

Double Deep Q-Network-Based Smart Channel Hopping for Reactive Jamming Mitigation in IoT Networks: A Reliability and Energy-Aware Approach

Enos Letsoalo¹, Topside Mathonsi², Tshimangadzo Tshilongamulenzhe³, Daniel du Duplesis⁴
Department of Information Technology, Tshwane University of Technology, Pretoria, South Africa^{1, 2, 3}
Department of Computer Science, University of Johannesburg, Johannesburg, South Africa⁴

Abstract—Reactive jamming attacks are prevalent in low power Internet of Things networks and are severely affecting the performance of IoT nodes. Existing anti-jamming studies focus primarily on communication reliability. Limited attention has been given to the joint optimization of packet delivery ratio (PDR), energy consumption, and network lifetime in mobile IoT networks under reactive jamming attacks. This study proposes a double deep Q-network (DDQN)-based smart channel hopping model (SCHM) that improves communication reliability while reducing energy expenditure and extending network lifetime. We use a Markov Decision Process (MDP) for modelling the interaction between the jammer and the legitimate IoT node. We develop a DDQN-based SCHM algorithm using the Python programming language. We implement the DDQN-based SCHM approach in the NS-3.41 environment. We evaluated the effectiveness of DDQN-based SCHM against Deep-Q-Network (DQN)-based hopping, non-hopping, and random hopping approaches. The results showed that our proposed DDQN-based SCHM approach achieved an average PDR of 99%, experienced the lowest energy consumption, and prolonged the network lifetime of IoT nodes. The proposed DDQN-based SCHM approach improves communication reliability and energy efficiency of IoT nodes.

Keywords—DDQN; SCHM; MDP; reactive jamming attacks; PDR; energy consumption; network lifetime

I. INTRODUCTION

Various application domains such as smart homes, smart cities, and environmental monitoring are beneficiaries of IoT network systems and sensor systems. IEEE 802.15.4 specifically benefits from IoT network systems [1], [2], [3]. For example, IEEE 802.15.4 enables users with wearable devices and ingestible sensors to be monitored remotely [4]. This helps healthcare professionals remotely monitor patients with chronic diseases for better disease management. IEEE 802.15.4 IoT networks are vulnerable to various attacks, specifically reactive jamming attacks that affect their performance severely.

IoT networks are more susceptible to various cyber-attacks because they inherit common characteristics of wireless networks, namely: open access and shared access. Exposure to several malicious attacks originates from factors such as IoT nodes having minimal or no security features embedded in them [5].

This is because IoT nodes have energy constraints imposed on them and have limited computational capabilities. Therefore,

the fact that IoT nodes utilize a shared transmission medium is also a concern [6].

During a reactive jamming attack, an attacker listens to activities in the channels and transmits high signal power to the active channel with the intention of disrupting communication [7]. A disrupted IoT node will try to retransmit the packets because of a lack of acknowledgment from the receiver. Rapid retransmission of packets can lead to a quick battery depletion of IoT nodes [8], [9]. This will further lead to a short lifetime of IoT nodes in the network. Furthermore, prolonged transmission delays and low PDR may be incurred. In IEEE 802.15.4 IoT networks, applications such as remote patient monitoring (RPM) with wearables or ingestible sensors will be left unmonitored due to jamming attacks, leading to poor chronic disease management [4].

Wireless sensor networks (WSNs) are applications of IoT networks and are susceptible to reactive jamming attacks. Sensor nodes are deployed in unattended areas and can be easily compromised. Reactive jamming attacks are frequently employed by attackers to disrupt normal operations of WSNs. The disruption of communication through jamming often leaves areas unmonitored in WSNs. The study by [10] suggested that jamming of communication channels can be mitigated by raising the transmission power of legitimate nodes. However, increasing power levels may lead to quick battery depletion of IoT nodes, leaving systems unmonitored [11].

