continuously executing the evolutionary optimization would also introduce unnecessary computational overhead. Therefore, the second decision layer employs a Deep Q-Network (DQN) to perform online configuration selection.

Reinforcement learning has recently been investigated for adaptive decision-making in blockchain-enabled IoV and related vehicular environments. For example, Alaghaa et al. employed multi-agent deep reinforcement learning to support adaptive coordination in a UAV-assisted IoV framework [9]. Other studies have explored reinforcement learning for optimizing blockchain consensus and decision-making under changing network conditions, including policy-based approaches for blockchain optimization [8, 19]. Deep reinforcement learning has also been investigated for adaptive consensus configuration and validator management [10, 11]. In contrast to approaches that directly optimize a particular consensus decision, the proposed framework uses the RL agent specifically to select among candidate configurations generated by the multi-objective optimization layer.

At decision step $t$, the DQN observes the current state of the IoV environment and selects one configuration from a fixed candidate action library. The state includes the main network conditions that may influence consensus behavior, namely vehicle density, transaction latency, network congestion, Sybil activity, DoS level, average trust, current $T_{\mathrm{PBFT}}$ latency and communication overhead.

The state is represented as

$$s_t = [D_t, L_t, C_t, S_t, D_t^{DoS}, T_t, P_t, M_t], \tag{13}$$

where $D_t$ denotes vehicle density, $L_t$ the average transaction latency, $C_t$ the average network congestion, $S_t$ the Sybil ratio, $D_t^{DoS}$ the DoS level, $T_t$ the average validator trust, $P_t$ the current $T_{\mathrm{PBFT}}$ latency and $M_t$ the current number of consensus messages.

The action space contains ten discrete actions:

$$a_t \in \{0, 1, \ldots, 9\}. \tag{14}$$

Each action corresponds to a complete candidate configuration containing a validator committee and its associated $T_{\mathrm{PBFT}}$ parameters. Configurations generated by NSGA-III are preserved when available. If fewer than ten configurations are produced, additional candidate configurations are introduced to maintain a fixed action-space size. These additional configurations are valid candidate configurations for the RL agent but are not necessarily Pareto-optimal.

The DQN therefore does not directly solve the multi-objective optimization problem. Instead, it learns a state-dependent policy for selecting among the available configurations. This separation reduces the complexity of the online decision process while preserving the diversity generated by the evolutionary optimization stage.

## 4.5 Integration with $T_{\mathrm{PBFT}}$

The selected validator committee and consensus parameters are subsequently applied to the $T_{\mathrm{PBFT}}$ consensus process. In the proposed framework, $T_{\mathrm{PBFT}}$ executes the consensus procedure, while the adaptive layer determines which validators participate and which configuration parameters are used.

At decision step $t$, the overall process can be summarized as

$$s_t \rightarrow a_t \rightarrow (C_t, \Theta_t) \rightarrow T_{\mathrm{PBFT}} \rightarrow \mathrm{Consensus}, \tag{15}$$

where $s_t$ is the observed network state, $a_t$ is the action selected by the DQN, $C_t$ is the selected validator committee, and $\Theta_t$ denotes the selected consensus configuration.

The configuration is defined as

$$\Theta_t = (k_t, \tau_t, \delta_t), \tag{16}$$

where $k_t$ is the committee size, $\tau_t$ is the trust threshold and $\delta_t$ is the $T_{\mathrm{PBFT}}$ timeout. The notation $\delta_t$ is used here to distinguish the timeout parameter from $T_t$, which denotes average validator trust in the RL state.

Once the configuration is applied, the consensus process is evaluated according to the resulting latency, communication overhead, computational cost, trust, security risk and consensus outcome. These measurements are returned to the adaptation layer and contribute to subsequent decisions. Meanwhile, NSGA-III periodically regenerates the candidate configuration space to account for changes in the IoV environment.

Several PBFT variants have been proposed to improve consensus efficiency in vehicular environments. For example, SV-PBFT introduces modifications to the PBFT process together with reputation-based mechanisms to improve the stability and efficiency of vehicle-to-vehicle blockchain transactions [7]. The proposed framework follows a different adaptation strategy by jointly optimizing the validator committee and the $T_{\mathrm{PBFT}}$ configuration according to multiple network objectives.

The resulting architecture therefore establishes a closed-loop adaptation mechanism in which NSGA-III provides a diverse set of candidate configurations and the DQN performs state-dependent selection during online operation.

## 4.6 Objective Functions, State, Action, and Reward

The interaction between the multi-objective optimizer and the RL agent can be summarized through four elements: the objective formulation, the observed state, the action space, and the reward function. The first three elements define the optimization and decision spaces, while the reward provides the feedback used by the DQN during learning.

### 4.6.1 Objective Functions

For each candidate solution $x$, NSGA-III evaluates six objectives:

$$F(x) = [f_1(x), f_2(x), f_3(x), f_4(x), f_5(x), f_6(x)]. \quad (17)$$

The first four objectives minimize consensus latency, communication messages, computational cost and security risk, respectively. The remaining two objectives maximize the average trust and availability of the selected validators. For compatibility with the minimization-based formulation, trust and availability are represented internally by their negative values.

