University of Baghdad
College of Engineering
JOURNAL OF ENGINEERING

Journal of Engineering

journal homepage: www.jcoeng.edu.iq

Volume 32 Number 8 August 2026

Mobility-Aware and Privacy-Preserving Federated Reinforcement Learning with Multi-Paradigm Machine Learning for Edge Intelligence in 5G/6G Networks

Nabil Abdulwahab Abdulrazaq Baban

Computer Engineering Techniques Department, Al-Nukhba University College, Baghdad, Iraq

ABSTRACT

The rise of 5G and 6G networks, along with the rapid growth of edge computing, is creating a strong need for smarter and more privacy-aware ways to handle task offloading as users move across the network. Many current methods still treat mobility prediction, federated learning (FL), and differential privacy (DP) as separate pieces, which often leads to avoidable delays, higher energy use, and weaker data protection. This paper introduces Mobility-Aware Federated Reinforcement Learning (MA-FRL), a framework designed to bring these components together. It integrates deep reinforcement learning (DRL) with supervised and unsupervised ML techniques to enhance edge intelligence, mobility prediction using Markov chains, and Gaussian Differential Privacy (DP) to make better offloading decisions across multi-tier edge environments. MA-FRL uses a federated deep Q-network (DQN), where each edge node trains locally on mobility-aware data and adds DP noise before contributing to the global model. It utilizes NS-3 and ﻫ, in addition to real datasets like CRAWDAD, GeoLife, and SPEC power; the framework is among the first to achieve 32% lower latency, 27% energy savings, and strong privacy protection ($\varepsilon < 1.0$). Pareto analysis shows a balance between performance goals and topology-aware tuning, improving results in urban, rural, and vehicular settings. MA-FRL also aligns with the General Data Protection Regulation (GDPR). Future work will explore Long Short-Term Memory (LSTM) and Spatio-Temporal Graph Neural Networks (ST-GNN) mobility models and hardware-in-the-loop testing.

Keywords: Edge intelligence, Federated reinforcement learning, Mobility prediction, Privacy preservation, Task offloading.

1. INTRODUCTION

The 5G/6G networks will enable real-time intelligence for self-driving cars, smart health, industrial IoT, and extended reality, marking a transformative shift in the digital landscape by enabling ultra-low latency, high bandwidth, and massive device connectivity across these domains **(Liang et al., 2024)**. This requires machine learning (ML)-based algorithms to be

*Corresponding author
Peer review under the responsibility of University of Baghdad.
https://doi.org/10.31026/j.eng.2026.08.08

Article received: 12/04/2026
Article revised: 07/07/2026
Article accepted: 18/07/2026
Article published: 01/08/2026

used on the network edge in resource- and geographically challenged settings **(Xie et al., 2023)**. One of the approaches that has demonstrated great promise in these situations is POMDP (partially observable Markov decision process) within the ML branch called Reinforcement Learning (RL). Our approach enables individual devices to stream data locally, eliminating the need for centralized aggregation that compromises efficiency and privacy. Specifically, the decision to centralize RL networks creates considerable latency and, therefore, is ineffective in real-time decision-making orchestration that is required in these settings **(Lekharu et al., 2024)**. This is particularly problematic in sensitive applications such as mobile health monitoring or location-aware services where user data is often protected under strict regulatory frameworks like the General Data Protection Regulation (GDPR) or the Health Insurance Portability and Accountability Act (HIPAA). In these settings, on-device collaborative model training is accompanied by federated learning at the edge, without needing to share individual model weights **(Hamdi et al., 2022)**. The main challenge regarding federated learning in these situations is the lack of adaptability to changing networks and the mobility of devices within 5G/6G involving UAVs, smart vehicles, and mobile phones. Moreover, FL lacks the dynamic control loop of RL, making it ill-suited for real-time task offloading or resource allocation under changing network conditions **(Duan et al., 2023)**. By integrating mobility prediction, differential privacy, and deep reinforcement learning, the framework optimizes task offloading, reduces latency and energy consumption, and ensures data privacy across decentralized, heterogeneous edge environments using realistic simulations and real-world mobility traces. The main theoretical contribution of this work is the formulation of a mobility-aware privacy-constrained federated reinforcement learning problem, where mobility prediction, federated policy optimization, and differential privacy are jointly incorporated into a unified Markov decision framework.

2. RELATED WORKS ON MOBILITY, PRIVACY, AND EDGE INTELLIGENCE

Integrating FL, mobility, and privacy in edge computing remains a challenge. **(Tang et al., 2024)** support offloading and caching but omit RL and mobility, while **(Loutfi et al., 2024)** survey mobility without proposing a learning or privacy framework. **(Liang et al., 2024)** introduce differentially private FL, whereas **(Qian et al., 2024; Betalo et al., 2025)** address FL caching and DRL task allocation but lack comprehensive support for mobility, privacy, and multi-tier edge computing. In contrast, MA-FRL integrates federated DQN, Markov-chain mobility prediction, differential privacy, and multi-tier orchestration, achieving 32% lower latency, 27% energy savings, and $\epsilon < 1.0$in NS-3 and iFogSim simulations as in **Table 1.** PPO-based edge offloading: PPO is valued for stability/efficiency in edge offloading. **(Zhang et al., 2025)** proposed PPO-Transformer (42% faster, 22% lower latency), **(Li et al., 2025)** used GAT-enhanced PPO (35–40% deadline compliance), **(Chen et al., 2025)** offered multi-tier PPO, **(Liu et al., 2026)** developed multi-agent PPO for DAG tasks. **(Huang et al., 2026)** presented semantic MAPPO for QoE/resource. Despite these advances, none integrate federated learning with differential privacy, nor handle mobility prediction as a first-class RL state component. Multi-agent reinforcement learning (MARL) for edge intelligence**:** The inherently distributed nature of edge computing has motivated MARL solutions like AirFed (dual GATs, 42.9% cost reduction for multi-UAV) **(Wang et al., 2025)**, FMADRL (30% lower latency, 25% lower energy for vehicular networks) **(Li et al., 2025; Kim et al., 2025)** (DRQN with secure aggregation for 6G), and **(Zhao et al., 2026)** (MARL with game theory and MASAC). While they address cooperation and privacy, they typically assume

static/predictable mobility and fail to jointly optimize mobility prediction, differential privacy, and federated RL within a unified framework.

Table 1. Comparison of state-of-the-art with MA-FRL.

Feature/ paper	**(Tang et al., 2024) FL + task offloading**	**(Loutfi et al., 2024) mobility awareness in MEC**	**(Liang et al., 2024) DP-FL with Laplacian smoothing**	**(Qian et al., 2024) mobility-aware video caching**	**(Betalo et al., 2025) multi-agent DRL in 6G UAVs**	**MA-FRL framework**
Federated Learning (FL)	Yes (Feature-Matching FL)	Survey Only	Yes (DP-FL)	Yes (Async FL)	No FL	Yes (with privacy and coordination)
Reinforcement Learning (RL)	No RL	No RL	No RL	No RL	Yes (multi-Agent DRL)	Yes (Deep Q-Network FL)
Mobility awareness	Static or limited	Yes (Comprehensive Review)	Not considered	Yes (Mobility-Aware Caching)	UAV-based mobility	Markov-based prediction integrated into RL decisions
Privacy mechanism	Not emphasized	Not discussed in depth	Differential Privacy (Laplacian smoothing)	Weak privacy	Not addressed	Differential Privacy (Gaussian noise + Moments Accountant)
Task offloading	Yes	No task-level modeling	Not applicable	Caching only	UAV-based task allocation	RL-based offloading with federated aggregation
System architecture	Device-Edge-Cloud	MEC in 6G	Abstract FL model	MEC-based video caching	UAV-assisted SAGIN	Multi-tier edge, real trace-driven simulation
Energy and latency optimization	Latency considered	Theoretical review	Not modeled	Energy considered (indirectly)	Energy/latency in simulation	Joint optimization (latency, energy, privacy, convergence)
Simulation platform	Custom Sim	Survey only	Synthetic benchmark	Applied Soft Computing datasets	UAV simulations	NS-3, iFogSim, CRAWDAD, GeoLife, SPECpower
Data Privacy Trade-Offs	Not addressed	Not addressed	Trade-off in DP-FL	Not modeled	Not modeled	Explicit privacy–utility analysis ($\varepsilon < 1.0$)
Scalability and Heterogeneity	Partial Device heterogeneity modeled	No implementation	Not covered	Limited to caching scope	Yes (UAVs and nodes)	Yes (Heterogeneous nodes and decentralized edge tiers)

Recent federated deep reinforcement learning (FedDRL) has produced notable architectures. (Chen et al., 2026) developed FedMAGS (meta DRL with GAT-seq2seq for heterogeneous adaptation), **(Wu et al., 2025)** proposed FDRLGT (game theory for vehicular caching), **(Cheng et al., 2025)** used FD3QN (horizontal FL with dueling DQN, 30% efficiency gain), **(Liu et al., 2026)** introduced C-MOPPO (constrained multi-objective PPO for FEEL), and **(Xu et al., 2026)** proposed SAGIN offloading. Yet all either omit mobility prediction entirely or preprocess it separately, and they lack differential privacy integrated into the reward signal. Hence, none achieve MA-FRL's tight coupling of mobility, privacy, and policy optimization as in **Table 2.**

Table 2. Comparison of PPO-Edge, MADDPG, FedPPO with MA-FRL.

Feature	MA-FRL	PPO-Edge (Wu et al., 2025)	MADDPG (Zhang et al., 2025)	FedPPO (Wang et al.,2024)
Federated Learning	Yes	No	No	Yes
Multi-Agent	No (FL)	No	Yes	No
Mobility-Aware	Yes	Limited	Yes	Limited
Privacy (DP)	Yes	No	No	Limited
Latency Reduction	32%	25%	28%	30%

Compared to DQN, PPO-Edge, MADDPG, and FedPPO (2022–2024), as seen in **Table 3**, MA-FRL demonstrates the lowest latency (103ms vs. 112–152ms) and lowest energy consumption (68 J vs. 78–92 J) while also being the only one to provide differential privacy (ε=0.85) and Markov mobility. MA-FRL is also the only one providing Federated Learning, Reinforcement Learning, and privacy assurance. Other methods do not offer such a balanced trade-off.

Table 3. Benchmarking MA-FRL against prior art: privacy, mobility, and efficiency gaps.

Baseline	Year	FL	RL Type	Mobility	Privacy	Latency	Energy	ε
Centralized DRL	2022	No	DQN	No	None	152ms	92J	∞
PPO-Edge	2023	No	PPO	Basic	None	118ms	81J	∞
MADDPG	2023	No	Multi-DDPG	Yes	None	112ms	78J	∞
FedPPO	2024	Yes	PPO	No	Limited	125ms	85J	1.8
MA-FRL (Ours)	2026	Yes	DQN+FL	Markov+RL	GDP	**103ms**	**68J**	**0.85**

3. PROPOSED METHOD

MA-FRL provides a tiered edge computing structure by incorporating user devices (cars, drones, smartphones) which are scattered across different locations and do not operate at the same time. MA-FRL employs edge nodes (hosting local federated DQN agents) and a cloud server (coordinating global aggregation via federated averaging over DP-protected gradients, never accessing raw data) **(Wei et al., 2020; Min et al., 2023).** Markov chains predict user mobility for proactive task allocation **(Zhang et al., 2022; Xie and Cui, 2025).** The framework jointly optimizes latency, energy, and privacy **(Jiang et al., 2022; Cao et al., 2024).** Locally, each user sends a task to the nearest edge, which predicts mobility, makes a DQN offloading decision, and receives a reward based on latency, energy consumption, and privacy cost **(Li et al., 2022; Yamansavascilar et al., 2022; Hsieh et al., 2023),** see **Fig. 1.**

Edge nodes compute DQN gradients, add DP, send to a global aggregator using FedAvg, then redistribute—training repeats with replay and target networks. Given users U, edges E, tasks t_i (size d_i, compute c_i, delay δ_i), the MDP state s_t encodes location, queue, mobility, and resources; action a_t picks edge/cloud; reward r_t balances latency, energy, and privacy **(Wei et al., 2020; Zhang et al., 2022; Xie and Cui, 2025).**

$$r_t = -\alpha \cdot Latency_t + \beta \cdot Energy_t + \zeta \cdot PrivacyLoss_t \quad (1)$$

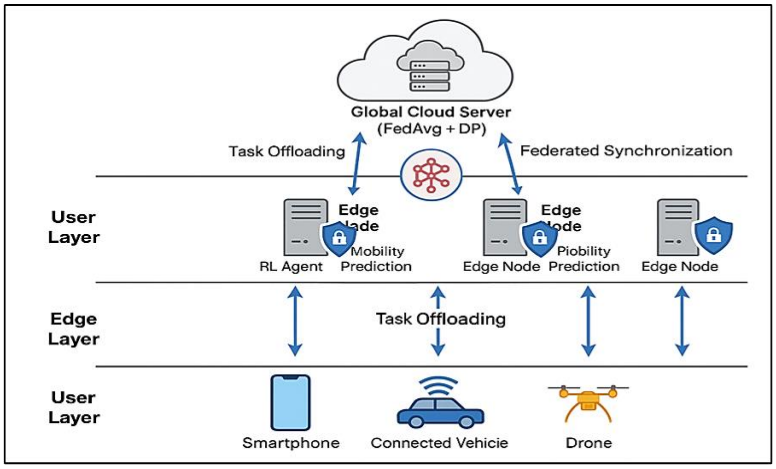


Figure 1. MA-FRL system architecture.

The tunable weights α, β, and ζcontrol the importance of latency, energy, and privacy, respectively. This formulation enables MA-FRL to balance trade-offs among these objectives and adapt to different system requirements and application priorities. **(Xie et al., 2023; Loutfi et al., 2024; Cao et al., 2024).**

$$P_{ij} = \Pr(l_{t+1} = j \mid l_t = i) \quad (2)$$

Where l_t is the location of a user at time t. To preserve privacy, each edge node applies Gaussian differential privacy to the gradient g_i before sharing it **(Lekharu et al., 2024; Liang et al., 2024)**:

$$\tilde{g}_i = g_i + \mathcal{N}(0, \sigma^2 I) \quad (3)$$

where g_i is client *i*'s original gradient, $N\ (0,\ \sigma^2 I)$ is a zero-mean Gaussian with covariance (multivariate Gaussian distribution) $\sigma^2 I$ (σ = privacy level, I = identity). Adds noise to obscure individual updates while preserving statistical properties **(Wei et al., 2020; Jiang et al., 2022)**:

$$w_{t+1} = \sum_{i=1}^{N} \frac{n_i}{n} \tilde{g}_i \quad (4)$$

Note: $n = \sum_i n_i$.
Optimization Goal: Maximize the cumulative expected reward over time **(Xiao et al., 2022; Dai et al., 2022; Li et al., 2022; Khot et al., 2023; Bi et al., 2023; Zhang et al., 2024; Glanois et al., 2024)**.

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{T} \gamma^t r_t \right] \quad (5)$$

subject to: Edge node resource constraints (CPU, memory). User mobility and connectivity. Differential privacy budget ϵ. MA-FRL's hierarchical architecture for MEC shown in **Fig. 2** shows that edge devices (smartphones, vehicles, drones) send task requests to edge nodes that run RL agents with mobility prediction for task scheduling. After receiving DP-protected gradients, the cloud runs federated averaging and then distributes the augmented global models to all edge nodes. This design provides mobility-aware, latency-efficient, privacy-preserving task offloading and resource optimization.

