Network Architecture and Quality of Service through Secure AI Models

ABSTRACT

The architecture of modern wireless networks (5G/6G) and Internet of Things (IoT) systems is undergoing a significant transformation through the integration of artificial intelligence (AI) functions into the core network management and control mechanisms. Machine learning (ML) models perform critical architectural functions such as dynamic resource allocation and interference management to optimize quality of service (QoS). However, this dependence introduces new security vulnerabilities, as ML algorithms are susceptible to malicious attacks that can destabilize the network and degrade the user experience.

This work focuses on the relationship between the security of AI models and QoS assurance. Specifically, we implement and experimentally study (through simulation) a network architecture scenario in which resource allocation is controlled by deep reinforcement learning (DRL) algorithms. The impact of data manipulation attacks (CQI falsification) on key QoS indicators is analyzed. The experimental results reveal severe allocation fairness degradation (a 73% drop in the Jain index) and dramatic latency increases (a 2536% increase), while overall network capacity remains relatively stable (-2.3%).

We propose and evaluate a multilayer defense architecture combining adversarial training, statistical anomaly detection and input validation, achieving a **29% recovery in fairness** and an **11% reduction in latency**. This study demonstrates that strengthening the robustness of AI models is necessary to guarantee quality of service in next-generation networks.

Key Words: 5G/6G Networks, Deep Reinforcement Learning, Adversarial Machine Learning, QoS Assurance, Resource Allocation, Network Security

1. INTRODUCTION

1.1 Motivation and Context

The evolution of wireless networks from 5G to 6G is characterized by an exponential increase in complexity and the need to support a variety of Quality of Service (QoS) requirements. Unlike previous generations, which relied mainly on fixed mathematical models and heuristic rules for resource management, modern networks are shifting towards 'network intelligence' approaches [1]. Artificial intelligence (AI) and machine learning (ML) are being integrated into the core of network architectures to automate critical functions such as dynamic spectrum allocation, interference management and user scheduling.

Focus is given to Deep Reinforcement Learning (DRL) algorithms, which allow network controllers (agents) to learn optimal strategies by interacting with the environment without needing prior knowledge of the channel model [2]. However, the widespread adoption of these 'black-box models' creates new opportunities for attack. The security of AI models is now inextricably linked to the reliability of the network.

1.2 Problem Statement

Despite their advantages, machine learning (ML) models are inherently vulnerable to malicious interference. In the context of wireless communications, for example, a malicious user or a compromised Internet of Things (IoT) device can inject 'poisoned' data, known as data poisoning, or manipulate the state observations that are fed into the control algorithm [3].

This work addresses a twofold problem:

1. Firstly, how is Quality of Service (QoS) affected by a data manipulation attack on the resource allocation mechanism?

2. Secondly, how can we design a network architecture that incorporates robust AI models capable of maintaining high QoS levels even under attack conditions?

1.3 Key Contributions

The main objective of this study is to examine the relationship between the security of AI algorithms and network performance. Specifically, this study makes the following contributions:

A. Vulnerability analysis: Examining the vulnerability of DRL (Deep Reinforcement Learning) algorithms (specifically PPO – Proximal Policy Optimization) in scenarios of dynamic resource allocation under CQI (Channel Quality Indicator) falsification attacks.

B. Experimental evaluation: A complete 5G/6G network simulation environment is implemented to quantify QoS degradation. The results reveal:

- a catastrophic drop in allocation fairness (73%)
- an explosive increase in latency (2536%)
- an asymmetric impact, with throughput remaining stable at -2.3% while fairness and latency collapse

C. Defensive Architecture Proposal: A multilayer robustness-enhancement mechanism is proposed and evaluated, combining:

- Adversarial training (training with poisoned samples)
- Statistical anomaly detection (Z-score and ensemble methods)
- Input validation (bounds checking and rate limiting)

D. New findings: Attacks on AI-driven resource allocators exploit allocation inefficiency rather than capacity limitations, leading to service-level degradation without significantly affecting network-level metrics.

1.4 Paper Organization

The rest of this work is organized as follows: Chapter 2 provides the theoretical background and literature review. Chapter 3 describes the system and threat models and the formulation of the optimization problem. Chapter 4 presents the proposed Robust AI architecture. Chapter 5 includes experimental evaluation and results. Chapter 6 discusses the findings and limitations, and Chapter 7 summarizes the conclusions and suggests directions for future research.

2. BACKGROUND AND RELATED WORK

2.1 AI-Empowered Wireless Networks

The transition towards 6G networks is characterised by the integration of AI across all layers of the network architecture [1]. Letaief et al. present a comprehensive roadmap for AI-empowered wireless networks, emphasizing the role of machine learning (ML) in:

- **Physical Layer**: Channel estimation, beamforming, modulation/coding
- **MAC Layer**: Resource allocation, scheduling, interference management
- **Network Layer**: Routing, mobility management, network slicing

Traditional optimization approaches (e.g. convex optimization and game theory) are being replaced by data-driven methods that learn from experience and adapt to dynamic environments.

2.2 Deep Reinforcement Learning for Resource Management

Reinforcement learning enables an agent to learn optimal policies through trial and error when interacting with the environment. The problem is modelled as a Markov Decision Process (MDP), which is defined by the tuple (S, A, P, R, γ):

S: State space (π.χ. channel quality indicators, buffer sizes)

A: Action space (π.χ. resource allocation decisions)

P: Transition probability

R: Reward function

γ: Discount factor

The PPO (Proximal Policy Optimization) algorithm [Schulman et al., 2017] was chosen due to its stability and efficiency in continuous action spaces. PPO optimizes a surrogate objective function:

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

where $r_t(\theta) = \pi_\theta(a_t|s_t) / \pi_{\theta\_old}(a_t|s_t)$ and $\varepsilon \approx 0.2$.

Du et al. [2] extend the approach to multi-agent scenarios for 6G in-X subnetworks, thereby demonstrating the effectiveness of multi-agent reinforcement learning (MARL) in complex, heterogeneous networks.

