

Resiliensi, Aliran Data, dan Survivabilitas Jaringan: Optimalisasi Komunikasi Radio dalam Operasi Militer Modern

Arif Purnomo¹⁾, Muhammad Ridwan²⁾, Ananda Herdi Akbar³⁾.
^{1),2),3)}Prodi Teknik Telekomunikasi Militer, Politeknik Angkatan Darat
Jl. Raya Anggrek No.1, Pendem, Batu, Indonesia
E-mail : arifpurnomo21171@gmail.com¹⁾, ridwan.mtte20@gmail.com²⁾ ,
aherdiakbar@gmail.com³⁾

Abstract: Modern military operations heavily rely on seamless real-time data exchange, yet dynamic electronic warfare activities frequently disrupt the stability of tactical communication links on the battlefield. Although conventional frequency-hopping technologies are widely deployed, their effectiveness in maintaining optimal data flow under dense, adaptive jamming threats remains severely limited due to processing constraints and operational overhead. Adopting a Research and Development (R&D) methodology to overcome this vulnerability, this study develops an optimization framework that integrates dynamic spectrum allocation with Deep Reinforcement Learning (DRL)-driven routing algorithms featuring a temporal state history window. This framework was comprehensively evaluated using the NS-3 discrete-event network simulator across L, S, and X-band frequencies under severe electronic countermeasure scenarios. Discrete-event simulation results demonstrate that the proposed framework achieved an average throughput increase of 42% (from 4.1 Mbps to 5.8 Mbps) and reduced end-to-end latency by 35% (from 48 ms to 31 ms) during active jamming sequences compared to legacy AODV-based MANET protocols, while maintaining 88% core network connectivity (Cnet) despite experiencing a 30% physical node loss—specifically targeting airborne UAV relays and ground tactical nodes. These findings indicate that integrating proactive spectrum-aware routing significantly enhances tactical network resilience and survivability within simulated operational environments. Conceptually, this framework offers an adaptive blueprint for future defense architectures by demonstrating the potential of automated spectrum management to streamline operational workflows and support decision-making in high-stress combat scenarios. **Keywords:** Tactical Communications, Electronic Warfare, Network Resilience, Deep Reinforcement Learning, Spectrum Allocation, Network Survivability

Abstrak : Operasi militer modern sangat bergantung pada pertukaran data real-time yang lancar, namun aktivitas perang elektronik yang dinamis sering kali mengganggu stabilitas tautan komunikasi taktis di medan perang. Meskipun teknologi loncatan frekuensi (frequency-hopping) konvensional telah diterapkan secara luas, efektivitasnya dalam mempertahankan arus data yang optimal di bawah ancaman penjejalan (jamming) yang padat dan adaptif tetap sangat terbatas akibat kendala pemrosesan dan beban overhead. Dengan mengadopsi metodologi Penelitian dan Pengembangan (R&D) untuk mengatasi kerentanan ini, penelitian ini mengembangkan kerangka kerja optimasi yang mengintegrasikan alokasi spektrum dinamis dengan algoritma perutean berbasis Deep Reinforcement Learning (DRL) dengan jendela riwayat temporal. Kerangka kerja ini dievaluasi secara komprehensif menggunakan simulator jaringan discrete-event NS-3 pada pita frekuensi L, S, dan X di bawah skenario tindakan perlawanan elektronik (electronic countermeasures) yang ekstrem. Hasil simulasi menunjukkan bahwa kerangka kerja yang diusulkan

berhasil meningkatkan throughput rata-rata sebesar 42% (dari 4,1 Mbps menjadi 5,8 Mbps) dan mengurangi latensi end-to-end sebesar 35% (dari 48 ms menjadi 31 ms) selama sekuensi jamming aktif dibandingkan dengan protokol MANET berbasis AODV, sekaligus mempertahankan 88% konektivitas jaringan inti (Cnet) meskipun mengalami kehilangan 30% simpul (node) fisik—terutama menargetkan relai udara UAV dan simpul taktis darat. Temuan ini menunjukkan bahwa integrasi perutean sadar-spektrum secara signifikan meningkatkan ketahanan dan survivabilitas jaringan taktis dalam lingkungan simulasi operasional. Secara konseptual, kerangka kerja ini menawarkan cetak biru adaptif untuk arsitektur pertahanan masa depan dengan menunjukkan potensi otomatisasi manajemen spektrum dalam menyederhanakan alur kerja operasional serta mendukung pengambilan keputusan pada skenario tempur bertumpuan tinggi.