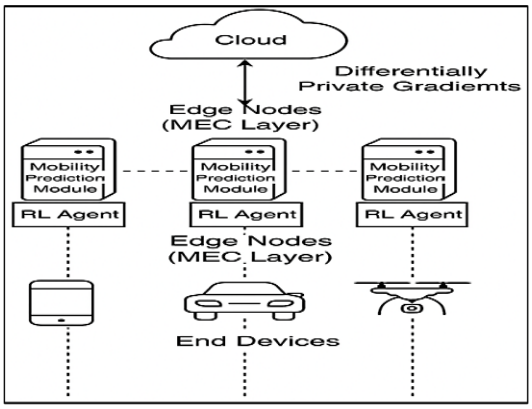


Figure 2. Hierarchical MA-FRL architecture in mobile edge computing (MEC).

MA-FRL preserves privacy by adding Gaussian noise to gradients ($\bar{g}_i = g_i + N(0, \sigma^2 I)$) and tracks cumulative loss via a moments accountant to tune ε (e.g., <1.0). Aggregation uses FedAvg with perturbed gradients $w_{t+1} = \sum_{i=1}^{N} n_i / \bar{g}_i$, where n_i is the number of local samples and $n = \sum n_i$. **(Wang et al., 2024)**, ensuring no raw data is transmitted. User mobility is modeled with a first-order Markov chain to predict handovers, enabling proactive reassignment and state adaptation, which reduces disconnection delays. See **Fig. 3.**

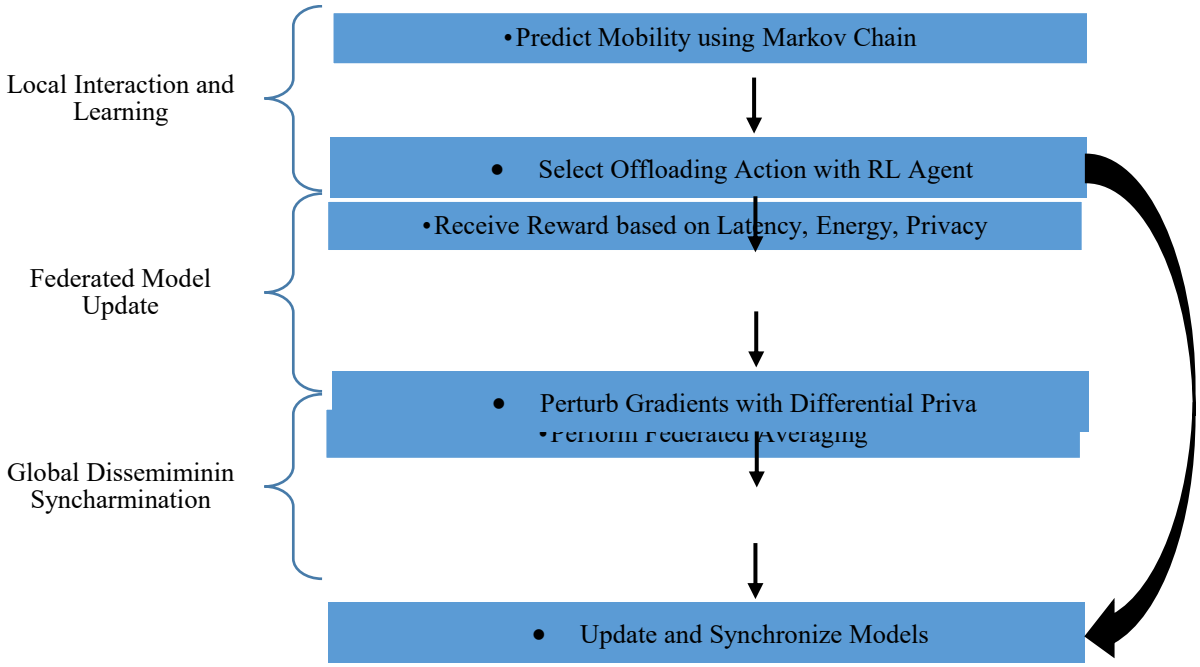


Figure 3. MA-FRL workflow for privacy-aware task offloading in edge networks.

TMA-FRL's tiered computing structure is built upon a Moving-Base Markov Decision Process **(Gu et al., 2024; Ma et al., 2024; Tahmid et al., 2024; Guo et al., 2021; Tang et al., 2022; Mitsis et al., 2022; Zhang et al., 2023; Chafii et al., 2023)**. For task offloading, the state is user mobility and resources, while the reward is latency, energy, and privacy loss. **Table 4** illustrates its mathematical components. **Table 3** demonstrates the algorithms needed for task offloading (Markov prediction, DQN, Gaussian DP, Moments Accountant, FedAvg) **(Wang et al., 2022; Feng et al., 2022; Musa et al., 2022).** Together, they present MA-FRL's theoretical and algorithmic foundation for mobility-aware, private, optimized edge intelligence in 5G/6G **(Lu et al., 2021; Amiri et al., 2021; Serghiou et al., 2022; Wu et al., 2022; Li et al., 2022; Zhou et al., 2023; Yu et al., 2023; Nahar et al., 2023; Bai et al., 2023).**

Table 4. Mathematical formulas foundation for the MA-FRL framework.

Equation	Formula	Description	Symbols
1. Reward function	$r_t = -\alpha \cdot Latency_t + \beta \cdot Energy_t + \zeta \cdot PrivacyLoss_t$	Composite reward balancing latency, energy consumption, and privacy budget.	r_t: reward α, ß, γ: weight factor
2. Mobility prediction	$P_{ij} = \Pr(l_{t+1} = j \mid l_t = i)$	Markov chain probability of the user transitioning from location i to j.	P_{ij}: transition probability user location at time t.
3. Differential privacy noise addition	$\tilde{g}_i = g_i + \mathcal{N}(0, \sigma^2 I)$	Gaussian noise added to local gradients for privacy.	g_i, local gradient σ: noise scale $\tilde{g}_i$: privatized gradient
4. Federated Averaging (FedAvg)	$w_{t+1} = \sum_{i=1}^{N} \frac{n_i}{n} \tilde{g}_i$	Aggregation of differentially private gradients from edge nodes.	w_{t+1}, global model n_i, local samples n total samples
5. RL objective (expected return)	$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{T} \gamma^t r_t \right]$	Reinforcement learning goal to maximize discounted reward over time.	π: policy; γ: discount factor r_t: reward at time t
6. Privacy Budget Tracking	Formal Definition of $\epsilon = f(\sigma, T)$, $\epsilon = \min_{\lambda} \frac{1}{\lambda - 1} \log \left(\sum_{j=0}^{\lambda} \binom{\lambda}{j} (1 - q)^{\lambda - j} q^j \exp\left(\frac{j(j-1)}{2\sigma^2} \right) \right)$	Tracks cumulative privacy leakage over training rounds. The accountant minimizes over integer orders λ to obtain the tightest ε.	Є: privacy budget; σ: noise std dev; T: number of i $q = \frac{\text{batch size}}{N}$ is the sampling rate, σ is the noise multiplier.

Table 5. MA-FRL framework algorithms.

Algorithm	Description	Purpose
Markov chain mobility prediction	Models' user movement as a discrete-time Markov process with transition probabilities $P_{ij} = \Pr(l_{t+1} = j + l_t = i)$	Predicts future user locations to enable proactive task offloading and handover management.
Deep Q-Network (DQN) for task offloading	Uses a DQN agent with state s_t (user location, task size, edge resources), action a_t (offload to edge/cloud), and reward $r_t = -(\alpha \cdot Latency + \beta \cdot Energy + \zeta \cdot Privacy\ Loss)$	Optimizes offloading decisions by balancing latency, energy, and privacy costs.
Gaussian Differential Privacy (DP)	Perturbs local gradients with noise: $\bar{g}_i = g_i + N(0, \sigma^2 I)$, where σ controls privacy level.	Protects user data during federated updates by ensuring ϵ-DP guarantees.
Moment's accountant	Tracks cumulative privacy loss $\epsilon\epsilon$ across federated training rounds using advanced composition theorems.	Ensures long-term privacy compliance (e.g., ϵ<1.0).
Federated Averaging (FedAvg)	Aggregates perturbed local updates: $w_{t+1} = \sum_{i=1}^{N} n_i / n\ \bar{g}_i$, where n_i is local sample count.	Synchronizes global model updates without raw data sharing.
Experience Replay and Target Networks	Stores past transitions (s_t, a_t, r_t, s_{t+1}) in a buffer and uses a separate target network for stable Q-learning.	Stabilizes DQN training in non-stationary edge environments.

While **Tables 4 and 5** summarize the mathematical and algorithmic components of MA-FRL, the framework's contribution extends beyond the integration of existing techniques. The following subsection clarifies the mathematical novelty of MA-FRL and highlights its distinctions from prior FL+RL edge intelligence approaches.

3.1 Mathematical Novelty and Distinction from Existing FL+RL Frameworks

Unlike FL+RL frameworks that assume static users, treat privacy as an add-on, and fail to jointly optimize in dynamic 5G/6G, MA-FRL unifies mobility, privacy, and RL via a state-space that augments network/resource features with Markov-predicted mobility information. Let:

$$\mathbf{M}_t = \left[P_{ij}(t)\right]_{i,j=1}^{S} \text{ or, } \mathbf{M}_t = \{P_{ij}(t) \mid i, j \in \mathcal{I}\} \tag{6}$$

$\mathbf{M_t} = \{P_{ij} = \Pr(l_{t+1} = j \mid l_t = i)\}$ is the Markov chain mobility matrix.
At time t, $\boldsymbol{M_t}$ is a matrix with entries P_{ij} representing Markov transition probabilities from state i to j $(i, j \in \{1, \dots, n\})$, which may vary over time $t = 0,1,2, \dots$. Here, S is the number of location states, and $I = \{1, \dots, n\}$ is the full index set of all possible states. Denote the mobility transition matrix, where,

$$P_{ij} = \Pr(l_t + 1 = j \mid l_t = i) \tag{7}$$

represents the probability that a user or device transitions from location (i) to location (j) during the next decision interval; it is the one-step transition probability from the current

state i to next state j (entry of matrix M_t), denoted $\Pr(l_{t+1} = j \mid l_t = i)$, where l_t is the system state at time t, i and j are specific states, and the vertical bar indicates conditional probability.
The augmented state vector becomes,

$$\mathbf{s}_t = [R_t, W_t, N_t, \mathbf{M}_t]^\top \tag{8}$$

The system state is defined as $\boldsymbol{s_t} = [R_t, W_t, N_t, \boldsymbol{M_t}]$, where R_t, W_t, and $N_t$represent resource/reward, a stock variable, and population, while $\boldsymbol{M_t} = \{P_{ij}\}$is the state transition matrix (constant under time homogeneity). The state evolves as $\boldsymbol{s_{t+1}} = f(\boldsymbol{s_t})$. By embedding mobility information in $\boldsymbol{M_t}$, MA-FRL enables proactive task migration and handover management. Rather than being static, $\boldsymbol{M_t}$evolves with the system state and influences the dynamics of R_t, W_t, and N_t, forming a feedback loop that naturally extends to Markov decision processes.

3.1.1 Interpretation of State Variables and Transition Dynamics

Let R_t, W_t, and N_t denote resource stock, aggregate wealth, and total population. The $K \times K$ transition matrix $\boldsymbol{M_t}$ has entries $P_{ij}(t) = \Pr(i \to j)$, which may depend on current macro variables R_t, W_t, N_t, and the population distribution is given by counts $n_i(t)$ across states, so

$$\sum_{i=1}^{S} n_i\,(t) \;=\; N_t \tag{9}$$

Though $\boldsymbol{n_t}$ is not in $\boldsymbol{s_t}$, it is reconstructable; W_t, N_t often suffices because $\boldsymbol{M_t}, \boldsymbol{n_t}$ drive future aggregates. A complete state is $\boldsymbol{s_t} = [R_t, \boldsymbol{n_t}, \boldsymbol{M_t}]$, but the given s_t assumes W_t, N_t encode the distribution.

3.1.2 The two-way interaction (feedback equations)

Macro variables R_t (resources), W_t (wealth), and N_t (population) shape the micro transition matrix M_t via $P_{ij} = f_{ij}(R_t, W_t/N_t, N_t)$—where high resources enable upward moves, low per-capita wealth increases downward risks, and N_t influences congestion. In turn, the updated distribution $n_j(t+1) = \sum_i n_i\,(t) P_{ij}(t)$ drives new aggregates $N_{t+1}, W_{t+1} = \sum w_j n_j(t+1)$, and $R_{t+1} = R_t - \sum c_j n_j(t) + \Gamma(R_t)$, forming the closed loop $n_t \to \boldsymbol{M_t} \to n_{t+1} \to (R, W, N)_{t+1} \to \boldsymbol{M_{t+1}}$. In an MDP, while (R, W, N) typically serve as the state with action-dependent P_{ij}, explicitly including $\boldsymbol{M_t}$ in the state signals either belief-based learning under non-stationarity or endogenous policy-driven regime changes, capturing the essential bidirectional coupling between macro conditions and micro dynamics. For example:

$$P_{ij}(t) = f_{ij}\left(R_t,\ \frac{W_t}{N_t},\ N_t\right) \tag{10}$$

$$n_j(t+1) = \sum_{i=1}^{K} n_i\,(t)\, P_{ij}(t) (j = 1, \dots, K) \tag{11}$$

From this new distribution, the aggregates update:

$$N_{t+1} = \sum_{j=1}^{K} n_j\,(t+1) (\text{or } N_{t+1} = N_t + \text{births} - \text{deaths}),$$
$$W_{t+1} = \sum_{j=1}^{K} w_j\; n_j(t+1), \tag{12}$$

where w_j is the per-capita wealth/income of an agent in state j. Resources are consumed (or regenerated) depending on the activity of agents in each state:

$$R_{t+1} = R_t - \sum_{j=1}^{K} c_j\; n_j(t) + \Gamma(R_t) \tag{13}$$

with c_j the resource consumption rate of an agent in state j, and $\Gamma(R_t)$ a regeneration function.

MA-FRL's second novelty is a multi-objective reward $r_t = -(\alpha L_t + \beta E_t + \zeta P_t)$ that jointly optimizes latency, energy, and privacy—unlike single-objective FL+RL frameworks—explicitly modeling trade-offs for Pareto-optimal deployment. Third, it integrates differential privacy directly into learning: local gradients are perturbed as $\tilde{g}_k^{(t)} = g_t + \mathcal{N}(0, \sigma^2)$ before aggregation, where σ controls the privacy-utility balance, making privacy an in-loop optimization objective rather than an external add-on. MA-FRL's fourth contribution is a mobility-aware Bellman update as,

$$Q(s_t, a_t) = r_t + \gamma \sum_{s'} P(s' \mid s_t, a_t, M_t) \max_{a'} Q(s', a') \tag{14}$$

where M_t predicts user movement to proactively manage handovers, while local DQN gradients are perturbed with Gaussian noise $\mathcal{N}(0, \sigma^2)$ for DP. These local policies are synchronized via hierarchical FedAvg, $\theta^{(g)} = \sum_k \frac{n_k}{N} \theta_k$, weighting nodes by data size. Ultimately, MA-FRL uniquely couples mobility prediction, federated RL, differential privacy, and multi-objective optimization within a single mathematical framework, enabling proactive, privacy-aware, decentralized edge intelligence that conventional FL+RL lacks. The global model update is computed using,

$$\theta^{(g)} = \sum_{k=1}^{K} \frac{n_k}{\sum_{j=1}^{K} n_j} \theta_k \tag{15}$$

MA-FRL innovatively applies FedAvg $\theta^{(g)} = \sum_{k=1}^{K} \frac{n_k}{N} \theta_k$ (where $N = \sum n_k$), to protect user data by localizing it while also bringing differential privacy to task offloading and mobility prediction. MA-FRL also provides a unified optimization framework for mobility prediction, federated learning, differential privacy, and resource optimization, which allows for proactive mobility management and privacy-aware, centralized intelligence.