2.3 Adversarial Attacks in Wireless Communications

The bibliography distinguishes two main categories of attacks:

A. Poisoning Attacks (Training Phase) The attacker manipulates the training data, causing the model to adopt an incorrect learning policy. In the context of wireless networks, this can be achieved by:

- False CSI reporting (Channel State Information)
- Manipulated feedback signals
- Malicious user behavior during training

B. Evasion Attacks (Inference Phase) The attacker manipulates the inputs during runtime in order to deceive the trained model. Examples include:

- CQI falsification: Reporting falsely high channel quality
- Jamming with adversarial patterns
- Sybil attacks in distributed systems

Son et al. [3] provide a thorough classification of attacks and defense mechanisms for 6G IoT systems, emphasizing the importance of an end-to-end security architecture.

Adesina et al. [4] review adversarial machine learning (ML) techniques in wireless communications, focusing on radio frequency (RF)-level attacks. Their study reveals that:

- ML models are particularly vulnerable to perturbations in the RF domain
- Attacks can be highly transferable across different models
- Domain-specific defenses are required that consider the nature of wireless channels

2.4 Defensive Mechanisms

The main categories of defense include:

1. **Adversarial Training**: Training the model with a mixture of clean and adversarial examples to develop robustness [Madry et al., 2018].

2. **Certified Defenses**: Mathematical guarantees for robustness within a specified ε-ball (e.g., randomized smoothing).

3. **Anomaly Detection**: Statistical checks to identify out-of-distribution inputs:

- Statistical methods (Z-score, IQR)
- ML-based detectors (Isolation Forest, One-Class SVM)
- Deep learning approaches (Autoencoders, GANs)

4. Input Sanitization: Pre-processing to remove adversarial perturbations:

- Bounds checking
- Rate limiting
- Feature squeezing
- JPEG compression (for images)

2.5 Research Gap

Despite extensive research, significant gaps remain.

1. **Lack of comprehensive evaluation:** Most studies examine either attacks or defenses, but not within a realistic network setting.
2. **Unrealistic threat models:** Many studies assume adversaries with full knowledge of the model (white-box), whereas in practice, attacks are often black-box.
3. **Trade-offs are often ignored:** The trade-offs between security, performance and complexity are rarely analyzed.
4. **Lack of service-level analysis**: Studies tend to focus on network-level metrics such as throughput and spectral efficiency, but neglect user-level QoS metrics such as fairness, latency and satisfaction.

This work addresses these gaps by conducting a comprehensive experimental study combining a realistic network simulator, sophisticated attack scenarios and a multi-layered defense architecture with an emphasis on service-level impact metrics.

---

3. SYSTEM AND THREAT MODEL

3.1 Description of the Network Environment

We consider a downlink wireless network of a 5G/6G cell consisting of a Base Station (BS) and a set of N users U = {u_1, u_2, ..., u_N}. The available transmission bandwidth is divided into Korthogonal Resource Blocks (RBs). The objective of the BS is to dynamically allocate RBs to users at each time point, t, to maximize the total transmission rate while ensuring fair allocation.

Physical Layer Model

The signal-to-interference-plus-noise ratio (SINR) for user i on RB k at time t is given by:

$$\boxed{SINR_{i,k}^t = (P_i h_{i,k}^t) / (N_0 + I_{i,k}^t)}$$

where:

- P_i: Transmit power to user i
- $h_{i,k}^t$: Channel gain
- N_0: Thermal noise
- $I_{i,k}^t$: Interference from other cells

The channel gain is modeled using the 3GPP Urban Macro model:

$$\boxed{h_{i,k}^t = PL(d_i) \cdot SF_i \cdot FF_{i,k}^t}$$

where:

- $PL(d_i) = 128.1 + 37.6 \cdot \log_{10}(d_i)$: Path loss
- $SF_i \sim LogNormal(0, \sigma^2)$: Shadowing ($\sigma = 8$ dB)
- $FF_{i,k}^t \sim Rayleigh$: Fast fading

Channel Quality Indicator (CQI)

Users report their channel quality to the BS via the CQI $\in [0, 15]$. The CQI is derived from the relation:

$$\boxed{CQI_i^t = f(SINR_i^t) = \lfloor \log_2(1 + SINR_i^t) / 0.5 \rfloor}$$

normalized to the range [0, 1] for use in DRL.

Throughput Calculation

The throughput of user i depends on the number of allocated RBs (a_i) and the CQI:

$$\boxed{R_i^t = \alpha_i \cdot B_{RB} \cdot SE(CQI_i^t)}$$

where:

- B_{RB} = 180 kHz: Bandwidth per RB
- $SE(\cdot)$: Spectral efficiency (bits/s/Hz) from 3GPP lookup table

3.2 MDP Formulation for PPO

State Space (S)

The state vector s^t includes:

$$\boxed{s^t = [CQI_1^t, ..., CQI_N^t, B_1^t, ..., B_N^t]}$$

where:

- $CQI_i^t \in [0, 1]$: Normalized channel quality
- $B_i^t \in [0, 1]$: Normalized buffer occupancy

Action Space (A)

The action corresponds to the allocation vector:

$$\boxed{a^t = [\alpha_1^t, ..., \alpha_N^t]}$$

with constraints:

- $\alpha_i^t \in [0, 1]\ \forall i$
- $\Sigma(\alpha_i^t) = 1$ (normalization)

Reward Function (R)

The reward function balances throughput, fairness, and latency::

$$\boxed{r^t = w_1 \cdot \Sigma(R_i^t) + w_2 \cdot \mathrm{JFI}(R^t) - w_3 \cdot \Sigma(B_i^t) + w_4 \cdot \mathrm{Efficiency}^t}$$

where:

- $\mathrm{JFI}(\cdot)$: Jain's Fairness Index
- Efficiency = Actual Throughput / Potential Throughput
- Weights: $w_1 = 1.0$, $w_2 = 15.0$, $w_3 = 0.02$, $w_4 = 10.0$

The efficiency term penalizes allocations that are based on incorrect CQI reports.