This research proposes a DDQN-based SCHM approach, which is applied to mitigating reactive jamming attacks in IoT networks. This study uses the MDP for the formulation of jammer and legitimate node interaction. DDQN-based SCHM server, which runs independently, is used for training the learning agent. This research combines IoT networks, specifically 802.15.4, PDR, energy consumption, channel hopping, and reinforcement learning in one NS-3.41 framework. The performance metrics of our proposed system are: PDR, energy consumption, network lifetime, and packet loss. The NS-3.41 testbed that was used includes nine IoT nodes, the access point as the sink node, and one reactive jammer. The effectiveness of the proposed SCHM approach was validated against the DQN-based hopping approach, non-hopping approach, and random hopping approach. The obtained simulation results demonstrated that the proposed DDQN-based SCHM has increased PDR by an average of 6%, reduced energy consumption by IoT nodes by an average of 77.5%, prolonged

the DDQN framework, namely the prediction Q-network $Q(s_t, a_t, \theta)$, the target Q-network $Q(s_t, a_t, \theta^-)$, and the memory buffer B.

The prediction network $Q(s_t, a_t, \theta)$ is used to approximate the Q-value; then the target Q-network $Q(s_t, a_t, \theta^-)$ and the memory buffer B are used to ensure that the DDQN is stable and that the target Q-value does not oscillate. The DDQN algorithm's advantage over the Q-learning algorithm is the ability to record the transition (s_t, a_t, r_t, s_{t+1}) , i.e., the current state, the action, the immediate reward, and the next state observed by the legitimate node, in a memory buffer B and then train the Q-network every epoch time slot. During Q-network training, buffered transitions are sampled randomly from a memory buffer and are used to compute the corresponding target Q-value. Table I gives a summary of the symbols and their description.

Algorithm 1: DDQN-Based SCHM Algorithm

1. Initialization of prediction Q-network Q with random weight θ , target Q-network Q^- with random weight θ^- and the replay memory B with capacity M
2. for episode = 1,2,3,...,E **do**
 Select two channels randomly from set C to initialize the state S, where N channels are used for communication and the other channel for hopping
3. for timeslot $t = 1$ to T **do**
 Choose a random action a_t with probability ϵ , otherwise choose $a_t = \text{argmax}_a Q(s_t, a, \theta)$
4. Execute action a_t and receive the immediate reward r_t , then obtain next state s_{t+1}
5. Save transition (s_t, a_t, r_t, s_{t+1}) in replay memory B
6. Select random samples (s_i, a_i, r_i, s_{i+1}) from replay memory B
7. Calculate target Q-value using the Q-value function below $Y_j = r_j + \gamma Q(s_{j+1}, \text{argmax}_a Q(s_{j+1}, a, \theta^-); \theta^-)$
8. The weights are then optimized by applying gradient descent using parameter θ to minimize the loss using the equation below
$$(y_j - Q(s_j, a_j; \theta))^2$$
9. Repeat steps 6 to 8.
10. Replace Q^- with Q after a fixed number of steps
11. end for
12. end for

TABLE I. SYMBOLS AND THEIR DESCRIPTION

Symbol	Description
$Q^*(s, a)$	Q-function given a state and action
Q_θ	Prediction network
Q_{θ^-}	Target network
t	timeslot
S	State space
A	Action space
s_t	State of the environment at timeslot t
a_t	Action taken at time slot t
r_t	Learning agent reward at time slot t

s_{t+1}	State of the environment at timeslot t+1
γ	Discount factor
ϵ	Epsilon

V. SIMULATION RESULTS

A. Simulation Setting

This research work implements an NS-3.41-based IoT simulation for evaluating reactive jamming and reinforcement-learning-based channel hopping. The simulation models mobile IoT nodes, a fixed access point, a reactive jammer, packet delivery ratio, packet loss, per-node energy consumption, and network lifetime. The reinforcement learning component is implemented separately in Python using a DDQN agent that selects channel-hopping actions based on network feedback from the NS-3.41 simulation. The plotting component is also separated into an independent Python script to compare non-hopping, random hopping, DQN-based hopping, and DDQN-based SCHM hopping scenarios. The contribution of this research is the integration of reactive jamming, channel hopping, DDQN decision-making, and performance evaluation into one experimental framework for studying communication reliability and energy efficiency in IoT networks under jamming attacks. RL parameters for simulation are shown in Table II.