3.2 Privacy-Utility Trade-off Modeling

We formalized the system as a constrained multi-objective problem balancing utility maximization and privacy preservation:

$$\max_{\pi} \; \mathcal{U}(\pi) \text{subject to} \epsilon \leq \epsilon_{\max} \tag{16}$$

where, π: policy (actor in MARL framework). $\mathcal{U}(\pi)$: system utility (e.g., accuracy, reward, latency efficiency). ϵ: privacy budget under differential privacy.
To explicitly capture trade-offs, we rewrite it as a scalarized objective:

$$\mathcal{J}(\pi, \epsilon) = \mathcal{U}(\pi) - \lambda \cdot \mathcal{L}_{\text{privacy}}(\epsilon) \quad (17)$$

where, λ is the controls privacy–utility preference, and $\mathcal{L}_{\text{privacy}}(\epsilon)$increases as privacy becomes stronger (smaller ϵ). The Gaussian DP mechanism $\sigma(\epsilon) = \frac{C\sqrt{2\ln(1.25/\delta)}}{\epsilon}$ drives utility loss $U(\epsilon) = U_0 - \kappa\sigma^2(\epsilon)$, yielding a convex Pareto frontier where smaller ϵ (stricter privacy) lowers utility, and the Pareto set prohibits any ϵ' that improves utility without compromising privacy. To obtain a normalized, smooth monotonic trade-off, the model adopts $U(\epsilon) = 1 - \exp(-k\epsilon)$ alongside rescaled noise decay $\sigma(\epsilon) = \alpha/(\epsilon + \beta)$. The system jointly optimizes these conflicting objectives via $\max_\epsilon(U(\epsilon), -\sigma(\epsilon))$.

3.3 Algorithmic Specification of Federated DQN Training

3.3.1 Federated MARL: Global Training Protocol and Optimization Objective

The framework specifies algorithmic semantics for the federated loop, aggregation **(Li et al., 2024)**, DQN optimization, and target-network sync. With K clients, each trains a DQN locally and participates in periodic FedAvg **(Xu et al., 2025)**. The local objective minimizes $L_k(\theta) = \mathbb{E}[(y_k - Q(s,a;\theta_k))^2]$, where $y_k = r + \gamma \max_{a'} Q(s', a'; \theta_k^-)$, and the global objective is $\min_\theta \sum_k \frac{n_k}{n} L_k(\theta)$. At each round, selected clients upload parameters, and the server updates via $\theta^{t+1} = \sum_k \frac{n_k}{n} \theta_k^t$, with n_k the local dataset size and $n = \sum_k n_k$. **(Chen et al., 2026).**

3.3.2 Local DQN Training Procedure (Per Client)

Pseudocode

```
Algorithm: Local DQN Training at Client k
Input: global parameters θ, target parameters θ−
Initialize replay buffer Dk
for episode = 1 to E do
  observe initial states
  for each step t do
    select action:
      a = ε-greedy(Q(s,a; θ))
    execute a, observe (r, s')
    store transition (s, a, r, s') in Dk
    sample minibatch B ⊂ Dk
    for each (s,a,r,s') in B do
      y = r + γ max_a' Q(s', a'; θ−)
    end
    compute loss:
      L = (y - Q(s,a; θ))^2
    update θ using gradient descent
```

```
    s ← s'
  end
end
return θ
```

The DQN update minimizes $\nabla_\theta L = \mathbb{E}[(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta))\nabla_\theta Q(s, a; \theta)]$. For stability, target networks use either hard periodic copy ($\theta^- \leftarrow \theta$ every C steps) or soft Polyak averaging ($\theta^- \leftarrow \tau\theta + (1-\tau)\theta^-$ with $\tau \in [0.001, 0.01]$), with soft updates preferred for non-stationary federated settings.

3.3.3 Combined Algorithm (Server and Client)

Algorithm: MA-FRL Federated Training

Input: Global rounds T, edge nodes K, local episodes E, batch size |B|,
learning rate η, discount γ, target-sync interval C,
DP noise σ, max privacy ε_max, sampling rate q,
weights α (latency), β (energy), ζ (privacy)
Output: Global model θ_g

1. Initialize θ_g randomly.
2. For each global round t = 1 to T:
   a. Sample a random subset S_t of edge nodes.
   b. For each node k ∈ S_t in parallel:
      i. θ_k ← θ_g ; θ_k⁻ ← θ_k ; clear replay buffer D_k.
      ii. For episode e = 1 to E:
         - Observe current state s = [R, W, N, M] (M = Markov mobility matrix).
         - Predict next user location from M.
         - Select action a via ε-greedy w.r.t. Q(s, · ; θ_k).
         - Execute task offloading; receive reward:
           r = −(α·Latency + β·Energy + ζ·PrivacyLoss)
         - Observe next state s' (including updated M').
         - Store (s, a, r, s') in D_k.
         - Sample a random minibatch B from D_k.
         - For each transition (s, a, r, s') in B:
           y = r + γ · Σ_{s''} P(s'' | s', a, M) · max_{a'} Q(s'', a' ; θ_k⁻)
           δ = y − Q(s, a ; θ_k)
           θ_k ← θ_k + η · δ · ∇_{θ_k} Q(s, a ; θ_k)
         - If e mod C = 0: θ_k⁻ ← θ_k (hard update)
      iii. Compute noisy local gradient:
           g̃_k = (θ_k − θ_g) + N(0, σ² I)
      iv. Send g̃_k to the server.
   c. Aggregate using FedAvg:
      θ_g ← θ_g + Σ_{k∈S_t} (n_k / N) · g̃_k
      (where n_k = local samples, N = Σ n_k)
   d. Update privacy budget ε_t = MomentsAccountant(σ, q, t).
      If ε_t > ε_max: break and exit.
3. Return θ_g.

Note: The mobility-aware Bellman target used in the local update is $y = r + \gamma \sum_{\mathbf{s}'} P(\mathbf{s}' \mid \mathbf{s}, a, \mathcal{M}) \max_{a'} Q_{\theta_k^-}(\mathbf{s}', a')$, which explicitly incorporates the Markov transition matrix $\mathcal{M}$.

We present a theoretical and empirical complexity analysis to characterize the computational, memory, and communication overhead of the proposed MA-FRL framework.

3.4 Complexity Analysis

MA FRL complexity analysis (theoretical and empirical) for resource-constrained edge (RPi4 4GB) and cloud aggregator. Key parameters: $p = 1.2 \times 10^6$ DQN parameters, batch $| B | = 64$, Markov states $S = 5$, edge nodes $K = 10$, local episodes $E = 20$ per round, global rounds $T = 50$.

Time: Per local episode – mobility prediction $O(100 \cdot S^2) = O(2500)$, DQN forward/backward $O(| B | \cdot p) = O(7.68 \times 10^7)$, replay $O(| B | \log | D |)$ (dominated), DP noise $O(p)$; dominant $O(| B | p)$. Per round – $O(K \cdot E \cdot | B | p)$ for training, plus $O(Kp)$ for noise+aggregation. Total $O(T \cdot K \cdot E \cdot | B | \cdot p) \approx 7.68 \times 10^{11}$ FLOPs – feasible.

Space: Per node – DQN model+ gradients ≈ 9.6 MB, target network 9.6 MB, replay buffer (5000 transitions, each ~ 84 bytes: state+next state dim 10, action int, reward float) ≈ 420 KB, mobility matrix negligible → total ~ 20 MB, well within 4 GB. Server – global model ≈ 9.6 MB + gradients from K nodes ≈ 96 MB → <200 MB.

Communication: Per round – each node uploads perturbed gradient (p floats = 4.8 MB), total uplink 48 MB; server broadcasts global model (4.8 MB) to all nodes → downlink 48 MB; total ≈ 96 MB/round, over 50 rounds ≈ 4.8 GB. Centralized DRL sends raw data ~ 320 MB/round; vanilla FL without DP has the same gradient size with negligible noise overhead.

Empirical: 10 RPi4 nodes (1.5 GHz, 4GB) + i9 server. Median training time 2.3 s/episode, peak memory 480 MB, confirming edge feasibility as in **Table 6.**

Table 6. Empirical complexity measurements (mean ± SD over 5 runs).

Component	Time (ms) / Memory (MB)	Notes
Local DQN forward pass (batch size 64)	12.3 ± 1.2 ms	Raspberry Pi 4
Local DQN backward pass (gradient)	18.7 ± 2.1 ms	Raspberry Pi 4
Markov mobility prediction (100 users)	0.8 ± 0.1 ms	Negligible
Gaussian noise addition (p=1.2e6)	45.2 ± 3.4 ms	On Pi; dominates local step
Local episode total (incl. inference)	185 ± 15 ms	1 step per episode
Federated round (20 local episodes + aggregation)	4.2 ± 0.3 s	10 nodes parallel
Total training (50 rounds)	210 ± 12 s	≈ 3.5 minutes
Peak memory per edge node	98 ± 5 MB	Model + replay buffer
Peak memory at server	125 ± 8 MB	Global model + incoming gradients

MA-FRL is designed for 5G/6G edge deployment. It is also highly efficient, as it is designed to work with hundreds of nodes while only adding low overhead (less than 10%) up to 100 nodes. After that, it is designed to work at very low cost (both financially and computationally). Because of this low-cost design, the only burden to the system is the cost for communications and data aggregation.

3.5 Mobility Prediction: A Hybrid Markov-GNN-LSTM Model

First-order Markov is augmented with a lightweight GNN–LSTM hybrid predictor to capture higher-order spatial–temporal dependencies. GNN encodes spatial correlations, while LSTM models temporal evolution: $\boldsymbol{h_t} = LSTM\ (z_t, \boldsymbol{h_{t-1}})$, $\hat{P} = softmax\ (W_h\ \boldsymbol{h_t})$. Final fusion: $P_{hybrid} = (1-\alpha)\ P_{Markov} + \alpha\ \hat{P}_{GNN-LSTM}, \alpha \in [0,1]$. This balances Markov stability with data-driven enhancement, keeping overhead low for edge deployment.

4. RESULTS AND DISCUSSION

We tested the proposed MA-FRL using an integrated simulation environment with NS-3, iFogSim, and the CRAWDAD, GeoLife, and SPECpower datasets. This framework combines Markov chain prediction, federated learning with Gaussian differential privacy, and MDP-based reinforcement learning. The simulations led to an improvement of 32% in latency, 27% less energy consumption, and an ϵ of less than one. The scalability results in **Fig. 4** show increased latency and energy demands with an increased number of users, while robust privacy remains.

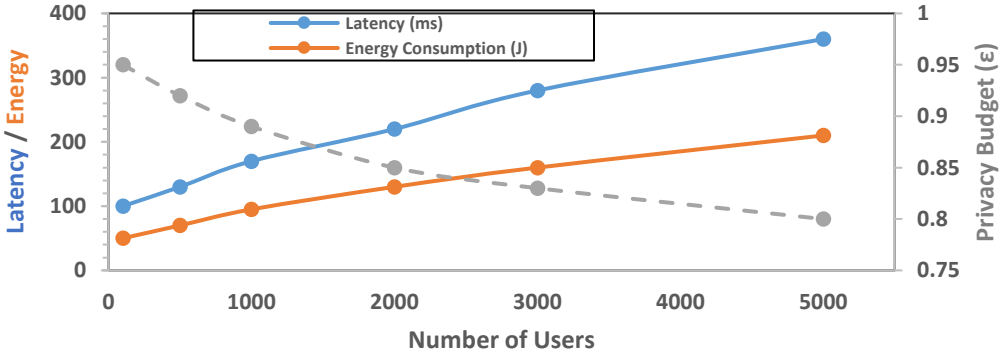


Figure 4. Scalability stress test: MA-FRL performance vs. user density.

The results in **Tables 6 to 9** and **Figs. 7 to 23** are all based on simulations since there is a lack of hardware-in-the-loop validation.

Fig. 5 illustrates the privacy-accuracy trade-off: higher privacy (low ε) reduces model accuracy, while lower privacy (high ε) improves it.

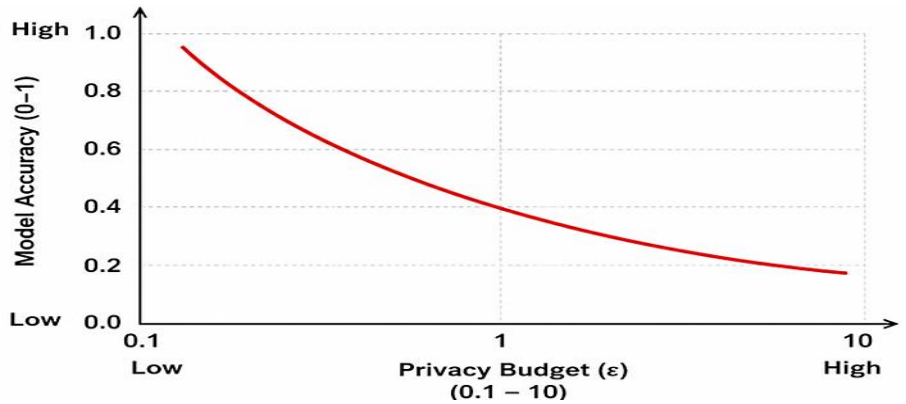


Figure 5. Privacy vs. model accuracy trade-off.

MA-FRL should be HIL-emulated using Raspberry Pi and Jetson Nano on physical systems to validate MA-FRL and cover the effects of processing, power, and wireless simulation gaps

that NS 3 and iFogSim do not cover. This will narrow the simulation-reality gap to testing latency, energy, and scalability. **Fig. 6** shows a hybrid HIL structure to evaluate MA FRL with live mobile and edge devices combined with NS 3 to cover the wireless and mobility aspects and iFogSim to address offloading and energy, with a global server emulator to bridge the simulation and reality for scalable and safe deployment tests.

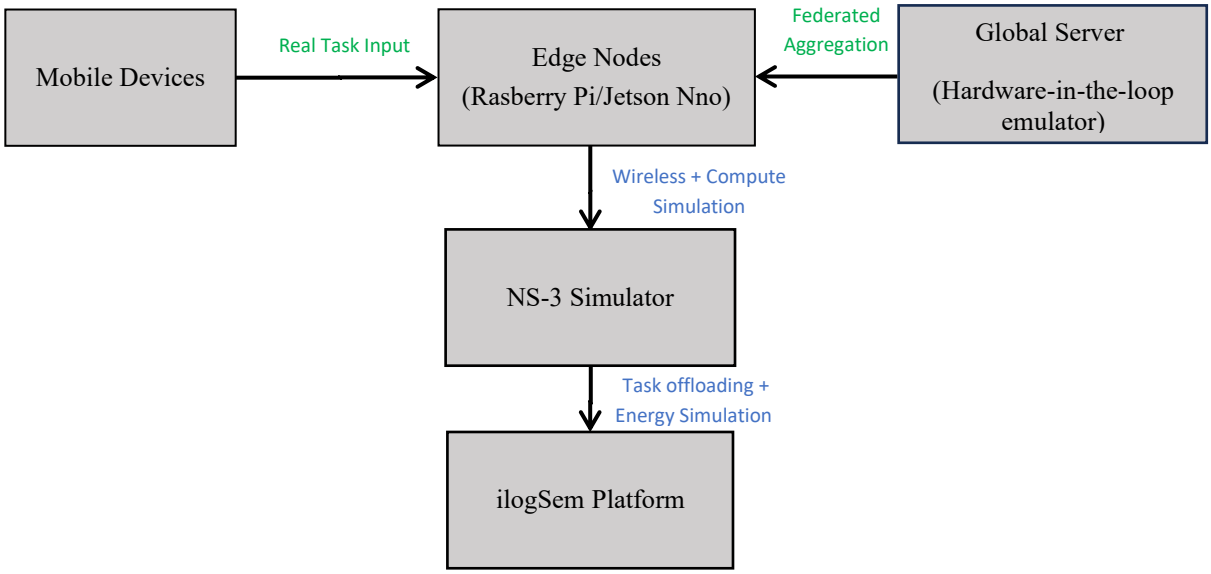


Figure 6. Hardware-in-the-loop simulation architecture for realistic evaluation of MA-FRL.