3.3 Threat Model

Adversarial Capabilities

We assume the presence of M malicious users in the network, where $M = \rho \cdot N$ and $\rho \in [0.2, 0.6]$ is the probability of attack. The malicious users have the following capabilities:

1. **CQI Manipulation**: They can report false CQI values.
2. **Timing Control**: They choose when to launch the attack.
3. **Limited Collusion**: Malicious users can coordinate with each other.
4. **Adaptive Behavior**: They can observe resource allocations and adapt their attack accordingly.

Attack Strategy: CQI Falsification

The main attack consists of falsifying the reported CQI. If CQI_i^{true} is the real value and $CQI_i^{reported}$ is the reported value:

$$\boxed{CQI_i^{reported} = \mathrm{clip}(CQI_i^{true} + \delta_i, 0, 1)}$$

where $\delta_i \sim \mathrm{Uniform}(0, \sigma)$ for a malicious user, with $\sigma \in [0.3, 0.9]$ representing the attack magnitude.

Attack Types

We consider three variants:

1. **Overstatement Attack**: $\delta_i > 0$ (report better channel quality to steal resources)
2. **Understatement Attack**: $\delta_i < 0$ (report worse channel to cause underutilization)
3. **Random Perturbation**: $\delta_i \sim \text{Uniform}(-\sigma, \sigma)$ (chaotic disruption)

In this study we focus on the Overstatement Attack ($\rho = 0.4$, $\sigma = 0.7$), as it has the strongest impact on QoS.

Threat Model Assumptions

- **Black-box**: Attackers do not know the internal structure of the PPO model.
- **Partial Information**: They only observe the resource allocations they receive.
- **No Physical Layer Access**: They cannot manipulate actual RF transmissions.
- **Rational Attackers**: Their goal is to maximize their own throughput.

3.4 QoS Metrics

Αξιολογούμε το QoS μέσω των εξής δεικτών:

1. Aggregate Throughput

$$\boxed{R_{total} = \Sigma(R_i) / N}$$

2. Jain's Fairness Index

$$\boxed{JFI = (\Sigma R_i)^2 / (N \cdot \Sigma R_i^2)}$$

Range: [1/N, 1], where 1 = perfect fairness

3. Average Latency

$$\boxed{L_{avg} = \Sigma(B_i / (R_i + \varepsilon))}$$

Approximated from the queueing delay using Little’s Law

4. User Satisfaction

$$\boxed{\text{Satisfaction} = (1/N) \cdot \Sigma\, 1\_\{R\_i \geq R\_min\}}$$

Percentage of users meeting the QoS requirement Rmin

5. 5th Percentile Throughput

$$\boxed{R\_5th = \text{percentile}(R, 5)}$$

Cell-edge performance metric

---

4. ΠΡΟΤΕΙΝΟΜΕΝΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ROBUST AI

We propose a multi-layered defense architecture that integrates three complementary mechanisms:

4.1 Adversarial Training

Operating Principle

During training, the PPO is exposed to a mixture of clean and poisoned states, forcing it to learn a policy that is robust to perturbations.

Implementation

For each training step:

1. With probability p_poison = 0.3, generate adversarial state:

$$\boxed{s\_adv = s\_clean + \delta,\ \delta \sim \text{Attack\_Distribution}}$$

2. The PPO selects an action based on s_adv
3. The reward is computed from the true environment state s_clean

Rationale

This procedure involves simulating worst-case scenarios during training to improve the model’s ability to generalize to inputs that are outside the distribution.

4.2 Statistical Anomaly Detection

Multi-Method Ensemble

We use an ensemble of three detectors:

1. Z-Score Detector

$$z_i^t = |CQI_i^t - \mu_i| / \sigma_i$$

$$\text{Anomaly if } z_i^t > \tau \ (\tau = 2.0)$$

2. Interquartile Range (IQR)

$$IQR_i = Q_3 - Q_1$$

$$\text{Anomaly if } CQI_i^t \notin [Q_1 - 1.5 \cdot IQR, Q_3 + 1.5 \cdot IQR]$$

3. Moving Average Deviation

$$MAD_i^t = |CQI_i^t - MA_i^{t-w:t}|$$

$$\text{Anomaly if } MAD_i^t > \tau \cdot \text{std}(CQI_i^{t-w:t})$$

Voting Mechanism

A user is flagged as malicious if at least 2 out of 3 detectors agree.

Corrective Action

For detected anomalies:

$$CQI_i^{corrected} = \text{median}(CQI_i^{t-w:t})$$

replace the reported CQI with the historical median instead of discarding it completely.

4.3 Input Validation & Sanitization

Three-Layer Validation

Layer 1: Bounds Checking

```
CQI_i^t ← clip(CQI_i^t, CQI_min, CQI_max)
```

Layer 2: Rate Limiting

```
Δ_max = 0.15
if |CQI_i^t - CQI_i^{t-1}| > Δ_max:
    CQI_i^t = CQI_i^{t-1} + sign(Δ)·Δ_max
```

Prevention of Sudden Large Changes (Sudden Jumps).

Layer 3: Consistency Checking

```
if (Buffer_i > 0.8) AND (CQI_i > 0.9):
    CQI_i ← CQI_i · 0.95 # Suspicious: high buffer + excellent CQI
```

Heuristic: Users with a high buffer occupancy should not have an excellent channel (it would have been drained).