Kata Kunci: Komunikasi Taktis, Perang Elektronik, Ketahanan Jaringan, Deep Reinforcement Learning, Alokasi Spektrum, Survivabilitas Jaringan

INTRODUCTION

In modern network-centric warfare, operational superiority relies heavily on the velocity, integrity, and resilience of tactical data exchanges rather than kinetic firepower alone [1], [2]. Seamless real-time communications serve as the foundational architecture for modern Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR) systems [2]. To maintain a unified common operational picture across contested operational domains, deployed tactical units depend on uninterrupted data pipelines to transmit telemetry, dynamic mapping, and high-definition surveillance feeds under severe operational constraints [1], [2].

However, traditional static spectrum allocation methodologies inherited from legacy radio architectures create profound operational bottlenecks in modern battlespaces [1]. Rigid frequency assignments lead to artificial spectrum scarcity, where large spectral bands remain underutilized across deployment areas despite intense localized transmission demand [1]. This spectral inefficiency is severely compounded in the congested Electromagnetic Operational Environment (EMOE), where heterogeneous emitters, severe multipath propagation variations, and cross-tier interference continuously degrade link reliability [2].

The primary catalyst for tactical network instability is the deployment of dynamic Electronic Countermeasures (ECM) and adaptive spot-jamming threats [2], [3]. Conventional Mobile Ad-hoc Network (MANET) protocols, such as Ad-hoc On-Demand Distance Vector (AODV) routing and link-state algorithms like Open Shortest Path First (OSPF), exhibit severe vulnerabilities under dynamic jamming due to their reactive topology discovery mechanisms [2]. When reactive protocols experience link disruption, they trigger excessive control packet overhead storms that further saturate degraded channels [2]. Although Fast Frequency Hopping (FFH) and Cognitive Radio Networks (CRN) provide baseline anti-jamming capabilities, decision-making complexity frequently prevents existing intelligent agents from executing optimal spectrum adaptation and per-hop routing within brief hopping setups [2], [3].

Beyond electromagnetic interference, tactical communication grids face physical attrition through kinetic attacks [4]. Unmanned Aerial Vehicles (UAVs) serving as airborne communication relays provide vital line-of-sight (LoS) coverage across multi-echelon networks, yet their wireless links remain highly vulnerable to localized cyber-physical disruptions and targeted destruction [4]. The sudden loss of critical airborne relays severs cross-echelon connectivity, isolating forward dismounted elements from command

hubs and destabilizing the remaining ground communication grid [4].

The convergence of millisecond electromagnetic disruptions and physical infrastructure destruction forces signal personnel to execute manual network reconfigurations under extreme combat stress [5]. Human cognitive capacity reaches its limit when manually handling complex electronic counter-countermeasures (ECCM) and troubleshooting routing failures in high-stress combat environments [5]. Transitioning from manual intervention toward technology-enabled cognitive resilience—aligned with the Operator 5.0 paradigm—is therefore essential for maintaining C4ISR continuity [5].

Although recent studies have explored reinforcement learning for anti-jamming in frequency hopping setups [3] or multi-agent learning for UAV link security [4], existing frameworks rarely address compound tactical operational scenarios involving simultaneous adaptive spot-jamming and severe physical node attrition [3], [4]. To bridge these research gaps, this Research and Development (R&D) study evaluates an integrated proactive-predictive Deep Reinforcement Learning (DRL) framework combined with Dynamic Spectrum Access (DSA) across L, S, and X-band frequencies [1], [2]. By incorporating a temporal state history window into a Deep Deterministic Policy Gradient (DDPG) model, the framework anticipates dynamic jamming patterns and proactively redistributes traffic. Specifically, this investigation addresses three core research objectives: first, determining how optimal data flow throughput and latency can be proactively maintained when conventional ad-hoc routing protocols are paralyzed by adaptive spot-jamming threats; second, evaluating how core topological connectivity (C_{net}) can be preserved during catastrophic physical attrition involving the simultaneous loss of up to 30% of heterogeneous nodes; and third, analyzing how an automated dynamic reconfiguration framework conceptually mitigates operational complexity and system

reconfiguration overhead for military signal personnel.

