A Reinforcement Learning–Enabled Digital Twin Framework for Mobile Network Infrastructure

Shahnaz Fatima[1]* , Mohammad Rafee Shaik[2], M.Ravi Sankar[3], K.Sathish Kumar[4], Tamalam Manogna[5], and Salar Mohammad[6]

[1] Department of Electronics and Communications and Engineering,
[2] Department of Artifitial Intelligence and Machine Learning,
[3] Department of Electronics and Communications Engineering,
Sasi Institute of Technology and Engineering, Tadepalligudem, A.P, India
[4] Department of Electrical and Electronics Engineering,
Vidya Jyothi Institute of technology, Telangana, India
[5] Department of Electronics and Communications and Engineering,
Andhra Loyola Institute of Engineering and Technology, Vijayawada, India
[6] Department of Data Science
Anurag University, Telangana, India
*Corresponding author: shahnaz1981fat@gmail.com

Abstract. Rising network complexity in next-generation systems underscores the shortcomings of traditional optimization methods in both modeling and algorithms. This paper introduces a digital twin for mobile network (DTMN) architecture tailored for 6G networks. To overcome the limitations of traditional optimization methods, the DTMN adopts a simulation–optimization framework. The network simulation engine provides feedback that validates optimizer decisions and is used to iteratively train ML-based optimizers. In practice, we implement a prototype system that leverages data-driven technologies for modelling network elements, environments and the BSs. Built upon the deployed DTMN prototype, the ML network optimizer provides ideal network configuration solutions. Experimental results show that our DTMN delivers solutions for optimizing networks in the real world.

Keywords: Digital Twin; mobile networks; reinforcement learning; network optimization.

1 Introduction

Mobile network optimization has traditionally relied on analytical modeling and algorithm design. In this paradigm, optimization problems are first abstracted into mathematical formulations, and appropriate techniques are applied to derive exact or approximate solutions. However, the increasing complexity of modern mobile networks has exposed two major limitations: difficulties in modeling and challenges in algorithmic scalability. Existing solutions often depend on engineering heuristics, and the complexity of next-generation networks and their

S. Mantena (ed.), *Proceedings of the Conference on Social and Sustainable Innovation in Technology & Engineering (SASI-ITE 2025)*, Advances in Intelligent Systems Research 199,
https://doi.org/10.2991/978-94-6463-940-7_3

Key Performance Indicators (KPIs) [1] further complicates the development of comprehensive models. On the algorithmic side, most optimization tasks are NP-hard, with computational demands that grow exponentially with system scale. Although traditional approaches attempt to reduce this cost through static network partitioning or simplified formulations, the resulting methods remain impractical in the vast, high-dimensional parameter spaces of modern networks.

To address these challenges, recent research has explored the use of digital twins in mobile networks. A digital twin is a virtual replica of a physical system that synchronizes with real-time data to reflect its structure, context, and behavior [2]. By leveraging this capability, digital twins enable continuous monitoring, simulation, and optimization, offering a promising pathway to overcome both modeling and algorithmic limitations.

In this work, we present a Digital Twin for Mobile Networks (DTMN) infrastructure designed for 6G. The prototype, developed using data-driven techniques, replicates the operation of mobiles, Base Station (BS) and the environment. An ML network optimizer integrated into the system enables adaptive and efficient network optimization. Experimental results demonstrate that the proposed DTMN effectively addresses the increasing complexity of mobile networks and delivers reliable optimization outcomes.

The remainder of this article is organized as follows. Section 2 presents the Configuration of the proposed DTMN System. Section 3 describes the individual components of DTMN. In Section 4 two representative wireless networking tasks based on ML network optimizer are analyzed. Section 5 provides evaluations on the prototype system . Section 6 concludes the article.

2 Configuration of DTMN

Digital twin [2], is a virtual imitation of real devices built through three key models: mirror, simulation, and integration. The mirror model focuses on monitoring and recording real-time and historical data of physical devices to create their digital replica. The simulation model enables the digital twin to mimic a physical device's behavior and test different configurations through scenario analysis. The integration model combines the physical and virtual products with their data, services, and connections, enabling two-way interaction and control of physical devices.

Figure 1 shows the integration model configuration of DTMN. Let at time step t, the models of mobiles and BSs are revised using data from physical devices after observing their behaviour. Next, the internal cycle, or simulation-optimization module, receives the data stream, where scenario analysis are performed. The optimizer iteratively works along simulation model to determine the best mobile network configurations, which the digital twin then applies to manage physical devices.Digital twins then put the configurations into action to manage the physical devices. The external cycle updates the mirror and simulation models of digital twins using feedback from real-world network performance at timestep t+1 .