TABLE II. RL PARAMETERS AND VALUES

Parameter	Value
Learning rate	1e-3
Gamma	0.98
buffer_size	20000
batch_size	64
target_update	100
Epsilon_start	1.0
Epsilon_min	0.05
Epsilon_decay	0.995

The performance metrics that were used in this research are energy consumption, PDR, network lifetime, and packet loss.

- Energy consumption – The amount of energy that the IoT node uses for transmitting or receiving packets. IoT nodes consume more energy during jamming attacks because of the need for retransmission when there is a lack of acknowledgment.
- PDR – Measures the number of packets that are successfully received at the destination in relation to the total packets sent at the source. Packet loss due to jamming attacks can significantly lead to poor PDR.
- Network lifetime – Refers to the duration of time in the LR-WPAN where all the nodes are active in the network during simulation and still have energy to participate in transmission of packets. This essentially measures the energy consumption of IoT nodes during the simulation.
- Packet loss – Occurs when data packets fail to reach the destination in the network. Interference caused by

reactive jamming attacks can result in severe packet loss in IoT networks.

B. Simulation Scenarios

Simulations are conducted under four scenarios, namely the non-hopping approach, which is a baseline, random hopping approach, DQN-based hopping approach, and DDQN-based SCHM approach, which is the proposed mitigation strategy.

- Non-channel hopping (baseline) where the IoT nodes use a single communication channel during transmission.
- Random channel hopping where the IoT nodes periodically switch the communication channels during transmission.
- DQN-based channel hopping, which uses a policy-based approach to send actions to the IoT nodes for switching to a new communication channel.
- DDQN-based SCHM approach where the learning agent receives network feedback from the network and selects actions and sends the signal to the IoT nodes for switching into the new communication channel.

The results are compared to assess the effectiveness of the proposed approach in improving network performance under reactive jamming conditions.

C. Results

When assessing the performance of four hopping strategies, the simulation period was configured in the NS-3.41 environment such that it runs for 220 seconds. Table III shows the simulation parameters and their values that were set on NS-3.41 environment during the simulation period.

TABLE III. SIMULATION PARAMETERS AND VALUES

Parameter	Value
BurstPackets	600
PacketSize	80 bytes
simulationTime	220 seconds
jamStartTime	20 seconds
jamEndTime	180 seconds
InitialEnergy	5.0 Joules
TransmitPowerW	60 Milliwatts
ReceivePowerW	50 Milliwatts
IdlePowerW	1 Milliwatts
FrequencyBand	2.4 GHz
NumOfChannels	16 (11 – 26)
JammerDetectionMetrics	RSS and PDR

To ensure methodological precision and improve the statistical reliability of the findings, each experimental scenario is repeated 20 times independently using the same simulation parameters. The results from these runs are aggregated by computing the mean PDR, while variation across runs is examined using appropriate dispersion measures such as standard deviation. Results are reported as mean $\pm$ standard

deviation over 20 independent simulation runs, as shown in statistical summary Table IV. The mean PDR is recorded in percentage, mean packet loss is recorded in number of packets lost, mean energy is recorded as total energy consumed, and mean network lifetime is recorded in seconds relative to total simulation time.

TABLE IV. STATISTICAL SUMMARY

Scheme	PDR (%)	Packet Loss (number)	Energy (J)	Network Lifetime (s)
Non-hopping	89 $\pm$ 0.05	212 $\pm$ 1.89	8.5 $\pm$ 0	132 $\pm$ 0.72
Random hopping	93 $\pm$ 1.28	134 $\pm$ 14.97	6.8 $\pm$ 1.39	172 $\pm$ 25.02
DQN	97 $\pm$ 1.25	5.4 $\pm$ 1.47	Mean $\pm$ SD	176 $\pm$ 26.12
DDQN	99 $\pm$ 0.57	9 $\pm$ 4.93	1.5 $\pm$ 0.47	220 $\pm$ 0

1) *PDR*: During the performance evaluation, the proposed SCHM approach, DQN hopping approach, random hopping approach, and non-hopping approach underwent mean PDR of 99%, 97%, 93%, and 89%, respectively, as displayed in Fig. 3. The proposed DDQN-based SCHM approach outperforms the DQN hopping approach by 2%. In addition, the proposed DDQN-based SCHM approach outperforms the random hopping approach by 6% PDR at the end of the simulation. Furthermore, the proposed DDQN-based SCHM approach outperforms the non-hopping approach by 10% PDR at the end of the simulation. The proposed approach significantly improves communication reliability as compared to non-hopping and random hopping approaches.

