AI-Based Security Mechanisms for Intelligent Routing in 6G Networks: A Systematic Literature Review

Mohd Hilal Muhammad[1]*, Muhammad Khairul Zharif Nor A'zam[2], Mohd Zhafri Mohd Zukhi[3]

[1,2] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Kedah, 08400 Merbok, Kedah, Malaysia

[3] School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

*** Corresponding author**

DOI: https://doi.org/10.47772/IJRISS.2026.101000163

Received: 13 October 2026; Accepted: 18 October 2026; Published: 30 October 2026

ABSTRACT

Sixth-generation (6G) networks are expected to integrate artificial intelligence (AI) into distributed and highly adaptive communication infrastructures. This convergence creates opportunities for intelligent routing, but also exposes routing decisions, learning models and trust relationships to complex security threats. This review critically examines how AI-based techniques interact with intrusion detection, adaptive route selection, decentralised trust, zero trust verification and physical layer protection across prospective 6G environments. The synthesis distinguishes security detection from secure routing control and relates both to device-to-device communication, edge networks, vehicular systems and AI-enabled radio access networks. The reviewed literature suggests that machine learning and deep learning primarily support anomaly recognition, whereas reinforcement learning addresses route adaptation; blockchain and zero-trust approaches contribute complementary integrity and policy functions.

However, evidence remains difficult to compare because studies investigate different network settings, threats, datasets and outcomes, with inconsistent reporting of latency, computation, communication cost and adversarial resilience. The appraisal distinguishes direct routing studies from enabling security-architecture research and avoids treating heterogeneous testbeds as directly comparable. It proposes a common reporting matrix covering detection quality, route reliability, latency, computational burden, signalling overhead, scalability and resilience under adversarial conditions. These dimensions clarify which claims require controlled benchmarking and which remain architectural propositions. The proposed synthesis therefore serves as both a conceptual account of the literature and a transparent evaluation agenda for future experiments. The study contributes an extended cross-layer conceptual framework, an evidence-linked comparison of representative approaches and a practical evaluation agenda for resource-constrained 6G settings. To substantiate its claim as a fully reproducible systematic review, the final submission must additionally disclose the executed database queries, search dates, screening decisions and verified PRISMA counts.

Keywords: 6G networks, intelligent routing, artificial intelligence, routing security, systematic literature review

INTRODUCTION

The transition from fifth-generation (5G) to sixth-generation (6G) mobile systems is expected to make network management more autonomous, distributed and context-aware. Proposed 6G scenarios combine highly mobile terminals, device-to-device (D2D) links, edge intelligence, open radio access networks (Open RAN) and space–air–ground communication. Consequently, routing must increasingly optimise paths under

changing conditions while managing trust, service continuity and the confidentiality of network information (Saad et al., 2020; Sanjalawe et al., 2025; Siriwardhana et al., 2021).

This creates a dual challenge. AI models can identify suspicious traffic and adapt routes, yet the models and their training data can themselves be manipulated. A route with favourable latency may traverse an untrusted node, while a stronger verification process may introduce computation and signalling costs that conflict with low-latency service requirements. Secure intelligent routing should therefore be viewed as a coupled problem of learning, threat detection, policy enforcement and architecture-specific deployment rather than as an isolated optimisation task (Lu et al., 2023; Sedjelmaci et al., 2024).

Existing surveys cover 6G security, AI-native network management, blockchain and physical-layer defences (Kumar et al., 2026; Ogenyi et al., 2025; Zuo et al., 2024). These reviews establish the technological landscape but do not always separate security functions that observe network behaviour from mechanisms that directly constrain forwarding decisions.

More importantly, findings from heterogeneous testbeds are often presented together without checking whether the reported performance measures are comparable. The resulting gap concerns the relationship among the security objective, AI technique, routing context and deployment cost not merely the presence of an AI algorithm.

The study addresses three research questions (RQs): RQ1: What security relevant AI techniques and architectural contexts are represented in the literature on intelligent 6G routing? RQ2: How do these techniques support threat detection, trust management and adaptive route selection, and what technical limitations emerge when they are compared? RQ3: Which thematic contributions and unresolved research directions are visible across representative publications, and what evidence is still needed to establish the field's maturity?