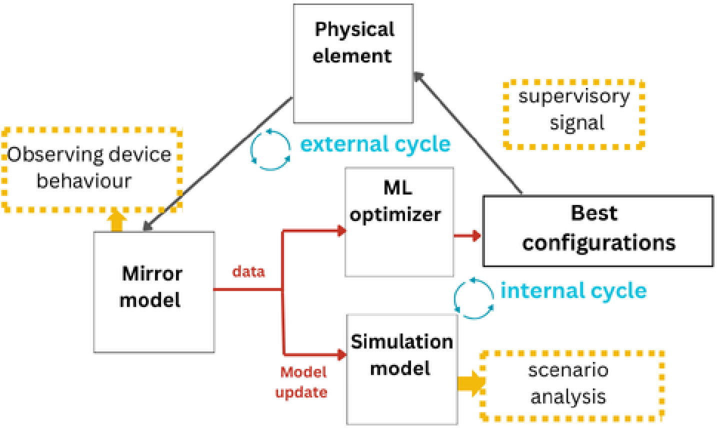

Fig. 1. Configuration of DTMN.

3 Twins of mobile network elements

Conventional network simulation platforms suffer from low fidelity and lack real-time evaluation due to simplified wireless communication models and the difficulty of deploying optimization algorithms. To address this, we adopt a digital twin approach—widely applied in aerospace, urban construction, and industrial operations—to mirror physical mobile network entities in virtual space. By collecting active data and modeling fundamental processes,the network is breakdown as mobiles, BSs and the environment, focusing on downlink transmission. Each element is virtualized using principles and empirical data, and together they form the DTMN system. This section details the digital twin design for each element.

3.1 Twins of Mobile Users

Mobiles follow different movement routes and generate varying traffic loads over time, but such data are typically not shared due to privacy concerns and the high expenses in data acquisition, resulting in the generative moving routes and communication traffic loads as alternatives.

For mobile users, the routes of outdoor movements are simulated using the path between the starting point and the target using the A^* method [4] as in [3] , while the locations of walkers and automobiles are determined by the Krauss method [5]. For indoor movements, the dynamics among mobiles, barriers, and arrival points are considered and jointly integrated [6]. For traffic load modeling, we employ a cross-scale tiered framework that incorporates clustering at the large-scale level and traffic source models at the fine-grained level.

3.2 Twins of BSs

BSs are central to mobile networks, supporting uplink and downlink communications through multiple access technologies. To build their digital twins, we configure them with realistic settings—GPS coordinates from real sites and communication attributes based on operator data. Core functions including multiple access, user association, scheduling, and mobility management are implemented.

When a user initiates a service, the Twin adapts communication attributes to meet QoS requirements, guided by the Shannon Equation. With many simultaneous users, efficient resource allocation becomes critical, requiring optimization algorithms. As baseline strategies, we use closest BS association and arbitrary sub-carrier access, while defaulting to fair power allocation if no optimization is applied.

The system operates in time steps of seconds, enabling dynamic reconfiguration of parameters and rapid testing of optimization solutions for resource allocation, offloading, energy efficiency, interference mitigation, and coverage enhancement. This bridges the gap between virtual digital twins and real mobile networks.

3.3 Twin of Wireless Environments

Accurate mobile network descriptions are vital for wireless system design and optimization, with channel state being the key factor. In real environments, transmitted signals attenuate through reflection, refraction, and diffraction, with Line-of-Sight (LoS) strongly influencing received strength. The path loss is

$$PL = L_f + L_l + L_s(t), \tag{1}$$

where L_f is free-space loss, L_l is shadowing from large objects (large-scale fading), and $L_s(t)$ is small-scale fading, often modeled by Rayleigh, Rician, or Nakagami distributions.

Channel models include stochastic (efficient but low fidelity) and deterministic (accurate but costly). Finite Difference Time Domain method resolves EM equations at fine scales, while ray tracing models LoS, reflection, diffraction, and scattering more efficiently. For digital twin environments, we combine real-world data from OpenStreetMap with ray tracing to model base station path loss.

