Slice Resource Management with MADDPG-Based Traffic Classification in 5G/B5G Networks

Chia-Yu, Yan
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
pooli0911@networklab.csie.ncu.edu.tw

Chih-Jou, Tai
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
taijustina@networklab.csie.ncu.edu.tw

Li-Der, Chou
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
cld@csie.ncu.edu.tw

***Abstract*—To address the limitations of 5G/B5G network slicing approaches in meeting diverse and high-reliability service demands, this paper proposes the Slice Resource Management based on Traffic Classification and Multi-agent Deep Deterministic Policy Gradient (SRM-TCM) framework. The SRM-TCM framework leverages Multi-agent Deep Deterministic Policy Gradient (MADDPG) for centralized training and dynamic resource scheduling, enabling precise traffic classification and adaptive slice management. This approach enhances resource utilization and ensures high Quality of Service (QoS), outperforming traditional methods that lack traffic-specific resource optimization.**

Keywords—B5G, Network Slicing, Reinforcement Learning, Resource Management, QoS

I. Introduction

The advent of 5G/Beyond 5G (5G/B5G) networks has revolutionized modern communication systems, offering unprecedented opportunities for innovation across various sectors. However, as the demand for high-speed, low-latency, and reliable services grows, effectively managing the diverse and dynamic requirements of network traffic has become a critical challenge. Network slicing, which allows the creation of virtualized slices tailored to specific applications, plays a pivotal role in addressing this issue. Each slice can be customized to meet the unique requirements of applications, such as low latency for real-time systems or high reliability for mission-critical services. Despite its potential, the dynamic allocation and scheduling of resources within slices remain a significant technical hurdle due to the highly variable nature of traffic demands.

Existing resource allocation methods, including Markov chains[1], big data analysis[2], and queuing theory[3], often fall short in adapting to the dynamic and unpredictable environment of 5G networks. These methods tend to either underutilize resources or fail to meet service-level agreements (SLAs), resulting in suboptimal Quality of Service (QoS)[4]. Furthermore, traditional slicing approaches typically treat traffic as homogeneous, leading to inefficient resource utilization and imbalanced service provisioning. These limitations underline the need for intelligent and adaptive solutions that can dynamically optimize resource allocation in response to real-time traffic demands.

To address these challenges, this paper proposes the Slice Resource Management based on Traffic Classification and Multi-Agent Deep Deterministic Policy Gradient (SRM-TCM) framework. Unlike conventional methods, the SRM-TCM framework integrates advanced artificial intelligence techniques to classify traffic patterns and optimize resource allocation dynamically. It employs Multi-Agent Deep Deterministic Policy Gradient (MADDPG) for centralized training and resource scheduling, enabling intelligent decision-making to enhance both resource utilization and QoS[5][6].

By combining traffic classification with intelligent resource management, the SRM-TCM framework represents a significant advancement in B5G network slicing. It not only enhances the adaptability and efficiency of resource allocation but also provides a scalable solution to meet the ever-growing demands of next-generation networks.

II. Background and Related Works

This section introduces the research background and basic knowledge of this paper, including 5G/B5G microservices, network slicing and reinforcement learning.

A. 5G/B5G Microservices

The microservices architecture decomposes complex systems into smaller, independent components, allowing for easier scalability, maintenance, and faulty isolation. In the context of 5G/B5G networks, this approach reduces system coupling, enhances reliability, and enables autonomous service management. These features make microservices an essential tool for improving network scalability and flexibility in dynamic environments.

B. Network Slicing

Network slicing allows the creation of virtualized networks tailored to diverse 5G application requirements[7], such as low latency, high bandwidth, and enhanced security. This segmentation supports key 5G scenarios defined by ITU[8], including Ultra-Reliable Low Latency Communications (uRLLC), Massive Machine Type Communications (mMTC), and Enhanced Mobile Broadband (eMBB). By enabling dynamic and flexible resource allocation, network slicing significantly improves network efficiency and ensures consistent Quality of Service (QoS) across applications[9][10].

C. Reinforcement Learning and MADDPG

Reinforcement Learning (RL) facilitates intelligent decision-making in dynamic environments[11] by enabling systems to learn and adapt autonomously. In this context, MADDPG serves as a robust approach for multi-agent systems, combining centralized training with decentralized execution. MADDPG enhances resource scheduling by enabling agents to predict and adapt to each other's behaviors, optimizing strategies based on real-time observations. This makes it particularly well-suited for complex tasks like resource allocation in network slicing, where multiple agents must cooperate to balance traffic demands and QoS requirements.