The resulting Pareto set consequently contains configurations representing differ-ent trade-offs among consensus efficiency, resource consumption, trust, security and validator availability. This formulation avoids imposing a single fixed preference on all operating conditions and instead provides a diverse candidate set for the online adaptation stage.

### 4.6.2 RL State

The RL state describes the observed condition of the IoV environment at decision step $t$:

$$s_t = [D_t, L_t, C_t, S_t, \rho^{DoS}, T_t, P_t, M_t]. \quad (18)$$

The state combines information about network conditions, attack levels, trust and the recent behavior of the consensus process. The variables are normalized in the implementation before being provided to the DQN so that features with different numerical scales can be processed consistently.

### 4.6.3 RL Action

The action space is defined as a fixed set of ten candidate configurations:

$$a_t \in \{0, 1, \ldots, 9\}. \quad (19)$$

Each action specifies a validator committee together with its associated $T_{\mathrm{PBFT}}$ parameters. Consequently, the DQN performs configuration selection rather than directly solving the multi-objective optimization problem.

### 4.6.4 Reward Function

The reward function evaluates the consequences of the configuration selected by the DQN. In the implementation, the reward combines normalized consensus latency,

communication overhead, computational cost, validator trust, security risk and the consensus outcome:

$$R_t = -1.5\hat{L}_t - 1.0\hat{M}_t - 0.5C_t + 1.5T_t - 2.5S_t + R_{\text{outcome}}, \tag{20}$$

where $\hat{L}_t$ denotes normalized consensus latency, $\hat{M}_t$ denotes normalized communication overhead, $C_t$ represents computational cost, $T_t$ denotes average validator trust, and $S_t$ represents security risk.

The outcome component is defined as

$$R_{\text{outcome}} = \begin{cases} +1, & \text{if consensus succeeds,} \\ -2, & \text{if consensus fails,} \\ -2, & \text{if the configuration is infeasible.} \end{cases} \tag{21}$$

Latency and communication overhead are normalized before being incorporated into the reward because they operate on different numerical scales. The resulting scalar feedback enables the DQN to learn a policy that considers several consensus requirements simultaneously rather than optimizing a single performance indicator.

Overall, the proposed framework combines the global search capability of NSGA-III with the state-dependent decisio n capability of DQN. NSGA-III periodically explores the configuration space and provides diverse candidate solutions, while the DQN selects among these configurations according to the observed IoV conditions. The resulting design provides an adaptive mechanism for balancing consensus efficiency, resource consumption, validator trust and security across changing network conditions.

# 5 Experimental Setup

This section presents the experimental methodology used to evaluate the proposed adaptive $T_{\text{PBFT}}$ framework. The evaluation considers variations in vehicle density, trust conditions, and adversarial activity, with particular attention to consensus latency, communication overhead, computational cost, consensus success, validator trust, security risk and view-change frequency.

The framework was implemented in Python using a synthetic IoV simulation environment. The simulator provides an analytical representation of the interactions between the IoV network, trust evaluation, validator selection, $T_{\text{PBFT}}$, NSGA-III, and the DQN agent. This controlled setting enables reproducible experiments under differ-ent network and attack conditions without requiring a packet-level network simulator or a physical blockchain deployment.

## 5.1 Simulation Environment

The simulated network consists of connected vehicles and Roadside Units (RSUs). Vehicles are the main participants in the consensus process, whereas RSUs provide infrastructure-related information used to model communication and computational conditions. Each vehicle is characterized by its position, speed, local trust, behavioral score and resulting trust value.

Network conditions evolve during the simulation according to vehicle density, communication delay, congestion, computational load, and attack conditions. Two RSUs are considered in all experiments. The number of vehicles is set to 30 in the Normal, Sybil and DoS scenarios and to 50 in the High Density and Combined Stress scenarios. Each experiment is conducted over 100 simulation steps.

To account for stochastic variations, each configuration is evaluated using five independent random seeds (42–46). Reported results are therefore expressed as mean ±standard deviation across the five runs.

## 5.2 IoV Network Configuration

The main simulation parameters are summarized in Table 1. The baseline network contains 30 vehicles and two RSUs. The High Density scenario increases the vehicle population to 50. The Sybil ratio is increased from 0.05 to 0.30 in the Sybil and Combined Stress scenarios, while a DoS level of 0.50 is introduced in the DoS and Combined Stress scenarios.

**Table 1** Main IoV simulation parameters.

| Parameter | Value |
|---|---|
| Number of vehicles | 30 or 50 |
| Number of RSUs | 2 |
| Baseline Sybil ratio | 0.05 |
| Sybil attack ratio | 0.30 |
| DoS level | 0.50 |
| Simulation horizon | 100 steps |
| Random seeds | 42, 43, 44, 45, 46 |

Communication latency is estimated from the distance between vehicles and RSUs and the current congestion level. DoS activity further increases the effective communication delay. Trust values are obtained using the simplified trust mechanism introduced in Section 4.2, which is inspired by the principles of EigenTrust rather than being a complete implementation of the original algorithm.

Sybil vehicles are assigned lower trust and behavioral values, allowing their impact on trust-aware validator selection to be evaluated as the proportion of potentially unreliable identities increases.

