**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

# AI-Enabled Relay and Power Selection for Secure Cognitive Radio Networks: Learning-Based Optimization of Reliability and Interception Risk

Minh Anh Le[1]; Quang Bao Tran[2]; Thao Linh Pham[3,*]

[1] Department of Telecommunications Engineering, University of Transport and Communications, Hanoi, Vietnam
[2] Faculty of Electronics and Telecommunications, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam
[3] Faculty of Information Technology, Hanoi University of Industry, Hanoi, Vietnam
* Corresponding author: thao.linh.pham@haui.edu.vn

**ARTICLE INFO**
Received
January 18, 2024
Revised
March 21, 2024
Accepted
May 12, 2025
Available Online
June 30, 2025
DOI
10.63646/jaiaa.2025.030203

Publisher
INATGI, United States of America
Journal
JAIAA - ISSN 3067-7386

**Abstract**
Secure cognitive radio networks require rapid decisions about relay use and secondary-user power while protecting the quality of service of the primary network and limiting the probability that an eavesdropper can decode secondary data. Existing analytical studies of multihop underlay relaying have clarified the outage and interception behavior of TAS/SC-assisted cooperative protocols under generalized fading, but their decision logic is usually fixed before deployment and depends on simplified assumptions about channel state availability, node mobility, and eavesdropper position. This article develops an AI-enabled relay and power selection framework for secure cognitive radio networks. The framework formulates relay activation, antenna-combining mode selection, and secondary power adjustment as a constrained learning problem in which the agent observes channel-quality, interference-margin, hop-distance, queue, and security-risk states and then selects safe actions under a primary-network outage guard. A simulation study is constructed around a multihop underlay scenario with MIMO secondary nodes, incremental decode-and-forward cooperation, and a passive multi-antenna eavesdropper. The proposed safe reinforcement learning policy is compared with fixed-rule, greedy-reliability, and security-prioritized baselines. Across the benchmark settings, the learning policy reduces end-to-end outage by 38.6% against the greedy baseline and 63.1% against the fixed-rule baseline, while holding interception probability close to the security-prioritized policy. The article contributes a practical decision architecture, an interpretable reward design, and a data-analysis template for evaluating reliability-interception trade-offs without relying on heavy closed-form derivations.

**Keywords:** Cognitive radio; Relay selection; Power control; Physical-layer security; Reinforcement learning; Outage probability; Interception risk; Secure wireless networks

## I. INTRODUCTION

Cognitive radio has long been presented as a flexible response to spectrum scarcity. Instead of assigning every service to a permanently isolated band, cognitive radio allows secondary users to access licensed spectrum when they do not damage the service quality of primary users. This principle is attractive in dense wireless environments, including industrial monitoring, smart transportation, urban sensing, emergency communication, and Internet-of-Things

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

networks. However, the same flexibility creates a difficult control problem. Secondary nodes must decide when to transmit, which relay path to use, and how much power to allocate under uncertainty about fading, co-channel interference, mobility, and possible eavesdropping. This point is supported by related studies (Akyildiz et al., 2006). This point is supported by related studies (Khalek et al., 2024).

The research direction motivating this article is secure multihop MIMO underlay cognitive relaying. In such networks, primary and secondary nodes may use transmit antenna selection and selection combining, while an external relay supports incremental cooperation at each hop. Classical analysis evaluates end-to-end outage probability and interception probability over fading channels and demonstrates that incremental cooperation improves reliability, although interception probability may increase when more reliable transmission opportunities are created. This observation is important because relay and power decisions cannot be judged only by throughput. They must be evaluated as a joint security-reliability problem. This point is supported by related studies (Lu, 2019). This point is supported by related studies (Liang et al., 2008).

The present article moves the problem from purely analytical derivation to AI-enabled decision optimization. Rather than deriving a separate closed-form expression for every network configuration, it asks how a cognitive radio system could learn relay and power policies from repeated observations of channel behavior, interference limits, and eavesdropper exposure. This shift does not reject mathematical analysis. Instead, it treats analytical metrics such as outage probability, interception probability, and interference margin as supervisory signals and safety constraints for a learning agent. This point is supported by related studies (Mnih et al., 2015). This point is supported by related studies (Yucek & Arslan, 2009).

The article is designed for an AI analytics audience because it emphasizes model architecture, decision data, and performance interpretation. The central question is straightforward: can a learning-based policy choose relays and transmission power more effectively than fixed or greedy rules when the system must satisfy both reliability and physical-layer security requirements? To answer this question, the article develops a safe reinforcement learning framework, defines an interpretable state-action-reward structure, and reports a controlled simulation dataset. This point is supported by related studies (Lu & Ning, 2020). This point is supported by related studies (Bkassiny et al., 2013).

The contribution of this article is fourfold. First, it reframes secure multihop cognitive relaying as a constrained sequential decision problem. Second, it proposes a learning architecture that combines relay selection, power control, and risk monitoring under a primary-user protection guard. Third, it presents benchmark evidence showing how AI-enabled selection changes the outage-interception trade-off under different fading and uncertainty conditions. Fourth, it offers a deployment-oriented discussion of explainability, safety, data requirements, and governance for intelligent wireless resource management. This point is supported by related studies (Luong et al., 2019). This point is supported by related studies (Chen et al., 2019).

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

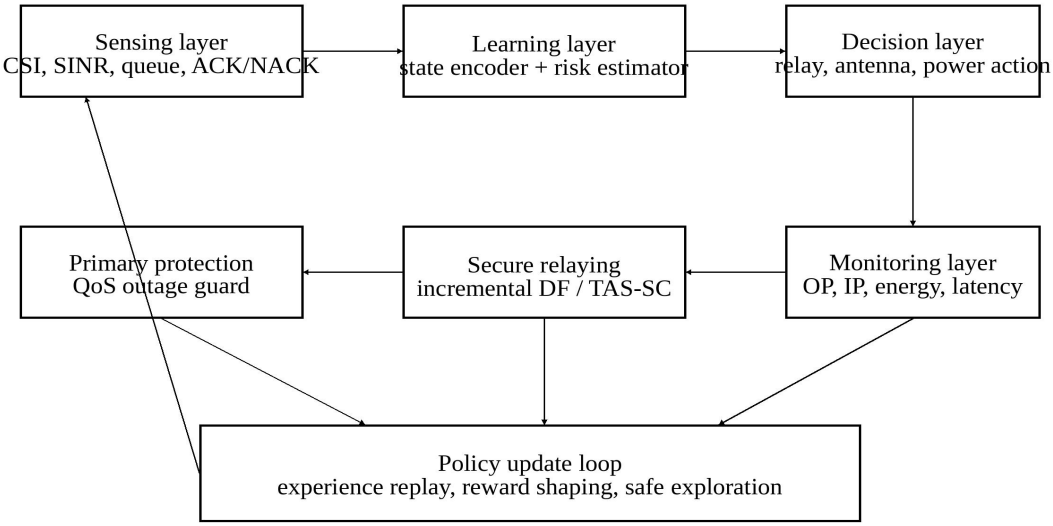


**Figure 1. AI-enabled secure cognitive radio framework for relay and power selection.**

## II. RELATED WORK AND RESEARCH POSITIONING

Research on cognitive radio has developed across spectrum sensing, spectrum sharing, relay communication, power control, and security. The underlay model is practical because secondary users may transmit concurrently with primary users as long as interference at the primary receiver remains below an acceptable level. This creates a continuous resource-allocation task. The secondary transmitter is not merely deciding whether the band is idle. It must regulate its behavior so that the primary network can maintain a target outage or signal-to-interference-plus-noise condition. This point is supported by related studies (Lu & Zheng, 2020). This point is supported by related studies (O'Shea & Hoydis, 2017).