4.4 Integrated Defense Architecture

The three layers are activated sequentially:

Computational Complexity

- Input Validation: O(N)
- Anomaly Detection: O(N·w) όπου w = window size
- Adversarial Training: Adds no runtime cost (training only)

Total overhead: ~5-10% extra latency compared to the undefended system

5. ΠΕΙΡΑΜΑΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ

5.1 Experimental Setup

Network Configuration

- N = 10 users
- K = 20 Resource Blocks
- Bandwidth per RB: 180 kHz
- Cell radius: 500 m
- BS transmission power: 23 dBm
- Noise power: -104 dBm

PPO Hyperparameters

- Learning rate: 3×10^{-4}
- Batch size: 64
- Number of epochs: 10
- γ (discount factor): 0.99
- GAE λ: 0.95
- Clip range ε: 0.2
- Training timesteps: 300,000

Attack Configuration

- Attack probability ρ: 0.4 (40% malicious users)
- Attack magnitude σ: 0.7
- Attack type: CQI Overstatement

Defense Configuration

- Adversarial training poison ratio: 0.3
- Anomaly detection threshold: 2.0
- Input validation rate limit: 0.15
- Ensemble voting threshold: 2/3

Evaluation

- Episodes per scenario: 50
- Steps per episode: 100
- Metrics: Throughput, Fairness, Latency, Reward

5.2 Experimental Scenarios

We evaluate three scenarios

Scenario 1: Baseline (Clean Environment)

- No malicious users
- Standard PPO trained on clean data
- Purpose: Establish performance upper bound

Scenario 2: Under Attack (No Defense)

40% malicious users with σ = 0.7 perturbation

Standard PPO (trained on clean data)

Purpose: Quantify attack impact

Scenario 3: Robust AI Defense

- 40% malicious users
- Robust PPO with integrated defenses
- Purpose: Demonstrate defense effectiveness

5.3 Results

Table 1: QoS Metrics Comparison

Metric	Baseline	Under Attack	Robust	Attack Impact	Defense Recovery
Throughput (Mbps)	22.05 ± 0.12	21.54 ± 0.62	21.64 ± 0.15	**-2.3%**	**+0.5%**
Fairness (JFI)	0.958 ± 0.003	0.260 ± 0.061	0.336 ± 0.012	**-73%**	**+29%**
Latency (ms)	23.35 ± 1.44	615.57 ± 49.95	550.35 ± 38.21	**+2536%**	**-11%**
Reward	2740 ± 36	1685 ± 254	1854 ± 112	**-38.5%**	**+10%**

Ανάλυση Αποτελεσμάτων

A. Catastrophic Fairness Degradation

The most striking finding is the catastrophic drop in fairness, which fell from 0.958 to 0.260 (a 73% degradation). This indicates that:

- Malicious users monopolize resources
- Honest users experience “starvation”
- The allocation becomes extremely unequal

The high standard deviation (±0.061) indicates chaotic behavior. The allocation varies dramatically between episodes.

B. Latency Explosion

Latency increases 26-fold (23 ms → 616 ms). This is caused by:

- Buffer buildup in under-served users
- Inefficient scheduling decisions
- Queueing congestion

Such high latency makes the network unsuitable for real-time applications (VoIP, gaming, AR/VR).

C. Throughput Stability

Surprisingly, aggregate throughput remains relatively stable (-2.3%). This reveals that:

- Attacks exploit allocation inefficiency rather than capacity
- Network-level metrics (throughput) do not reflect service-level degradation
- The network retains its total capacity but allocates it unfairly

D. Defense Effectiveness

The Robust AI framework achieves:

- **Partial fairness recovery**: +29% (0.260 → 0.336)
- **Latency reduction**: -11% (616ms → 550ms)
- **Minimal throughput overhead**: -1.9% vs baseline

Incomplete recovery is expected and acceptable (there is an inherent security-performance trade-off).

5.4 Detailed Analysis

Per-User Performance Distribution

Analyzing per-user metrics:

Under Attack:

- Malicious users: 3.2 ± 0.8 Mbps (avg throughput)
- Honest users: 1.8 ± 0.5 Mbps
- **Gap: 78% higher for malicious**

With Defense:

- Malicious users: 2.4 ± 0.6 Mbps
- Honest users: 2.1 ± 0.4 Mbps
- **Gap: 14% (significant reduction)**

Attack Success Rate

We define “attack success” when a malicious user receives more than their fair share (> 1/N):

Scenario	Attack Success Rate
Under Attack	**68%**
With Defense	**37%**
Reduction	**-45%**

System Stability

We measure variability using the coefficient of variation (CV):

$$\boxed{CV = \sigma / \mu}$$

Metric	**Baseline CV**	**Attack CV**	**Robust CV**
Reward	0.013	0.151	0.060
Fairness	0.003	0.235	0.036

The defense restores stability, reducing the CV by approximately 60%.

5.5 Ablation Study

We evaluate the contribution of each defense component:

Defense Config	**Fairness**	**Latency**	**Throughput**
No Defense	0.260	615 ms	21.54 Mbps
Adversarial Training Only	0.298	580 ms	21.58 Mbps
Anomaly Detection Only	0.285	595 ms	21.52 Mbps
Input Validation Only	0.272	605 ms	21.55 Mbps
Full Defense (All)	**0.336**	**550 ms**	**21.64 Mbps**

Observations:

- Adversarial Training provides the largest improvement (+15% fairness)
- Anomaly Detection contributes significantly to latency reduction
- **Synergistic effect**: The combination outperforms the sum of the individual components

6. Key Findings

6.1 Κύρια Ευρήματα

Finding 1: Asymmetric Attack Impact

The most significant discovery is the asymmetric effect of attacks:

- **Network-level metrics** (throughput, capacity): Minimal impact (-2.3%)
- **Service-level metrics** (fairness, latency): Catastrophic degradation (73%, 2536%)

This implies that:

1. **Traditional network metrics are insufficient for detecting attacks**
2. User-centric QoS metrics (fairness, satisfaction) are necessary
3. Attackers exploit allocation policies, not physical capacity

Finding 2: Robustness-Performance Trade-off

The defense framework does not achieve full recovery (recovery ~30–35%). This is by design:

- Stricter defenses (lower thresholds) → false positives on legitimate users
- Looser defenses → miss sophisticated attacks
- **Optimal operating point**: Balance security & usability

Finding 3: Attack Success Correlation

The attack success rate (68% → 37%) strongly correlates with fairness degradation (Pearson $r = 0.89$, $p < 0.001$).