4 ML network optimizer

Much of the decision-making in mobile networks is expressed as a numerical optimization problem but traditional methods scale poorly and struggle with real-time, large-scale scenarios. They also require clear, simplified objective functions, limiting applicability. In contrast,Reinforcement learning (RL, a machine learning framework, imposes no strict objective form, leverages generalization from diverse training samples, and offers fast inference, enabling efficient resource allocation for large-scale problems [7]. In this study, we focus on two representative wireless networking tasks where RL shows particular promise: wireless network resource allocation (RA) [8] and the base station sleep mechanism [9]. Both are examined using synthetic traffic data that follows realistic diurnal variations and statistical properties reported in [10]. The following subsections describe the application of RL to these two tasks in detail.

4.1 RL Based Dynamical Wireless Network Resource Allocation

RA in wireless Networks is difficult due to non convexity and high dimentinality in optimization. Dynamic wireless RA does optimal user association, choice of resource block, and power allocation for improved service quality. With many mobile users and numerous base stations, we model the problem as a Constrained Markov Decision Process (CMDP) [11] as shown in Figure 2 with the reward as total data rate with included cost as demand satisfaction and constraints as one-to-one user–base station/block mapping and power limits.

To address complex multi-step and single-step constraints, we introduce single-step demand, enabling sequential decision-making and constraint decoupling. A safe RL method Primal–Primal–Dual Policy Optimization (P3O) [12] solves the CMDP: user association and block selection are framed as bipartite graph matching with edge weights from neural network–predicted rewards and costs, solved via the Hungarian algorithm [13]. Power allocation is handled by a policy network trained with an RL method similar to Deep Deterministic Policy Gradient (DDPG) [14] but with an explicit computational action-value form. This framework maximizes throughput while ensuring demand satisfaction under strict constraints.

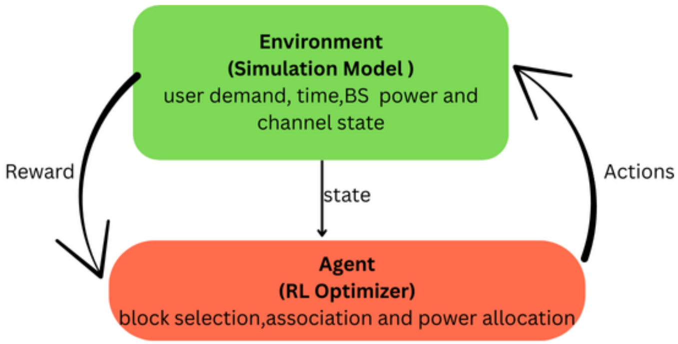

Fig. 2. RL based interaction in Resource allocation problem.

4.2 RL Based Sleep Mechanism

BSs are key components of mobile networks, but since user demand varies periodically, shutting down base stations or their internal cells can save energy.

Cells are organized in clusters, where RL agents determine if they remain operational, dormant, or powered down during a time instant, with the BS switching on if atleast one cell is active. load is distributed in the cluster proportional to capacity with Markov Decision Process [15], as summarized in Table 1.

Table 1. RL based interaction in Sleep mechanism problem

Component	Definition
State	Timestep, recent traffic demand, device parameters, and status (cells/BSs)
Action	Vector to set the status of cell.
Reward	Negative total power consumption (RRUs, BBUs, air conditioners) plus switching costs.
Transition	Traffic changes according to demand data; control status updates based on chosen actions.

Since direct control of all cells is infeasible, we assign an agent to each cell, sharing policy/action-value networks and experiences. Each agent's reward is the sum of its own and its base station's power cost. Policy network inputs include timestep, recent traffic, device parameters, and control states, while the action-value network uses two masks: (i) whether demand can be met if a cell sleeps/closes, and (ii) whether the base station can close if the cell closes. These masks capture interdependencies between agents. The networks are trained using a method similar to Mean Field RL, but with masks providing more accurate coordination than mean actions.

5 Evaluation of DTMN on the Prototype System

To assess the performance of DTMN, a prototype is developed that models the intricate dynamics of mobiles and BSs while continuously tracking their states.

5.1 System Deployment

For demonstration, we use a 3×3 km^2 area with mobiles, lanes and 170 BSs, along with wireless transmission parameters as summarized in Table 2.

The resource allocation is modeled as a three-step process. First, a mobile user accesses the available base stations in sorted order. Next, the eligibility of each BS is verified based on its available resources and allocated to the user accordingly. User–base station associations are formed through iterative steps, with base stations, users, their attributes, and the ML optimizer comprising the prototype system. During optimization, the network updates states and rewards from the previous step and feeds them to the optimizer, which selects the next action. Each episode contains fixed-length steps, with states and rewards input at the beginning of each step and new rewards computed by the simulator.