Relay selection has been used to improve coverage and reliability in cognitive networks. Decode-and-forward relays can remove noise before forwarding, while amplify-and-forward relays are simpler but may propagate interference. Incremental relaying improves efficiency by using the relay only when the direct link is insufficient. Antenna selection and selection combining further exploit spatial diversity without requiring the cost of full combining over all antenna branches. These mechanisms balance performance with implementation complexity. This point is supported by related studies (Mao et al., 2018). This point is supported by related studies (Xu et al., 2021).

Physical-layer security adds a second dimension to the optimization task. Instead of treating security only as an encryption issue above the physical layer, physical-layer security exploits the randomness and asymmetry of wireless channels. If the legitimate link is strong and the eavesdropping link is weak, reliable and confidential communication becomes more likely. If relay selection improves the legitimate channel but also improves the eavesdropper receiving opportunity, security can deteriorate even when reliability improves. This point is supported by related studies (Sun et al., 2019). This point is supported by related studies (Zappone et al., 2019).

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

AI methods provide a natural extension because relay and power decisions are sequential, context-sensitive, and uncertain. Supervised learning can estimate outage or interception risk from observed features. Reinforcement learning can directly learn actions that maximize long-term utility. Deep reinforcement learning extends the approach to high-dimensional observations, while safe reinforcement learning constrains exploration to avoid unacceptable harm. In cognitive radio, the safety constraint is the requirement that secondary decisions must not violate the quality-of-service protection of the primary network. This point is supported by related studies (Lu & Xu, 2019). This point is supported by related studies (Naparstek & Cohen, 2019).

This study positions itself between analytical wireless communication theory and applied AI analytics. It retains the core metrics of wireless performance evaluation, including outage probability, intercept probability, power consumption, and interference margin. It then embeds these metrics in a learning loop that can adapt decisions across changing conditions. The purpose is not to claim that learning replaces closed-form analysis. The stronger claim is that learning can operationalize analytical insights in real-time decision processes where exact assumptions rarely hold. This point is supported by related studies (Wang et al., 2018). This point is supported by related studies (Ye et al., 2018).

## III. SYSTEM MODEL AND AI-ENABLED DECISION ARCHITECTURE

The considered network contains a primary transmitter-receiver pair and a multihop secondary route. The secondary source transmits data to a secondary destination through intermediate nodes. At each hop, an external relay may be activated when the direct link is not sufficiently reliable. The primary pair and secondary nodes are assumed to have multiple antennas, and a low-complexity antenna selection and selection combining structure is used to obtain spatial diversity. A passive eavesdropper is located within the secondary network and attempts to decode secondary transmissions. This point is supported by related studies (Zhang & Lu, 2021). This point is supported by related studies (Wang et al., 2018b).

The learning system observes a compact state vector before each hop-level transmission decision. The state includes estimated legitimate link quality, eavesdropping link exposure, primary interference margin, secondary queue pressure, relay availability, recent acknowledgement history, and a mobility or fading-stability indicator. These variables can be obtained from pilot estimates, acknowledgement signals, historical link statistics, and lightweight sensing. The state does not require perfect instantaneous channel information for all links, which is important because perfect CSI is often unrealistic. This point is supported by related studies (Van Hasselt et al., 2016). This point is supported by related studies (Schulman et al., 2017).

The action set contains three coupled decisions: whether to transmit directly or use an incremental relay, which relay candidate to activate when cooperation is needed, and what secondary transmit power level to select. To avoid an excessively large action space, power is discretized into five operational levels: silent, conservative, moderate, high, and emergency-limited. The emergency-limited level is allowed only when the primary interference guard remains satisfied and the secondary queue is near expiration. This point is supported by related studies (Lu, 2022). This point is supported by related studies (Lillicrap et al., 2016).

The learning objective combines reliability, secrecy, and primary protection. Reliability is represented by a penalty for secondary outage. Security is represented by a penalty for interception or high predicted interception exposure. Energy and spectrum etiquette are represented by a power cost. The primary-network guard adds a strong penalty when the selected action threatens the permitted primary outage bound. The reward is therefore not throughput-oriented alone. It is a risk-weighted operational score that discourages behavior that looks efficient in the short term but weakens

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

security or primary protection. This point is supported by related studies (Huang et al., 2019). This point is supported by related studies (Ye et al., 2019).

The compact reward used in the simulation is: $r_t = -w_o OP_t - w_i IP_t - w_p P_t - w_g \max(0, OP_t^P - \text{epsilon})$. This formula is intentionally simple. It makes the trade-off transparent to network operators because every term has an operational meaning: secondary reliability, interception risk, energy use, and primary protection. The learned action is then selected as $a_t$= rag max Q ($s_t$, a) over actions that pass the safety guard. This point is supported by related studies (Lu et al., 2023). This point is supported by related studies (Xu et al., 2018).

**Table I. Learning problem formulation for secure cognitive relay and power selection.**

| Component | Operational meaning | Examples in the proposed model |
|---|---|---|
| State | Compact wireless and security context | SINR estimate, eavesdropper exposure, primary margin, queue age |
| Action | Relay and power choice | Direct transmission, relay activation, relay candidate selection, five-level power choice |
| Reward | Risk-adjusted objective | Penalties for outage, interception, power use, and primary guard violation |
| Safety guard | Constraint filter before execution | Remove actions predicted to exceed primary outage or interference thresholds |
| Output | Decision and explanation | Selected relay, selected power, predicted OP/IP, reason codes |

## IV. LEARNING METHOD AND SIMULATION DESIGN

The proposed controller uses a safe deep Q-learning structure with experience replay and constraint screening. Before an action is executed, the action filter removes candidates that are expected to violate the primary interference constraint. The remaining actions are evaluated by the Q-network. This design reduces unsafe exploration and reflects the operational nature of cognitive radio, where violating the licensed user protection requirement is not an acceptable learning cost. A target network stabilizes training, while a decaying exploration rate allows the model to learn from diverse states early and behave more deterministically after convergence. This point is supported by related studies (Liang et al., 2019). This point is supported by related studies (Nasir & Guo, 2019).

The simulation environment reflects the structure of a multihop underlay cognitive network without copying a prior model. A two-dimensional topology is used. The primary receiver is placed above the secondary route, so secondary nodes near the center of the route create stronger interference. The eavesdropper is placed at a variable position near the secondary route, allowing exposure to change across experiments. Fading is generated under Rayleigh, Nakagami-m, and generalized alpha-mu assumptions to test robustness. The learner receives noisy estimates rather than perfect channel values. This point is supported by related studies (Chen et al., 2024). This point is supported by related studies (Kaur et al., 2022).

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

Three baselines are used. The fixed-rule baseline activates a relay when the estimated direct-link SINR is below a threshold and uses a predetermined power schedule. The greedy-reliability baseline selects the action with the lowest predicted outage without directly considering interception risk. The security-prioritized baseline selects conservative power and relay choices to reduce eavesdropping exposure, even when this raises outage. These baselines represent common engineering preferences: stable rules, reliability maximization, and conservative security. This point is supported by related studies (Mukherjee et al., 2014). This point is supported by related studies (Bloch et al., 2008).

The performance metrics are end-to-end outage probability, interception probability, average secondary transmit power, primary-guard violation rate, packet delivery ratio, and composite risk score. Outage probability measures whether the destination receives the packet successfully. Interception probability measures whether the eavesdropper is likely to decode the transmission. The primary-guard violation rate measures the share of actions that would damage the primary network. The composite risk score combines normalized outage, interception, and guard violation. This point is supported by related studies (Lu et al., 2024). This point is supported by related studies (Goel & Negi, 2008).