The review offers three contributions. First, it organises representative research around the pathway from AI technique to security function, routing context and measurable outcome. Second, it distinguishes reported design intentions from demonstrated technical results, thereby avoiding unsupported cross-study rankings. Third, it proposes a reproducible benchmarking agenda and an expanded conceptual framework that explicitly includes adversarial robustness, energy, latency, privacy and governance. The strength of quantitative bibliometric claims remains contingent upon access to the final screened Scopus corpus and its metadata.

METHODOLOGY

Review design and reporting standard

This study is using Scopus-based systematic literature review assisted by Scopus AI. The reporting requirements is conducted according to PRISMA 2020 for transparent literature searches (Page et al., 2021; Rethlefsen et al., 2021). PRISMA is a reporting guideline, not evidence that screening has occurred; compliance must be demonstrated by an auditable record of queries, dates, decisions and counts. The original submitted document did not contain those records. Accordingly, the procedures and fields below should be reconciled against the authors' actual search export and screening log before they are presented as completed methods.

Preliminary Identification

Scopus is retained as the principal bibliographic database because the submitted review explicitly relied on its indexed records, abstracts and citation metadata. The search includes the terms, dictionaries verified identification equivalent expression, AI-related security functions and intelligent routing in prospective 6G and relevant beyond-5G systems. The search from the following bolean string produced 189 articles in total.

Table 1: Search Strings

Scopus
TITLE-ABS-KEY(("6G" OR "sixth generation" OR "beyond 5G" OR "B5G") AND (rout* OR "path selection" OR "traffic steering" OR "network orchestration") AND (secur* OR trust OR "intrusion detection" OR "anomaly detection" OR adversarial OR spoof* OR privacy) AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "reinforcement learning" OR "federated learning" OR blockchain OR "zero trust")) AND PUBYEAR > 2019 AND PUBYEAR < 2027

Screening and Eligibility

Screening activity was conducted to search for potentially relevent research that related to the predefined research questions. The first stage of screening excluded 135 articles and the second stage asess the remaining 54 articles through the screening selection with inclusive and exclusive criterion (see table 2). The main document assessed are the articles, while other are excluded. The article must be written in english, starting 2021 to the latest year. The research area are computer sciences and engineering with artificial intelligent. Lastly, no publication is exluded due to duplication.

Table 2: Screening Selection Criterion

Dimension	**Include**	**Exclude**
Language	English	Non-english
Published Year	2021 to the latest	<2021
Sources Type	Journals (research articles ONLY)	Conference proceeding
Document Type	Articles	Letter, Review, Conference, Note
Research Area	Computer Sciences and Engineering	Non Computer Sciences and Engineering

Next is to conduct the final examination for the eligibility of the related publication. During the eligibility screening, 54 articles were examined throuhg careful examination of the titles and substantive content to verify their relevances to the study focus. Finally, only 8 articles choosen into final discussion.

Data extraction, Synthesis and the Role of Scopus AI

A study-level extraction table should capture the full reference, publication year, document type, AI method, security task, threat model, routing architecture, dataset/testbed, experimental baseline, evaluation measures and limitations. Critical appraisal should distinguish reported measurements from claims made only in introductions or abstracts. Where studies do not provide sufficiently comparable numerical results, a narrative comparative synthesis is more defensible than averaging accuracy or latency values across dissimilar experiments.

For transparency, Scopus AI-generated topic suggestions, concept-map nodes and researcher recommendations should be logged separately from author-created codes. A reviewer should be able to identify which categories were proposed by the system, which were accepted or altered by the authors, and which included studies support each final theme. Scopus AI output should not be the sole basis for identifying

"leading" researchers; influence assessments require a specified corpus, disambiguated author profiles and a declared indicator.