Table 2. System Parameters of wireless network prototype

Parameter	**Value**
Area	$3 \times 3\, km^2$
Users count	14,000
Lanes count	1,500
BSs count	170
Channel number	45
Noise density	$-174\, dBm/Hz$
Bandwidth	$1.8 \times 10^5\, Hz$
Transmit power	Outdoor: 30 dBm, Indoor: 24 dBm

5.2 Experimental Analysis

We evaluate the effectiveness of the proposed RL algorithm by formulating the problem settings for dynamic wireless RA and sleep mechanisms, and compare the heuristic-based and RL-based approaches.

Dynamical Wireless Network RA We consider that each base station is equipped with 170 resource blocks. The users are assumed to move along straight-line routes at uniform speeds, with their starting and ending points randomly generated. Sampling is done for each users total demand at equal intervals over 20 timesteps. The agent trained is responsible for allocating resources to users while ensuring that the Satisfaction Ratio (SR), defined as the fraction of users whose achieved data rate meets or exceeds their required data rate, remains above 94%

For comparison, two benchmark strategies are considered: the equal allocation method, in which the total demand is distributed evenly across all steps, and the data rate maximization method, which allocates resources solely to maximize the total throughput while disregarding individual user demands. The BSs aggregate acheived data rate delivered to all users and the SR is shown in Figure 3. Our RL based resource allocation acheives a data rate of 4.35e5 with SR of 94.28%, which is good in terms of both performance indices compared to the two benchmark methods.

Sleep Mechanism The RL method, which involves periodic regulation of cell status by the agent to conserve power, is evaluated against two benchmark approaches. The first does not employ any heuristic, while the second, the algorithm proposed in [16], selects the minimum number of cells within a cluster based on their loads. The experimental results of a weeks traffic load and power consumption at BSs are demonstrated in Figure 4. Our method consumes the least power and better fits the traffic flow trend.

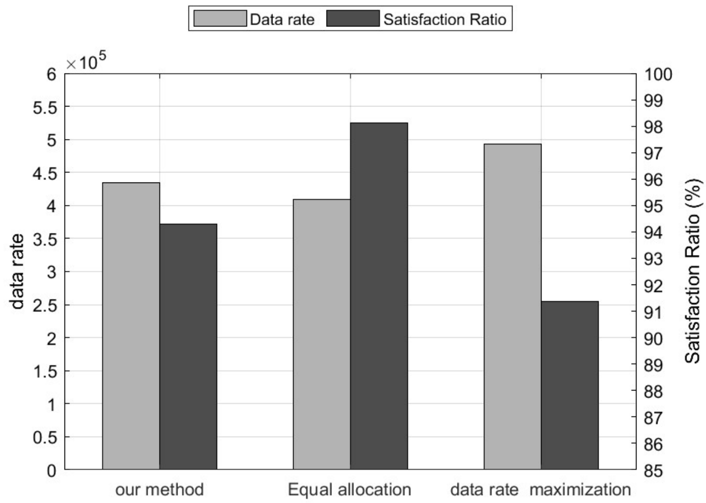

Fig. 3. Aggregate data rate and SR for different RA schemes .

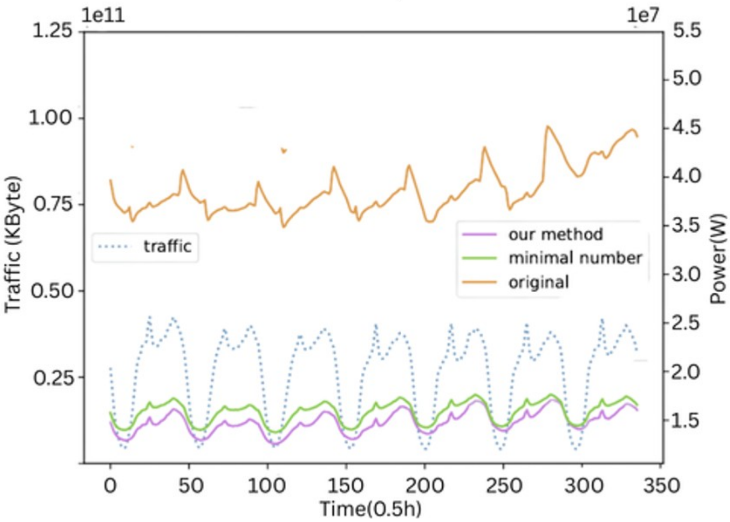

Fig. 4. weeks traffic load and power consumption at BSs.