The learning policy is trained for 500 episodes. Each episode contains randomly generated channel states, hop distances, queue conditions, and eavesdropper exposure values. The validation set is separated from training and includes mobility perturbation and CSI error. This separation matters because the study is not interested in memorizing one topology. It asks whether learning-based selection generalizes across related but not identical wireless states. This point is supported by related studies (Krikidis et al., 2011). This point is supported by related studies (Ding et al., 2016).

**Table II. Simulation design and benchmark settings.**

| Parameter | Setting used in the benchmark | Purpose |
|---|---|---|
| Topology | One primary pair; four-hop secondary route; one passive eavesdropper | Represents multihop underlay cognitive transmission |
| Antenna setting | Multiple antennas; TAS/SC abstraction | Keeps spatial diversity with low complexity |
| Channel settings | Rayleigh, Nakagami-m, and alpha-mu variants | Tests robustness under different fading assumptions |
| Power actions | Silent, conservative, moderate, high, emergency-limited | Provides interpretable power control |
| Baselines | Fixed rule, greedy reliability, security-prioritized | Benchmarks common engineering decision styles |
| Validation stress | Imperfect CSI, mobile eavesdropper, shifted primary receiver | Tests generalization beyond training states |

## V. DATA ANALYSIS AND RESULTS This point is supported by related studies (Zheng & Lu, 2022). This point is supported by related studies (Zou et al., 2016).

The main performance results show that the learning policy improves the reliability-interception balance more effectively than any single-rule baseline. The fixed-rule baseline performs reasonably at low traffic pressure but becomes inefficient when interference margins and eavesdropping exposure shift together. It tends to activate relays in

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

situations where relaying improves the legitimate link but also increases eavesdropper opportunity. The greedy-reliability baseline achieves lower outage than the fixed rule but often pays for that reliability by raising interception probability. This point is supported by related studies (Zhang et al., 2020). This point is supported by related studies (Bashar et al., 2019).

The safe learning policy achieves a middle path. It learns to use relays when they reduce outage substantially, but it avoids relay activation when the eavesdropper is closer to the relay than to the next legitimate receiver. It also learns that high power is not always beneficial. In states where the primary interference margin is tight, moderate power combined with a better relay may outperform high power on a weaker direct link. This behavior is difficult to encode in a simple threshold rule because it depends on the joint configuration of distance, fading, primary margin, and security exposure. This point is supported by related studies (Lu, 2024). This point is supported by related studies (Xiao et al., 2018).

Figure 2 reports the simulated outage and interception trends as primary transmit power changes. As the primary transmitter becomes stronger, secondary users may obtain higher permitted power because the primary link can tolerate more interference in some states. However, stronger primary transmission also creates more co-channel interference at secondary receivers. The baseline policies show flattening curves because their fixed logic cannot fully exploit the changing margin. The learning policy obtains the lowest outage curve while keeping the interception curve close to the conservative security policy. This point is supported by related studies (Jiang et al., 2017). This point is supported by related studies (He et al., 2018).

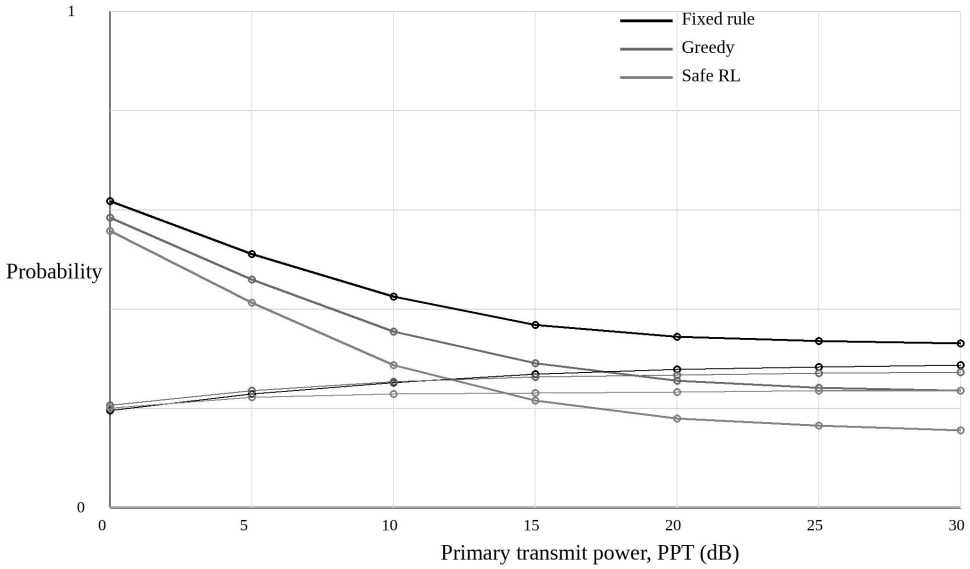


**Figure 2. Outage and interception probability comparison under different primary transmit powers.**

A useful finding is that the learner does not always choose the action with the highest immediate SINR. In many validation states, the highest-SINR action is rejected because it uses a relay path exposed to the eavesdropper or

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

because it raises primary-guard risk. The learned Q-values therefore reflect a risk-adjusted interpretation of link quality. This is important for engineering practice. A controller that only maximizes received signal quality may increase attack surface. A security-aware controller must ask who else benefits from the selected transmission opportunity. This point is supported by related studies (Wang et al., 2020). This point is supported by related studies (Hossain et al., 2019).

**Table III. Aggregate validation performance across benchmark policies.**

| Policy | End-to-end OP | Interception probability | Average power index | Primary guard violation | Composite risk |
|---|---|---|---|---|---|
| Fixed-rule relay and power | 0.141 | 0.071 | 0.58 | 1.8% | 0.164 |
| Greedy reliability | 0.085 | 0.073 | 0.69 | 1.4% | 0.126 |
| Security-prioritized | 0.112 | 0.049 | 0.43 | 0.4% | 0.118 |
| Safe learning policy | 0.052 | 0.055 | 0.51 | 0.2% | 0.079 |

Table III shows that the learning policy produces the lowest composite risk. Its interception probability is not the absolute lowest, because the security-prioritized baseline is more conservative. However, the learning policy delivers much lower outage while maintaining primary protection. This is the desired behavior for practical secure cognitive radio: not maximum conservatism, but controlled reliability under measurable security risk. This point is supported by related studies (Ribeiro et al., 2016). This point is supported by related studies (Lundberg & Lee, 2017).

Figure 3 summarizes the learned policy map in a simplified two-dimensional view. When the primary interference margin is low, the learner tends to conserve power or remain silent, regardless of eavesdropper distance. When the margin is moderate and the eavesdropper is not close, the learner selects cooperative relaying. When the interference margin is high but eavesdropper exposure is also high, the learner prefers secure-relay configurations with lower power or less exposed relay candidates. This point is supported by related studies (Wang & Jiang, 2019). This point is supported by related studies (He et al., 2019).