RESULTS: CRITICAL SYNTHESIS OF REPRESENTATIVE RESEARCH

Threat Identification to Security-Constrained Route Adaptation

Across the cited research, the first important distinction is between systems that detect a security event and those that change routing behaviour. Supervised and deep learning approaches typically learn patterns associated with suspicious traffic or node activity. Detection can inform routing but does not by itself ensure that a secure path is selected. By contrast, reinforcement-learning approaches can adapt route selection to evolving conditions, although the reward function, observation quality and safety constraints determine whether security is genuinely prioritised (Rahman & Hossain, 2022; Lu et al., 2023; Gottam & Kar, 2026).

Gottam and Kar (2026) provide a routing-specific example by combining learning and graph-based modelling for secure D2D operation. Their design is more directly related to secure route control than broad surveys of 6G security, but results from a particular simulated environment should not be generalised to arbitrary 6G topologies. Singh et al. (2026) present a reinforcement-learning-based secure mobile ad hoc routing protocol; this is informative as adjacent evidence, not automatically an empirical validation within a 6G deployment. These distinctions help prevent the review from treating every AI-based security paper as a direct test of intelligent 6G routing.

Trust, Continuous Verification and Complementary Protection

A second research stream emphasises trust and policy enforcement. Li et al. (2020) discuss blockchain-supported data integrity for AI applications in 6G, while Zuo et al. (2024) review broader blockchain–AI integration. These works show why distributed record integrity may be useful in decentralised settings, but blockchain does not automatically verify that a routing decision is optimal or that its underlying training data are trustworthy. Consensus cost, storage growth and network signalling therefore matter when such mechanisms are placed near resource-constrained routing agents.

Sedjelmaci et al. (2024) propose an AI-assisted zero-trust 6G security framework with hierarchical defence agents. The approach is notable because it links adaptive intelligence to continuous verification; however, the policy-enforcement path must be examined to determine whether authentication steps can satisfy routing latency and control-plane requirements. Physical-layer protection offers a distinct security boundary: Lu et al. (2023) examine reinforcement-learning-based cross-layer security and privacy, whereas Ara and Kelley (2024) address the ambitions of native physical-layer security. These functions complement but cannot replace route-level identity, integrity and adversarial validation.

Routing Architectures Change the Security Problem

A technique cannot be judged independently of its operating context. D2D and mobile ad hoc systems are sensitive to local trust, changing neighbours and rapid route turnover. Edge-intelligent and industrial Internet of Things (IIoT) scenarios must balance local inference against device capabilities and communication cost (Rahman & Hossain, 2022). Vehicular networks add stringent mobility and access-control demands (Zhou et al., 2022), while aerial systems require adaptation to changing link availability (Bilen et al., 2022). Open RAN introduces additional control interfaces and distributed software components that must be governed securely (Soltani et al., 2022).

The central result is not that one architecture is uniformly more secure than another. Instead, the same AI technique may offer different benefits and impose different costs depending on where it is executed, which information it receives, how frequently the topology changes and which security policy it is required to satisfy. Figure 1 makes these relationships explicit.

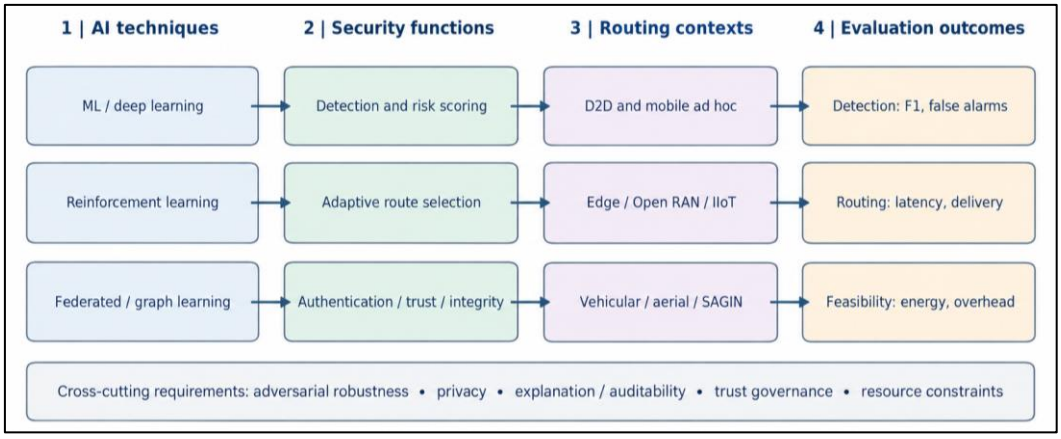


Figure 1: Expanded cross-layer framework linking AI techniques, security functions, routing contexts and measurable outcomes