## 5.3 $T_{PBFT}$ Configuration

The consensus layer uses a simplified analytical model of $T_{\mathrm{PBFT}}$. The model captures the effects of committee size, communication delay, congestion, DoS conditions, and timeout on consensus performance. It is intended as a controlled abstraction for evaluating the adaptive configuration strategy rather than as a full distributed blockchain implementation.

For a committee of $k$ validators, the basic communication overhead is modeled as

$$M_{\mathrm{PBFT}} = 3k^2, \tag{22}$$

with an additional $k^2$ messages when a view change occurs. Consensus latency is estimated from network delay, congestion, DoS conditions, and committee size. A view change is triggered when the resulting base latency exceeds the configured timeout.

Consensus is considered successful when the resulting latency remains within twice the configured timeout. This formulation provides a consistent basis for comparing the different validator and consensus configurations under identical simulated conditions.

## 5.4 NSGA-III Configuration

NSGA-III is used to periodically generate candidate validator and consensus configurations. The decision vector is

$$\mathbf{x} = [p_1, p_2, \ldots, p_N, k, \tau, T], \tag{23}$$

where $p_i$ denotes the priority of vehicle $i$, $k$ the committee size, $\tau$ the trust threshold and $T$ the consensus timeout.

Vehicles satisfying the trust threshold are eligible for selection, after which the optimized priorities and trust values determine the validator committee. The optimization considers six objectives: consensus latency, communication overhead, computational cost, security risk, negative validator trust, and negative validator availability.

The main NSGA-III parameters are given in Table 2.

**Table 2** NSGA-III configuration.

| Parameter | Value |
|---|---|
| Algorithm | NSGA-III |
| Number of objectives | 6 |
| Reference-direction partitions | 4 |
| Population size | max(120, number of reference directions) |
| Number of generations | 60 |
| Committee size $k$ | Optimized variable |
| Trust threshold $\tau$ | Optimized variable |
| Timeout $T$ | Optimized variable |

The optimization is performed periodically rather than at every simulation step. The resulting non-dominated configurations are provided to the RL layer for online selection.

## 5.5 Reinforcement Learning Configuration

The online adaptation layer uses a Deep Q-Network (DQN) to select a consensus configuration according to the current network state. Reinforcement learning has previously been investigated for adaptive decision-making in blockchain and IoV environments [8–11]. In particular, Alaghaa et al. explored multi-agent reinforcement learning in a blockchain-enabled UAV-assisted IoV environment [9]. In the present work, the DQN is used specifically to select among candidate consensus configurations generated by NSGA-III.

The state contains eight variables:

$$s_t = [d_t, l_t, c_t, r_t^{\text{Sybil}}, r_t^{\text{DoS}}, T_t, L_t^{\text{PBFT}}, M_t^{\text{PBFT}}], \tag{24}$$

where $d_t$ denotes vehicle density, $l_t$ average transaction latency, $c_t$ congestion, $r_t^{\text{Sybil}}$ the Sybil ratio, $r_t^{\text{DoS}}$ the DoS level, $T_t$ average validator trust, $L_t^{\text{PBFT}}$ current PBFT latency, and $M_t^{\text{PBFT}}$ PBFT message overhead.

The action space contains ten discrete actions. Each action represents a complete consensus configuration. NSGA-III-generated configurations are retained first; if fewer than ten candidates are available, the action library is completed with deterministic configurations. These additional configurations are not assumed to be Pareto-optimal.

The DQN uses an MLP policy with two hidden layers of 128 neurons. The main training parameters are reported in Table 3.

**Table 3** DQN training configuration.

| Parameter | Value |
|---|---|
| RL algorithm | DQN |
| Policy | MLP |
| Hidden layers | 128–128 |
| Learning rate | 0.0005 |
| Replay buffer size | 20,000 |
| Learning starts | 1,000 steps |
| Batch size | 64 |
| Discount factor $\gamma$ | 0.99 |
| Training frequency | 4 |
| Gradient steps | 1 |
| Exploration fraction | 0.30 |
| Initial exploration $\epsilon$ | 1.00 |
| Final exploration $\epsilon$ | 0.05 |
| Target update interval | 500 |
| Training timesteps | 50,000 |
| Random seed | 42 |

The reward function combines consensus efficiency, resource consumption, trust, security and the consensus outcome:

$$R_t = -1.5\hat{L}_t - 1.0\hat{M}_t - 0.5C_t + 1.5T_t - 2.5S_t + R_{\text{outcome}}, \tag{25}$$

where $\hat{L}_t$ and $\hat{M}_t$ are normalized latency and message overhead, respectively , while $C_t$, $T_t$, and $S_t$ denote computational cost, average validator trust and security risk.

The outcome component is defined as

$$R_{\text{outcome}} = \begin{cases} +1, & \text{if consensus succeeds,} \\ -2, & \text{if consensus fails,} \\ -2, & \text{if the configuration is infeasible.} \end{cases} \tag{26}$$

## 5.6 Attack Scenarios

Five scenarios are considered to evaluate the framework under progressively different operating and security conditions.