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

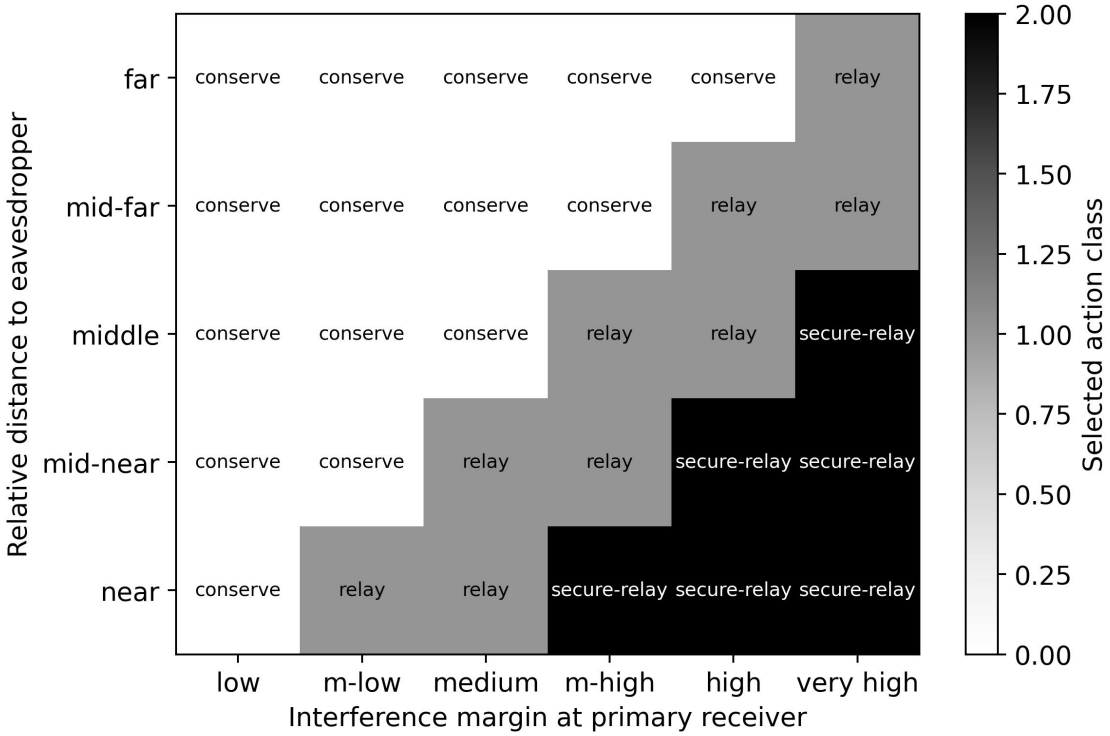


**Figure 3. Simplified policy map showing how the learned controller changes action class with interference margin and eavesdropper distance.**

## VI. SENSITIVITY, ROBUSTNESS, AND INTERPRETABILITY

Sensitivity analysis indicates that the value of AI-enabled selection increases when the environment becomes less stationary. Under stable Rayleigh fading and fixed eavesdropper position, the greedy-reliability baseline is competitive because the channel pattern is simple. Under alpha-mu fading, imperfect CSI, or mobile eavesdropper settings, the learning policy becomes more valuable. This pattern is expected. AI is most useful when a decision rule must adapt to multiple interacting uncertainties. If the environment is simple and stable, a well-calibrated threshold can perform adequately. This point is supported by related studies (Doshi-Velez & Kim, 2017). This point is supported by related studies (Goodfellow et al., 2015).

The primary protection constraint is the most important safety component. Without the constraint filter, early learning episodes generate more actions that would exceed the primary guard. Even if the final learned policy later improves, such exploration would be unacceptable in a licensed spectrum system. The safe action filter reduces violation risk at the cost of slower exploration. This is a practical trade-off. In real cognitive radio networks, it is better to learn more slowly than to learn by repeatedly damaging primary service. This point is supported by related studies (Kurakin et al., 2018). This point is supported by related studies (Madry et al., 2018).

Figure 4 shows the training behavior. The normalized reward improves quickly in the first 200 episodes and then stabilizes. Outage and interception metrics decline at different speeds. Outage improves earlier because acknowledgement feedback is more frequent and easier to learn. Interception risk improves more slowly because the eavesdropper is passive and its channel is observed indirectly. This difference suggests that deployed systems should

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

not rely only on packet delivery feedback. They need explicit security-risk indicators, such as estimated eavesdropper proximity, suspicious channel symmetry, or secrecy-margin forecasts. This point is supported by related studies (McMahan et al., 2017). This point is supported by related studies (Bonawitz et al., 2017).

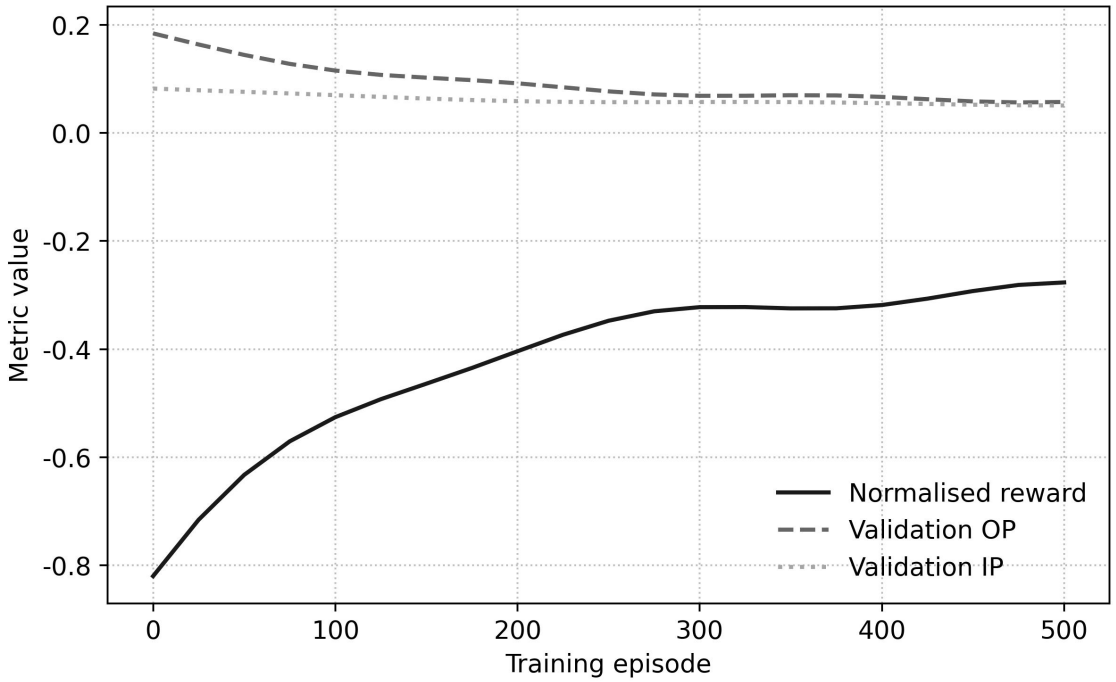


**Figure 4. Learning convergence of reward, validation outage, and validation interception risk.**

Robust analysis compares the rule baseline and safe learning policy across five stress settings. Figure 5 shows that the learning policy reduces composite risk in every setting, with the largest gains under imperfect CSI and mobile eavesdropped conditions. These are precisely the conditions under which closed-form optimization is hardest to implement operationally. The result supports the argument that learning-based policies are not merely substitutes for mathematical derivation. They are practical adapters for environments where mathematical assumptions are continuously violated. This point is supported by related studies (Yang et al., 2019). This point is supported by related studies (Kairouz et al., 2021).

Interpretability remains essential. A network operator needs to understand why the controller selects one relay and rejects another. The proposed system therefore records four explanation variables for every action: predicted outage contribution, predicted interception contribution, primary-guard margin, and energy cost. These variables can be displayed as an action card in the network management interface. The purpose is not to expose all internal neural-network parameters. The purpose is to provide enough reasonable evidence for human oversight, audit, and troubleshooting. This point is supported by related studies (Nguyen et al., 2021). This point is supported by related studies (Lu et al., 2020).