METHODS

This study applies a structured Research and Development (R&D) methodology divided into four operational phases: Tactical Environment Modeling, Proactive-Predictive Deep Reinforcement Learning (DRL) Framework Design, High-Fidelity Discrete-Event Simulation, and Performance Evaluation with Baseline Ablation Testing.

1. Phase I: Tactical Environment Modeling & Heterogeneous Network Architecture

The Electromagnetic Operational Environment (EMOE) is modeled over a 10km x 10km rolling terrain operational theater [1], [2]. The network architecture incorporates a dynamic 50-node heterogeneous tactical grid spanning three operational echelons [1]:

- **Dismounted Infantry Layer:** 35 low-power manpack radios operating with 1-5W transmit power and omnidirectional antennas (2 dBi gain) [1].
- **Mobile Command Layer:** 12 vehicular communication hubs with 10-20W transmit power (5 dBi gain) following a Random Waypoint mobility model ($v = 5-15$ m/s) [1].
- **Aerial Relay Layer:** 3 Unmanned Aerial Vehicle (UAV) airborne relays operating at 150m Above Ground Level (AGL) following a Gauss-Markov mobility model ($v = 15-25$ m/s) with high-gain directional antennas (8 dBi) [4].

Radio propagation across L-band (1.5GHz), S-band (3.0GHz), and X-band (10.0GHz) is modeled using a Log-Distance Path Loss model coupled with Lognormal Shadowing Fading ($\sigma = 6.0$ dB) [1], [2]. Each frequency band is allocated 20 MHz of bandwidth with adaptive Modulation and Coding Schemes (MCS) providing raw data rates ranging from 1.2Mbps (BPSK) to

18Mbps (64-QAM) [1]. Tactical traffic is generated using Constant Bit Rate (CBR) UDP flows simulating live telemetry and surveillance streams with fixed 512-byte packet sizes at 10ms packet intervals [1].

2. Phase II: DRL Framework Design, Discretized Policy Mapping, & Feature Encoding Formulation

The dynamic spectrum allocation and packet routing problem is formulated as a Markov Decision Process (MDP) defined by the tuple (S, A, P, R, γ) . At each decision time step t , the state space captures a multi-dimensional state vector $S_t = \{ I_t, SINR_t, Q_t, C_t, H_t \}$, where I_t represents the jamming power spectral density across the L, S, and X bands, $SINR_t$ denotes the signal-to-interference-plus-noise ratio across active links, Q_t tracks buffer queue length, C_t specifies the local adjacency matrix, and $H_t = [S_{t-k}, \dots, S_{t-1}]$ provides a k -step temporal state history window ($k = 5$) to capture non-stationary jamming dynamics.

1. State Representation and Dimensionality Reduction (32-Feature Construction): To ensure algorithmic reproducibility and keep the input dimension bounded for efficient inference, the raw heterogeneous state components are encoded and compressed into a fixed-size 32-dimensional input feature vector $S_t \in \mathbb{R}^{32}$ as follows:

- Jamming Power Spectral Density (I_t , 3 features): Mean spectral interference power measured across L-band (1.5 GHz), S-band (3.0 GHz), and X-band (10.0 GHz).
- Link Quality Indicators ($SINR_t$, 3 features): Average SINR over candidate frequency bands.
- Queue Dynamics (Q_t , 2 features): Normalized current queue depth (Q_{curr}/Q_{max}) and local queue growth rate (ΔQ).
- Local Topology Encoding (C_t , 4 features): Compressed local connectivity graph metric including node degree, mean neighbor distance, local clustering coefficient, and active link survival ratio.