Comparative Assessment and Evidence Boundaries

Table 3 organises representative publications by their most defensible contribution. It does not report numerical winners because the submitted evidence does not include a harmonised extraction of datasets, baselines or comparable effect sizes. "Evidence boundary" describes the limits of a study's relevance to secure intelligent routing; it is not a claim that the individual article omitted every listed evaluation metric.

Table 3: Comparative evidence for representative AI-based security and routing approaches.

Representative Study	Technique / Function	Routing Relevance	Evidence Boundary
Gottam & Kar (2026)	Reinforcement learning and graph-based modelling for secure path selection.	Direct: D2D routing in proposed 6G settings.	Experiment-specific baselines; verify latency, attack scenarios and inference cost from full text.
Singh et al. (2026)	Reinforcement-learning protocol for adaptive, secure routing.	Adjacent: mobile ad hoc networks.	Transfer to 6G needs explicit justification; validate attack and mobility settings.
Rahman & Hossain (2022)	Deep-learning-assisted, software-defined network security.	Adjacent: 6G IIoT security architectures.	Detection and control integration should be separated from route optimisation.
Sedjelmaci et al. (2024)	AI and zero trust with hierarchical defence agents.	Architecture-level 6G security.	Policy and verification overhead require routing-specific measurement.
Li et al. (2020)	Blockchain-supported data security for 6G AI applications.	Indirect: integrity/trust supporting network decisions.	Does not by itself demonstrate secure next-hop or end-to-end path selection.

Lu et al. (2023)	Reinforcement-learning-based physical/cross-layer security.	Cross-layer 6G security and privacy.	Link-layer improvement cannot be equated to route-level integrity.
Zhou et al. (2022)	Fine-grained access control in 6G vehicular transport.	Adjacent: vehicular security and access governance.	Evaluate routing effect separately from authentication effectiveness.
Soltani et al. (2022)	Security of open and AI-enabled radio access networks.	Architectural context: Open RAN.	Broad security discussion does not provide a routing-specific benchmark.

A fair comparison must examine both defensive effectiveness and operational feasibility. For example, high intrusion-detection accuracy cannot establish overall superiority if inference time, false alarms or energy use are not measured. Similarly, reduced routing delay under normal traffic does not demonstrate adversarial robustness. Table 4 defines the minimum evaluation dimensions needed for meaningful cross-study comparison.

Table 4: Recommended technical measures for consistent cross-study evaluation

Dimension	**Comparable Reporting Requirement**	**Justification**
Threat detection	Precision, recall, F1 score, false-positive rate; dataset class balance.	Avoids interpreting accuracy alone as strong security.
Routing performance	End-to-end delay, delivery ratio, throughput, route discovery/recovery time.	Shows whether detection and enforcement improve forwarding outcomes.
Computation and energy	Inference/training time, memory footprint, energy per decision.	Indicates feasibility on edge nodes or constrained terminals.
Communication cost	Signalling bytes, model-update traffic and control-message frequency.	Exposes overhead from federated learning and distributed verification.
Scalability / mobility	Node count, topology churn, mobility rate, convergence time.	Tests robustness beyond a fixed or small topology.
Adversarial resilience	Attack model, attack success rate, poisoning/evasion tests, recovery behaviour.	Tests the security of both routing decisions and the AI model.
Reproducibility	Dataset access, implementation details, baselines and uncertainty estimates.	Makes performance results independently interpretable.

Thematic Contributors and Emerging Research Directions

The original manuscript listed identifying topic experts as a major objective but did not present the names or criteria used. Table 5 therefore reports verifiable thematic contributors represented in the cited literature. This is a contribution map, not a bibliometric ranking or a claim that these researchers are the most influential in

Scopus. An influence ranking should only be added after author-profile disambiguation and citation counting from the final included corpus.