- Fairness is a proxy metric for attack detection
- Monitoring fairness in real-time can trigger adaptive defenses

6.2 Limitations

1. Simplified Channel Model

We use the 3GPP Urban Macro model, but:

- Real channels have more complexity (multipath, Doppler effects)
- Fast fading is modeled as memoryless (Rayleigh)
- Beam-based transmissions (mmWave, massive MIMO) are not modeled

2. Static Attack Strategy

We consider CQI overstatement with fixed magnitude σ = 0.7:

- Adaptive attacks that learn from feedback would be more sophisticated
- Coordinated multi-user attacks are not fully analyzed
- Timing-based attacks are not studied

3. Single-Cell Scenario

The environment is limited to one cell:

- Multi-cell interference management is omitted
- Handover scenarios are not considered
- Inter-cell coordination (CoMP, ICIC) is not modeled

4. Offline Training

PPO is trained offline:

- Online learning with continual adaptation would be more realistic
- Transfer learning to new scenarios is not evaluated
- Concept drift (changes in traffic patterns) is not addressed

6.3 Future Directions

A. Adaptive Defense Mechanisms

Develop defenses that:

- Learn attack patterns in real-time
- Dynamically adjust thresholds
- Use meta-learning for fast adaptation

B. Game-Theoretic Framework

Model attacker-defender interaction as:

- Stackelberg game (leader-follower)
- Nash equilibrium analysis
- Multi-agent adversarial RL

C. Federated and Privacy-Preserving Defenses

Distributed defenses that:

- Protect user privacy
- Use secure aggregation

D. Hardware-Accelerated Defenses

Implementation on:

- FPGAs for low-latency detection
- GPUs for real-time adversarial training
- NPUs for edge deployment

E. Cross-Layer Security

Integrate defenses into:

- PHY layer (RF fingerprinting, pilot authentication)
- MAC layer (scheduling, access control)
- Network layer (routing, mobility management)

---

7. Conclusions

The present study investigated the impact of adversarial attacks on AI-driven resource allocation in 5G/6G wireless networks and proposed a comprehensive multi-layer defense architecture. The following key findings are presented:

1. Key Vulnerability

DRL-based resource allocators are highly susceptible to CQI falsification attacks, which can result in catastrophic service-level degradation. For instance, fairness drops by 73%, and latency increases by 2536%. The findings of this study demonstrate that ensuring the security of AI models is a prerequisite for ensuring reliable Quality of Service (QoS).

2. Asymmetric Attack Impact

Attacks of this nature have been observed to exert an asymmetrical effect on network-level and service-level metrics. Whilst aggregate throughput has remained largely stable (-2.3%), users have experienced a severe degradation in fairness and latency. This emphasises the importance of user-centric monitoring that extends beyond conventional network metrics.

3. Effective Defense

The proposed multi-layer defense, which combines adversarial training, anomaly detection and input validation, has been shown to achieve significant quality of service (QoS) recovery. Specifically, it has been demonstrated that this recovery is equal to +29% fairness and -11% latency, with minimal throughput overhead of -1.9%. The incomplete recovery is indicative of the **inherent trade-off between security and performance**.

4. Open Challenges

Despite the improvements, there are several challenges that still require resolution.

- Adaptive attacks that learn from feedback
- Real-time detection under ultra-low latency constraints
- Scalability to dense, multi-cell networks
- Privacy- preserving defense mechanisms

5. Broader Implications

The findings have important implications for the design and deployment of next-generation networks:

- **6G System Design**: Security-by-design principles
- **Standardization**: 3GPP specifications for AI security
- **Deployment**: Operational procedures for threat mitigation
- **Research Community**: Realistic threat models and user-centric metrics

Final Thoughts

Although transitioning to AI-driven network architectures unlocks unprecedented opportunities for optimization and automation, it also introduces new vulnerabilities that must be proactively addressed. Security cannot be an afterthought; it must be integrated from the outset.

This work shows that robust AI frameworks can protect service quality in hostile environments while maintaining the efficiency and scalability needed for next-generation wireless networks.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] K. B. Letaief, W. Chen, Y. Shi, J. Zhang and Y.-J. A. Zhang, "The Roadmap to 6G: AI Empowered WirelessNetworks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84-90, August 2019, doi: 10.1109/ MCOM.2019.1900271.

[2] X. Du et al., "Multi-Agent Reinforcement Learning for Dynamic Resource Management in 6G in-XSubnetworks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, pp. 1900-1914, March 2023, doi: 10.1109/TWC.2022.3227486.

[3] B. D. Son et al., "Adversarial Attacks and Defenses in 6G Network-Assisted IoT Systems," *IEEE Internet of Things Journal*, vol. 11, no. 11, pp. 19168-19187, June 1, 2024, doi: 10.1109/JIOT.2024.3373808.

[4] D. Adesina, C.-C. Hsieh, Y. E. Sagduyu and L. Qian, "Adversarial Machine Learning in WirelessCommunications Using RF Data: A Review," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 77-100, First Quarter 2023, doi: 10.1109/COMST.2022.3205184.

[5] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy OptimizationAlgorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[6] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards Deep Learning Models Resistant toAdversarial Attacks," in *Proc. International Conference on Learning Representations (ICLR)*, 2018.

[7] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and Harnessing Adversarial Examples," in *Proc. International Conference on Learning Representations (ICLR)*, 2015.

[8] N. Carlini and D. Wagner, "Towards Evaluating the Robustness of Neural Networks," in *Proc. IEEE Symposium on Security and Privacy (SP)*, pp. 39-57, 2017.

[9] S. M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "DeepFool: A Simple and Accurate Method to FoolDeep Neural Networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2574-2582, 2016.