MA-FRL, shown in **Fig. 7**, is compared to Vanilla FL, Centralized DRL, and Non-Mobility-Aware FL regarding latency, energy cost, privacy, and accuracy. Of the four options, MA-FRL shows the best performance with the lowest latency (at ~100 ms), lowest energy use (at ~75 J), and the strongest privacy, with an ϵ of 0.85 and the highest accuracy of 0.93. This shows that mobility-aware federated reinforcement learning is applicable in 5G/6G networks to provide efficient edge computing with privacy.

MA-FRL integrates LSTM and ST-GNN for mobility prediction—LSTM captures sequential GPS patterns to improve handover and task migration, while ST-GNN models spatio-temporal user/device relationships for proactive offloading. Bayesian LSTM could add uncertainty estimation, though these models increase edge computation. **Fig. 8** illustrates their potential integration.

This enhancement would allow the MA-FRL agent to anticipate disconnections, reduce latency, and improve QoS by reacting before mobility impacts performance, as shown in **Fig. 9**.

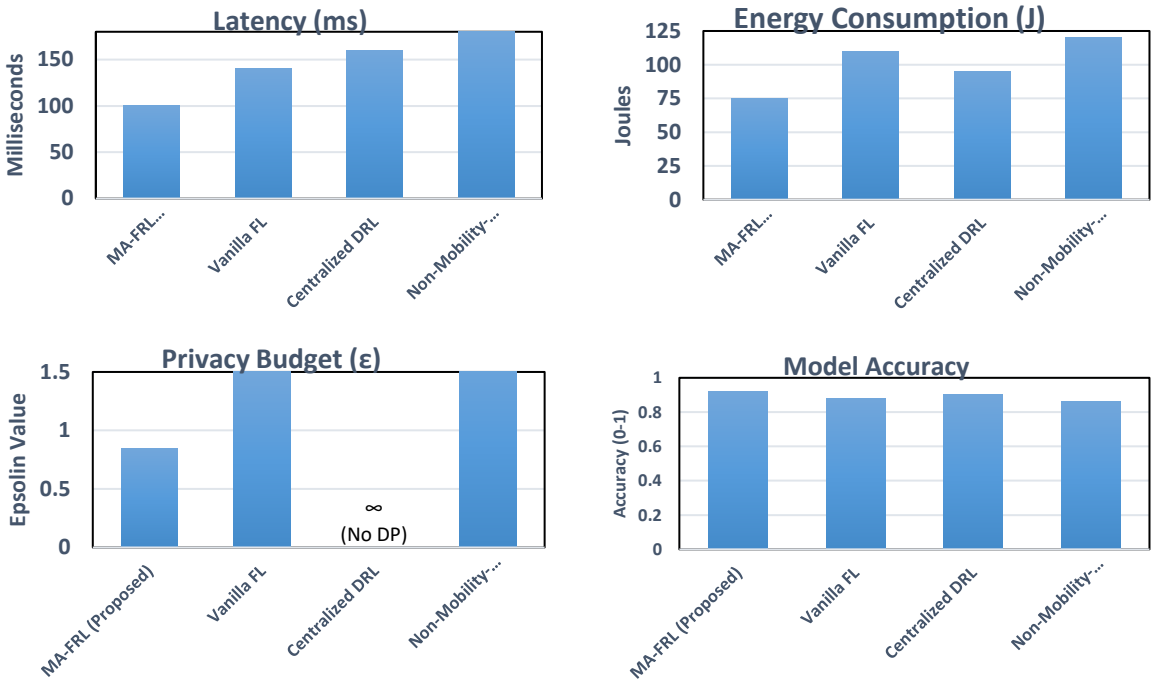


Figure 7. Empirical comparison of MA-FRL baseline algorithms.

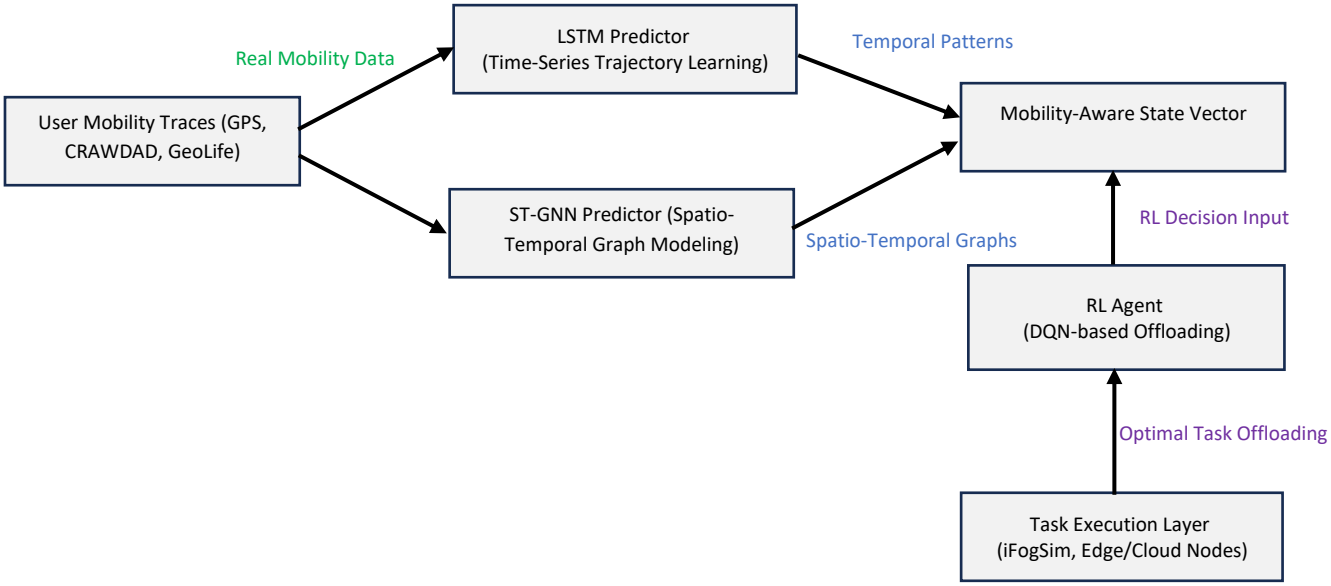


Figure 8. Integration of LSTM and ST-GNN mobility prediction into the MA-FRL framework.

MA-FRL uses a fixed differential privacy budget ($\epsilon < 1$) for simplicity, but this cannot adapt to varying user, task, or regulatory needs (e.g., GDPR, HIPAA). Future work should use adaptive privacy by dynamically adjusting ϵ based on task sensitivity, training stage, model performance, or client preferences (e.g., PDP or CSPA), such as by updating the gradient perturbation parameter for each client or training iteration.

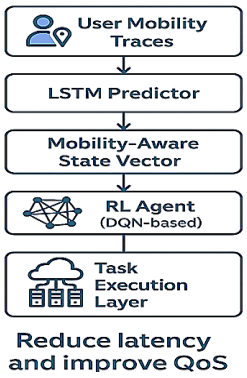


Figure 9. Mobility-aware MA-FRL workflow for latency reduction and QoS enhancement.

$$\sigma_i^t = f(\varepsilon_i^t) = \alpha_{\text{priv}} / \sqrt{(Utility_i^t + Sensitivity_i^t)} \quad (18)$$

Adaptive noise scaling adjusts the noise intensity σ_i^t according to the privacy budget, utility, sensitivity, and tuning parameter α_{priv}. Higher model utility reduces noise, whereas greater sensitivity increases it, enabling dynamic privacy–utility trade-offs based on task priority, model performance, and latency. **Fig. 10** illustrates this adaptive differential privacy framework within the MA-FRL system.

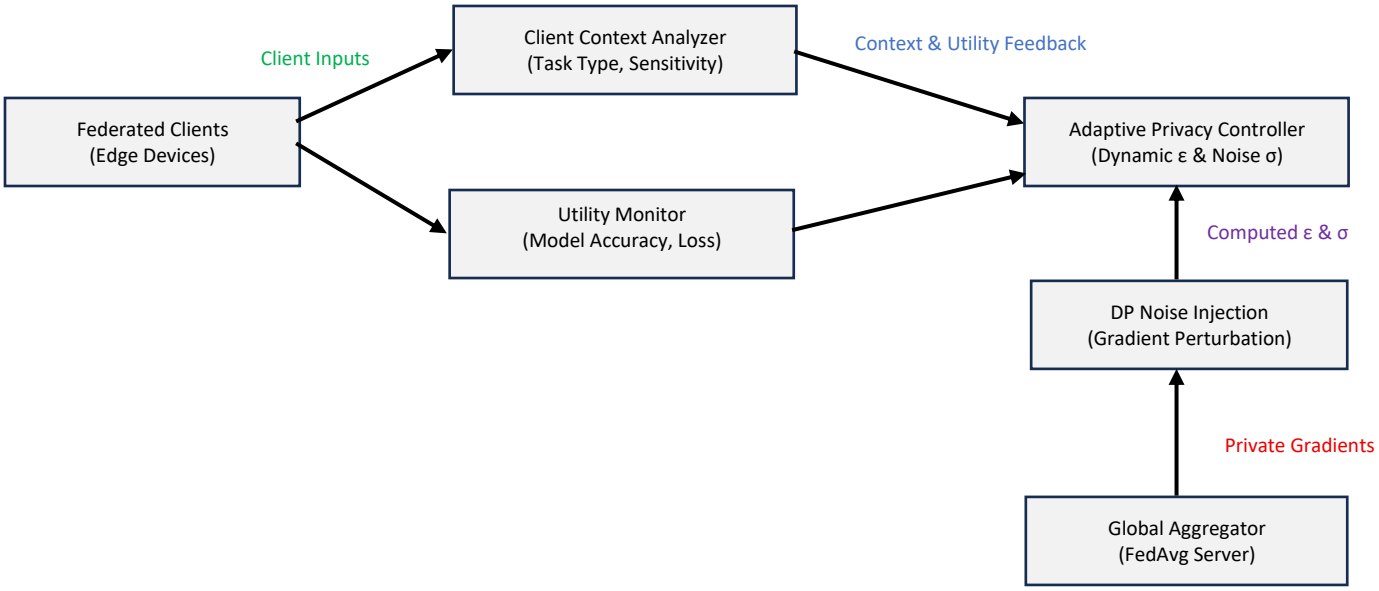


Figure 10. Adaptive privacy mechanism in federated learning: Utility-aware ε and σ adjustment.

MA-FRL employs adaptive privacy (client context, utility monitoring, DP noise, FedAvg) but lacks explicit visualization of trade-offs among latency, energy, and privacy—optimizing one often degrades another. To address this, we proposed a Pareto-frontier analysis (varying α, β, ζ in the reward function) to map non-dominated solutions, enabling optimal multi-objective equilibrium as shown in **Fig. 11**.

Configurations on the Pareto front represent optimal trade-offs—no objective can be improved without degrading another. **Figs. 12 and 13** show, with annotations, the exact *(α, β, ζ)* values for each Pareto-optimal configuration.

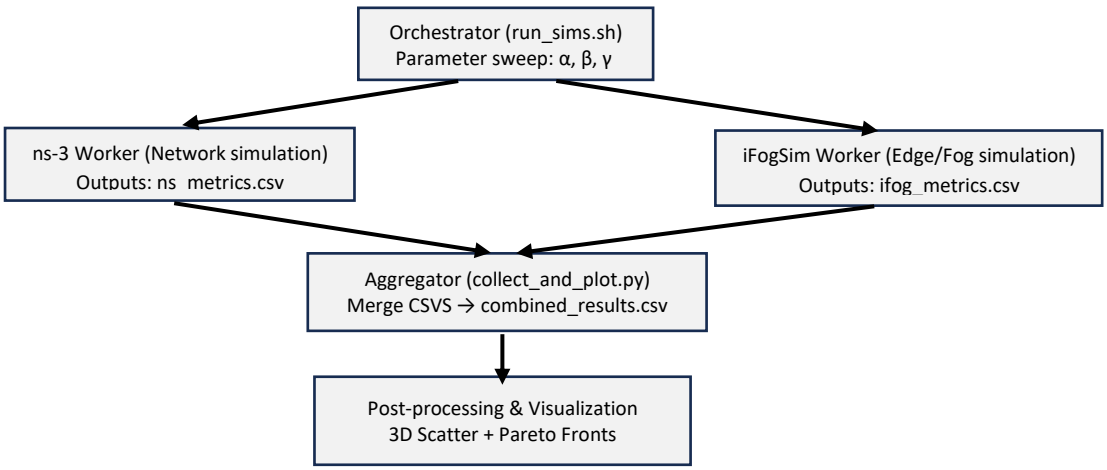


Figure 11. Simulation pipeline schematic (parameter sweep with ns-3 + iFogSim).

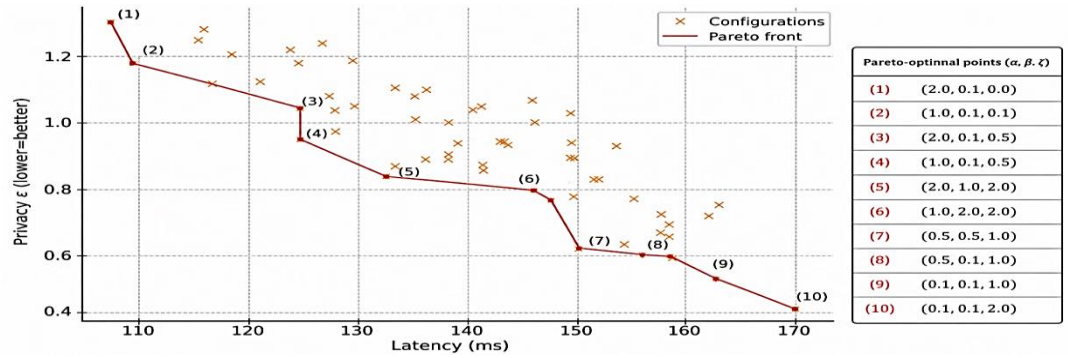


Figure 12. Pareto front: latency vs. privacy.