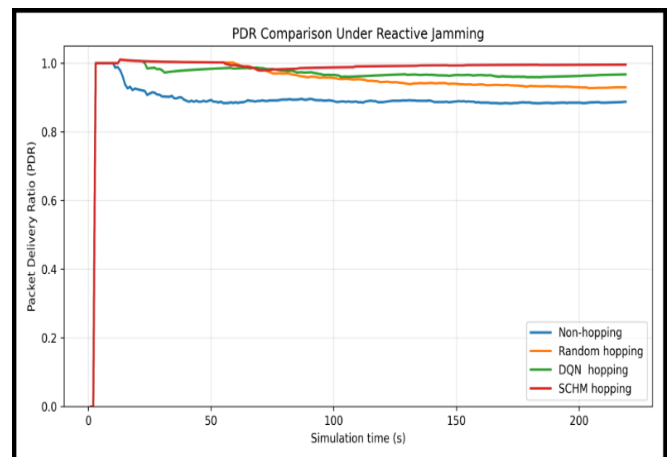


Fig. 3. PDR comparison under reactive jamming attack.

2) *Energy consumption per node*: All nodes, including the AP, are initially allocated the same energy budget of 5 Joules. During the performance evaluation, the average energy consumption per node for the proposed DDQN-based SCHM approach, DQN hopping approach, random hopping approach, and non-hopping approach is 1.5 Joules, 5.4 Joules, 6.8 Joules, and 8.5 Joules, respectively, as displayed in Fig. 4. Compared to the DQN hopping approach, the proposed DDQN-based

SCHM approach reduced energy consumption by 72.2%. Compared to the random hopping approach, the proposed SCHM approach reduced energy consumption by 77.9%. Compared to the non-hopping approach, the proposed DDQN-based SCHM approach reduced energy consumption by 82.3%. The proposed DDQN-based SCHM approach significantly improved the energy consumption of IoT nodes by the end of the simulation period. Random hopping and non-hopping strategies did not utilize energy efficiently, as shown in Fig. 4. All the nodes depleted the entire 5 joules of energy allocated to each of them at the end of the simulation.

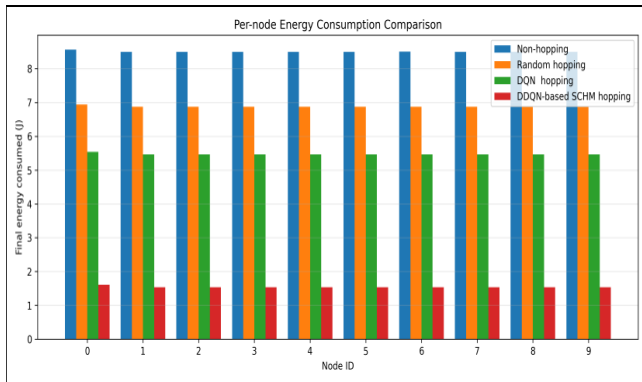


Fig. 4. Energy consumption comparison under reactive jamming attack.

3) *Network lifetime*: During the performance evaluation, the DQN hopping approach, random hopping approach, and non-hopping approach recorded the mean network lifetime of 176 seconds, 172 seconds, and 132 seconds of the simulation time, respectively. The proposed SCHM approach did not record any death of IoT nodes; all IoT nodes are alive by end of simulation period, hence the mean network lifetime recorded is 220 seconds. Compared to the DQN hopping approach, the proposed SCHM approach extended the network lifetime by 20% during the simulation period. Compared to the random hopping approach, the proposed SCHM approach extended the network lifetime by 21.8% during the simulation period. Compared to the non-hopping approach, the proposed SCHM approach extended the network lifetime by 40% during the simulation period. The proposed SCHM approach experienced the longest span in terms of network lifetime, as displayed in Fig. 5. DQN hopping, random hopping, and non-hopping strategies experienced short to medium durations of network lifetime. IoT nodes' energy is depleted before the simulation time expires, as shown in Fig. 5. The proposed SCHM approach extends network lifetime compared to DQN hopping, random hopping, and non-hopping approaches.