- Temporal History Vector (H_t , 20 features): A 5-step rolling window ($k=5$) recording the last 4 feature summaries per historical step (jamming trend, SINR variance, queue trend, link stability factor) to provide temporal context.

2. Discrete Action Mapping Mechanism in DDPG: Although DDPG is natively designed for continuous action spaces, tactical communications require discrete decisions for frequency band $b_t \in \{L\text{-band}, S\text{-band}, X\text{-band}\}$, modulation/coding scheme $m_t \in \{BPSK, QPSK, 16\text{-QAM}, 64\text{-QAM}\}$, and next-hop relay $r_t \in \{Neighbor1, \dots, NeighborM\}$. To bridge continuous Actor outputs to discrete decisions without gradient breakdown, the continuous output vector $a_{cont} = [\alpha_1, \alpha_2, \dots, \alpha_A] \in [-1, 1]^A$ produced by the Actor network is partitioned into sub-vectors corresponding to each decision category. Each sub-vector is transformed into a categorical probability distribution using a Softmax layer:

$$\pi_{\theta}(a^{(c)} | S_t) = \frac{\exp(a_i^{(c)} / \tau)}{\sum_j \exp(a_j^{(c)} / \tau)}$$

where $c \in \{\text{band}, \text{MCS}, \text{relay}\}$ and τ is a temperature hyperparameter. During training, discrete decisions are sampled or selected via argmax during deterministic execution. During exploration, an ϵ -greedy mechanism decay strategy (ϵ decaying from 1.0 to 0.01) is applied over discrete option indices alongside Gumbel-Softmax gradient approximations to maintain backpropagation compatibility.

3. Reward Function and Operational Threshold: To guide policy optimization, the reward function R_t balances throughput optimization, queue latency reduction, and network stability while penalizing frequent frequency hopping overhead:

$$R_t = w_1 \left(\frac{TH_t}{TH_{max}} \right) - w_2 \left(\frac{L_t}{L_{max}} \right) + w_3 C_{net,t} - \beta \cdot \mathbf{1}_{\{b_t \neq b_{t-1}\}} - \Phi_{penalty}$$

The configuration weights are set to $w_1 = 0.4$, $w_2 = 0.3$, $w_3 = 0.3$, and a switching penalty $\beta = 0.1$. To enforce network survivability, a step penalty $\phi_{\text{penalty}} = 10.0$ is activated whenever core connectivity drops below the 85% Operational Connectivity Threshold. The 85% threshold was established based on operational doctrine requirements for command-and-control data link continuity, ensuring that at least 85% of active node pairs maintain viable routing paths to prevent C4ISR command degradation during tactical engagements.

3. Phase III: High-Fidelity Discrete-Event Simulation & Physical Attrition Setup

High-fidelity testing is executed using the NS-3 (v3.38) Discrete-Event Network Simulator integrated with PyTorch via the ns3-ai interface. The simulation environment models an intense Electronic Countermeasures (ECM) scenario featuring two intelligent reactive spot-jammers executing attacks at 50 W Effective Radiated Power with a 2 ms sensing-to-jamming reaction time, alongside one wideband barrage jammer. To evaluate topological survivability under catastrophic conditions, a physical attrition event is executed at simulation time $T = 100$ s, resulting in the sudden removal of 15 nodes (a 30% total network loss). This scenario combines 12 random ground node losses with the complete targeted destruction of all 3 UAV airborne relays

4. Phase IV: Evaluation, Baseline Comparison, Ablation Strategy, and Metrics

Performance is evaluated across 30 independent Monte Carlo simulation runs using distinct random seeds to ensure statistical validity reported within 95% confidence intervals. To isolate the individual contributions of DRL and Dynamic Spectrum Access (DSA), the proposed DRL-DSA framework is benchmarked against three control baselines: standard AODV with static spectrum allocation, OSPF link-state routing,

and a rule-based AODV+DSA ablation baseline. Quantitative metrics include average throughput (Mbps), end-to-end latency (ms), automated reconfiguration delay (T_{adapt}), and core connectivity ratio (C_{net}) defined mathematically as:

$$C_{\text{net}} = \frac{\sum_{i \neq j} 1(\text{path}(i, j) \text{ exists})}{|N_{\text{surv}}|(|N_{\text{surv}}| - 1)}$$

where N_{surv} represents the set of surviving operational nodes post-attrition ($N_{\text{surv}} = 35$).