6 Conclusion

The increasing complexity of next-generation networks uncovers the limitations of conventional optimization methods in both modeling and algorithmic design. This article introduces a 6G DTMN architecture composed of internal and external cycles. The internal cycle consists of an ML optimizer which operates repeatedly on the simulation module and refines decisions using feedback. Such a simulation–optimization paradigm can effectively address the modeling and algorithmic limitations of conventional methods. Following the deployment of a network configuration in the real system, the external cycle updates both the mirror and simulation models according to the observed network performance. In practice, we develop a network digital twin prototype that accurately represents mobile network elements (e.g., users and base stations), wireless environments, and overall network performance. Building on the DTMN prototype, an ML optimizer was developed. The experimental findings indicate that the proposed DTMN infrastructure can adapt to complex mobile network environments and offers effective solutions for network optimization.

References

1. Fatima, S., Kondamuri, S. (2025). Machine learning in future wireless communications: Innovations, applications,and open challenges. Wireless Personal Communications, 140(2), 1269–1313.https://doi.org/10.1007/s11277-025-11770
2. Grieves M (2022) Intelligent digital twins and the development and management of complex systems. Digital Twin 2:8. https://doi.org/10.12688/DIGITALTWIN.17574.1
3. Yuan Y, Ding J, Wang H, et al (2022) Activity Trajectory Generation via Modeling Spatiotemporal Dynamics. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 22:4752–4762. https://doi.org/10.1145/3534678.3542671
4. Wang H, Qi X, Lou S, et al (2021) An Efficient and Robust Improved A* Algorithm for Path Planning. Symmetry 2021, Vol 13, Page 2213 13:2213. https://doi.org/10.3390/SYM13112213
5. Adamidis FK, Mantouka EG, Vlahogianni EI (2020) Effects of controlling aggressive driving behavior on network-wide traffic flow and emissions. International Journal of Transportation Science and Technology 9:263–276. https://doi.org/10.1016/J.IJTST.2020.05.003
6. Yan D, Ding G, Huang K, et al (2024) Enhanced Crowd Dynamics Simulation with Deep Learning and Improved Social Force Model. Electronics 2024, Vol 13, Page 934 13:934. https://doi.org/10.3390/ELECTRONICS13050934
7. Pecioski D, Gavriloski V, Domazetovska S, Ignjatovska A (2023) An overview of reinforcement learning techniques. 12th Mediterranean Conference on Embedded Computing, MECO 2023. https://doi.org/10.1109/MECO58584.2023.10155066
8. Liang H, Zhang W (2021) A Survey and Taxonomy of Resource Allocation Methods in Wireless Networks. Journal of Communications and Information Networks 6:372–384. https://doi.org/10.23919/JCIN.2021.9663102
9. Yu Y, Kong F, Chen D (2018) Research on Base Station Sleeping Mechanism of User Connections in Ultra Dense Network. Wirel Pers Commun 102:79–93. https://doi.org/10.1007/S11277-018-5828-9

10. Wang X, Wang Z, Yang K, et al (2024) A Survey on Deep Learning for Cellular Traffic Prediction. Intelligent Computing 3:. https://doi.org/10.34133/ICOMPUTING.0054
11. Singh R, Gupta A, Shroff NB (2023) Learning in Constrained Markov Decision Processes. IEEE Trans Control Netw Syst 10:441–453. https://doi.org/10.1109/TCNS.2022.3203361
12. Mu C, Liu S, Lu M, et al (2024) Autonomous spacecraft collision avoidance with a variable number of space debris based on safe reinforcement learning. Aerosp Sci Technol 149:109131. https://doi.org/10.1016/J.AST.2024.109131
13. Li Y, Miao N, Ma L, et al (2023) Transformer for object detection: Review and benchmark. Eng Appl Artif Intell 126:. https://doi.org/10.1016/j.engappai.2023.107021
14. Tan H (2021) Reinforcement Learning with Deep Deterministic Policy Gradient. Proceedings - 2021 International Conference on Artificial Intelligence, Big Data and Algorithms, CAIBDA 2021 82–85. https://doi.org/10.1109/CAIBDA53561.2021.00025
15. Wang X, Yang Z, Chen G, Liu Y (2023) A Reinforcement Learning Method of Solving Markov Decision Processes: An Adaptive Exploration Model Based on Temporal Difference Error. Electronics 2023, Vol 12, Page 4176 12:4176. https://doi.org/10.3390/ELECTRONICS12194176
16. Peng C, Lee SB, Lu S, et al (2011) Traffic-driven power saving in operational 3G cellular networks. Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM 121–132