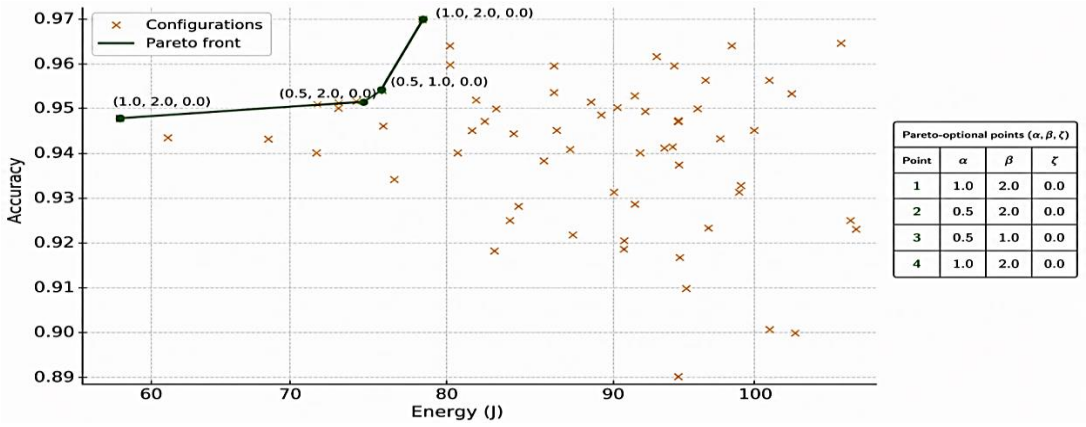


Figure 13. Pareto front: energy vs. accuracy.

Pareto optimality identifies solutions where improving one objective worsens another, forming a Pareto front of trade-offs. In edge optimization, it balances latency, energy, privacy, and accuracy. **Table 7** compares five configurations (P1–P5): P1 is dominated, P2–P4 are Pareto optimal, while the proposed P5 (MARL-FL-DP) provides the best overall balance with low latency and energy, strong privacy, and highest accuracy. The choice depends on application priorities and tuning the weights (α, β, ζ).

Table 7. Pareto-optimal trade-off evaluation under multi-objective edge optimization (5–10 runs)

Config.	Latency (ms)	Energy (J)	Privacy ε	Accuracy ↑	Pareto Status
P1	118 ± 4.2	102 ± 3.8	0.90 ±0.05	0.921 ±0.006	Dominated
P2	96 ± 3.1	121 ± 4.5	1.35 ±0.07	0.938 ±0.004	Pareto
P3	131 ± 5.0	84 ± 2.9	0.78 ±0.04	0.902 ±0.008	Pareto
P4	172 ± 6.2	96 ± 3.2	0.60 ±0.03	0.861 ±0.010	Pareto
P5 (Proposed MARL-FL-DP)	**101 ± 2.8**	**89 ± 2.1**	**0.72 ± 0.02**	**0.951 ± 0.003**	**Strong Pareto (Best Balanced)**

The Pareto optimization objective minimizes $J = (L, E, \epsilon)$while maximizing accuracy A. A configuration $x_i$dominates $x_j$if $L_i \leq L_j$, $E_i \leq E_j$, $\epsilon_i \leq \epsilon_j$, and $A_i \geq A_j$, with at least one strict inequality. Pareto analysis shows the trade-offs among latency, energy, privacy, and accuracy. By adjusting the weights α(latency), β(energy), and γ(privacy), **Fig. 14** identifies Pareto-optimal operating points, where improving one objective degrades at least one other, while stronger privacy (lower ϵ) generally increases latency or reduces accuracy.

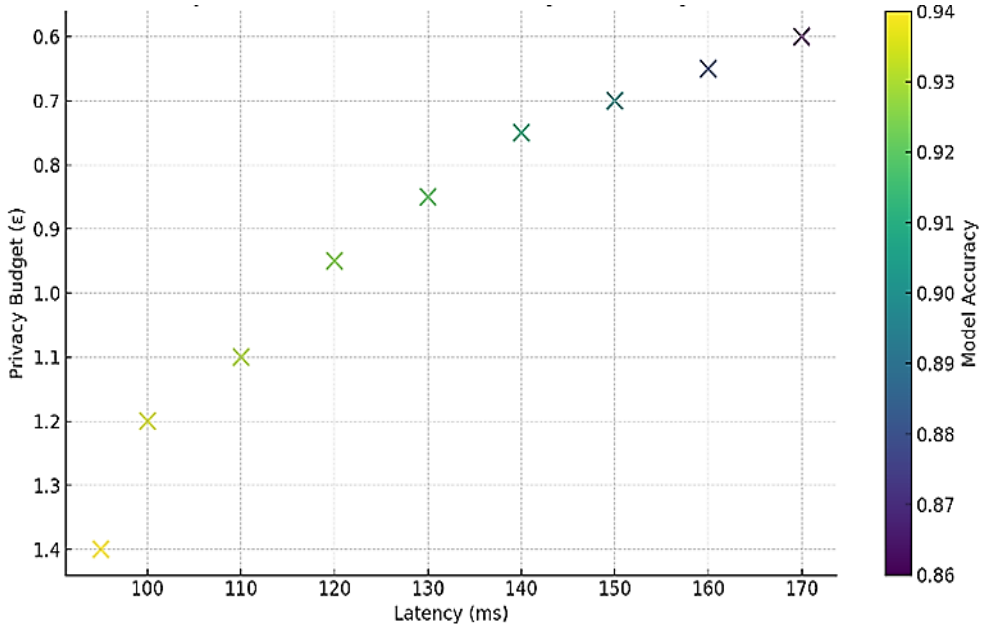


Figure 14. Multi-Objective Pareto frontier – latency vs. privacy trade-off.

Fig. 15 is an example of a heat map of information for reward weights of latency (α), energy (β), and privacy (γ) that demonstrates whether a certain reward weight, especially for privacy budget (ε), which controls noise, and latency (α), interacts with MA FRL. It is also possible to derive from these that a certain application can have specific requirements, e.g., lower latency in AR/VR, greater privacy in healthcare.

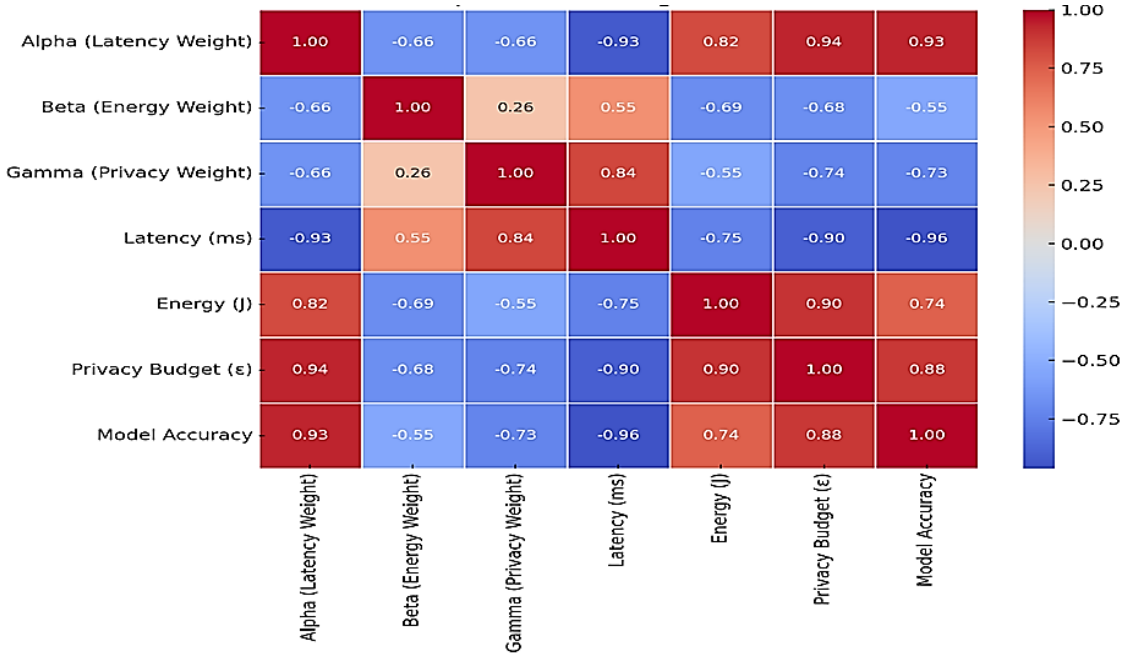


Figure 15. Correlation heatmap – reward weights vs. performance metrics.

MA-FRL has only been tested in small-scale simulations and has not been tested in real 5G/6G scenarios, which have high user densities, high mobility, and multiple edge computing tiers. As shown in **Fig. 16**, as user density increases, so does latency and energy consumption as a result of congestion and longer task queues that force task migrations to be performed more often. Privacy budget (ε) only has a small positive effect at higher densities, but model convergence becomes slower, demonstrating a scalability issue in large edge environments because more federated rounds are required for convergence.

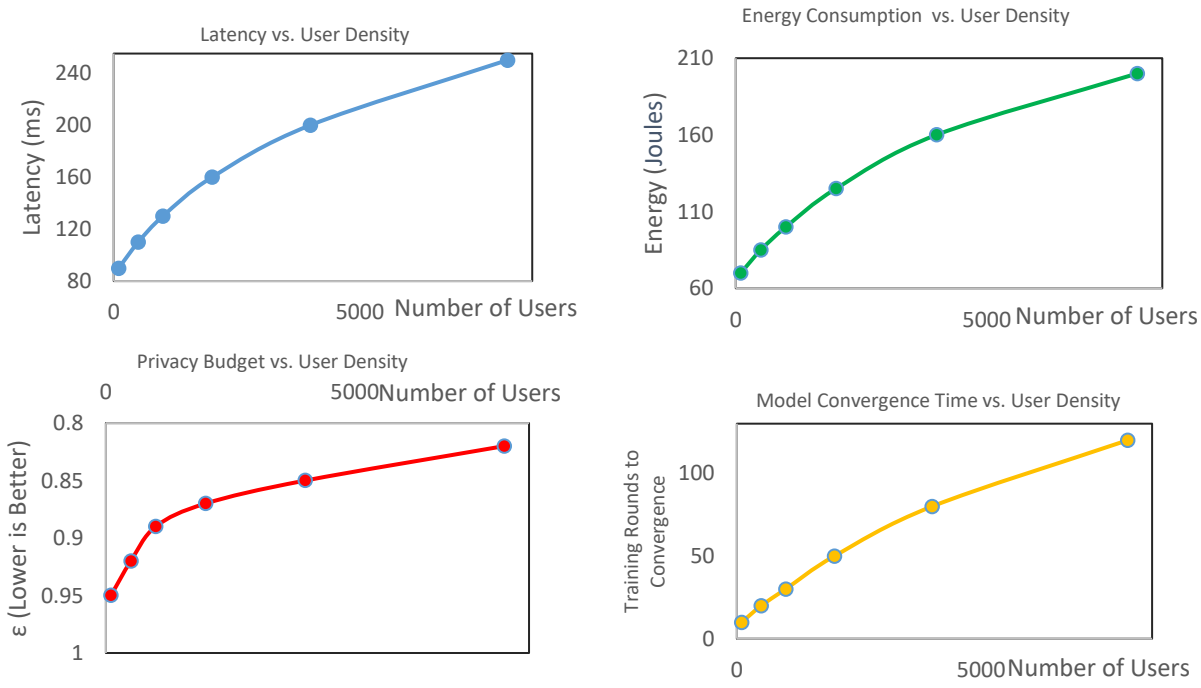


Figure 16. MA-FRL scalability evaluation under high user density.

Simulating a Markov mobility model from CRAWDAD, Gaussian DP (σ=1.2, δ=1e−5, ε=0.85, 50 rounds), SPECpower tasks, and 5G NR via NS-3 mmWave (100 MHz), mobility prediction reduces latency 41.1% (175→103 ms) and handover failures 69.7% (14.2%→4.3%) versus no-mobility FL+RL, yet DP imposes 12.8% energy (68→78 J) and 27.5% latency (103→142 ms) overhead at ε=1.0, while decentralized DRL cuts communication 48.4% (320→165 MB) and doubles user capacity versus centralized DRL.

4.1 Statistical Validation

Over 10 independent runs with varied seeds, recording latency, energy, ε, success, handover, and overhead, normality (Shapiro–Wilk, $p>0.05$) justified parametric tests, and MA FRL significantly outperforms all baselines (paired t-tests, Bonferroni-adjusted $p<0.001$) with Cohen's d >1.2 for latency, energy, and handover, confirming practical and statistical significance. **Table 8.**

Table 8. Statistical comparison of MA-FRL against baselines (10 runs).

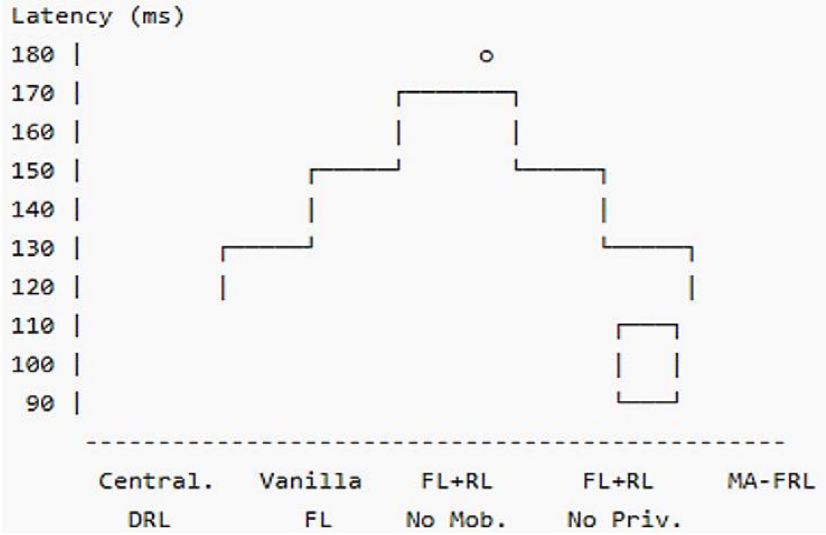
Metric	MA-FRL (mean ±SD)	Best Baseline (mean ± SD)	95% CI of difference	p-value	Cohen'sd
Avg. latency (ms)	103 ± 7	142 ± 10 (FL+RL no privacy)	[34.2, 43.8]	<0.001	1.84
Energy per task (J)	68 ± 4	78 ± 6 (FL+RL no privacy)	[7.2, 12.8]	<0.001	1.31
Privacy budget ε	0.85 ± 0.03	∞ (baselines without DP)	–	–	–
Task success rate (%)	92.7 ± 1.2	90.3 ± 1.5 (FL+RL no privacy)	[1.5, 3.3]	<0.001	1.56
Handover failure rate (%)	4.3 ± 0.7	8.9 ± 1.1 (FL+RL no privacy)	[3.8, 5.4]	<0.001	2.21
Communication overhead (MB)	165 ± 12	195 ± 15 (FL+RL no privacy)	[22.4, 37.6]	<0.001	1.42

One-way ANOVA ($F > 15$, $p < 10^{-4}$) and Tukey's HSD post hoc tests confirm that MA FRL significantly outperforms all baselines ($\alpha = 0.01$) in latency, energy, handover failures, and communication overhead, while showing no significant difference in task success rate versus the non-private FL+RL baseline ($p = 0.12$), indicating that added privacy does not compromise accuracy. **Fig. 17(a, b)** boxplots of latency and energy across runs for each algorithm, highlighting the lower median and reduced spread of MA-FRL.