4) *Packet loss*: During the performance evaluation, the mean number of packet losses for the proposed DDQN-based SCHM approach, DQN hopping approach, random hopping approach, and non-hopping approach are 9 packets, 38 packets, 137 packets, and 212 packets, respectively, as demonstrated in Fig. 6. Compared to the DQN hopping approach, the proposed DDQN-based SCHM approach reduced packet loss by 76.3% during the simulation period. Compared to the random hopping

approach, the proposed SCHM approach reduced packet loss by 93.4% during the simulation period. Compared to the non-hopping approach, the proposed DDQN-based SCHM approach reduced packet loss by 95.7% during the simulation period. The proposed DDQN-based SCHM approach experienced the lowest number of packet losses during the simulation, as shown in Fig. 6. The DQN hopping strategy experienced the lowest number of packet losses during the simulation. The random hopping strategy experienced a medium number of packet losses during the simulation. The non-hopping strategy experienced the highest number of packet losses, as displayed in Fig. 6.

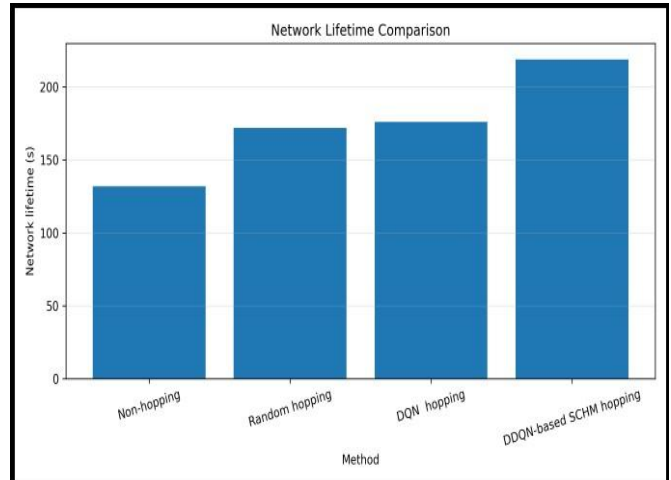


Fig. 5. Network lifetime comparison under reactive jamming attack.

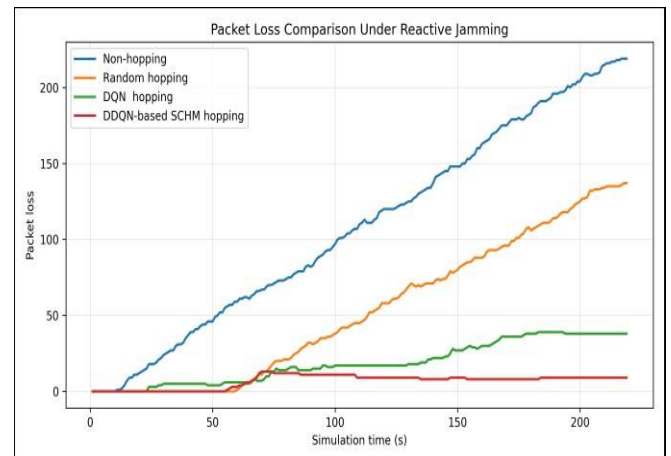


Fig. 6. Packet loss comparison under reactive jamming attack.

D. Discussion

The results demonstrate that the DDQN-based SCHM approach significantly improves communication reliability under reactive jamming conditions. Compared to the baseline scenario without channel hopping, the proposed approach achieves a higher PDR and reduces packet loss. This is shown in Fig. 3 and Fig. 6.

Furthermore, the DDQN-based SCHM agent exhibits improved energy efficiency by avoiding heavily jammed

channels, thereby reducing unnecessary retransmissions. This is shown in Fig. 4. This leads to a slower rate of energy depletion across IoT nodes and extends overall network lifetime. This is shown in Fig. 5.

When compared to random channel hopping, the DDQN-based SCHM approach shows more stable performance and faster adaptation to jammer behavior. These findings are consistent with existing literature, which highlights the advantages of deep reinforcement learning in dynamic and adversarial wireless environments.

