Advanced International Journal for Research (AIJFR)

E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

AI-Driven Multi-Agent Routing for 6G Networks

Seema Rani

Assistant Prof, Computer Science, GCG Gurugram

Abstract

The sixth generation (6G) of communication networks is envisioned to deliver ubiquitous intelligence, extremely low latency, and seamless connectivity across terrestrial, aerial, and satellite domains. These goals demand a fundamental transformation in how routing decisions are made. Traditional deterministic algorithms are unable to adapt to the rapidly changing, complex, and multi-layered 6G environment. Artificial Intelligence (AI), particularly Reinforcement Learning (RL), offers a pathway for routing mechanisms that can self-learn, self-correct, and self-optimize without human intervention. This paper presents a conceptual and theoretical discussion on AI-driven multi-agent routing for 6G networks, emphasizing the principles of distributed learning, federated intelligence, and knowledge-defined networking (KDN). Simplified theoretical models and possible architectural directions are provided, along with open challenges for future research.

Keywords: 6G networks, intelligent routing, multi-agent systems, reinforcement learning, federated learning, knowledge-defined networking, graph-based optimization.

1. Introduction

6G networks are expected to interconnect billions of heterogeneous devices through a highly dynamic infrastructure that includes terrestrial base stations, unmanned aerial vehicles (UAVs), low-Earth-orbit (LEO) satellites, and mobile edge computing nodes. These systems will require flexible and intelligent routing mechanisms that can handle fluctuating link conditions, node mobility, and real-time service requirements. Conventional routing protocols—such as OSPF, DSR, or AODV—operate on static rules and cannot adapt quickly to the non-linear changes in topology typical of 6G environments. In contrast, AI-driven routing uses machine learning (ML) and reinforcement learning (RL) to continuously improve path selection through feedback-based optimization. Each network node is transformed into an intelligent "agent" capable of perceiving its surroundings and making routing decisions collaboratively. Such cooperation forms the basis of multi-agent routing, where distributed intelligence emerges from the local actions of many autonomous nodes. This distributed, learning-based approach aligns with the overall philosophy of Knowledge-Defined Networking (KDN) — a system where decisions at all layers are guided by AI models rather than manual configuration.

2. Related Work

The idea of intelligent routing is not entirely new, but its application to 6G introduces novel dimensions. Bouchmal *et al.* [1] provided an overview of classical and quantum machine learning models for optimizing routing within software-defined 6G networks, highlighting the scalability benefits of hybrid AI approaches. Urgelles *et al.* [2] proposed a quantum-based routing optimization framework that minimizes delay and energy simultaneously using approximate optimization.

Liu *et al.* [3] introduced a self-evolving 6G architecture with autonomous decision-making, and Aktas [4] explored integrated terrestrial–satellite routing mechanisms that can dynamically allocate resources based on channel state information (CSI). These studies converge on the premise that future routing must combine learning, cooperation, and adaptation.

However, while the frameworks are promising, there is still limited theoretical understanding of how agents interact, how learning converges, and how stability is maintained when hundreds of agents continuously update their routing policies.

3. Theoretical Foundations

Routing in 6G can be modeled as a **Markov Decision Process (MDP)** where an agent's goal is to learn the best policy that maximizes cumulative rewards.

At any time t, the agent observes a state s_t (such as signal strength, delay, or queue length), chooses an action a_t (such as selecting the next hop), and receives a reward $R(s_t,a_t)$. Over time, the agent refines its decision policy π to maximize long-term performance:

$$\pi^* = \arg\max_{\pi} E_{\pi}\left[\sum_t \gamma^t R(s_t, a_t)\right]$$

Here, γ is the discount factor controlling how much future rewards influence present decisions. This equation reflects how each agent learns to balance **short-term goals** (e.g., minimizing immediate delay) and **long-term goals** (e.g., improving energy efficiency).

In **multi-agent systems**, each node not only learns from its own environment but also indirectly from the behavior of neighboring agents. Thus, the environment becomes non-stationary, and learning must be coordinated. Multi-agent reinforcement learning (MARL) addresses this by introducing *shared policies* and *joint rewards*, ensuring cooperation rather than competition among agents.

To stabilize learning, methods such as *centralized training with decentralized execution (CTDE)* are used. In CTDE, agents are trained collectively using shared state information but operate independently during actual deployment. This hybrid setup allows high performance without excessive communication overhead.

4. Proposed Multi-Agent Routing Framework

In a practical 6G system, multi-agent routing can be organized hierarchically. At the **local level**, each node (e.g., base station or UAV) makes rapid routing decisions using its local RL model.