Table 5: Thematic contributors represented in the reviewed references (not ranked)

Authors	Thematic Contribution
Saad, Bennis & Chen (2020)	Foundational perspective on prospective 6G systems and open research challenges.
Lu and colleagues (2023)	Reinforcement learning for physical/cross-layer security and privacy.
Sedjelmaci, Tourki & Ansari (2024)	AI-supported zero-trust approach to 6G network security.
Zuo and colleagues (2024)	Synthesis of blockchain and AI concepts in 6G networks.
Gottam & Kar (2026)	Learning-based security-oriented D2D routing in 6G scenarios.
Kumar and colleagues (2026)	Broad review of AI's defensive and offensive implications for 6G.

DISCUSSION

The review shows the value of separating proposed security architecture from demonstrated secure-routing performance. A detection system may recognise anomalous traffic without altering paths; a reinforcement-learning controller may adapt quickly but depend on vulnerable observations or reward signals; a blockchain or zero-trust layer may improve accountability while consuming bandwidth or computation. These are complementary functions, not interchangeable solutions. Integrating them requires a well-defined control loop: observe network state, assess threats, evaluate trust and policy, select an admissible path, and monitor the consequences of the decision.

A prominent limitation of the published landscape is the inconsistent boundary between "6G-specific" evaluation and results obtained from generic mobile ad hoc, industrial or vehicular networks. Because commercial 6G infrastructures are not mature, transferable findings often come from simulation, emulation or prospective architectural studies. This makes explicit threat models and external-validity statements especially important. Reported advantages should be qualified by the scenario tested, not extended without evidence to a fully operational 6G system.

CONCLUSION

Security-aware intelligent routing in 6G requires more than the addition of an AI classifier to a conventional network. It involves coordinated detection, constrained path selection, continuous verification and architecture-sensitive evaluation. Representative research indicates that machine learning and deep learning contribute threat perception, reinforcement learning supports adaptive route decisions, and blockchain, zero-trust and physical-layer approaches provide complementary security controls. Yet these approaches cannot be ranked reliably without comparable experiments and explicit reporting of threat models, latency, energy, signalling and adversarial resilience.

The principal conceptual contribution of this review is a cross-layer explanation of how AI techniques are translated into security functions within specific routing environments and assessed against common technical

constraints. Its practical contribution is an evidence-linked comparison and evaluation agenda that makes the required trade-offs explicit. Future work should prioritise reproducible, adversary-aware and resource-sensitive assessments of hybrid security mechanisms. For the review itself, completion of PRISMA counts and an auditable Scopus-derived bibliometric analysis is a necessary final step before the manuscript can support strong claims about the research landscape and influential contributors.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Universiti Teknologi Mara (UiTM) Kedah Branch for providing the necessary resources and support throughout the course of this research. Special appreciation is extended to colleagues and peers who contributed valuable insights and constructive feedback, which greatly enhanced the quality of this paper.

The author(s) declare that generative AI or AI-assisted technologies were used during the preparation of this manuscript. The tool(s) were used: [Gemini, ChatGPT and Scopus AI]. The tool(s) were used for language editing, grammar improvement, translation, and data analysis. All AI-assisted content was reviewed and verified by the author(s), who take full responsibility for the accuracy, originality, integrity, and final content of the manuscript.

Conflict of Interest

The authors have no conflicts of interest.