1. **Normal:** 30 vehicles, a Sybil ratio of 0.05 and no DoS activity. “‘
2. **High Density:** 50 vehicles, a Sybil ratio of 0.05 and no DoS activity.
3. **Sybil Attack:** 30 vehicles, a Sybil ratio of 0.30 and no DoS activity.
4. **DoS Attack:** 30 vehicles, a Sybil ratio of 0.05 and a DoS level of 0.50.
5. **Combined Stress:** 50 vehicles, a Sybil ratio of 0.30 and a DoS level of 0.50. “‘

Sybil attacks are particularly relevant in vehicular networks because multiple malicious identities can affect trust and participation decisions. Existing security mechanisms have therefore investigated authentication, pseudonym management and protection against Sybil-related threats in vehicular communications [12, 20]. In this study, the attack is modeled by assigning lower trust and behavioral values to Sybil vehicles, allowing its effect on trust-aware consensus selection to be measured.

Blockchain-based trust management has also been investigated for vehicular networks through combinations of trust evaluation, consensus mechanisms and malicious-node detection [13, 21]. Other blockchain-based IoV studies have addressed secure data sharing, authentication and protection against malicious or unauthorized entities [22]. These studies motivate the evaluation of trust-aware consensus mechanisms under adversarial conditions. Accordingly, the DoS scenario models network degradation through an increased communication delay, while the Combined Stress scenario simultaneously introduces high density, Sybil identities and DoS conditions.

## 5.7 Baseline Methods

Three configurations are compared in the experiments: Fixed $T_{\text{PBFT}}$, NSGA-III Only, and Hybrid NSGA-III+DQN.

### *Fixed $T_{\text{PBFT}}$.*

The fixed baseline uses $k = 5$, $\tau = 0.5$, and $T = 10$. The validator committee is selected at the beginning of the experiment and remains unchanged throughout the simulation.

### *NSGA-III Only.*

This configuration uses NSGA-III to periodically generate and select consensus con-figurations without the DQN layer. It isolates the contribution of multi-objective optimization.

### *Hybrid NSGA-III+DQN.*

The proposed configuration combines periodic NSGA-III optimization with online DQN-based configuration selection. NSGA-III generates the candidate configuration space, while the DQN selects a configuration according to the current network state.

No independent DQN-only baseline is considered because such a configuration is not implemented in the experimental framework.

## 5.8 Evaluation Metrics

The experimental comparison is based on the following metrics.

***Consensus latency.***

The time required by the analytical $T_{\mathrm{PBFT}}$ model to complete the consensus process. Lower values indicate lower consensus delay.

***Communication overhead.***

The number of consensus messages generated by the validator committee, including additional messages caused by view changes.

***Computational cost.***

The estimated processing burden associated with the selected consensus configuration and current network conditions.

***Consensus success rate.***

The percentage of simulation steps for which the consensus process satisfies the defined success condition.

***Validator trust.***

The average trust value of the selected validator committee.

***Security risk.***

The estimated security exposure associated with the selected validator configuration and the presence of potentially malicious or low-trust nodes.

***View-change rate.***

The percentage of simulation steps in which a consensus timeout is exceeded and a view change is triggered.

***Average reward.***

The mean RL reward obtained during evaluation. This metric is interpreted jointly with the individual performance measures because it combines several objectives into a single feedback signal.

For each scenario, the three configurations are evaluated for 100 simulation steps and repeated with five independent seeds (42–46). The results reported in Section 6 are presented as mean ± standard deviation, providing both the average behavior and the variability across runs.

# 6 Results and Discussion

The experimental results are reported as mean ± standard deviation over five independent random seeds. The analysis focuses on consensus performance under dif-ferent network densities and attack conditions, as well as on the contribution of multi-objective optimization and reinforcement learning.

## 6.1 Overall Performance under Normal and High-Density Conditions

The Normal and High Density scenarios were first considered to evaluate the effect of increasing the number of participating vehicles. As shown in Fig. 2, increasing the network size from 30 to 50 vehicles produces only a moderate increase in latency for the adaptive approaches. Under Normal conditions, the average latency is 21.64 ± 5.75 for Fixed T-PBFT, 12.48±2.77 for NSGA-III Only and 13.14±3.29 for Hybrid NSGA-III+DQN. Under High Density, these values become 23.20 ±5.80, 14.05 ±4.11, and 14.38 ±4.53, respectively.