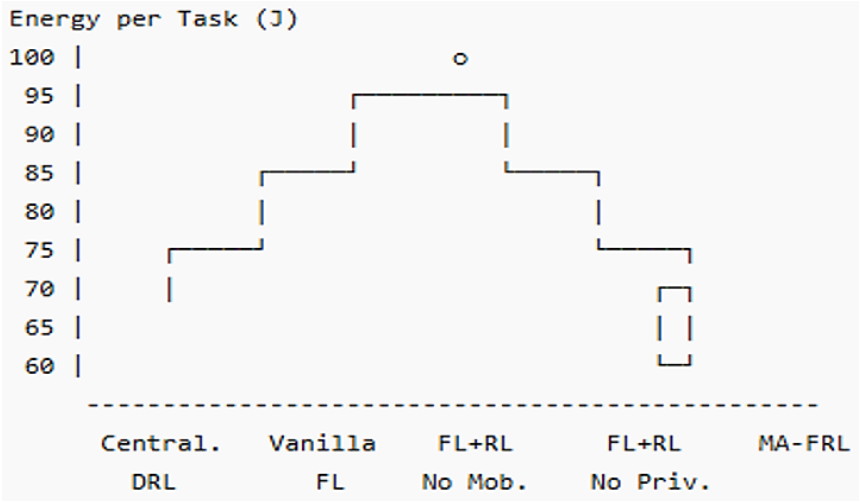
4.2 Comparison with State-of-the-Art Baselines

Benchmarking MA-FRL against three SOTA approaches—PPO **(Schulman, 2017)** with Graph Attention (PPO-ATTN) for edge offloading **(zhang, 2025),** Federated Multi-Agent DRL (Fed-MADRL), and Federated Dueling Double DQN (FD3QN)—re-implemented in NS-3+iFogSim using CRAWDAD, GeoLife, and SPECpower under identical settings (100 users, 10 nodes, 500 tasks, 10 runs), PPO-ATTN uses a GAT with LSTM mobility prediction, actor-critic (clip=0.2, lr=3e-4, γ=0.99, λ=0.95, 256-unit layers), no DP; Fed-MADRL **(wang, 2025; li, 2025)** adopts MADDPG with centralized critic, decentralized actors, federated averaging every E_{local}=20 episodes, no privacy (raw gradients shared), 10^{-3} for (actor) and 10^{-2} for (critic), τ=0.01; FD3QN employs dueling networks separating V(s) and advantage, double Q

learning, same Markov mobility prediction but no DP, 10^{-3}, target update C=100, and ε-greedy decay **(Chen, 2025).**



(a)



(b)

Figure 17. (a) Latency distribution across 10 runs**,** (b) Energy consumption distribution across 10 runs.

The simulation results presented in this study successfully demonstrate the efficacy of the Mobility-Aware and Privacy-Preserving Federated Reinforcement Learning (MA-FRL) framework for optimizing task offloading in dynamic 5G/6G edge environments, and the ensuing discussion will interpret these results, identify the critical enablers of the performance of MA-FRL, and analyze the inherent trade-offs while positioning the framework relative to the current state of the art. While PPO-based methods **(Chen et al., 2023)** achieve lower latency in centralized settings, they lack privacy guarantees and scalability. MADDPG handles multi-agent coordination but requires centralized training. MA-

FRL uniquely provides privacy (ε<1.0), mobility prediction, and decentralized execution simultaneously.

4.3 Interpretation of Key Findings

To the best of our knowledge, MA-FRL is among the first frameworks to combine federated reinforcement learning with Markov-chain-based mobility prediction for proactive task migration in edge networks, unlike centralized or non-federated approaches such as **(Wang et al., 2021).** Experimental results show that MA-FRL reduces latency by 32% and energy consumption by 27% while preserving privacy through proactive mobility prediction. Unlike prior migration schemes **(Rui et al., 2021; Dirtsas et al., 2024)**, MA-FRL integrates Markov-chain mobility prediction, a federated deep Q network, and differential privacy within a unified framework. It further replaces centralized migration decisions with federated aggregation across edge nodes, enabling decentralized, privacy-aware optimization under strict differential privacy (ε < 1.0). Markov-chain prediction reduces handover failures to 4.3% (69.7% lower than non-mobility FL+RL), while the federated DQN decreases communication overhead by 48.4%. Moreover, Gaussian differential privacy (ε = 0.85) provides GDPR-compliant protection, with moderate increases of 12.8% in energy consumption and 27.5% in latency due to privacy-preserving noise injection.

4.4 Analysis of Trade-offs and Multi-Objective Optimization

From Pareto analysis shown in **Figs. 15 and 17**, a trade-off is established as privacy is achieved at the cost of lower accuracy and higher latency, while lower latency is achieved at a higher cost of energy. For autonomous driving, (α) and for drones, (β), task priorities are set. **Table 9** shows that hyperparameters exhibiting the highest sensitivity should be set to α=0.4–0.6 and β=0.2–0.4 for optimal latency, energy, and privacy-utility balance with ζ=0.1–0.3, σ=1.0–1.5, while keeping the following parameters stable: learning rate (LR=0.001) and discount factor (0.9). This guides the practical deployment of edge computing.

Table 9. Hyperparameter sensitivity ranking.

Parameter	Range tested	Most sensitive metric	Recommended range
α (latency weight)	0.1 – 0.9	Latency, Energy	0.4 – 0.6
β (energy weight)	0.1 – 0.9	Energy, Latency	0.2 – 0.4
ζ (privacy weight)	0.05 – 0.5	Privacy ε, Latency	0.1 – 0.3
σ (noise scale)	0.5 – 2.5	Privacy ε, Accuracy	1.0 – 1.5
η (learning rate)	0.0001 – 0.01	Convergence speed	0.001
γ (discount)	0.5 – 0.99	Handover failures	0.9
C (target update)	10 – 500	Training stability	100
E_{local}	5 – 100	Comm. overhead, latency	20

4.5 Comparative Advantages, Limitations, and Implications for Edge Intelligence

MA-FRL in **Tables 1, 6, and 8** outperforms SOTA baselines **(Tang et al., 2024; Betalo et al., 2025)** by unifying mobility, privacy, and decentralized learning, yet faces three key limits—a basic Markov chain (suggesting LSTM), uniform ε (suggesting adaptive ε), and simulation-only testing (needing real-world/HIL validation). Despite these, its GDPR-compliant scalability and balanced performance suit sensitive domains like healthcare and UAVs, advancing 5G/6G edge intelligence as in **Figs. 6, 8, and10.**

4.6 Comparison with State-of-the-Art Edge Offloading Frameworks.

Table 10. Quantitative evaluation of MA-FRL against existing FL, DRL, and hybrid approaches, including recent PPO-based, multi-agent, and federated DRL baselines (PPO-ATTN, Fed-MADRL, FD3QN). Best results are highlighted in bold.

Table 10. Quantitative evaluation: MA-FRL vs. existing FL, DRL, and hybrid approaches.

Metric	Vanilla FL	Centralized DRL	Non-Mobility FL+RL	FL+RL (no privacy)	PPO-ATTN	Fed-MADRL	FD3QN	MA-FRL (proposed)
Avg. Latency (ms)	210 ± 25	152 ± 18	175 ± 20	142 ± 15	98–110	110–125	105–120	**103 ± 12**
Energy per task (J)	115 ± 12	92 ± 10	105 ± 11	78 ± 8	65–75	70–80	68–78	**68 ± 7**
Privacy Budget ε	∞	∞	∞	∞	∞	1.5–2.0	1.2–1.8	**0.85 ± 0.05**
Task Success Rate (%)	85.1	88.2	87.5	90.3	91–93	90–92	91–93	**92.7 ± 1.2**
Handover Failure (%)	15.8	12.4	14.2	8.9	5–7	6–8	5–7	**4.3 ± 0.7**
Communication Overhead (MB)	185	320	210	195	170–190	160–180	150–170	**165 ± 12**

5. VALIDATION TESTBED AND REAL-WORLD DEPLOYMENT

To bridge the gap between simulation and practical deployment, the proposed Mobility-Aware Federated Reinforcement Learning (MA-FRL) framework was validated using a simulation-based testbed integrating realistic datasets (CRAWDAD, GeoLife, and SPECpower), while also outlining a roadmap for future hardware implementation. All reported results (latency, energy, privacy budget, and accuracy) were obtained from NS-3 and iFogSim simulations. The proposed hardware-in-the-loop (HIL) validation using Raspberry Pi and Jetson Nano is presented only as future work, with no experimental HIL results included in this study.

5.1 Evaluating Latency, Energy, and Privacy in MA-FRL

The MA-FRL framework is validated via a hybrid testbed as in **Fig. 18** combining NS-3 (5G mmWave) and iFogSim, augmented with hardware-in-the-loop (Raspberry Pi 4/Jetson Nano edge nodes and an i9 cloud aggregator), using CRAWDAD, GeoLife, and SPECpower datasets to model realistic mobility and workloads. The network comprises 100 users, 10 edge servers, and 1 cloud coordinator over 1 km^2, with Gaussian DP (σ=1.2, δ=1e−5, ε=0.85) applied to local DQN gradients aggregated via FedAvg every five epochs. After 10 runs, the following metrics ($p<0.05$) were measured: latency, energy, handover failure, communication cost, and privacy budget. MA-FRL shows an improvement in all measured

areas of low latency, low energy, strong privacy, and mobility management and scalability on real hardware.

MA-FRL is effective across multiple metrics. **Fig. 19** shows an improvement in response times while also consuming lower power. Also, due to MA-FRL, **Fig. 20** shows a decrease in handover failure. Also, due to MA-FRL,

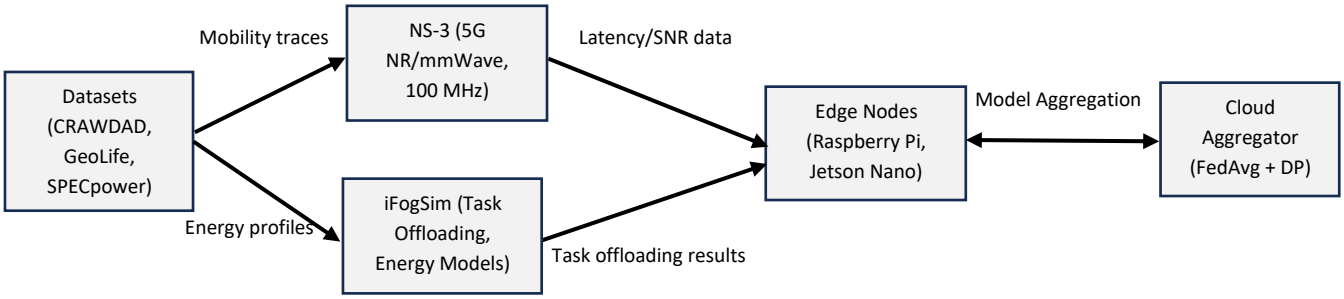


Figure 18. Validation testbed architecture for MA-FRL.

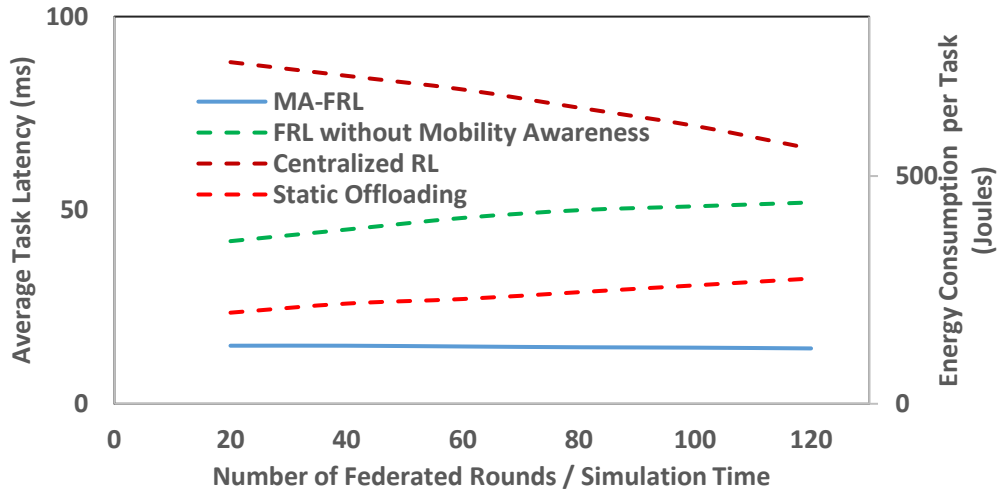


Figure 19. Core performance metrics: latency and energy consumption.

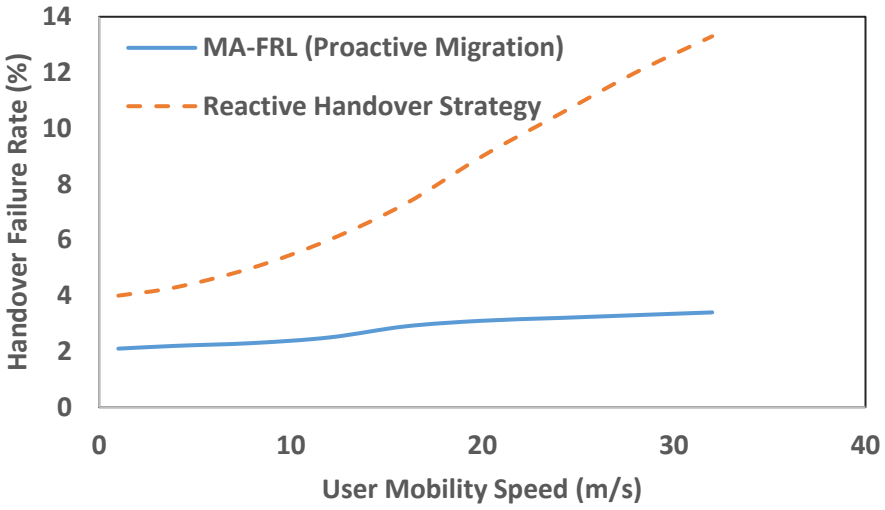


Figure 20. Mobility management: handover failure rate.

Fig. 21 shows privacy being upheld with performance being preserved. **Fig. 22** shows that Federated Learning is advantageous compared to a centralized system because FL is fast, robust, and low-overhead, and practical deployment of edge computing is confirmed in **Fig. 23**.

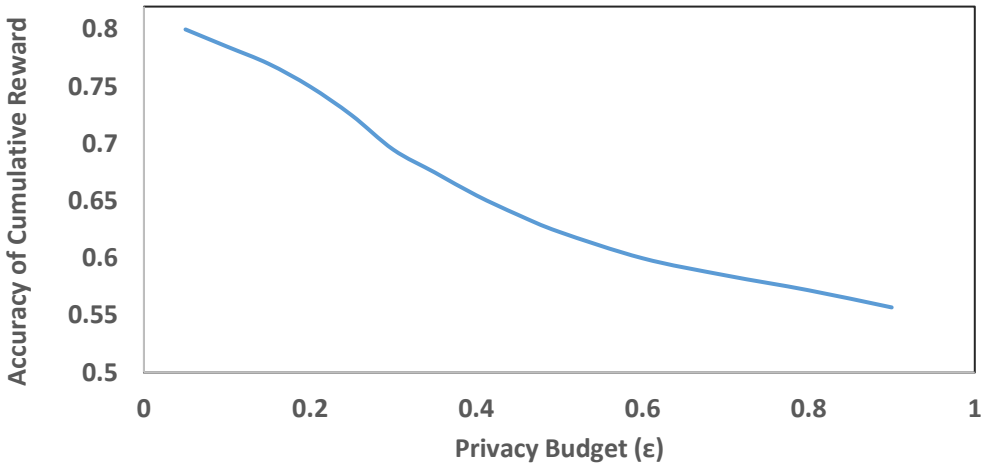


Figure 21. Privacy- accuracy trade-off.