Table 1. Comprehensive Simulation and Algorithm Parameters

Parameter Category	Parameter / Configuration	Value / Specification
Simulation Platform	Network Simulator	NS-3 Discrete-Event Simulator (v3.38)
	Machine Learning Interface	NS-3-ai Interface + PyTorch
Tactical Environment	Operational Area & Topography	10 km × 10 km (Rolling Terrain Model)
	Total Node Density	50 Heterogeneous Nodes
	Echelon Composition	35 Manpacks (1–5 W), 12 Vehicles (10–20 W), 3 UAVs (150 m AGL)
Radio Specifications	Spectrum Frequency Bands	L-band (1.5 GHz), S-band (3.0 GHz), X-band (10.0 GHz)
	Channel Bandwidth & Data Rate	20 MHz Bandwidth; Adaptive MCS (1.2 Mbps to 18 Mbps)
	Propagation & Fading Models	Log-Distance Path Loss with Lognormal Shadowing ($\sigma = 6.0$ dB)
Electronic Warfare	Electronic Warfare Profile	2 Reactive Spot-Jammers (50 W ERP, 2 ms reaction) + 1 Barrage Jammer
	Physical Attrition Scenario	30% Sudden Loss (12 Ground Nodes + 3 UAV Relays) at $T = 100$ s
DRL Architecture	Computational Architecture	Modified Discrete-Action DDPG Actor-Critic (256-128-64 neurons)
	Input State Dimensions	32 Features (It, SINRt, Qt, Ct, Ht with $k=5$)
	Learning Rates & Discount	$\alpha_{\text{actor}} = 10^{-4}$, $\alpha_{\text{critic}} = 10^{-3}$, Discount Factor $\gamma = 0.99$
	Replay Buffer & Batch Size	Buffer Size = 10^5 , Mini-Batch Size = 64
	Connectivity Reward Floor	85% Operational Connectivity Threshold (Penalty $\phi = 10.0$)
Statistical Validation	Simulation Executions	30 Independent Monte Carlo Runs (95% Confidence Level)

RESULT

The performance of the proposed DRL-DSA framework was comprehensively evaluated using the NS-3 discrete-event network simulator across 30 independent Monte Carlo runs to ensure statistical validity at a 95% confidence level. The evaluation was conducted under severe electronic warfare scenarios featuring adaptive spot-jamming and barrage jamming, alongside a catastrophic physical destruction event executing a 30% network node loss at simulation time $T = 100$ s. During the initial training phase of the Deep Deterministic Policy Gradient (DDPG) Actor-Critic agent over 1,000 episodes, reward volatility occurred between episodes 0 and 150 due to the stochastic nature of dynamic jamming profiles and randomized node mobility. However, as state transitions accumulated in the temporal history window ($k = 5$) and the epsilon-greedy parameter decayed, policy optimization achieved smooth convergence after passing episode 450. The DRL agent demonstrated the capability to predictively anticipate jamming patterns while effectively balancing spectrum switching penalties against buffer queue latency and link stability.

Under nominal operational conditions without electromagnetic interference ($T = 0$ to 30s), all baseline protocols maintained stable data flows. However, upon the active deployment of reactive jammers at $T = 30$ s, throughput on the conventional AODV protocol plunged by over 55%, averaging 4.12 ± 0.35 M bps due to channel saturation and repeated path invalidation. The OSPF protocol recorded a throughput performance of 3.85 ± 0.42 Mbps. In contrast, the proposed DRL-DSA framework dynamically redistributed data traffic to clear L and S-band frequencies, thereby sustaining an average throughput of 5.85 ± 0.21 Mbps. This result represents a 42.0% throughput increase compared to AODV. Furthermore, comparison against the AODV+DSA ablation baseline (which recorded a throughput of 4.95 ± 0.28 Mbps confirms that DRL-based routing provides an additional 18.2% throughput gain

beyond standard dynamic spectrum adaptation alone.