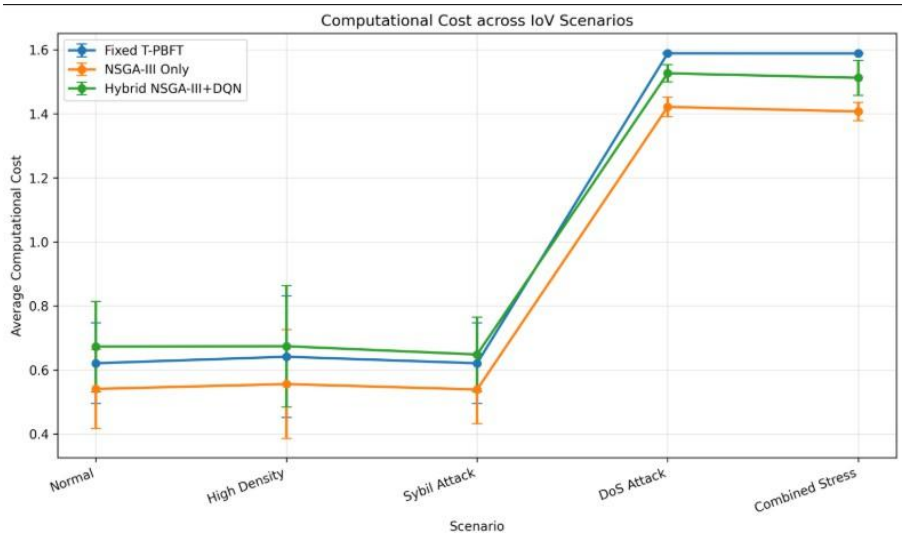



**Fig. 2** Average consensus latency across the evaluated IoV scenarios. Error bars represent ± standard deviation over five independent seeds.

A similar trend is observed for consensus success. Fixed T-PBFT achieves only 27.40% ± 25.38% success under Normal conditions and 23.00% ± 21.00% under High Density. In contrast, NSGA-III Only reaches 97.20%±5.72% and 94.00%±6.96%, while Hybrid NSGA-III+DQN reaches 94.60% ±8.71% and 94.40% ±7.96%, respectively (Fig. 3).

These results indicate that the adaptive configurations maintain relatively stable consensus performance as network density increases. The fixed configuration is

more sensitive to the changing network conditions, whereas the adaptive approaches periodically modify the validator and consensus configuration.

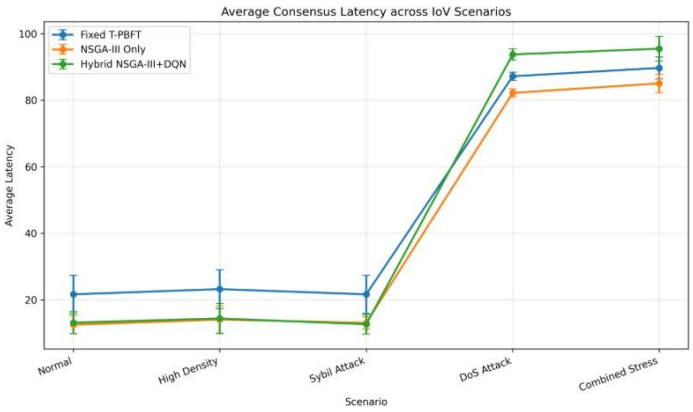



**Fig. 3** Consensus success rate across the evaluated IoV scenarios. Error bars represent ± standard deviation over five independent seeds.

## 6.2 Consensus Stability and Communication Overhead

View-change frequency provides an indication of consensus stability. Under Normal conditions, Fixed T-PBFT reaches a view-change rate of 72.60% ± 25.38%, compared with only 2.80% ± 5.72% for NSGA-III Only and 5.40% ± 8.71% for Hybrid NSGA-III+DQN. Under High Density, the corresponding rates are 77.00% ± 21.00%, 6.00% ± 6.96% and 5.60% ±7.96%.

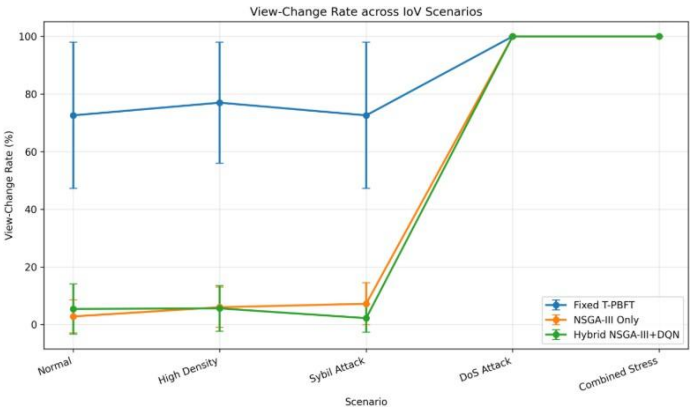



**Fig. 4** View-change rate across the evaluated IoV scenarios. Error bars represent ± standard deviation over five independent seeds.

Communication overhead shows a different trade-off. NSGA-III Only generates 38.37 ± 5.48 messages under Normal conditions and 37.86 ± 4.83 under High Density, compared with 93.15 ± 6.35 and 94.25 ± 5.25 for Fixed T-PBFT. Hybrid NSGA-III+DQN generates 109.95 ±14.00 and 98.25 ±10.45 messages, respectively.

The lower overhead obtained by NSGA-III Only is consistent with the smaller committees selected by the optimization process. In contrast, the DQN may select configurations with higher communication cost because its reward jointly considers several objectives rather than minimizing message overhead alone.