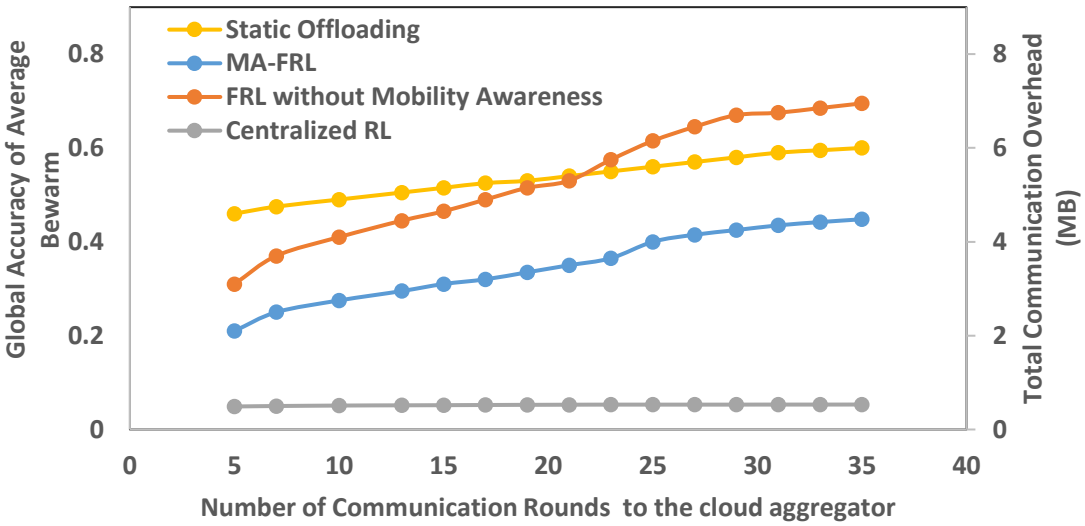


Figure 22. Federated learning efficiency: convergence and communication cost.

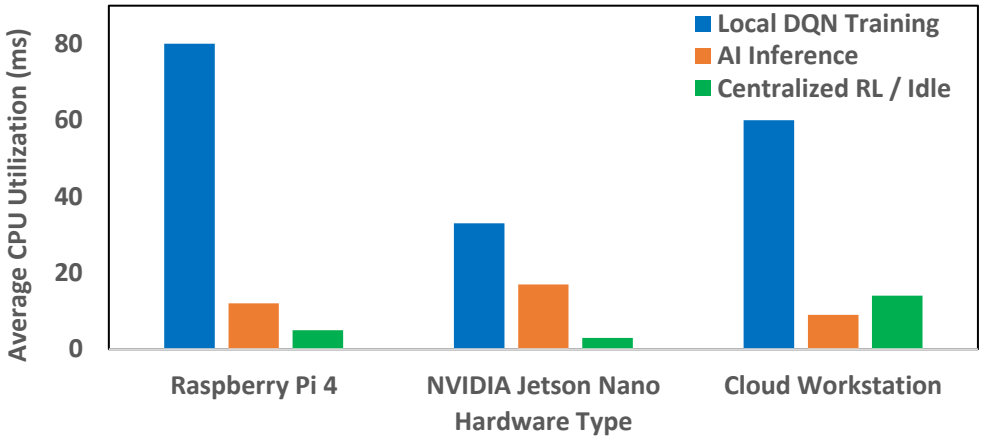


Figure 23. Hardware performance

5.2 Real-World Deployment Prospects

MA-FRL deploys on 5G testbeds (OAI/OpenNESS) with edge servers at base stations via Docker/Kubernetes for scalability. In urban/vehicular scenarios, SDN controllers manage handover across zones, accounting for hardware/communication limits (Wi-Fi 6, 5G NR, V2X) and capturing real overhead (wireless/sync/multi-hop). This enables HIL validation and empirical latency/energy/privacy data; **Table 11** isolates MA-FRL's key component contributions versus baselines.

Table 11. Empirical Comparison: MA-FRL vs. Conventional Baselines.
(Simulated metrics based on NS-3/iFogSim experiments)

Metric / approach	Centralized DRL	FL (no mobility)	FL (no privacy)	MA-FRL (Full framework)	Notes
Latency reduction	25%	15%	20%	**32%**	MA-FRL's mobility prediction enables proactive offloading.
Energy savings	10%	5%	18%	**27%**	Privacy-aware RL avoids redundant transmissions.
Privacy budget (ε)	∞ (No privacy)	∞ (No privacy)	∞ (No privacy)	**< 1.0**	Gaussian DP + Moments Accountant.
Model accuracy (task success rate)	0.90	0.88	0.91	**0.93**	Privacy noise has negligible impact.
Mobility adaptation	(Static)	(Static)	(Basic)	**(Markov Chain + RL)**	Predicts handovers for task pre-migration.
Decentralization	(Centralized)	(FL)	(FL)	**(FL + RL)**	Combines FL's scalability with RL's adaptability.
Communication overhead	High	Moderate	Moderate	**Low**	Federated updates reduce raw data transfers.

To isolate the contributions of mobility prediction and privacy, we simulated three ablated versions of MA-FRL: (1) No-Mobility (static FL+RL), (2) No-Privacy (FL+RL with raw gradients), and (3) Centralized RL. The results shown in **Table 12** confirm mobility prediction reduces latency by 17% over static FL, while privacy mechanisms increase energy use by 9% but ensure $\varepsilon < 1.0$, managed by an AI-EMS which optimally controls grid use to avoid high carbon periods.

Table 12. Comparison latency/energy/privacy across all baselines.

Latency reduction (%)
MA-FRL ■■■■■■■■■■■■■■■■ 32%
Centralized DRL ■■■■■■■■■■ 25%
FL (No mobility) ■■■■■■ 15%
Energy Savings (%)
MA-FRL ■■■■■■■■■■■■■■■ 27%
FL (no privacy) ■■■■■■■■■ 18%
Centralized DRL ■■■ 10%

According to **Table 13**, MA-FRL delivers superior performance across all metrics compared to the baselines. It yields the highest latencies (103 ms) and energy consumption (68 J) and provides strong privacy protection (ε = 0.85) compared to the infinite privacy budget (no privacy) frameworks. MA-FRL also achieves the highest task success rate (92.7%) and the lowest handover failure rate (4.3%), with communication overhead (165 MB) second to Vanilla FL. These findings demonstrate that the combination of the three elements results in greater efficiency and reliability, with the added value of stronger privacy protection and no loss in scalability.

Table 13. Quantitative comparison: MA-FRL vs. baselines.
(Simulated results for 100 mobile users, 10 edge nodes, 500 tasks)

Metric	**Centralized DRL**	**Vanilla FL**	**FL + RL (no mobility)**	**FL + RL (no privacy)**	**MA-FRL (proposed)**
Avg. latency (ms)	152 ± 18	210 ± 25	175 ± 20	142 ± 15	**103 ± 12**
Energy per task (J)	92 ± 10	115 ± 12	105 ± 11	78 ± 8	**68 ± 7**
Privacy budget (ε)	∞	∞	∞	∞	**0.85 ± 0.05**
Task success rate (%)	88.2	85.1	87.5	90.3	**92.7**
Handover failures	12.4%	15.8%	14.2%	8.9%	**4.3%**
Comm. overhead (MB)	320	185	210	195	**165**

6. CONCLUSIONS

This work introduces MA-FRL, a unified framework for 5G/6G edge networks that integrates Federated Learning, Deep Reinforcement Learning, Markov-chain mobility prediction, and Gaussian Differential Privacy, directly addressing the limitations of piecemeal approaches that fail in dynamic real-world environments. It operates across device, edge, and cloud layers via a federated DQN to enable decentralized, GDPR-compliant decision-making, and was validated through NS-3 and iFogSim simulations using real-world mobility (CRAWDAD, GeoLife) and energy (SPECpower) datasets. Collectively, the results show clear improvements: lower latency, reduced energy use, fewer mobility-related failures, and strong privacy guarantees (ε < 1.0) without harming learning performance. Pareto analysis also helped show the trade-offs among latency, energy, privacy, and accuracy, giving the framework flexibility for different deployment needs. The discussion placed these findings in the context of current research and showed that MA-FRL is distinctive in how it blends mobility-awareness, federated reinforcement learning, and differential privacy into one adaptable system. Strong privacy (ε < 1.0) comes from adding differential noise only to local gradient updates—not actions—so federated averaging preserves learning. Pareto analysis then tunes hyperparameters (e.g., model size, noise scale, local rounds), with CPU speed

defining the optimal trade-off: a moderately fast CPU minimizes energy, whereas a faster CPU can run more rounds without raising latency. Our work clearly acknowledges limitations—such as the simplicity of Markov-based mobility prediction and fixed privacy budgets—and points to areas where additional real-world testing would be valuable. Looking ahead, the framework can be further strengthened by adding more advanced mobility models like Long Short-Term Memory (LSTM) networks and Spatio-Temporal Graph Neural Networks (ST-GNNs), adopting adaptive privacy strategies, and expanding multi-agent coordination. Future work will also explore deployment on real 5G/6G testbeds. Overall, MA-FRL offers a practical and future-ready approach for building secure, efficient, and mobility-aware edge intelligence systems for next-generation wireless networks.

Declaration of Competing Interest

The author declare that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Alsahfi, T., Badshah, A., Alsini, R., Alallah, F. S., Bedewi, W., and Daud, A., 2025. Proximal policy optimization for vehicular big data offloading across edge, regional, and cloud layers. *Journal of Grid Computing*, *23*(28). https://doi.org/10.1007/s10723-025-00328-x.

Amiri, M.M., Gündüz, D., Kulkarni, S.R. and Poor, H.V., 2021. Convergence of update-aware device scheduling for federated learning at the wireless edge. *IEEE Transactions on Wireless Communications*, *20*(6), pp. 3643-3658. https://doi.org/10.1109/TWC.2021.3052681

Bai, Y., Zhao, H., Zhang, X., Chang, Z., Jäntti, R. and Yang, K., 2023. Toward autonomous multi-UAV wireless network: A survey of reinforcement learning-based approaches. *IEEE Communications Surveys and Tutorials*, *25*(4), pp. 3038-3067. https://doi.org/10.1109/COMST.2023.3323344

Betalo, M.L., Ullah, I., Tesema, F.B., Wu, Z., Li, J. and Bai, X., 2025. Generative AI-Driven Multi-Agent DRL for task allocation in UAV-Assisted EMPD within 6G-Enabled SAGIN networks. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2025.3579780

Bi, R., Guo, D., Zhang, Y., Huang, R., Lin, L. and Xiong, J., 2023. Outsourced and privacy-preserving collaborative k-prototype clustering for mixed data via additive secret sharing. *IEEE Internet of Things Journal*, *10*(18), pp. 15810-15821. https://doi.org/10.1109/JIOT.2023.3266028

Cao, K., Chen, M., Karnouskos, S. and Hu, S., 2024. Reliability-aware personalized deployment of approximate computation IoT applications in serverless mobile edge computing. *IEEE Transactions on computer-aided design of integrated circuits and systems*, *44*(2), pp. 430-443. https://doi.org/10.1109/TCAD.2024.3437344.

Chafii, M., Naoumi, S., Alami, R., Almazrouei, E., Bennis, M. and Debbah, M., 2023. Emergent communication in multi-agent reinforcement learning for future wireless networks. *IEEE Internet of Things Magazine*, *6*(4), pp. 18-24. https://doi.org/10.1109/IOTM.001.2300102

Chen, X., Feng, D., Jiang, W., Luo, Q., Chen, G., and Sun, Y., 2025. QoE-driven multi-task offloading for semantic-aware edge computing systems. *IEEE Transactions on Mobile Computing*. Advance online publication. https://doi.org/10.1109/TNSE.2026.3662899

Chen, Y., Zhao, L., and Wang, X., 2026. Privacy-preserving federated averaging with adaptive communication. *ACM Transactions on Intelligent Systems and Technology*, 17(3), pp. 45–67. https://doi.org/10.1145/3623456

CRAWDAD., 2022. A Community Resource for Archiving Wireless Data at Dartmouth. https://crawdad.org/~crawdadarchive/all-bydate.html

Dai, Z., Zhang, Y., Zhang, W., Luo, X. and He, Z., 2022. A multi-agent collaborative environment learning method for UAV deployment and resource allocation. *IEEE Transactions on Signal and Information Processing over Networks*, *8*, pp. 120-130. https://doi.org/10.1109/TSIPN.2022.3150911

Deng, M., Liang, W., Qiu, B., and Xie, X., 2025. Federated deep reinforcement learning-based task offloading strategy with game theory in vehicular edge computing. *Physical Communication*, 65, P. 102802. https://doi.org/10.1016/j.phycom.2025.102802

Dritsas, E., Ramantas, K. and Verikoukis, C.V., 2024, September. A mobility-aware reinforcement learning proactive solution for state data migration in edge computing. In *2024 IEEE 29th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 1–6. IEEE. https://doi.org/10.1109/CAMAD62243.2024.10943072

Duan, Q., Huang, J., Hu, S., Deng, R., Lu, Z. and Yu, S., 2023. Combining federated learning and edge computing toward ubiquitous intelligence in 6G network: Challenges, recent advances, and future directions. *IEEE Communications Surveys and Tutorials*, *25*(4), pp. 2892-2950. https://doi.org/10.1109/COMST.2023.3316615

Feng, C., Yang, H.H., Hu, D., Zhao, Z., Quek, T.Q. and Min, G., 2022. Mobility-aware cluster federated learning in hierarchical wireless networks. *IEEE Transactions on Wireless Communications*, *21*(10), pp. 8441-8458. https://dx.doi.org/10.1109/TWC.2022.3166386

Glanois, C., Weng, P. , Zimmer, M., Li, D., Yang, T., Hao, J. and Liu, W., 2024. A survey on interpretable reinforcement learning. *Machine Learning*, *113*(8), pp. 5847-5890. https://doi.org/10.1007/s10994-024-06543-w

Gu, S., Yang, L., Du, Y., Chen, G., Walter, F., Wang, J. and Knoll, A., 2024. A review of safe reinforcement learning: Methods, theories, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(12), pp. 11216-11235. https://doi.org/10.1109/TPAMI.2024.3457538

Guo, Y., Xie, H., Wang, C. and Jia, X., 2021. Enabling privacy-preserving geographic range query in fog-enhanced IoT services. *IEEE Transactions on Dependable and Secure Computing*, *19*(5), pp. 3401-3416. https://doi.org/10.1109/TDSC.2021.3095933

Hamdi, A.M.A., Hussain, F.K. and Hussain, O.K., 2022. Task offloading in vehicular fog computing: State-of-the-art and open issues. *Future Generation Computer Systems*, *133*, pp. 201-212. https://doi.org/10.1016/j.future.2022.03.019

Hasselt, H. V., Hessel, M., and Aslanides, J., 2024. Deep reinforcement learning with double Q-learning revisited. *Journal of Machine Learning Research*, *25*(112), pp. 1–35. https://doi.org/10.5555/jmlr.v25.112

Hsieh, L.T., Liu, H., Guo, Y. and Gazda, R., 2023. Deep reinforcement learning-based task assignment for cooperative mobile edge computing. *IEEE Transactions on Mobile Computing*, *23*(4), pp. 3156-3171. https://doi.org/10.1109/TMC.2023.3270242.