D. Related Works

Previous studies have leveraged RL and AI-based methods for 5G network slicing and resource management. However, many approaches either lack traffic classification or rely on static resource allocation strategies, leading to inefficiencies and suboptimal QoS. Compared to these methods, the proposed SRM-TCM framework integrates traffic classification with MADDPG to dynamically allocate resources, significantly enhancing network performance while addressing diverse application demands[12].

III. System Architecture

In this section, the proposed SRM-TCM will be introduced, as well as the strategies used when intensive learning is taking place. It includes an introduction to the system architecture, environment definition, state space, action space, reward function and replay buffer.

A. System Architecture

SRM-TCM consists of two main components: the AI Traffic Classifier and the MADDPG-based Resource Management Mechanism[13]. The AI Traffic Classifier[14][15] analyzes network packets to classify traffic into Light, Hybrid, and Heavyweight types, corresponding to 5G/B5G slices A, B, and C, enabling more accurate and efficient resource allocation. The MADDPG-based Resource Management Mechanism dynamically adjusts bandwidth allocation through its Actor, optimizing strategies based on real-time network load to enhance resource utilization and user experience. The bandwidth requirements and overall system architecture are depicted in Figures 1 and 2.

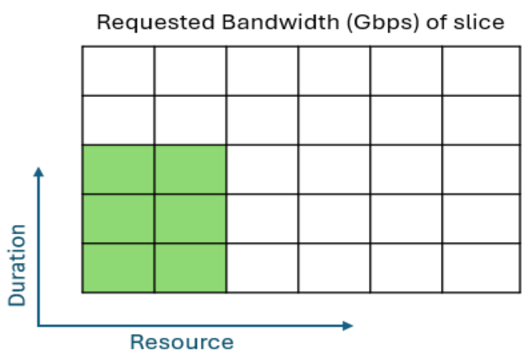

Fig. 1. Resource Allocation for Slice Requests

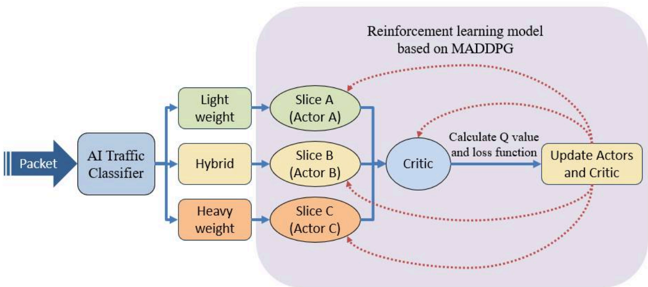

Fig. 2. SRM-TCM framework

B. Environment Definition

The proposed SRM-TCM framework employs a modified MADDPG model to enable multiple agents to collaborate and share experiences[16] in the learning environment. Unlike traditional MADDPG, which assigns a separate Critic for each Actor, this framework uses a centralized Single-Critic approach to evaluate all Actors, reducing computational complexity and improving training efficiency. The centralized Critic evaluates Q-values for different states, while Actors dynamically manage resource slices based on policy functions. Figure 3 provides the flowchart of reinforcement learning for resource management. This setup ensures better coordination among agents and efficient resource allocation.

To encourage exploration during the early training stages, an ε-Greedy strategy[17] is employed, gradually shifting from exploration to exploitation as training stabilizes. Actors allocate resources by selecting actions that optimize the overall system performance, balancing resource availability and slice demands.

The state space includes information about current idle resources, requested resources, and pending demands for each slice, ensuring a comprehensive understanding of the network environment. The action space allows Actors to decide whether to accept or reject slice requests, request additional resources, or release unused ones, guided by a reward function. The reward function encourages optimal resource utilization by rewarding successful slice requests and penalizing idle resources.

A replay buffer is used to store experiences, enabling random sampling during training to reduce correlation between consecutive samples. This mechanism enhances training efficiency and prevents overfitting, ensuring robust performance in dynamic network environments[18].

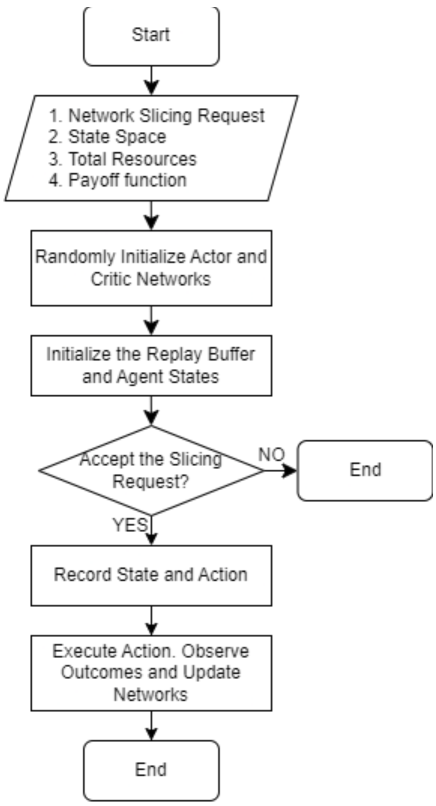

Fig. 3. Reinforcement Learning Flowchart for the Proposed SRM-TCM Mechanism in Network Slicing Resource Management