REFERENCES

1. Ara, I., & Kelley, B. (2024). Physical layer security for 6G: Toward achieving intelligent native security at Layer-1. IEEE Access, 12, 82800–82824. https://doi.org/10.1109/ACCESS.2024.3413047
2. Bilen, T., Canberk, B., Sharma, V., Fahim, M., & Duong, T. Q. (2022). AI-driven aeronautical ad hoc networks for 6G wireless: Challenges, opportunities, and the road ahead. Sensors, 22(10), Article 3731. https://doi.org/10.3390/s22103731
3. Gottam, S. R., & Kar, U. N. (2026). A hybrid quantum-meta reinforcement learning and graph attention transformer approach for low-latency and a secure D2D routing in 6G cellular networks. IEEE Access, 14, 25017–25037. https://doi.org/10.1109/ACCESS.2026.3664972
4. Kumar, R., Dutta, J., Vamsi, N., Sankararao Varri, U., & Puthal, D. (2026). Next-generation security in the 6G era: The role of AI in safeguarding future networks. IEEE Access, 14, 17347–17380. https://doi.org/10.1109/ACCESS.2025.3650208
5. Li, W., Su, Z., Li, R., Zhang, K., & Wang, Y. (2020). Blockchain-based data security for artificial intelligence applications in 6G networks. IEEE Network, 34(6), 31–37. https://doi.org/10.1109/MNET.021.1900629
6. Lu, X., Xiao, L., Li, P., Ji, X., Xu, C., Yu, S., & Poor, H. V. (2023). Reinforcement learning-based physical cross-layer security and privacy in 6G. IEEE Communications Surveys & Tutorials, 25(1), 425–466. https://doi.org/10.1109/COMST.2022.3224279
7. Ogenyi, F. C., Baki, M. M., Ogenyi, A. C., Yusuf, Y. O., Ali, M. M., & Haseeb, K. (2025). A comprehensive review of AI-native 6G: Integrating semantic communications, reconfigurable intelligent surfaces, and edge intelligence for next-generation connectivity. Frontiers in Communications and Networks, 6, Article 1655410. https://doi.org/10.3389/frcmn.2025.1655410
8. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71
9. Rahman, M. A., & Hossain, M. S. (2022). A deep learning assisted software-defined security architecture for 6G wireless networks: IIoT perspective. IEEE Wireless Communications, 29(2), 52–59. https://doi.org/10.1109/MWC.006.2100438

10. Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. Systematic Reviews, 10, Article 39. https://doi.org/10.1186/s13643-020-01542-z
11. 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
12. Sanjalawe, Y. K., Fraihat, S., Al-E'mari, S., Abu-Alhaj, M., Makhadmeh, S. N., & Alzubi, E. (2025). A review of 6G and AI convergence: Enhancing communication networks with artificial intelligence. IEEE Open Journal of the Communications Society, 6, 2308–2355. https://doi.org/10.1109/OJCOMS.2025.3553302
13. Sedjelmaci, H., Tourki, K., & Ansari, N. (2024). Enabling 6G security: The synergy of zero trust architecture and artificial intelligence. IEEE Network, 38(3), 171–177. https://doi.org/10.1109/MNET.2023.3326003
14. Singh, S. B., Rizvi, M. A., Saxena, K., Gupta, R., Tripathi, A. N., & Dewangan, N. K. (2026). An adaptive, energy-efficient and secure routing protocol for zone-related mobile ad-hoc networks using reinforcement learning. Scientific Reports, 16, Article 3002. https://doi.org/10.1038/s41598-025-32918-7
15. Siriwardhana, Y., Porambage, P., Liyanage, M., & Ylianttila, M. (2021). AI and 6G security: Opportunities and challenges. In 2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit) (pp. 616–621). IEEE. https://doi.org/10.1109/EuCNC/6GSummit51104.2021.9482503
16. Soltani, S., Shojafar, M., Taheri, R., & Tafazolli, R. (2022). Can open and AI-enabled 6G RAN be secured? IEEE Consumer Electronics Magazine, 11(6), 11–12. https://doi.org/10.1109/MCE.2022.3205145
17. Zhou, Z., Gaurav, A., Gupta, B. B., Lytras, M. D., & Razzak, I. (2022). A fine-grained access control and security approach for intelligent vehicular transport in 6G communication systems. IEEE Transactions on Intelligent Transportation Systems, 23(11), 22195–22206. https://doi.org/10.1109/TITS.2021.3106825
18. Zuo, Y., Guo, Y., Nie, L., Yang, P., He, D., Luo, X., & Wu, D. O. (2024). A survey of blockchain and artificial intelligence for 6G wireless networks. IEEE Communications Surveys & Tutorials, 26(2), 982–1040. https://doi.org/10.1109/COMST.2023.3315374