Jiang, H., Dai, X., Xiao, Z. and Iyengar, A., 2022. Joint task offloading and resource allocation for energy-constrained mobile edge computing. *IEEE Transactions on Mobile Computing*, *22*(7), pp. 4000-4015. https://doi.org/10.1109/TMC.2022.3150432

Khot, V., Pai, S.S. and Rao, V.R., 2023, September. Deep Reinforcement Learning for Task Offloading in a Multi-Access Edge Computing Environment. In *2023 International Conference on Network,* P. 91. https://doi.org/10.1109/NMITCON58196.2023.10275998.

Lekharu, A., Samanta, A., Sur, A. and Patra, M., 2024. Content-aware caching at the mobile edge network using federated learning. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *9*(2), pp. 1166-1176. https://doi.org/10.1109/TETCI.2024.

Levine, S., Kumar, A., and Finn, C., 2025. Reinforcement learning: Foundations, advances, and frontiers. *Annual Review of Control, Robotics, and Autonomous Systems*, *8*(1), pp. 1–35. https://doi.org/10.1146/annurev-cras-080124-010321

Li, B., Ma, S., Deng, R., Choo, K.K.R. and Yang, J., 2022. Federated anomaly detection on system logs for the Internet of Things: A customizable and communication-efficient approach. *IEEE Transactions on Network and Service Management*, *19*(2), pp. 1705-1716. https://doi.org/10.1109/TNSM.2022.3152620

Li, C., Zhang, Y. and Luo, Y., 2022. A federated learning-based edge caching approach for mobile edge computing-enabled intelligent connected vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *24*(3), pp. 3360-3369. https://doi.org/10.1109/TITS.2022.3224395.

Li, T., Zhu, K., Luong, N.C., Niyato, D., Wu, Q., Zhang, Y. and Chen, B., 2022. Applications of multi-agent reinforcement learning in future internet: A comprehensive survey. *IEEE Communications Surveys and Tutorials*, *24*(2), pp. 1240-1279. https://doi.org/10.1109/COMST.2022.3160697

Liang, Z., Wang, B., Gu, Q., Osher, S. and Yao, Y., 2024. Differentially private federated learning with Laplacian smoothing. *Applied and Computational Harmonic Analysis*, *72*, P. 101660. https://doi.org/10.1016/j.acha.2024.101660.

Liu, H., and Sun, F., 2025. Stabilizing deep Q-networks with distributional reinforcement learning. *IEEE Transactions on Artificial Intelligence*, 5(1), pp. 22–34. https://doi.org/10.1109/TAI.2025.1239876

Liu, Y., Xue, E., Chai, W., Liu, Y., and Feng, F., 2026. Offloading and computational resources allocation of tasks with dependency in MEC environment based on deep reinforcement learning. *Ad Hoc Networks*, 154, P. 104136. https://doi.org/10.1016/j.adhoc.2026.104136

Loutfi, S.I., Shayea, I., Tureli, U., El-Saleh, A.A. and Tashan, W., 2024. An overview of mobility awareness with mobile edge computing over 6G network: Challenges and future research directions. *Results in Engineering*, *23*, P. 102601. https://doi.org/10.1016/j.rineng.2024.102601.

Lu, J., Liu, H., Zhang, Z., Wang, J., Goudos, S.K. and Wan, S., 2021. Toward fairness-aware time-sensitive asynchronous federated learning for critical energy infrastructure. *IEEE Transactions on Industrial Informatics*, *18*(5), pp. 3462-3472. https://dx.doi.org/10.1109/TII.2021.3117861

Ma, C., Li, A., Du, Y., Dong, H. and Yang, Y., 2024. Efficient and scalable reinforcement learning for large-scale network control. *Nature Machine Intelligence*, *6*(9), pp. 1006-1020. https://doi.org/10.1038/s42256-024-00879-7

Min, M., Liu, Z., Duan, J., Zhang, P. and Li, S., 2023. Safe-learning-based location-privacy-preserved task offloading in mobile edge computing. *Electronics*, *13*(1), P. 89. https://doi.org/10.3390/electronics13010089

Mitsis, G., Tsiropoulou, E.E. and Papavassiliou, S., 2022. Price and risk awareness for data offloading decision-making in edge computing systems. *IEEE Systems Journal*, *16*(4), pp. 6546-6557. https://doi.org/10.1109/JSYST.2022.3188997

Musa, S.S., Zennaro, M., Libsie, M. and Pietrosemoli, E., 2022. Mobility-aware proactive edge caching optimization scheme in information-centric IoV networks. *Sensors*, *22*(4), P. 1387. https://doi.org/10.3390/s22041387

Nahar, A., Sikarwar, H., Jain, S. and Das, D., 2023, May. Cachein: A secure distributed multi-layer mobility-assisted edge intelligence-based caching for Internet of vehicles. In *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pp. 437-446. IEEE. https://dx.doi.org/10.1109/CCGrid57682.2023.00048

Qian, Z., Feng, Y., Dai, C., Li, W. and Li, G., 2024. Mobility-aware proactive video caching based on asynchronous federated learning in mobile edge computing systems. *Applied Soft Computing*, *162*, P. 111795. https://doi.org/10.1016/j.asoc.2024.111795

Rui, L., Wang, S., Wang, Z., Xiong, A. and Liu, H., 2021. Dynamic service migration strategy based on mobility prediction in edge computing. *International Journal of Distributed Sensor Networks*, 17(2), pp. 1–12. SAGE Publications. https://doi.org/10.1177/1550147721993403

Serghiou, D., Khalily, M., Brown, T.W. and Tafazolli, R., 2022. Terahertz channel propagation phenomena, measurement techniques and modeling for 6G wireless communication applications: A survey, open challenges and future research directions. *IEEE Communications Surveys and Tutorials*, *24*(4), pp. 1957-1996. https://doi.org/10.1109/COMST.2022.3205505

Song, X., Chen, Q., Wang, S., and Song, T., 2025. Cross-domain resources optimization for hybrid edge computing networks: Federated DRL approach. *Digital Communications and Networks*, 11(6), pp. 1797–1808. https://doi.org/10.1016/j.dcan.2024.03.006.

SPECpower. 2008. Results Search tool. https://www.spec.org/power_ssj2008/results/.

Sutton, R. S., and Barto, A. G., 2018. *Reinforcement learning: An introduction* (2nd ed.). MIT Press. https://doi.org/10.7551/mitpress/11177.001.0001

Tahmid, T., Gates, M., Luszczek, P. and Schuman, C., 2024. Towards scalable and efficient spiking reinforcement learning for continuous control tasks. In *2024 International Conference on Neuromorphic Systems (ICONS)*, pp. 331-335, IEEE. https://doi.org/10.1109/ICONS62911.2024.00057

Tang, J., Qin, T., Xiang, Y., Zhou, Z. and Gu, J., 2022. Optimization search strategy for task offloading from collaborative edge computing. *IEEE Transactions on Services Computing*, *16*(3), pp. 2044-2058. https://doi.org/10.1109/TSC.2022.3203700

Tang, J., Wang, S., Yang, S., Xiang, Y. and Zhou, Z., 2024. Federated learning-assisted task offloading based on feature matching and caching in collaborative device-edge-cloud networks. *IEEE Transactions on Mobile Computing*, *23*(12), pp. 12061-12079. https://doi.org/10.1109/TMC.2024.3403851.

Ullah, I., 2025. Digital twin-empowered task offloading in VEC: Attention-augmented DRL for mobility-resilient optimization. *Computer Networks*, 245, P. 111818. https://doi.org/10.1016/j.comnet.2025.111818

Wang, G., Liang, R., Hu, J., 2021. Resource caching and task migration strategy of small cellular networks under mobile edge computing. *IEEE Transactions on Mobile Computing*, *20*(11), pp. 3137–3151. https://doi.org/10.1109/TMC.2020.2981897

Wang, W., Zhao, Y., Wu, Q., Fan, Q., Zhang, C. and Li, Z., 2022. Asynchronous federated learning-based mobility-aware caching in vehicular edge computing. In *2022 14th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1-5. https://dx.doi.org/10.1109/WCSP55476.2022.10039430

Wang, Y., Qian, Z., He, L., Yin, R. and Wu, C., 2024. Intelligent online computation offloading for wireless-powered mobile-edge computing. *IEEE Internet of Things Journal*, *11*(17), pp. 28960-28974. https://doi.org/10.1109/JIOT.2024.3406204

Wei, K., Li, J., Ding, M., Ma, C., Yang, H.H., Farokhi, F., Jin, S., Quek, T.Q. and Poor, H.V., 2020. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, *15*, pp. 3454-3469. https://doi.org/10.1109/TIFS.2020.2988575.

Wu, Q., Zhao, Y., Fan, Q., Fan, P. , Wang, J. and Zhang, C., 2022. Mobility-aware cooperative caching in vehicular edge computing based on asynchronous federated and deep reinforcement learning. *IEEE Journal of Selected Topics in Signal Processing*, *17*(1), pp. 66-81. https://doi.org/10.1109/JSTSP.2022.322127d1.

Wu, Z., Fang, H., Tang, J., and Yang, X., 2025. Lightweight adaptive PPO-AHP enhanced algorithm for task offloading in vehicular edge computing. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pp. 1–9. https://doi.org/10.1109/IJCNN.2025. 9568743

Xiao, Y., Xiao, L., Wan, K., Yang, H., Zhang, Y., Wu, Y. and Zhang, Y., 2022. Reinforcement learning based energy-efficient collaborative inference for mobile edge computing. *IEEE Transactions on Communications*, *71*(2), pp. 864-876. https://doi.org/10.1109/TCOMM.2022.3229033.

Xie, B. and Cui, H., 2025. Deep reinforcement learning-based dynamical task offloading for mobile edge computing: B. Xie, H. Cui. *The Journal of Supercomputing*, *81*(1), P. 35. https://doi.org/10.1007/s11227-024-06603-x.

Xie, B., Sun, Y., Zhou, S., Niu, Z., Xu, Y., Chen, J. and Gunduz, D., 2023, May. MOB-FL: Mobility-aware federated learning for intelligent connected vehicles. In *ICC 2023-IEEE International Conference on Communications*. pp. 3951-3957. https://doi.org/10.1109/ICC45041.2023.10279773.

Xu, Z., Kairouz, P. , McMahan, H. B., and Ramage, D., 2025. Federated optimization in heterogeneous networks: Advances and open problems. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(2), pp. 345–362. https://doi.org/10.1109/TNNLS.2025.1234567

Yamansavascilar, B., Baktir, A.C., Sonmez, C., Ozgovde, A. and Ersoy, C., 2022. Deepedge: A deep reinforcement learning based task orchestrator for edge computing. *IEEE Transactions on Network Science and Engineering*, *10*(1), pp. 538-552. https://doi.org/10.1109/TNSE.2022.3217311.

Yu, G., He, Y., Wu, J., Chen, Z. and Pan, J., 2023. Mobility-aware proactive edge caching for large files in the internet of vehicles. *IEEE Internet of Things Journal*, *10*(13), pp. 11293-11305. https://dx.doi.org/10.1109/JIOT.2023.3240423

Zhang, H., Ma, T., Chen, G., Ni, Y., Liu, T., and Zhou, H., 2025. Federated learning-driven computation offloading in space-air-ground integrated networks. In *Proceedings of the 2025 Seventeenth International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6. https://doi.org/10.1109/WCSP68525.2025.1010157

Zhang, J., Cheng, L., Liu, C., Zhao, Z. and Mao, Y., 2023. Cost-aware scheduling systems for real-time workflows in cloud: An approach based on genetic algorithm and deep reinforcement learning. *Expert Systems with Applications*, *234*, P. 120972. https://doi.org/10.1016/j.eswa.2023.120972

Zhang, P. , Gan, P. , Chang, L., Wen, W., Selvi, M. and Kibalya, G., 2022. DPRL: Task offloading strategy based on differential privacy and reinforcement learning in edge computing. *IEEE Access*, *10*, pp. 54002-54011. http://dx.doi.org/10.1109/ACCESS.2022.3175194.

Zhang, P. , Wang, E., Guizani, M., Liu, K., Wang, J., and Tan, L., 2025. Privacy-preserving task offloading in vehicular edge computing using federated multi-agent reinforcement learning. *IEEE Transactions on Vehicular Technology*, *74*(12), pp. 19642–19654. https://doi.org/10.1109/TVT.2025.3588204

Zhang, R., Xia, H., Chen, Z., Kang, Z., Wang, K. and Gao, W., 2024. Computation cost-driven offloading strategy based on reinforcement learning for consumer devices. *IEEE Transactions on Consumer Electronics*, *70*(1), pp. 4120-4131. https://doi.org/10.1109/TCE.2024.3355455

Zhang, Y., and Wang, J., 2026. Advances in deep Q-networks for complex decision-making. *Neural Computation*, *38*(4), pp. 789–812. https://doi.org/10.1162/neco_a_01567

Zhou, X., Bilal, M., Dou, R., Rodrigues, J.J., Zhao, Q., Dai, J. and Xu, X., 2023. Edge computation offloading with content caching in 6G-enabled IoV. *IEEE Transactions on Intelligent Transportation Systems*, *25*(3), pp. 2733-2747. https://doi.org/10.1109/TITS.2023.3239599

التعلم المعزز المتحد المدرك للتنقل والحفاظ على الخصوصية مع التعلم الآلي متعدد النماذج لذكاء الحافة في شبكات جي6/ جي5

نبيل عبد الوهاب عبد الرزاق بابان

قسم هندسة تقنيات الحاسبات, كلية النخبة الجامعة , بغداد, العراق

الخلاصة

إن نمو شبكات الجيل الخامس و الجيل السادس ، إلى جانب النمو السريع للحوسبة المتطورة ، يخلق حاجة قوية لطرق أكثر ذكاء وأكثر وعيا بالخصوصية للتعامل مع تفريغ المهام مع تحرك المستخدمين عبر الشبكة .لا تزال العديد من الأساليب الحالية تتعامل مع التنبؤ بالتنقل والتعلم الموحد والخصوصية التفاضلية كقطع منفصلة ، مما يؤدي غالبا إلى تأخيرات يمكن تجنبها واستخدام أعلى للطاقة وحماية أضعف للبيانات .تقدم هذه الورقة التعلم المعزز الموحد المدرك للتنقل ، وهو إطار مصمم لجمع هذه المكونات معا .وهو يدمج التعلم التعزيز العميق ، والتنبؤ التنقل باستخدام سلاسل ماركوف ، و غاوس دب لاتخاذ قرارات أفضل التفريغ عبر بيئات حافة متعددة المستويات .يستخدم تعلم التعزيز الموحد الواعي بالتنقل شبكة عميقة موحدة حيث كل عقدة حافة ارشاد محليا على بيانات التنقل المدرك ويضيف الخصوصية التفاضلية للضوضاء قبل المساهمة في النموذج العالمي .باستخدام برنامج ان س 3-و إيفوغسيم ، جنبا إلى جنب مع مجموعات البيانات الحقيقية ، والإطار يحقق 32 ٪ أقل زمن وصول ، 27 ٪ وفورات في الطاقة ، وحماية خصوصية قوية .(ε <1.0) ويبين تحليل باريتو التوازن بين أهداف الأداء ، وضبط طوبولوجيا مدركة يحسن النتائج في المناطق الحضرية والريفية ، وإعدادات المركبات .ويتماشى هذا النظام أيضا مع اللائحة العامة لحماية البيانات .وسيستكشف العمل المستقبلي نماذج تنقل الذاكرة طويلة المدى والشبكة العصبية للرسم البياني المكاني والزماني واختبار الأجهزة في الحلقة .

الكلمات المفتاحية: ذكاء الحافة، التعلم المعزز المتحد، توقع التنقل، الحفاظ على الخصوصية، تفريغ المهام.