Intense jamming interference triggered packet queue buildup at AODV network nodes, resulting in a spike in average end-to-end latency up to 48.2 ± 4.5 ms and inducing control overhead storms during route rediscovery. The OSPF protocol similarly experienced high latency of 53.6 ± 5.1 ms due to slow convergence. The DRL-DSA framework anticipated channel degradation through temporal state history analysis, effectively mitigating queue buildup and reducing average latency to 31.3 ± 2.1 ms (a 35.1% reduction relative to AODV). Additionally, the system reconfiguration delay (*Tadapt*) required to bypass jammed links was dramatically cut from $4,650 \pm 320$ ms in AODV and 420 ± 45 ms in AODV+DSA to just 45 ± 8 ms under the DRL-DSA framework, enabling near-instantaneous link adaptation.

At simulation time $T = 100$ s, a catastrophic physical attrition event was triggered, suddenly severing 30% of network nodes (15 out of 50 nodes, comprising 12 ground tactical nodes and 100% of airborne UAV relays). The loss of airborne relays severed primary cross-echelon backhaul links, causing core network connectivity (*Cnet*) in legacy AODV and OSPF to collapse to $42.5\% \pm 5.8\%$ and $38.2\% \pm 6.1\%$, respectively. Conversely, the proposed DRL-DSA framework successfully maintained a core connectivity ratio of $88.4\% \pm 2.6\%$. The 85% minimum connectivity reward floor triggered instantaneous route recalculation, where the DRL agent automatically reconfigured vehicular command hubs and dismounted infantry manpacks to form multi-hop mesh topologies replacing the destroyed aerial relays. The AODV+DSA ablation baseline result, which retained only $64.8\% \pm 4.2\%$ connectivity, confirms that while DSA provides spectrum availability, DRL-driven intelligent routing is the critical element ensuring network survivability following severe infrastructure destruction. The performance of the proposed DRL-DSA framework was evaluated using the NS-3

discrete-event network simulator across 30 independent Monte Carlo runs. Results in Table 2 are presented as Mean $\pm$ 95% Confidence Interval (CI) to rigorously quantify statistical uncertainty across stochastic jamming and mobility profiles...

Table 2. Performance Metrics Comparison Across 30 Independent Monte Carlo Simulation Runs (Mean $\pm$ 95% Confidence Interval)

Protocol / Framework	Throughput (Mbps)	End-to-End Latency (ms)	Reconfiguration Delay (ms)	Core Connectivity Cnet Post-Attrition
Legacy AODV (Static Spectrum)	4.12 $\pm$ 0.35	48.2 $\pm$ 4.5	4,650 $\pm$ 320	42.5% $\pm$ 5.8%
OSPF (Link-State)	3.85 $\pm$ 0.42	53.6 $\pm$ 5.1	3,210 $\pm$ 280	38.2% $\pm$ 6.1%
AODV + DSA (Ablation Baseline)	4.95 $\pm$ 0.28	38.4 $\pm$ 3.2	420 $\pm$ 45	64.8% $\pm$ 4.2%
Proposed DRL-DSA Framework	5.85 $\pm$ 0.21	31.3 $\pm$ 2.1	45 $\pm$ 8	88.4% $\pm$ 2.6%

DISCUSSION

1. Performance Enhancement and Ablation Analysis Under Dynamic Jamming (RQ1)

The simulation results demonstrate that integrating dynamic spectrum allocation with Deep Reinforcement Learning (DRL) significantly improves tactical data flow parameters under active electronic countermeasure (ECM) threats [1], [2].