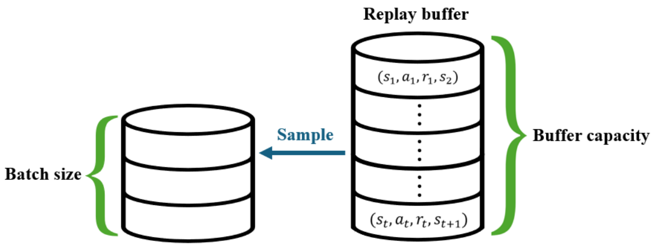

Fig. 4. Experience Replay Buffer Sampling Diagram

IV. Experiments and Discussions

This section presents the results and discussion of the experiments, including the training process, results and performance evaluation of the experiments, and the evaluation of the trained models in a simulated environment. To visualize the slice usage and performance metrics, Grafana[19] and Prometheus[20] were used as tools.

A. *Performance Evaluation of Slice Resource Allocation*

The performance of the proposed SRM-TCM mechanism was evaluated through simulations under two scenarios: low-resource requests with 150 packets per unit time and high-resource requests with 240 packets per unit time. Key performance metrics included idle Global Session Manager (GSM) resources, and average rewards. Results were compared against DQN and random strategies, and the experimental parameters are summarized in Table 1.

TABLE I. Experimental Parameters Table

Critic Learning Rate	1×10^{-3}
Actor Learning Rate	1×10^{-3}
Buffer Capacity	1×10^{4}
Batch Size	64
Gamma (γ)	0.95
System Time	10

1) *Reward Variations During Training*

The SRM-TCM mechanism demonstrated superior reward trends. In low-resource conditions, SRM-TCM achieved an average reward of 0.968, slightly outperforming DQN with 0.947 and significantly surpassing the random strategy with 0.765. Under high-resource conditions, SRM-TCM achieved an average reward of 0.941, exceeding DQN's 0.883 and the random strategy's 0.531. Figures 5 and 6 illustrate these trends.

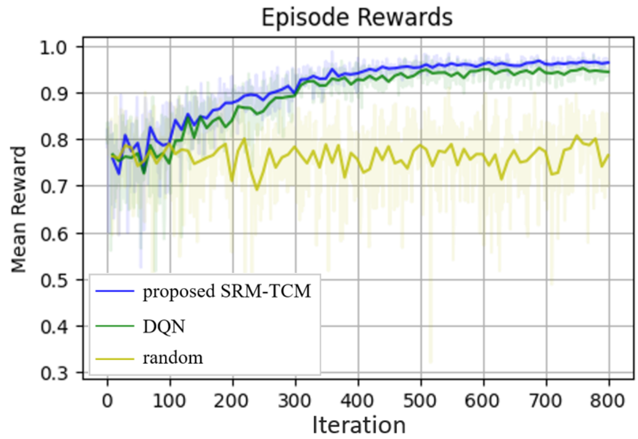

Fig. 5. Average Reward of Training Process with Low Resource Requests

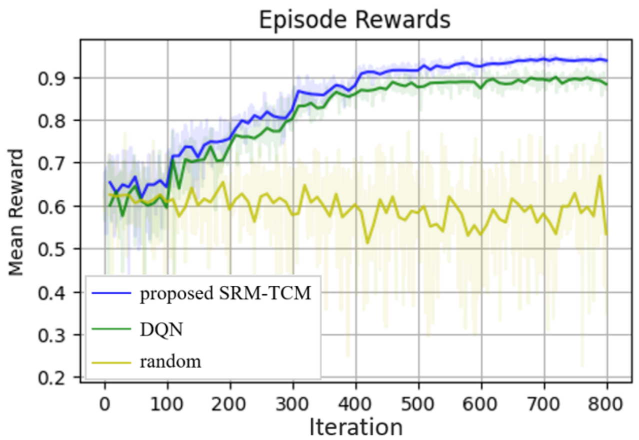

Fig. 6. Average Reward of Training Process with High Resource Requests

2) *Comparison of Idle GSM Resources*

SRM-TCM minimized resource wastage effectively across all scenarios. In low-resource conditions, idle resources averaged 0.91 for SRM-TCM, 0.52 for DQN, and 8.92 for the random strategy. Under high-resource conditions, idle resources were further reduced to 0.62 for SRM-TCM, compared to 0.89 for DQN and 17.09 for the random strategy. Figures 7 and 8 illustrate these results.