This research study compares the proposed DDQN-based SCHM approach with the RL algorithm proposed by [16] and the DDQN algorithm proposed by [18]. The results of the study by [16] show that DRL-based hopping outperforms classical Q-learning and is robust even when jammer behaviour is unknown. The results of the study by [18] show that DDQN fixes overestimation problems, better stability and performance and DDQN performs better and more stable than DQN. The key validation factors are listed in Table V.

TABLE V. VALIDATION FACTORS

Metric	Paper [16]	Paper [18]	This research study
Throughput / PDR	Yes	Yes	Yes
Channel selection policy	RL-based	DDQN	DDQN
Jamming type	Probabilistic	Sweeping	Reactive
Performance improvement	DRL better than Q-learning	Fix overestimation	DDQN-based vs baseline
Limitations	Q-value overestimation	Did not evaluate energy or lifetime	Evaluates PDR, packet loss, energy, network lifetime and convergence

Overall, the proposed DDQN-based SCHM approach demonstrates the effectiveness of intelligent channel selection in enhancing both reliability and energy efficiency in IoT networks under jamming attacks.

VI. CONCLUSION

This research evaluated the performance of the proposed DDQN-based SCHM approach under reactive jamming attacks against DQN-based hopping, non-hopping, and random hopping approaches. The proposed DDQN-based SCHM approach was implemented using NS-3.41 with the IEEE 802.15.4 protocol, simulating an IoT network. The simulation results showed that the DDQN-based SCHM approach achieves better convergence in terms of reward, PDR, and loss. During the performance evaluation the following metrics were used namely: PDR, packet loss, energy consumption, and network lifetime. The simulation results demonstrated that the DDQN-based SCHM approach can improve PDR and reduce packet loss as compared to DQN-based hopping, non-hopping, and random hopping approaches. The simulation results showed that the proposed DDQN-based SCHM has improved communication reliability of the 802.25.4 IoT network. The results are consistent with similar studies that utilized DDQN.

Furthermore, the simulation results showed that the DDQN-based SCHM approach can significantly reduce energy consumption and prolong the network lifetime as compared to DQN-based hopping, non-hopping, and random hopping approaches. The simulation results showed that the proposed DDQN-based SCHM has addressed the shortcomings of the existing literature in similar studies. The existing literature has given little attention to energy consumption and network lifetime of IoT nodes and has focused on detection and communication reliability. As future work, the proposed DDQN-based SCHM approach can include performance improvements in QoS and channel access to reduce collisions. Future extension of the study will also evaluate the performance of the proposed approach in a scenario where multiple reactive jammers exist. In addition, future work can adopt the solution and evaluate it in multichannel jamming, UAV, and satellite navigation systems.