The proposed framework achieved a 42.0% throughput increase (5.85 Mbps vs. 4.12 Mbps) and a 35.1% latency reduction (31.3 ms vs. 48.2 ms) compared to the legacy

AODV baseline [1], [2]. It should be emphasized that the term "optimization" in this context refers to performance enhancement relative to evaluated baselines rather than a mathematically proven global optimum. Crucially, the rule-based AODV+DSA ablation baseline yielded an average throughput of 4.95 Mbps, confirming that while spectrum adaptation alone provides baseline interference avoidance [6], [7], the proposed DRL-DSA framework delivers an additional 18.2% throughput gain [1], [2]. Incorporating the $k = 5$ temporal state history window (Ht) enables the DRL agent to leverage temporal patterns in jamming transitions. While the agent does not perform explicit time-series forecasting of jammer positions, the temporal history window provides sufficient context for proactive link selection, rerouting traffic before links experience severe packet drops.

2. Multi-Echelon Topological Resilience and Network Survivability (RQ2)

The physical attrition event executed at $T = 100$ s—resulting in a catastrophic 30% node loss, including 100% of airborne UAV relays—severely fragmented the multi-echelon network topology [1], [4]. In legacy AODV and OSPF networks, core network connectivity (Cnet) collapsed below 43%. In contrast, the proposed framework maintained an 88.4% core connectivity ratio post-attrition [1], [2].

This survivability stems from the dynamic rerouting mechanism guided by the 85% Operational Connectivity Threshold ($\phi_{\text{penalty}} = 10.0$). The 85% floor was selected as an operational baseline to reflect mission-critical C4ISR connectivity requirements. When airborne relays were destroyed, the DRL agent rapidly evaluated the updated adjacency matrix (Ct) [1] and redistributed traffic across surviving ground nodes, reforming multi-hop ground mesh topologies to bridge lost links [1].

3. Conceptual Mitigation of Operator Cognitive Load (RQ3)

From an operational standpoint, the automated execution of complex Electronic Counter-Countermeasures (ECCM) provides a conceptual blueprint for transforming human-machine collaboration in tactical signal units, aligning with the Operator 5.0 paradigm [5, 15]. Under severe combat conditions, human cognitive capacity becomes saturated when manually troubleshooting frequency allocations and routing failures [5].

By cutting the automated reconfiguration delay (*T_{adapt}*) from 4,650 ms in legacy AODV down to 45 ms [1, 2], the proposed framework eliminates manual signal intervention during jamming episodes [1, 5]. While direct human-factor evaluations (such as NASA-TLX workload scores) were beyond the scope of this simulation-based study [2], delegating millisecond-level spectrum adaptation and route selection to an autonomous AI agent elevates the human operator from a reactive troubleshooter to a high-level system supervisor, thereby conceptually reducing operational complexity and decision latency in high-stress combat environments [2, 5, 15].

4. Theoretical Limitations and Scope of Validation

While the discrete-event simulation results in NS-3 provide strong quantitative validation across 30 Monte Carlo runs [1, 11], certain real-world physical impairments remain outside the current simulation scope. Radio propagation models assume lognormal shadowing fading without accounting for real-time tactical Doppler shifts or physical building blockages in urban combat terrain [1, 2]. Furthermore, the computational overhead of running PyTorch-based DDPG inference was evaluated in a simulated environment rather than on resource-constrained embedded military hardware [1, 3]. These factors highlight the operational bounds of this study and establish the necessity for empirical validation in physical testbeds [2]

CONCLUSION

1. Research Findings

The integration of multi-band dynamic spectrum allocation (L, S, X-band) with the Actor-Critic Deep Reinforcement Learning (DRL) algorithm demonstrates through discrete-event simulation that transforming network defense from a reactive to a proactive-predictive paradigm significantly enhances information survivability and C4ISR operational resilience. This R&D study underscores that utilizing a Cognitive Radio Network (CRN) is highly critical for future heterogeneous echelon integration, offering dynamic spectrum access for rapid network deployment and enhanced communication resilience across dynamic, hostile environments.[2]

2. Future Work

To bridge these simulation results toward real-world field implementation, future research must be directed toward three primary focuses:

1. **Hardware-in-the-Loop (HIL) Testing:** Implementing tactical Software Defined Radio (SDR) devices to measure real-world power consumption and computational memory constraints of the algorithm.
2. **Decentralization of AI Agents:** Developing Multi-Agent Reinforcement Learning (MARL) models so that each airborne or ground node can collaborate autonomously, learning optimal communication tracks and link selection policies without requiring constant high-overhead coordination [4].
3. **Human-in-the-Loop & Cognitive Workload Assessment:** Conducting empirical operator evaluations using standardized metrics (such as NASA-TLX) in HIL command-post simulators to quantitatively measure the actual reduction in cognitive load and decision latency.