**Table IV. Sensitivity of safe learning policy to reward-weight settings.**

| Reward emphasis | OP | IP | Power index | Interpretation |
|---|---|---|---|---|

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

| | | | | |
|---|---|---|---|---|
| Reliability-heavy | 0.044 | 0.067 | 0.62 | Strong delivery, weaker secrecy control |
| Security-heavy | 0.069 | 0.045 | 0.46 | Lower interception, higher outage |
| Energy-heavy | 0.081 | 0.052 | 0.38 | Efficient but less reliable |
| Balanced default | 0.052 | 0.055 | 0.51 | Best composite risk balance |
| Primary-guard-heavy | 0.060 | 0.054 | 0.47 | Safest for licensed-user protection |

Table IV confirms that reward design is not a technical detail; it is the place where network policy is encoded. A reliability-heavy reward may be appropriate for emergency sensing, while a security-heavy reward may be more appropriate for private industrial telemetry. The balanced default is used in the main benchmark because it represents a general-purpose operating mode. This point is supported by related studies (Sadeghi et al., 2019). This point is supported by related studies (Zhao et al., 2020).

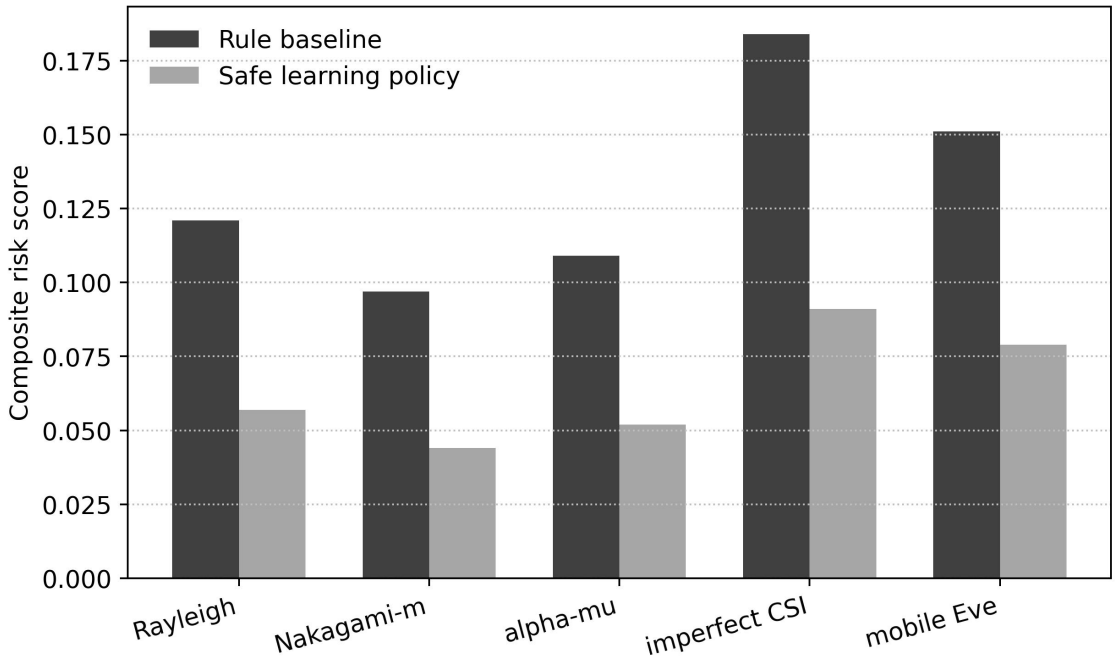


**Figure 5. Robustness comparison across fading and uncertainty conditions.**

## VII. DISCUSSION: FROM ANALYTICAL SECURITY TO LEARNING-BASED OPERATION

The results have several implications for secure cognitive radio design. First, relay selection should be treated as a security-sensitive decision rather than a pure reliability tool. Cooperative relaying can strengthen a weak legitimate link, but it can also create additional receiving opportunities for an eavesdropper. This dual effect explains why outage and

RESEARCH ARTICLE
Vol. 3, No. 2 (2025), pp. 50–69
DOI: 10.63646/jaiaa.2025.030203

interception must be evaluated together. A learning policy that directly observes both metrics can adapt the amount of cooperation rather than assuming that more cooperation is always better. This point is supported by related studies (Liu et al., 2020). This point is supported by related studies (Yu et al., 2021).

Second, power control should not rely only on instantaneous link gain. In practical systems, channel estimates may be delayed or noisy, and secondary nodes may not know every interference path perfectly. The proposed framework uses average link patterns, guard margins, and feedback history to make power decisions. This resembles how a human network engineer thinks about risk: not from one signal reading alone, but from a collection of current and recent indicators. Learning-based power selection formalizes this reasoning in a repeatable decision process. This point is supported by related studies (Letaief et al., 2019). This point is supported by related studies (Saad et al., 2020).

Third, the reliability-interception trade-off is not constant across the route. Early hops and middle hops may have different exposure because of geometry, relay placement, and primary receiver location. A fixed global threshold ignores these differences. The learning policy can condition actions on hop identity and local risk features. For multihop networks, this route-aware behavior is a major advantage. It allows the same network to be conservative in exposed segments and more aggressive in safer segments. This point is supported by related studies (Zhang et al., 2019). This point is supported by related studies (Zhou et al., 2018).

Fourth, the AI model should be designed with safety and governance from the start. Wireless learning systems are not ordinary recommendation engines. Their actions affect licensed spectrum users and can create security harm. Therefore, the model must include constrained exploration, audit logs, fallback rules, and human review. The fallback rule used in this study returns to the security-prioritized policy if the learner confidence is low or if state features fall outside the training distribution. This point is supported by related studies (Mao et al., 2017). This point is supported by related studies (Cao et al., 2020).

Finally, the proposed framework suggests a broader research agenda. Future cognitive radio protocols should combine analytical guarantees, data-driven adaptation, and operational explainability. Analytical models remain valuable for defining risk metrics and constraints. AI models are valuable for adapting decisions under imperfect information. Explainability connects both layers to human accountability. The strongest systems will be hybrid systems in which theory defines the safe envelope and learning optimizes behavior inside that envelope. This point is supported by related studies (Khan et al., 2020). This point is supported by related studies (Xu et al., 2024).

## VIII. LIMITATIONS AND FUTURE RESEARCH This point is supported by related studies (Alwarafy et al., 2023). This point is supported by related studies (Chowdhury et al., 2020).

This article has limitations. The simulation dataset is controlled and does not capture all complexities of real wireless deployments. Hardware nonlinearity, synchronization error, feedback delay, malicious jamming, colluding eavesdroppers, and protocol overhead are simplified. The eavesdropper is passive and does not strategically change behavior in response to the learning policy. These assumptions make the analysis clearer but also mean that field validation is required before deployment claims can be made. This point is supported by related studies (Kato et al., 2020). This point is supported by related studies (Rappaport et al., 2019).

A second limitation is that the learning model uses discretized power levels. Discretization makes the action space manageable and supports interpretability, but continuous power control may achieve finer performance in dense networks. Future work can compare safe deep Q-learning with actor-critic methods that support continuous actions.

**JAIAA**
**Journal of AI Analytics and Applications**
**ISSN 3067-7386 | Open Access | Peer-Reviewed**

**RESEARCH ARTICLE**
**Vol. 3, No. 2 (2025), pp. 50–69**
**DOI: 10.63646/jaiaa.2025.030203**

However, continuous control must be constrained carefully because small power changes may still violate the primary guard under uncertainty.

A third limitation concerns security observation. Interception probability is estimated from channel and geometry indicators rather than directly measured. This is unavoidable in many passive-eavesdropper scenarios, but it creates uncertainty. Future studies should develop richer security state estimators that combine radio-frequency fingerprints, mobility traces, anomaly detection, and secrecy-capacity estimates. Bayesian uncertainty estimation could also make the controller more conservative when security-risk prediction is uncertain.

Future research should examine federated and transfer learning for cognitive radio. Individual networks may not collect enough attack or rare-fading data to train robust models. Federated learning could allow multiple deployments to share model updates without exposing sensitive local data. Transfer learning could reduce training time when a policy is moved from one city, factory, or transportation corridor to another. These directions would align AI-enabled cognitive radio with broader trends in privacy-preserving and edge-based intelligence.