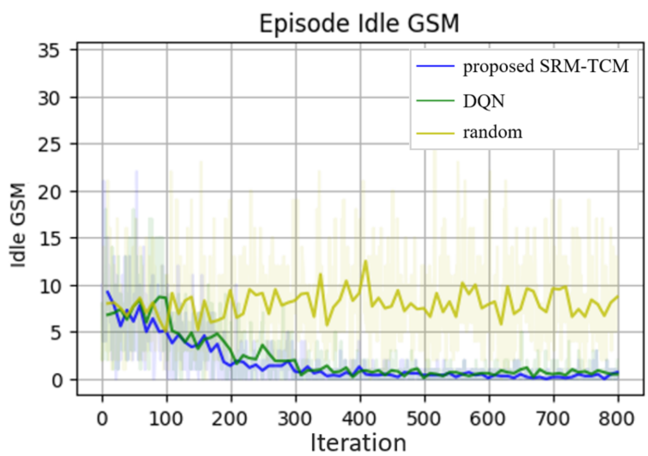

Fig. 7. Number of Idle GSMs under Low Resource Request Scenario

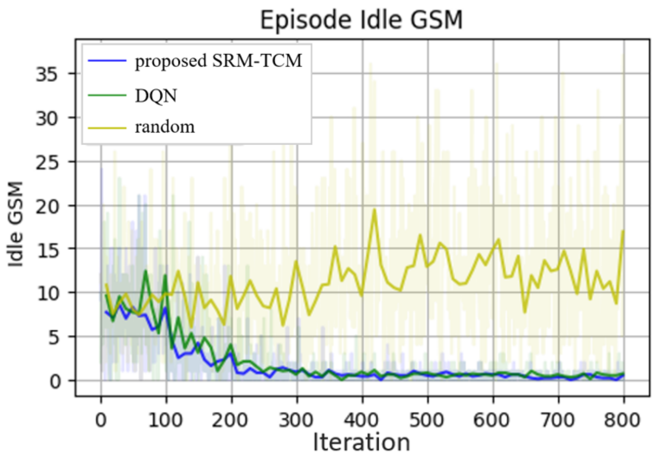

Fig. 8. Number of Idle GSMs under High Resource Request Scenario

B. *Monitor and Visualize Slice Usage*

Prometheus and Grafana were used for monitoring and visualizing resource usage and performance metrics. These tools provided real-time insights into resource allocation efficiency aduring experiments.

1) *Usage of Slice Resources*

Figure 11 shows the cumulative resource usage across time intervals, highlighting the total number of slice resources utilized during the experiment. It provides an overview of how effectively the system allocated resources to meet the demand, with higher utilization indicating better resource efficiency. The results suggest that SRM-TCM efficiently manages the available resources, minimizing underutilization.

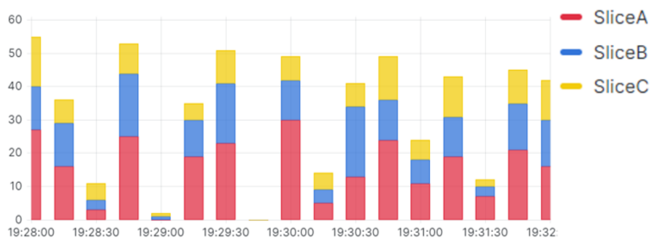

Fig. 9. Total number of slice resources used

2) *Number of Unused GSM-Requested Resources*

Figure 12 presents the number of unused GSM-requested resources over time. It highlights how well the SRM-TCM mechanism anticipates future resource needs, keeping idle resources to a minimum. The model maintains a near-zero idle resource level, ensuring optimal resource allocation and reducing waste, which is a key metric for improving system efficiency.

Fig. 10. Number of unused resources that GSM assigns to slice

V. CONCLUSIONS

The SRM-TCM framework effectively addresses the challenges of dynamic resource allocation in 5G/B5G networks by combining MADDPG-based traffic classification with adaptive slice management. Experimental results show that SRM-TCM outperforms traditional methods, improving resource utilization and minimizing idle resources. This highlights its potential to optimize QoS in complex network environments. Future work will focus on integrating advanced AI techniques and expanding its application to diverse use cases.