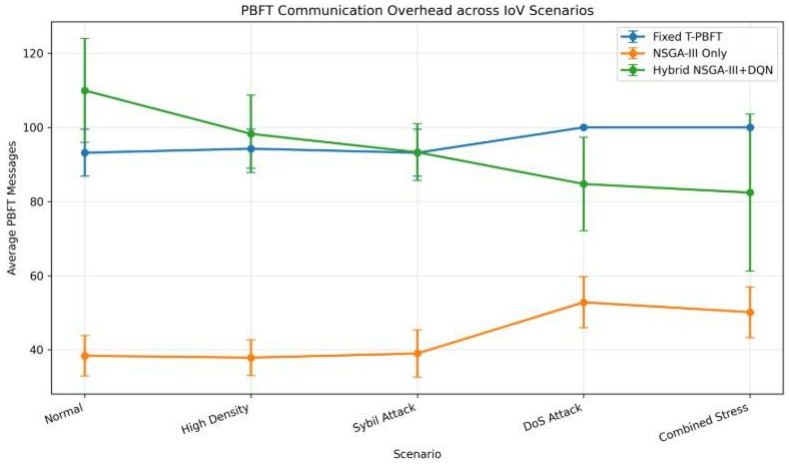



**Fig. 5** Average PBFT communication overhead across the evaluated IoV scenarios. Error bars represent ± standard deviation over five independent seeds.

## 6.3 Behavior under Sybil Attacks

The Sybil scenario evaluates the effect of increasing the proportion of low-trust identities. Sybil attacks are an important concern in vehicular networks because multiple identities may affect trust and participation decisions; related studies have therefore investigated authentication and pseudonym-based protection mechanisms [12]. Blockchain-based trust management combined with malicious-node detection has also been investigated for vehicular networks [13].

In the simulated Sybil scenario, NSGA-III Only achieves a success rate of 92.80% ± 7.29%, while Hybrid NSGA-III+DQN reaches 97.80% ± 4.92%. Fixed T-PBFT remains at 27.40% ±25.38%. The corresponding latencies are 21.64 ±5.75, 13.06 ±1.88, and 12.66 ±2.95.

The average validator trust remains high for all configurations, with approximately 0.926 for Fixed T-PBFT and NSGA-III Only and 0.920 for the Hybrid configuration. This suggests that the trust-aware selection mechanism generally favors higher-trust vehicles despite the presence of Sybil identities.

However, the security-risk metric remains zero in this scenario. This should not be interpreted as complete Sybil detection, since the current simulation represents

Sybil behavior through reduced trust values rather than through an explicit detection mechanism. The results therefore evaluate the effect of trust-aware selection under the modeled attack conditions.

## 6.4 Behavior under DoS and Combined Stress

The DoS and Combined Stress scenarios produce substantially more challenging conditions. In both cases, all three configurations reach a 100% view-change rate and a 0% consensus success rate.

Under DoS, the average latency is $87.19 \pm 1.23$ for Fixed T-PBFT, $82.19 \pm 1.23$ for NSGA-III Only and $93.76 \pm 1.67$ for Hybrid NSGA-III+DQN. Under Combined Stress, the corresponding values increase to $89.71 \pm 3.28$, $85.06 \pm 2.81$ and $95.49 \pm 3.68$. NSGA-III Only also produces lower communication overhead, with $52.80 \pm 6.86$ messages under DoS and $50.14 \pm 6.84$ under Combined Stress, compared with 100 messages for Fixed T-PBFT. However, this reduction does not restore consensus
availability.

These results highlight an important limitation of the current framework. Multi-objective adaptation can reduce latency and communication cost under severe network degradation, but validator selection and consensus parameter tuning alone are insufficient to maintain successful consensus under the considered DoS intensity.

## 6.5 Baseline Comparison and Discussion

The comparison between Fixed T-PBFT, NSGA-III Only, and Hybrid NSGA-III+DQN shows that the main improvement comes from moving from a fixed configuration toward adaptive multi-objective optimization. Under Normal and High Density conditions, NSGA-III Only provides high consensus success with lower latency and substantially lower communication overhead than the fixed baseline.

The Hybrid configuration provides an additional state-dependent selection mech-anism. Its effect is scenario-dependent: under the Sybil scenario, it achieves a higher success rate than NSGA-III Only, whereas it generally introduces higher communi-cation and computational costs. This reflects the multi-objective nature of the DQN reward, which does not optimize a single performance metric.

The results should nevertheless be interpreted within the scope of the simulation. The evaluation uses a synthetic Python environment and an analytical abstraction of $T_{\mathrm{PBFT}}$ rather than a full distributed blockchain or packet-level vehicular network simulator. The trust mechanism is a simplified approach inspired by EigenTrust principles and the attack models represent controlled simulation conditions. In addition, the experiments rely on five random seeds, which provides an indication of variability but does not constitute extensive statistical validation.

Overall, the experiments show that adaptive consensus configuration improves the simulated performance of $T_{\mathrm{PBFT}}$ under non-DoS conditions, while the contribution of reinforcement learning depends on the network scenario. Severe DoS conditions remain a limitation and indicate the need for additional mechanisms specifically designed to preserve consensus availability under strong network disruption.

# 7 Conclusion and Future Work

Blockchain-based consensus in Internet of Vehicles (IoV) environments must operate under continuously changing network conditions, node trustworthiness and security threats. Previous studies have investigated trust-aware consensus, improved PBFT mechanisms, and learning-based adaptation to address these challenges [1–3, 6, 8]. However, jointly adapting validator selection and consensus parameters remains challenging in highly dynamic vehicular environments.