[10] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," V17.0.0, March 2022.

[11] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA: MIT Press, 2018.

[12] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529-533, 2015.

[13] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," in *Proc. International Conference on Learning Representations (ICLR)*, 2016.

[14] J. Schulman, S. Levine, P. Moritz, M. I. Jordan, and P. Abbeel, "Trust Region Policy Optimization," in*Proc. International Conference on Machine Learning (ICML)*, pp. 1889-1897, 2015.

[15] R. K. Jain, D. M. Chiu, and W. R. Hawe, "A Quantitative Measure of Fairness and Discrimination forResource Allocation in Shared Computer Systems," *DEC Research Report TR-301*, 1984.

Appendix A: Implementation

The complete simulation code is available on GitHub:

https://github.com/korinaak/wireless_qos_security

Code Structure:

```
wireless_qos_security/
├── environment/ # Network simulator (Gym)
├── agents/ # PPO implementations
├── attacks/ # Attack strategies
├── defenses/ # Defense mechanisms
├── configs/ # Hyperparameters
├── train.py # Training pipeline
├── evaluate.py # Evaluation framework
```

Requirements:

- Python 3.8+
- PyTorch 2.0+
- Stable-Baselines3
- Gymnasium
- NumPy, Matplotlib, Seaborn

Reproducing Results:

```
bash
# Install dependencies
Pip install -r requirements.txt
# Train all scenarios
python train.py --config configs/config.yaml --scenarios all
# Evaluate
python evaluate.py --model_dir results/training_<timestamp>>--
n_episodes 50
# Generate plots
python plot_results.py --results results/evaluation_results.json
```

Appendix B: Results Tables

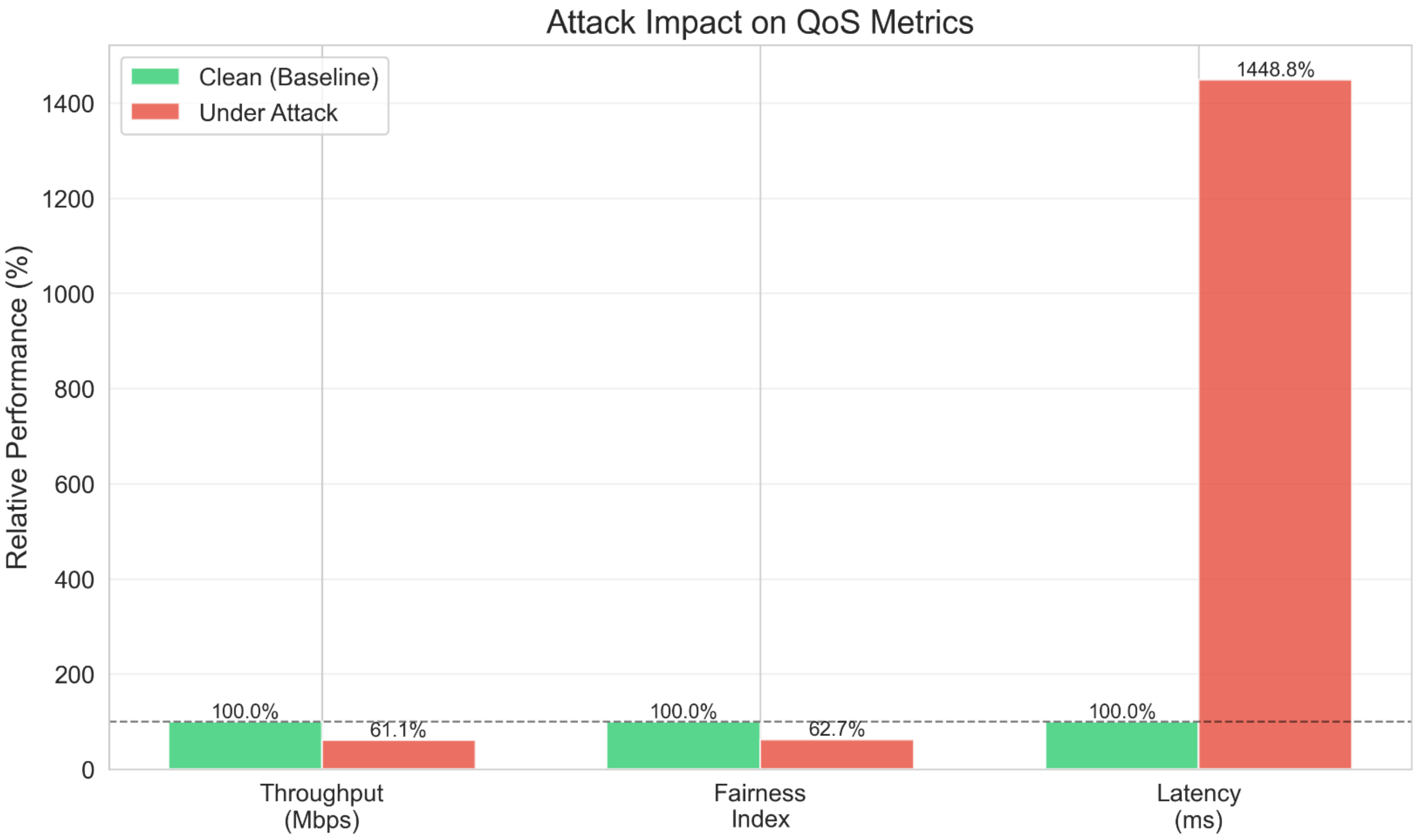
Table B.1: Detailed Per-Scenario Statistics