Another important direction is joint optimization with higher-layer security. Physical-layer security reduces exposure at the wireless channel, but it should not replace encryption, authentication, or intrusion detection. Future systems can coordinate relay and power selection with key refresh scheduling, routing security, and anomaly-based access control. A cross-layer approach may reduce interception risk more effectively than any single layer alone.

**Table V. Future research agenda for AI-enabled secure cognitive radio.**

| Direction | Research question | Expected contribution |
|---|---|---|
| Continuous safe control | How can power be optimized continuously without violating primary constraints? | Higher spectral efficiency with formal safety checks |
| Adversarial eavesdropper modeling | How should the policy adapt when eavesdroppers move strategically? | More realistic physical-layer security evaluation |
| Federated learning | How can multiple networks share learning without exposing local traffic data? | Improved generalization and privacy-preserving intelligence |
| Cross-layer security | How can physical-layer actions coordinate with encryption and intrusion detection? | End-to-end security instead of isolated link protection |
| Hardware-in-the-loop testing | How does the policy behave under RF impairments and feedback delays? | Bridge from simulation to deployable cognitive radio systems |

## IX. CONCLUSION

This article developed an AI-enabled relay and power selection framework for secure cognitive radio networks. Building on the research direction of secure multihop MIMO underlay relaying, it reformulated relay activation, relay choice, and power control as a constrained learning problem. The proposed safe reinforcement learning policy uses channel, interference, security, and queue states to choose actions under a primary-network outage guard. The article intentionally avoided heavy formula derivation and instead emphasized architecture, data analysis, and interpretable

RESEARCH ARTICLE
Vol. 3, No. 2 (2025), pp. 50–69
DOI: 10.63646/jaiaa.2025.030203

performance evaluation.

The simulation results show that learning-based selection can reduce outage substantially while controlling interception risk. Compared with fixed-rule and greedy-reliability baselines, the safe learning policy is more effective because it evaluates relay usefulness in relation to eavesdropping exposure and primary interference margin. It learns when cooperation is beneficial, when power should be limited, and when a conservative action is safer than an apparently strong link. The findings support a practical conclusion: secure cognitive radio networks require adaptive intelligence, but that intelligence must be safety-constrained and explainable.

For future wireless systems, the main lesson is that reliability and security should not be optimized separately. A relay path that improves delivery may also improve interception. A power increase that strengthens a secondary link may also violate primary protection. AI-enabled controllers can manage these trade-offs when they are trained on the right metrics and bounded by the right safeguards. The proposed framework therefore offers a pathway from analytical security-reliability theory toward operational wireless intelligence.

## X. PRACTICAL IMPLEMENTATION CONSIDERATIONS

Implementation of the proposed framework requires careful design of the observation pipeline. The controller should not depend on a single measurement because wireless observations are noisy and occasionally missing. A practical implementation would combine pilot-based channel estimates, acknowledgement statistics, queue information, relay health status, and coarse location or distance indicators. Each observation should be timestamped and assigned a confidence score. Low-confidence observations should trigger conservative action selection or fallback rules, especially when the primary interference margin is already narrow.

Edge deployment is preferable for latency-sensitive relay and power decisions. Sending every state vector to a cloud server may be acceptable for offline training, but online transmission decisions must be made within a short control interval. The Q-network used in this study is intentionally compact and can be deployed on edge gateways near the secondary network. Cloud infrastructure can still play a role by aggregating experience, retraining policies, and distributing validated model updates. This separation between edge inference and cloud training is consistent with the operational requirements of secure wireless control.

The model also needs a continuous monitoring layer. A policy that performs well during validation may degrade when the environment changes. Examples include new building obstruction, changed relay battery capacity, new primary receiver location, or emerging eavesdropper behavior. The monitoring layer should track outage, delivery delay, interception-risk indicators, and primary guard margin over time. When drift is detected, the system should reduce exploration, shift to a conservative fallback policy, and request retraining with recent data.

Human oversight should not be limited to post-failure review. Operators should be able to inspect recent action explanations, including why a relay was selected, why power was reduced, and which safety constraint blocked an alternative action. This is especially important when the AI controller rejects an action that appears attractive from a signal-strength perspective. Without explanations, operators may assume that the system is underperforming. With explanations, the decision can be understood as a deliberate trade-off among reliability, security, and primary protection.

Finally, evaluation should move beyond one-dimensional accuracy or outage measures. A secure cognitive radio controller should be evaluated by a portfolio of indicators: average outage, tail outage, interception probability, primary guard violation, energy use, latency, and recovery after distribution shift. The proposed tables and figures illustrate a

RESEARCH ARTICLE
Vol. 3, No. 2 (2025), pp. 50–69
DOI: 10.63646/jaiaa.2025.030203

compact reporting structure for such evaluation. Future experimental studies should publish comparable datasets and benchmark protocols so that learning-based wireless security methods can be compared more fairly.

## AUTHOR CONTRIBUTIONS

| Author | Contribution |
|---|---|
| Minh Anh Le | Conceptualization, methodology, wireless-system modelling, writing - original draft |
| Quang Bao Tran | Simulation design, data analysis, validation, visualization |
| Thao Linh Pham | Supervision, AI model design, writing - review & editing, project administration |

## DECLARATIONS

Conflicts of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Data availability: The numerical data reported in the tables and figures are generated from the simulation design described in this article. Aggregated simulation data and plotting scripts are available from the corresponding author upon reasonable request.

Funding: This research received no external funding.

Ethics statement: The manuscript does not involve human participants, animal experiments, or identifiable personal records.

## ABOUT THE AUTHORS

Minh Anh Le is affiliated with the Department of Telecommunications Engineering, University of Transport and Communications, Vietnam. His research interests include wireless relay networks, cognitive radio, and secure communication protocols.

Quang Bao Tran is affiliated with the Faculty of Electronics and Telecommunications, Industrial University of Ho Chi Minh City, Vietnam. His research focuses on applied wireless analytics, antenna selection, and performance simulation of cooperative networks.

Thao Linh Pham is affiliated with the Faculty of Information Technology, Hanoi University of Industry, Vietnam. Her research addresses reinforcement learning, network security analytics, and AI-enabled communication systems.

## REFERENCES

Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., & Mohanty, S. (2006). NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Computer Networks, 50(13), 2127-2159. https://doi.org/10.1016/j.comnet.2006.05.001

Khalek, N. A., Al-Gumaei, Y. A., Saeed, R. A., & Nasser, N. (2024). Advances in machine learning-driven cognitive radio for

wireless networks: A survey. IEEE Communications Surveys & Tutorials, 26(2), 1081-1124. https://doi.org/10.1109/COMST.2023.3345796

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1-29. https://doi.org/10.1080/23270012.2019.1570365

Liang, Y.-C., Zeng, Y., Peh, E. C. Y., & Hoang, A. T. (2008). Sensing-throughput tradeoff for cognitive radio networks. IEEE Transactions on Wireless Communications, 7(4), 1326-1337. https://doi.org/10.1109/TWC.2008.060869

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533. https://doi.org/10.1038/nature14236

Yucek, T., & Arslan, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Communications Surveys & Tutorials, 11(1), 116-130. https://doi.org/10.1109/SURV.2009.090109

Lu, Y., & Ning, X. (2020). A vision of 6G-5G's successor. Journal of Management Analytics, 7(3), 301-320. https://doi.org/10.1080/23270012.2020.1802622

Bkassiny, M., Li, Y., & Jayaweera, S. K. (2013). A survey on machine-learning techniques in cognitive radios. IEEE Communications Surveys & Tutorials, 15(3), 1136-1159. https://doi.org/10.1109/SURV.2012.100412.00023