This work presented an adaptive T-PBFT framework that combines a simplified trust evaluation mechanism inspired by the principles of EigenTrust, NSGA-III multi-objective optimization, and DQN-based online configuration selection. The proposed formulation jointly considers consensus latency, communication overhead, computational cost, security risk, validator trust and availability. Unlike a fixed consensus configuration, the proposed approach allows both the validator committee and T-PBFT parameters to be adapted according to the current network conditions.

The framework was evaluated using a Python-based IoV simulation under Normal, High Density, Sybil, DoS and Combined Stress scenarios. Five independent random seeds were considered for each scenario, and the proposed configurations were compared with Fixed T-PBFT and NSGA-III Only. The results show that adaptive configuration selection substantially improves the simulated consensus behavior under non-DoS conditions, particularly in terms of latency, consensus success, and communi-cation overhead. NSGA-III Only generally provides low resource consumption, while the Hybrid NSGA-III+DQN configuration introduces an additional state-dependent adaptation mechanism. Its benefit is scenario-dependent and is particularly visible in the Sybil scenario, where it achieves a higher mean success rate than NSGA-III Only. The experiments also reveal important limitations. Under the considered DoS and Combined Stress conditions, all evaluated configurations reach a 100% view-change rate and a 0% consensus success rate. Thus, although adaptive selection can modify the resource and latency characteristics of the simulated consensus process, it is not sufficient to maintain consensus availability under severe network disruption. In addi-tion, the current Sybil model relies on reduced trust values rather than a dedicated attack-detection mechanism. The security results should therefore be interpreted as an evaluation of trust-aware validator selection rather than as a direct measure of
Sybil detection capability.

Several limitations should be considered when interpreting these results. The evaluation relies on a synthetic Python simulation and an analytical abstraction of T-PBFT rather than a full distributed blockchain or packet-level vehicular network simulator. The trust mechanism is also a simplified formulation inspired by EigenTrust principles rather than a complete implementation of the original protocol. Finally, the use of five random seeds provides an initial assessment of variability but does not establish broad statistical generalization.

## Future Work

Future work will first focus on evaluating the proposed framework in more realistic IoV environments. Integration with established mobility and network simulation platforms

would allow realistic vehicle mobility, V2V/V2I communication, packet-level network behavior and blockchain execution to be considered.

A second direction concerns trust and security. The current trust mechanism could be extended by incorporating direct and indirect trust, temporal trust evolution, reputation propagation and explicit malicious-node and Sybil detection mechanisms. This would allow the security component to be evaluated using dedicated attack-detection metrics, in addition to the current security-risk formulation [3, 4].

Another priority is improving resilience to severe DoS/DDoS attacks. Since the current experiments show that adaptive validator selection and parameter tuning alone cannot preserve consensus under strong disruption, future work could investi-gate attack-aware admission control, resource allocation, traffic filtering, or adaptive consensus mechanisms. These mechanisms could be integrated into the multi-objective optimization process to explicitly account for consensus availability under attack.

The reinforcement learning component could also be extended to larger configuration spaces. Alternative approaches such as PPO and other actor–critic methods could be investigated when the number of vehicles, validators, or adjustable consensus parameters increases. Recent studies have demonstrated the potential of reinforcement learning for adaptive blockchain consensus and adversarial environments [8, 10, 11]. Finally, future experiments should consider larger vehicle populations, more diverse mobility conditions, different attack intensities, simultaneous attacks and a larger number of independent runs. Comparing NSGA-III with other multi-objective algorithms such as NSGA-II and MOEA/D could also provide additional insight into the
optimization component.

Overall, the results indicate that combining multi-objective optimization with reinforcement learning provides a promising framework for adapting T-PBFT consensus in dynamic IoV environments. At the same time, the observed limitations under severe DoS conditions highlight the need for more attack-aware and realistic mechanisms. Future work will therefore focus on improving scalability, trust management, attack resilience and validation in realistic IoV environments.

# References


[1] Surapaneni, P., Bojjagani, S., Bharathi, V.C., Morampudi, M.K., Maurya, A.K., Khan, M.K.: A systematic review on blockchain-enabled internet of vehicles (biov): Challenges, defenses, and future research directions. IEEE Access **12**, 123529–123560 (2024) https://doi.org/10.1109/ACCESS.2024.3453433

[2] Bitok, H.J., Wang, M., Desmond, D.: Consensus on the internet of vehicles: A systematic literature review. World Electric Vehicle Journal **16**(11), 616 (2025) https://doi.org/10.3390/wevj16110616

[3] Razafimanjato, M., Saad, M.M., Kim, D.: Blockchain-based trust management systems in the internet of vehicles: A comprehensive survey. ICT Express **11**(6), 1265–1285 (2025) https://doi.org/10.1016/j.icte.2025.09.012

[4] Alalwany, E., Mahgoub, I.: Security and trust management in the internet of vehicles (iov): Challenges and machine learning solutions. Sensors **24**(2), 368 (2024) https://doi.org/10.3390/s24020368

[5] Du, Z., Zhang, J., Fu, Y., Huang, M., Liu, L., Li, Y.: A scalable and trust-value-based consensus algorithm for internet of vehicles. Applied Sciences **13**(19), 10663 (2023) https://doi.org/10.3390/app131910663