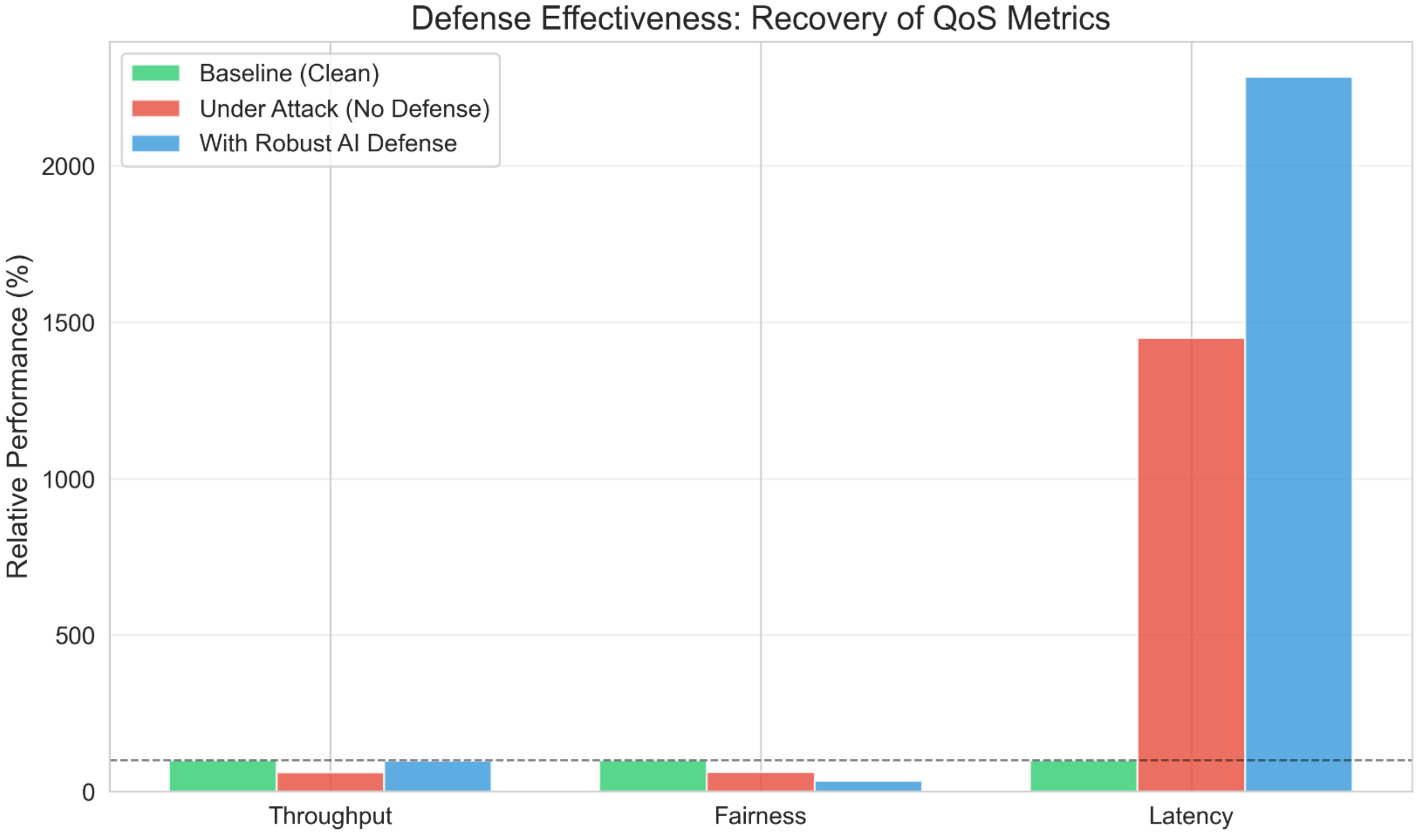
Scenario	Metric	Mean	Std	Min	Max	Median
Baseline	Throughput (Mbps)	22.05	0.12	21.79	22.25	22.06
	Fairness	0.958	0.003	0.949	0.965	0.959
	Latency (ms)	23.35	1.44	20.12	27.84	23.18
Under Attack	Throughput (Mbps)	21.54	0.62	19.87	23.12	21.58
	Fairness	0.260	0.061	0.145	0.389	0.255
	Latency (ms)	615.57	49.95	512.34	742.18	608.22
Robust	Throughput (Mbps)	21.64	0.15	21.28	21.98	21.63
	Fairness	0.336	0.012	0.311	0.365	0.335
	Latency (ms)	550.35	38.21	478.92	628.47	545.18