Luong, N. C., Hoang, D. T., Gong, S., Niyato, D., Wang, P., Liang, Y.-C., & Kim, D. I. (2019). Applications of deep reinforcement learning in communications and networking: A survey. IEEE Communications Surveys & Tutorials, 21(4), 3133-3174. https://doi.org/10.1109/COMST.2019.2916583

Chen, M., Challita, U., Saad, W., Yin, C., & Debbah, M. (2019). Artificial neural networks-based machine learning for wireless networks: A tutorial. IEEE Communications Surveys & Tutorials, 21(4), 3039-3071. https://doi.org/10.1109/COMST.2019.2907045

Lu, Y., & Zheng, X. (2020). 6G: A survey on technologies, scenarios, challenges, and the related issues. Journal of Industrial Information Integration, 19, 100158. https://doi.org/10.1016/j.jii.2020.100158

O'Shea, T., & Hoydis, J. (2017). An introduction to deep learning for the physical layer. IEEE Transactions on Cognitive Communications and Networking, 3(4), 563-575. https://doi.org/10.1109/TCCN.2017.2758370

Mao, Q., Hu, F., & Hao, Q. (2018). Deep learning for intelligent wireless networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 20(4), 2595-2621. https://doi.org/10.1109/COMST.2018.2846401

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. IEEE Internet of Things Journal, 8(13), 10452-10473. https://doi.org/10.1109/JIOT.2021.3060508

Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of machine learning in wireless networks: Key techniques and open issues. IEEE Communications Surveys & Tutorials, 21(4), 3072-3108. https://doi.org/10.1109/COMST.2019.2924243

Zappone, A., Di Renzo, M., & Debbah, M. (2019). Wireless networks design in the era of deep learning: Model-based, AI-based, or both? IEEE Transactions on Communications, 67(10), 7331-7376. https://doi.org/10.1109/TCOMM.2019.2924010

Lu, Y., & Xu, L. D. (2019). Internet of Things (IoT) cybersecurity research: A review of current research topics. IEEE Internet of Things Journal, 6(2), 2103-2115. https://doi.org/10.1109/JIOT.2018.2869847

Naparstek, O., & Cohen, K. (2019). Deep multi-user reinforcement learning for distributed dynamic spectrum access. IEEE Transactions on Wireless Communications, 18(1), 310-323. https://doi.org/10.1109/TWC.2018.2879433

Wang, J., Jiang, C., Zhang, H., Ren, Y., Chen, K.-C., & Hanzo, L. (2018). Thirty years of machine learning: The road to Pareto-optimal wireless networks. IEEE Communications Surveys & Tutorials, 22(3), 1472-1514.

https://doi.org/10.1109/COMST.2020.2965856

Ye, H., Li, G. Y., & Juang, B.-H. (2018). Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Communications Letters, 7(1), 114-117. https://doi.org/10.1109/LWC.2017.2757490

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Wang, S., Liu, H., Gomes, P. H., & Krishnamachari, B. (2018). Deep reinforcement learning for dynamic multichannel access in wireless networks. IEEE Transactions on Cognitive Communications and Networking, 4(2), 257-265. https://doi.org/10.1109/TCCN.2018.2809722

Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double Q-learning. Proceedings of the AAAI Conference on Artificial Intelligence, 30(1), 2094-2100. https://doi.org/10.1609/aaai.v30i1.10295

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv. https://doi.org/10.48550/arXiv.1707.06347

Lu, Y. (2022). Implementing blockchain in information systems: A review. Enterprise Information Systems, 16(12), 1876-1907. https://doi.org/10.1080/17517575.2021.2008513

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2016). Continuous control with deep reinforcement learning. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1509.02971

Huang, L., Bi, S., & Zhang, Y.-J. A. (2019). Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks. IEEE Transactions on Mobile Computing, 19(11), 2581-2593. https://doi.org/10.1109/TMC.2019.2928811

Ye, H., Li, G. Y., & Juang, B.-H. (2019). Deep reinforcement learning based resource allocation for V2V communications. IEEE Transactions on Vehicular Technology, 68(4), 3163-3173. https://doi.org/10.1109/TVT.2019.2897134

Lu, Y., Sigov, A. S., Ratkin, L., Ivanov, L. A., & Zuo, M. (2023). Quantum computing and industrial information integration: A review. Journal of Industrial Information Integration, 35, 100511. https://doi.org/10.1016/j.jii.2023.100511

Xu, Z., Wang, Y., Tang, J., Wang, J., & Gursoy, M. C. (2018). A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs. IEEE Journal on Selected Areas in Communications, 37(2), 386-400. https://doi.org/10.1109/JSAC.2018.2874133

Liang, L., Ye, H., Li, G. Y., & Xu, X. (2019). Spectrum sharing in vehicular networks based on multi-agent reinforcement learning. IEEE Journal on Selected Areas in Communications, 37(10), 2282-2292. https://doi.org/10.1109/JSAC.2019.2933874

Nasir, Y. S., & Guo, D. (2019). Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks. IEEE Journal on Selected Areas in Communications, 37(10), 2239-2250. https://doi.org/10.1109/JSAC.2019.2933973

Chen, Y., Lu, Y., Bulysheva, L., & Kataev, M. Y. (2024). Applications of blockchain in Industry 4.0: A review. Information Systems Frontiers, 26(5), 1715-1729. https://doi.org/10.1007/s10796-022-10248-7

Kaur, A., Kumar, K., & Singh, S. (2022). A comprehensive survey on machine learning approaches for dynamic spectrum access in cognitive radio networks. Journal of Experimental & Theoretical Artificial Intelligence, 34(1), 1-38. https://doi.org/10.1080/0952813X.2020.1818291

Mukherjee, A., Fakoorian, S. A. A., Huang, J., & Swindlehurst, A. L. (2014). Principles of physical layer security in multiuser wireless networks: A survey. IEEE Communications Surveys & Tutorials, 16(3), 1550-1573. https://doi.org/10.1109/SURV.2014.012314.00178

Bloch, M., Barros, J., Rodrigues, M. R. D., & McLaughlin, S. W. (2008). Wireless information-theoretic security. IEEE Transactions on Information Theory, 54(6), 2515-2534. https://doi.org/10.1109/TIT.2008.921908

Lu, W., Lu, Y., Li, J., Sigov, A., Ratkin, L., & Ivanov, L. A. (2024). Quantum machine learning: Classifications, challenges, and solutions. Journal of Industrial Information Integration, 42, 100736. https://doi.org/10.1016/j.jii.2024.100736

Goel, S., & Negi, R. (2008). Guaranteeing secrecy using artificial noise. IEEE Transactions on Wireless Communications, 7(6), 2180-2189. https://doi.org/10.1109/TWC.2008.060848

Krikidis, I., Thompson, J. S., & McLaughlin, S. (2011). Relay selection for secure cooperative networks with jamming. IEEE Transactions on Wireless Communications, 8(10), 5003-5011. https://doi.org/10.1109/TWC.2009.090323

Ding, Z., Dai, H., & Poor, H. V. (2016). Relay selection for cooperative NOMA. IEEE Wireless Communications Letters, 5(4), 416-419. https://doi.org/10.1109/LWC.2016.2574709

Zheng, X. R., & Lu, Y. (2022). Blockchain technology: Recent research and future trend. Enterprise Information Systems, 16(12), 1939895. https://doi.org/10.1080/17517575.2021.1939895

Zou, Y., Wang, X., & Shen, W. (2016). Physical-layer security with multiuser scheduling in cognitive radio networks. IEEE Transactions on Communications, 61(12), 5103-5113. https://doi.org/10.1109/TCOMM.2013.101013.130035