At the **regional level**, multiple nodes share their experiences or model parameters with a controller through **Federated Reinforcement Learning (FRL)**. At the **global level**, the network aggregates these local models to update a shared "knowledge plane" that guides large-scale optimization.

The federated update process can be expressed as:

$$\theta \leftarrow \sum_i \frac{n_i}{N} \Delta\theta_i'$$

where θ is the global model, $\Delta\theta_i$ is the model update from the i_{th} agent, n_i is the data volume from that agent, and N is the total data across all agents. This update ensures that learning is proportional to each node's contribution, reducing bias and maintaining scalability.

The main advantage of such architecture lies in privacy preservation, as raw data (such as channel statistics or user mobility patterns) is not exchanged—only model updates are shared. Moreover, federated multi-agent routing enables adaptive policy transfer, where high-level learning from one region can be reused in another.

5. Additional Theoretical Insights

The combination of RL and graph-based reasoning is a key theoretical advancement in intelligent routing. 6G networks can naturally be represented as dynamic graphs, where nodes are vertices and communication links are edges. Graph Neural Networks (GNNs) can capture the relational structure between nodes and predict link states or traffic flow.

An RL agent using GNN features can better infer the global context of the network, leading to improved routing stability. The integration of GNN-based policy functions in MARL frameworks is expected to enable topology-aware routing — a crucial step for dense and mobile 6G networks. Another theoretical focus is reward design. For example, instead of a single performance metric, multi-objective reward functions can be defined as weighted sums:

$$R_t = w_1 \times \mathrm{Throughput}_t - w_2 \times \mathrm{Delay}_t - w_3 \times \mathrm{Energy}_t$$

This simple formulation ensures that learning does not optimize one metric (like speed) at the expense of another (like power). Such reward shaping strategies guide routing decisions toward balanced performance.

6. Proposed Research Directions

The theoretical foundations open several paths for deeper exploration:

A. Stability and Convergence of MARL Models

Mathematical convergence proofs for distributed RL under dynamic topologies remain limited. Future research should analyze how agent policies converge under stochastic link and node variations.

B. Energy-Aware Reinforcement Learning

Introduce low-power RL agents capable of maintaining performance under constrained computational and energy budgets, critical for UAVs and IoT sensors in 6G.

C. Explainable and Transparent AI for Routing

Develop explainable AI (XAI) models that interpret how decisions are made by the routing agent. This will be crucial for ensuring accountability and trust in autonomous routing systems.

D. Cross-Domain Routing Frameworks

Explore integration between terrestrial, aerial, and satellite routing through unified decision models that dynamically adapt policies based on latency and coverage requirements.

E. Graph-Based and Knowledge-Defined Learning Planes

Build frameworks where graph neural networks (GNNs) feed into knowledge-defined networking (KDN) controllers, creating an AI-driven control plane that proactively reconfigures routes.

F. Federated Learning Scalability

Investigate how federated learning can scale efficiently to tens of thousands of agents without significant delay or communication burden.

Conclusion

AI-driven multi-agent routing forms the cornerstone of intelligent, adaptive, and self-organizing 6G networks. This paper emphasized the theoretical models underpinning such systems — from reinforcement learning to federated intelligence and graph-based optimization. Future research must establish rigorous stability analyses, energy-efficient algorithms, and explainable models that ensure trust and transparency. The ultimate goal is a fully autonomous routing system that continuously learns, adapts, and evolves with the changing network environment — enabling 6G to realize its vision of ubiquitous intelligent connectivity.

References

1. O. Bouchmal, Y. Sadeg, B. Benmammar, D.-E. Meddour, “From Classical to Quantum Machine Learning: Survey on Routing Optimization in 6G Software-Defined Networking,” Frontiers in Communications and Networks, 2023, vol. 4, 1220227.
2. H. Urgelles, P. Picazo-Martinez, D. Garcia-Roger, J.F. Monserrat, “Multi-Objective Routing Optimization for 6G Communication Networks Using a Quantum Approximate Optimization Algorithm,” Sensors, 2022, vol. 22, no. 19, pp. 7570.
3. B. Liu, J. Chen, Z. Su, Q. Wu, Q. Du, “The Framework of 6G Self-Evolving Networks and Autonomous Decision-Making: Challenges and Opportunities,” Applied Sciences, 2021, vol. 11, no. 19, pp. 9353.
4. F. Aktas, “Routing Challenges and Enabling Technologies for 6G Networks: Satellite Networks and Integrated Routing Techniques,” Technologies, 2025, vol. 13, no. 6, pp. 245.
5. R.P. Marinho, “An Energy-Aware and Intelligent Increasing Coverage Area Routing Protocol for 6G,” Computer Communications, 2023, vol. 198, pp. 178–188.