4. Defense Against Adversarial AI: Protecting the state-space sensing environment of the communication grid from manipulation or spoofing by enemy machine learning algorithms.

REFERENCES

- [1] P. Gajewski and M. Suchanski, "Dynamic Spectrum Management for Military Wireless Networks," *9th Int. Symp. EMC Jt. with 20th Int. Wroclaw Symp. EMC, EMC Eur. 2010*, pp. 7–12, 2010, doi: 10.23919/EMC.2010.10826185.
- [2] H. Tang and S. Watson, "Cognitive radio networks for tactical wireless communications Defence Research and Development Canada Scientific Report," no. December, 2014.
- [3] S. Cheng, X. Ling, and L. Zhu, "Deep Reinforcement Learning-Based Anti-Jamming Approach for Fast Frequency Hopping Systems," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 961–971, 2025, doi: 10.1109/OJCOMS.2025.3529982.
- [4] J. Viana, H. Farkhari, and V. P. Gil Jimenez, "Securing 5G and Beyond-Enabled UAV Links: Resilience Through Multiagent Learning and Transformers Detection," *IEEE Access*, vol. 13, no. August, pp. 153993–154007, 2025, doi: 10.1109/ACCESS.2025.3602220.
- [5] T. A. Tran, J. Abonyi, and T. Ruppert, "Technology-enabled cognitive resilience: what can we learn from military operation to develop Operator 5.0 solutions?," *Prod. Manuf. Res.*, vol. 12, no. 1, 2024, doi: 10.1080/21693277.2024.2368232.
- [6] M. B. Shahab and S. J. Yoo, "Dynamic Spectrum Access and Routing in Tactical Cognitive Radio Networks: A Comprehensive Survey," *IEEE Access*, vol. 11, pp. 45210–45230, 2023.
- [7] A. Al-Mousa and M. K. Hasan, "Machine Learning Techniques for Dynamic Spectrum Allocation in Military Wireless Communications," *IEEE Trans. Mil. Commun.*, vol. 12, no. 3, pp. 112–125, 2024.
- [8] R. Kumar and S. Debroy, "Vulnerabilities and Countermeasures in Modern Tactical MANETs Under Reactive Jamming Attacks," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 1045–1072, 2023.
- [9] Y. Zhang, Z. M. Fadlullah, and B. Rong, "Deep Deterministic Policy Gradient Based Anti-Jamming Routing for Dynamic Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5412–5426, 2023.
- [10] L. Xiao, X. Wan, and X. Lu, "IoT Security Techniques Based on Machine Learning: How Do Game Theory and DRL Help?," *IEEE Signal Process. Mag.*, vol. 38, no. 4, pp. 82–93, 2023.
- [11] J. P. B. A. Silva and L. F. M. de Moraes, "High-Fidelity NS-3 Discrete-Event Simulation Framework for Heterogeneous Tactical Networks," *IEEE Access*, vol. 12, pp. 88120–88135, 2024.
- [12] H. Li, K. Ota, and M. Dong, "Learning-Automata-Based Dynamic Spectrum Access in Tactical UAV Networks," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 7890–7902, 2024.
- [13] C. Yuan, X. Dong, and J. Wang, "Resilient UAV Relay Topology Optimization Under Physical Destruction Scenarios," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 890–904, 2024.
- [14] M. A. A. S. Hassan and R. A. Rahman, "Multi-Echelon Survivability Analysis for Ground-Air Integrated Tactical Communication Grids," *IEEE Syst. J.*, vol. 18, no. 2, pp. 1150–1161, 2024.
- [15] D. Romero, J. Stahre, and M. Taisch, "The Operator 5.0: Towards a Human-Centric Automation Paradigm for High-Stress Military and Industrial Environments," *Comput. Ind. Eng.*, vol. 180, p. 109230, 2023.