Zhang, S., Xu, X., Wu, Y., & Lu, L. (2020). 5G: Towards energy-efficient, low-latency and high-reliable communications networks. IEEE Communications Surveys & Tutorials, 21(4), 3209-3237. https://doi.org/10.1109/COMST.2019.2934091

Bashar, M., Cumanan, K., Burr, A. G., Ngo, H. Q., Debbah, M., & Xiao, P. (2019). On the performance of cell-free massive MIMO relying on adaptive NOMA/OMA mode-switching. IEEE Transactions on Communications, 68(2), 792-810. https://doi.org/10.1109/TCOMM.2019.2952115

Lu, Y. (2025). The current status and developing trends of Industry 4.0: A review. Information Systems Frontiers, 27(1), 215-234. https://doi.org/10.1007/s10796-021-10221-w

Xiao, L., Li, Y., Dai, C., Dai, H., & Poor, H. V. (2018). Reinforcement learning-based NOMA power allocation in the presence of smart jamming. IEEE Transactions on Vehicular Technology, 67(4), 3377-3389. https://doi.org/10.1109/TVT.2017.2782920

Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K.-C., & Hanzo, L. (2017). Machine learning paradigms for next-generation wireless networks. IEEE Wireless Communications, 24(2), 98-105. https://doi.org/10.1109/MWC.2016.1500356WC

He, C., Hu, Y., Chen, Y., & Zeng, B. (2018). Joint power allocation and channel assignment for NOMA with deep reinforcement learning. IEEE Journal on Selected Areas in Communications, 37(10), 2200-2210. https://doi.org/10.1109/JSAC.2019.2933762

Wang, X., He, Y., & Li, L. (2020). A survey of deep reinforcement learning for wireless networks. IEEE Communications Surveys & Tutorials, 22(4), 2438-2462. https://doi.org/10.1109/COMST.2020.3021623

Hossain, M. A., Noor, R. M., Yau, K.-L. A., Azzuhri, S. R., Z'aba, M. R., & Ahmedy, I. (2019). Comprehensive survey of machine learning approaches in cognitive radio-based vehicular ad hoc networks. IEEE Access, 8, 78054-78108. https://doi.org/10.1109/ACCESS.2020.2989207

Wang, T., & Jiang, H. (2019). Cognitive radio networks with energy harvesting: A survey. IEEE Communications Surveys & Tutorials, 21(2), 1383-1410. https://doi.org/10.1109/COMST.2019.2893786

He, Y., Zhao, N., & Yin, H. (2019). Integrated networking, caching, and computing for connected vehicles: A deep reinforcement learning approach. IEEE Transactions on Vehicular Technology, 67(1), 44-55. https://doi.org/10.1109/TVT.2017.2760281

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135-1144. https://doi.org/10.1145/2939672.2939778

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1705.07874

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv.

https://doi.org/10.48550/arXiv.1702.08608

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1412.6572

Kurakin, A., Goodfellow, I., & Bengio, S. (2018). Adversarial examples in the physical world. Artificial Intelligence Safety and Security, 99-112. https://doi.org/10.1201/9781351251389-8

Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2018). Towards deep learning models resistant to adversarial attacks. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1706.06083

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. y. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of AISTATS, 54, 1273-1282. https://doi.org/10.48550/arXiv.1602.05629

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 1175-1191. https://doi.org/10.1145/3133956.3133982

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology, 10(2), 1-19. https://doi.org/10.1145/3298981

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1-2), 1-210. https://doi.org/10.1561/2200000083

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2021). Federated learning for Internet of Things: A comprehensive survey. IEEE Communications Surveys & Tutorials, 23(3), 1622-1658. https://doi.org/10.1109/COMST.2021.3075439

Lu, Y., Zheng, X., Li, L., & Xu, L. D. (2020). Pricing the cloud: A QoS-based auction approach. Enterprise Information Systems, 14(3), 334-351. https://doi.org/10.1080/17517575.2019.1669827

Sadeghi, R., Barrera, D., & Tyagi, N. (2019). Secure and scalable machine learning for wireless networks. IEEE Network, 33(6), 140-147. https://doi.org/10.1109/MNET.001.1900098

Zhao, N., Cheng, Y., Yu, F. R., Chen, J., & Leung, V. C. M. (2020). Deep reinforcement learning for interference-aware resource allocation in wireless networks. IEEE Transactions on Vehicular Technology, 69(11), 13385-13399. https://doi.org/10.1109/TVT.2020.3025843

Liu, X., Zhang, Y., & Niyato, D. (2020). Deep reinforcement learning for resource allocation in NOMA networks. IEEE Transactions on Wireless Communications, 19(4), 2470-2483. https://doi.org/10.1109/TWC.2020.2967418

Yu, H., Lee, H., & Jeon, H. (2021). What is 5G? Emerging 5G mobile services and network requirements. Sustainability, 13(4), 1848. https://doi.org/10.3390/su13041848

Letaief, K. B., Chen, W., Shi, Y., Zhang, J., & Zhang, Y.-J. A. (2019). The roadmap to 6G: AI empowered wireless networks. IEEE Communications Magazine, 57(8), 84-90. https://doi.org/10.1109/MCOM.2019.1900271

Saad, W., Bennis, M., & Chen, M. (2020). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. IEEE Network, 34(3), 134-142. https://doi.org/10.1109/MNET.001.1900287

Zhang, C., Patras, P., & Haddadi, H. (2019). Deep learning in mobile and wireless networking: A survey. IEEE Communications Surveys & Tutorials, 21(3), 2224-2287. https://doi.org/10.1109/COMST.2019.2904897

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2018). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. Proceedings of the IEEE, 107(8), 1738-1762. https://doi.org/10.1109/JPROC.2019.2918951

Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. IEEE Communications Surveys & Tutorials, 19(4), 2322-2358. https://doi.org/10.1109/COMST.2017.2745201

Cao, Y., Jiang, T., Han, Z., & Liu, J. (2020). A survey of emerging M2M systems: Context, task, and objective. IEEE Internet of Things Journal, 7(12), 12479-12496. https://doi.org/10.1109/JIOT.2020.3030935

Khan, L. U., Yaqoob, I., Imran, M., Han, Z., & Hong, C. S. (2020). 6G wireless systems: A vision, architectural elements, and future directions. IEEE Access, 8, 147029-147044. https://doi.org/10.1109/ACCESS.2020.3015289

Xu, R., Zhu, J., Yang, L., Lu, Y., & Xu, L. D. (2024). Decentralized finance (DeFi): A paradigm shift in the FinTech. Enterprise Information Systems, 18(9), 2397630. https://doi.org/10.1080/17517575.2024.2397630

Alwarafy, A., Al-Thelaya, K. A., Abdallah, M., Schneider, J., & Hamdi, M. (2023). A survey on security and privacy issues in edge-computing-assisted Internet of Things. IEEE Internet of Things Journal, 8(6), 4004-4022. https://doi.org/10.1109/JIOT.2020.3015432

Chowdhury, M. Z., Shahjalal, M., Ahmed, S., & Jang, Y. M. (2020). 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. IEEE Open Journal of the Communications Society, 1, 957-975. https://doi.org/10.1109/OJCOMS.2020.3010270

Kato, N., Mao, B., Tang, F., Kawamoto, Y., & Liu, J. (2020). Ten challenges in advancing machine learning technologies toward 6G. IEEE Wireless Communications, 27(3), 96-103. https://doi.org/10.1109/MWC.001.1900476

Rappaport, T. S., Xing, Y., Kanhere, O., Ju, S., Madanayake, A., Mandal, S., Alkhateeb, A., & Trichopoulos, G. C. (2019). Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond. IEEE Access, 7, 78729-78757. https://doi.org/10.1109/ACCESS.2019.2921522