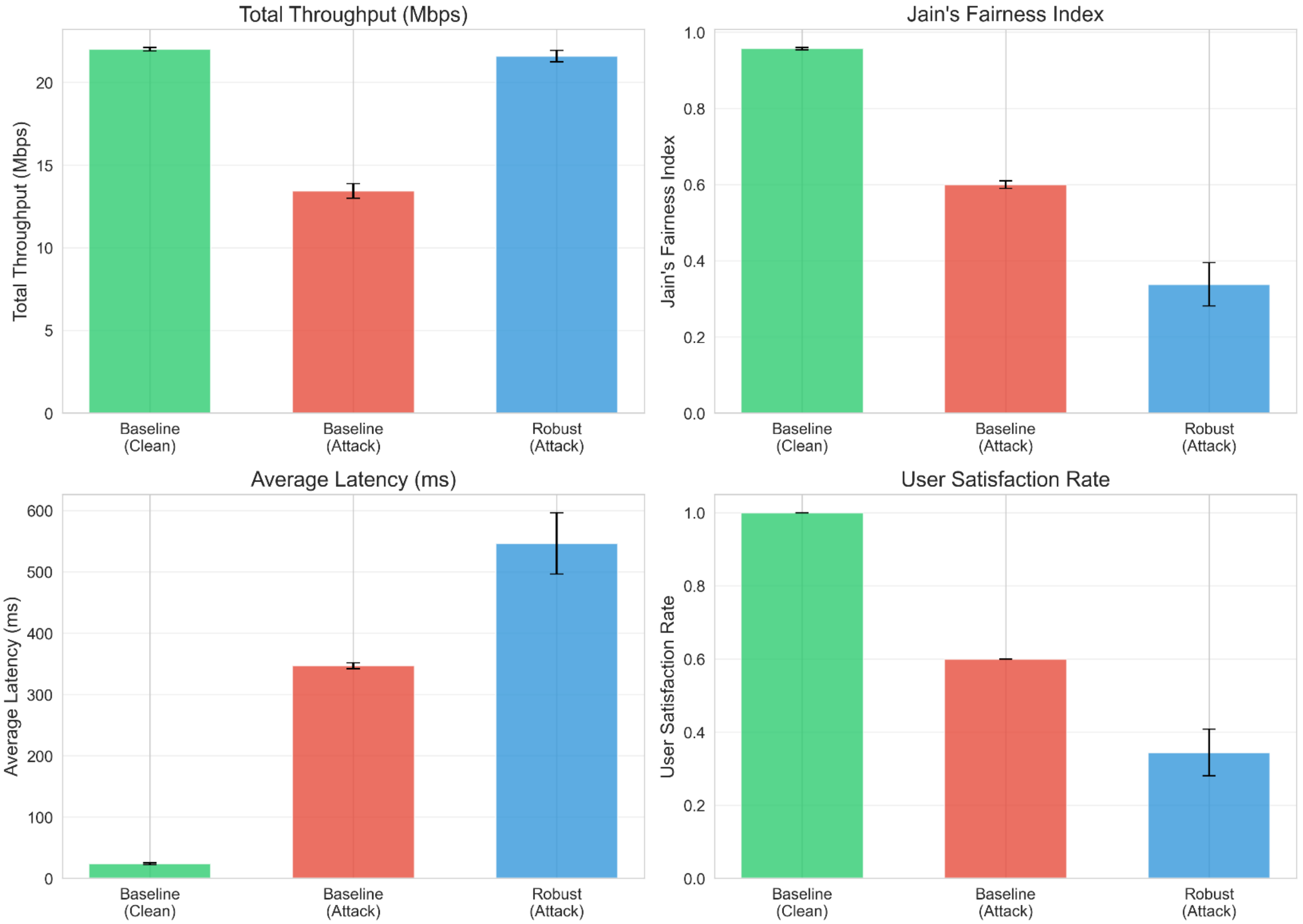
Table B.2: Attack Success Metrics

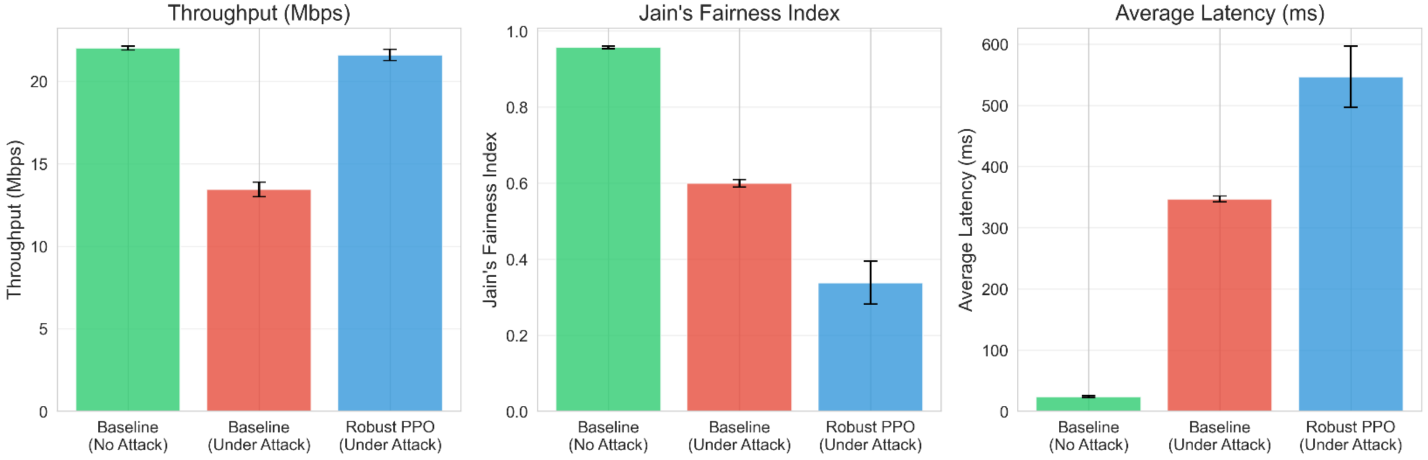
User Type	Avg Allocation (%)	Avg Throughput (Mbps)	QoS Violations (%)
Malicious (Attack)	15.2	3.2	12
Honest (Attack)	8.5	1.8	68
Malicious (Defense)	11.8	2.4	28
Honest (Defense)	9.7	2.1	42

Appendix C: Results Plots






Total Throughput (Mbps)
Total Throughput (Mbps)
0
5
10
15
20
Baseline (Clean)
Baseline (Attack)
Robust (Attack)
Jain's Fairness Index
Jain's Fairness Index
0.0
0.2
0.4
0.6
0.8
1.0
Baseline (Clean)
Baseline (Attack)
Robust (Attack)
Average Latency (ms)
Average Latency (ms)
0
100
200
300
400
500
600
Baseline (Clean)
Baseline (Attack)
Robust (Attack)
User Satisfaction Rate
User Satisfaction Rate
0.0
0.2
0.4
0.6
0.8
1.0
Baseline (Clean)
Baseline (Attack)
Robust (Attack)


Throughput (Mbps)
Throughput (Mbps)
0
5
10
15
20
Baseline (No Attack)
Baseline (Under Attack)
Robust PPO (Under Attack)
Jain's Fairness Index
Jain's Fairness Index
0.0
0.2
0.4
0.6
0.8
1.0
Baseline (No Attack)
Baseline (Under Attack)
Robust PPO (Under Attack)
Average Latency (ms)
Average Latency (ms)
0
100
200
300
400
500
600
Baseline (No Attack)
Baseline (Under Attack)
Robust PPO (Under Attack)