[6] Xu, G., Bai, H., Xing, J., Luo, T., Xiong, N.N., Cheng, X., Liu, S., Zheng, X.: Sg-pbft: A secure and highly efficient distributed blockchain pbft consensus algorithm for intelligent internet of vehicles. Journal of Parallel and Distributed Computing **164**, 1–11 (2022) https://doi.org/10.1016/j.jpdc.2022.01.029

[7] Suo, Y., Wang, Y., Luo, S., Yang, Q., Zhao, J.: Sv-pbft: An efficient and sta-ble blockchain pbft improved consensus algorithm for vehicle-to-vehicle energy transactions. Journal of Internet Technology **23**(6), 1191–1201 (2022) https:
//doi.org/10.53106/160792642022112306003

[8] Zou, Y., Jin, Z., Zheng, Y., Yu, D., Lan, T.: Optimized consensus for blockchain in internet of things networks via reinforcement learning. Tsinghua Science and Technology **28**(6), 1009–1022 (2023) https://doi.org/10.26599/TST.2022.
9010045

[9] Alaghaa, A., Kadadhab, M., Mizouni, R., Singh, S., Bentahar, J., Otrok, H.: Uav-assisted internet of vehicles: A framework empowered by reinforcement learning and blockchain. Vehicular Communications **52**, 100874 (2025) https://doi.org/10.1016/j.vehcom.2025.100874

[10] Villegas-Ch, W., Govea, J., Gutierrez, R.: Optimizing consensus in blockchain with deep and reinforcement learning. Emerging Science Journal **9**(4), 1886–1908 (2025) https://doi.org/10.28991/ESJ-2025-09-04-08

[11] Gutierrez, R., Villegas-Ch, W., Govea, J.: Adaptive consensus optimization in blockchain using reinforcement learning and validation in adversarial environments. Frontiers in Artificial Intelligence **8**, 1672273 (2025) https://doi.org/10.3389/frai.2025.1672273

[12] Zhang, C., Chang, X., Jia, X., Zhang, S., Wang, X.: Sybil-resistant and privacy-preserving authentication based on short-term pseudonym for internet of vehicles.
Scientific Reports **15**, 44966 (2025) https://doi.org/10.1038/s41598-025-28854-1

[13] Ullah, A., Ullah, Z., Rizvi, S.S., Shah, I.A., Kwon, S.J.: Decentralized trust opti-mization in vanets: A blockchain-driven hybrid pos-pbft architecture for enhanced security and energy-efficient communication. PLOS ONE **20**(12), 0338618 (2025) https://doi.org/10.1371/journal.pone.0338618

[14] Si, X., Li, M., Yao, Z., Zhu, W., Liu, J., Zhang, Q.: An efficient and secure

blockchain consensus protocol for internet of vehicles. Electronics **12**(20), 4285 (2023) https://doi.org/10.3390/electronics12204285

[15] Shen, X., Ma, R.: A blockchain solution for the internet of vehicles with better filtering and adaptive capabilities. Sensors **25**(4), 1030 (2025) https://doi.org/10. 3390/s25041030

[16] Shamili, F.S., Gopi, R.: Optimizing iov cloud trust with adaptive blockchain and reinforcement learning. Scientific Reports **15**, 32850 (2025) https://doi.org/10. 1038/s41598-025-00951-1

[17] Chai, H., Leng, S., He, J., Zhang, K.: Digital twin empowered lightweight and efficient blockchain for dynamic internet of vehicles. Digital Communications and Networks **10**(6), 1698–1707 (2024) https://doi.org/10.1016/j.dcan.2023.08.004

[18] Zhao, F., Cheng, R., Li, C., Su, Z., Liang, G., Yang, C.: A proof-of-multiple-state consensus mechanism for mobile nodes in internet of vehicles. Electronics **13**(8), 1553 (2024) https://doi.org/10.3390/electronics13081553

[19] Dutta, A., Rafin, N.I., Dewan, M.A.A., Alam, M.G.R.: Robb: Recurrent proximal policy optimization reinforcement learning for optimal block formation in bit-coin blockchain network. IEEE Access **12**, 31287–31311 (2024) https://doi.org/ 10.1109/ACCESS.2024.3369896

[20] Alsaeed, A., Almowuena, S., Alyahya, A.N.: Baaiov: A blockchain-based authen-tication and authorization framework for secure and reliable internet of vehicles communication. IEEE Access **13**, 150821–150837 (2025) https://doi.org/10.1109/ ACCESS.2025.3601003

[21] Bandarapu, S.R., Bilal, M., Chatterjee, P., Cheema, A.M., Rashid, J., Kim, J.: Blockchain-based federated learning framework for malicious node detection in internet of vehicles (iov) networks using fog and cloud computing. Journal of King Saud University - Computer and Information Sciences **37**(6), 120 (2025) https://doi.org/10.1007/s44443-025-00134-y

[22] Wang, L., Guan, C.: Improving security in the internet of vehicles: A blockchain-based data sharing scheme. Electronics **13**(4), 714 (2024) https://doi.org/10. 3390/electronics13040714