REFERENCES

- [1] Fourati, S. Khssibi, T. Val, H. Idoudi, a. Van Den Bossche, and L.A. Saidane, "Comparative Study of IEEE 802.15. 4 and IEEE 802.15. 6 for WBAN-based", CANet, November 2015
- [2] Iova, F. Theoleyre, T. Watteyne, and T. Noel, "the Love-hate Relationship between IEEE 802.15. 4 and RPL". IEEE Commun. Mag. 55(1), P.188–194, 2016.
- [3] Ahmed, H. Rahman, and M.I. Hussain, "A Comparison of 802.11 Ah and 802.15. 4 for IoT". ICT Express 2(3), P. 100–102, 2016.
- [4] C. Fourati, and S. Said, "Remote Health Monitoring Systems Based on Bluetooth Low Energy (BLE) Communication Systems" ICOST 2020, LNCS 12157, Pp. 41–54, 2020.
- [5] Zahra, Y.S. Bostanci, and M. Soy Turk, "Real-Time Jamming Detection in Wireless IoT Networks", IEEE Access, Volume 11, 2023, P. 70425 – 70442.
- [6] Mukherjee, "Physical-layer Security in the Internet of Things: Sensing and Communication Confidentiality under Resource Constraints," Proc. IEEE, Vol. 103, No. 10, Pp. 1747–1761, Oct. 2015.
- [7] Abdullahi, Y. Baashar, H. Alhussain, a. Alwadain, N. Aziz., L.F. Capretz, and S.J. Abdulkadir, "Detecting Cybersecurity Attacks in Internet of Things Using Artificial Intelligence Methods: a Systematic Literature Review", Electronics, Vol. 11, No. 2, P. 198, 2022.
- [8] Chen, Y. Li, D. Xu, and L. Xiao, "DQN-based Power Control for IoT Transmission against Jamming". IEEE 87th Vehicular Technology Conference (VTC Spring), 2018.
- [9] Chkirbene, R. Hamila, and a. Erbad, "Secure Wireless Sensor Networks for Anti-Jamming Strategy Based on Game Theory". 2023 International Wireless Communication and Mobile Computing (IWCMC), 2023.
- [10] Vadlamani, B. Eksioglu, H. Medal, and a. Nandi, "Jamming Attacks on Wireless Networks: a Taxonomic Survey," International Journal of Production Economics, Vol. 172, Pp. 76–94, 2016.
- [11] Strasser, B. Danev, and S. C. Capkun, "Detection of Reactive Jamming in Sensor Networks," ACM Transactions on Sensor Networks (TOSN), Vol. 7, No. 2, P. 16, 2010.
- [12] Xiao, D. Jiang, D. Xu, H. Zhu, Y. Zhang, and H. V. Poor, "Two-dimensional Antijamming Mobile Communication Based on Reinforcement Learning," IEEE Transactions on Vehicular Technology, Vol. 67, Pp. 9499–9512, Oct. 2018.
- [13] Liu, Y. Xu, L. Jia, Q. Wu, and a. Anpalagan, "Antijamming Communications Using Spectrum Waterfall: a Deep Reinforcement Learning Approach," IEEE Communications Letters, Vol. 22, Pp. 998–1001, May 2018.
- [14] Gao, Z. Qin, X. Jing, Q. Ni, and S. Jin, "Antiintelligent UAV Jamming Strategy via Deep Q-networks," IEEE Transactions on Communications, Vol. 68, Pp. 569–581, Jan. 2020.
- [15] Thamilarasu and R. Sridhar, "Game Theoretic Modeling of Jamming Attacks in Ad Hoc Networks," Proceedings of 18th International

- Conference on Computer Communications and Networks (ICCCN), Pp. 1–6, 2009.
- [16] Qi, H. Zhang, X. Qi, and M. Peng, “Deep Reinforcement Learning Based Hopping Strategy for Wideband Anti-Jamming Wireless Communications”. *IEEE Transactions on Vehicular Technology*, Vol. 73, No. 3, March 2024.
- [17] Thrun and a. Schwartz, “Issues in Using Function Approximation for Reinforcement Learning,” 10 1993.
- [18] K. Nguyen, D.H. N. Nguyen, N.H. Tran, and D. Brunnenmeyer, “Deep Reinforcement Learning Approaches against Jammers with Unequal Sweeping Probability Attacks”. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, Volume 12, Issue 4, Pages 1 -9, 2025.
- [19] Abdolkhani, N.A. Khalek, and W. Hamouda, “Deep Reinforcement Learning for EH-Enabled Cognitive-IoT under Jamming Attacks”. *IEEE Internet of Things Journal*, Vol. 11, No. 24, Pages 40800 – 40813, December 2024.
- [20] Slimeni, F., B. Scheers, Z. Chtourou, and V. Le Nir, “Jamming Mitigation in Cognitive Radio Networks Using a Modified Q-learning Algorithm”, in *Proc. IEEE Int. Conf. Mil. Commun. Inf. Syst.*, 2015, Pp. 1–7.
- [21] Fährmann, D., Jorek, N., Damer, N., Kirchbuchner, F., & Kuijper, F. (2022). Double Deep Q-Learning With Prioritized Experience Replay for Anomaly Detection in Smart Environments. *IEEE Access*. Volume 10.
- [22] Huynh, N., Nguyen, D.N., Hoang, D.T., and Dutkiewicz, E. (2021). Defeating Reactive Jammers with Deep Dueling-based Deception Mechanism. *ICC 2021 - IEEE International Conference on Communications*.