ACKNOWLEDGEMENT

This work was supported in part by the National Science and Technology Council (NSTC), Taiwan, under Grant no. 111-2221-E-008-062-MY2, 111-2221-E-008-063-MY3 and 113-2221-E-008-110-MY3.

REFERENCES

[1] H. Fourati, R. Maaloul and L. Chaari. A Survey of 5G Network Systems: Challenges and Machine Learning Approaches. International Journal of Machine Learning and Cybernetics, 2021.

[2] J. Kaur, M. A. Khan, M. Iftikhar, M. Imran and Q. E. U. Haq. Machine Learning Techniques for 5G and Beyond. IEEE Access, vol.9, 2021.

[3] B. Han, V. Sciancalepore, X. Costa-Pérez, D. Feng and H. D. Schotten. Multiservice-Based Network Slicing Orchestration with Impatient Tenants. IEEE Transactions on Wireless Communications, 2020.

[4] S. Saibharath, S. Mishra and C. Hota. Joint QoS and Energy-Efficient Resource Allocation and Scheduling in 5G Network Slicing. Computer Communications, 2023.

[5] C. C. Liu and L. D. Chou. B5G Network Slice Management via Staged Reinforcement Learning. IEEE Access, vol. 11, 2023.

[6] R. Lowe, Y. Wu, A. Tamar, J. Harb, P.Abbeel and I. Mordatch. Multi-Agent Actor- Critic for Mixed Cooperative-Competitive Environments. Arxiv Computer Science, 2017.

[7] D. C. Li, H. H. Tu and L. D. Chou. Cross-layer Detection and Defense Mechanism against DDoS and DRDoS Attacks in Software-defined Networks using P4 Switches. Computers and Electrical Engineering, vol.118, 2024.

[8] F. H. Tseng, Y. M. Jheng, L. D. Chou, H. C. Chao and V. C.M. Leung. Link-Aware Virtual Machine Placement for Cloud Services based on Service-Oriented Architecture. IEEE Transactions on Cloud Computing, 2020/

[9] W. Patel and P. Tripathy. eMBB, URLLC, and mMTC 5G Wireless Network Slicing: A Communication-Theoretic View. 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), 2023.

[10] D. C. Li, P. H. Chen and L. D. Chou. GAP4NSH: A Genetic Service Function Chaining with Network Service Header for P4-based Software-defined Networks. Journal of Supercomputing, vol.79, no. 10, 2023.

[11] D. C. Li, M. R. Mauluna and L. D. Chou. NNSplit-SØREN: Supporting the Model Implementation of Large Neural Networks in a Programmable Data Plane. Computer Networks, vol. 222, 2023.

[12] B. Zhang, V. H. Nguyen, T. H. Dinh, Q.-V. Pham and D. N. Nguyen. DDPG-E2E: A Novel Policy Gradient Approach for End-to-End Communication Systems. arxiv Computer Science, 2024.

[13] F. H. Tseng, X. J. Chen, L. D. Chou, H. C. Chao and S. P. Chen. Support vector machine approach for virtual machine migration in cloud data center. Multimeida Tools and Applications, 2015.

[14] M. K. N. Immadisetti, M. A P and S. M.. Automate Allocation of Secure Slice in Future Mobile Networks using Machine Learning. International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021.

[15] C. C. Liu, Y. Chang, C. W. Tseng, Y. T. Yang, M. S. Lai and L. D. Chou. SVM-based Classification Mechanism and its Application in SDN Networks. Proceedings of 2018 10th International Conference on Communication Software and Networks (ICCSN 2018), 2018.

[16] C. Park, W. J. Yun, J. P. Kim, T. K. Rodrigues, S. Park, S. Jung and J. Kim. Quantum Multiagent Actor–Critic Networks for Cooperative

Mobile Access in Multi-UAV Systems. IEEE Internet of Things Journal, vol.10, 2023.

[17] S. E. Jaghaoui, A. K. Elmaid and A. B. Lmah. Enhancing the Traveling Salesman Problem Solutions with Reinforcement Learning: A Variant Exploration-Exploitation Approach Beyond ε-Greedy. 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA), 2023.

[18] M. Nauman, M. Bortkiewicz, M. Ostaszewski, P. Miłoś, T. Trzciński and M. Cygan. Overestimation, Overfitting, and Plasticity in Actor-Critic: the Bitter Lesson of Reinforcement Learning. arxiv Computer Science, 2024.

[19] "Grafana,"[Online].Available:https://github.com/grafana/grafana.

[20] "Promethues,"[Online].Available:https://github.com/prometheus/prometheus.