Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 13, Issue. 1, January 2024, pg.82 – 87

SliceOptiAI: Smart Resource Allocation for Seamless Network Slicing

Dr. Alex Mathew
Department of Cybersecurity & Bethany College, USA
amathew@bethanywv.edu

DOI: https://doi.org/10.47760/ijcsmc.2024.v13i01.006

Abstract— An AI-driven model for resource allocation in network slicing is examined in this research paper. The model's algorithm comprises three stages, each with its specific algorithm. Resource allocation begins with reservation, where the controller reserves minimum resources for each slice. The second stage is autonomous radio resource management, mainly focusing on AI model training and decision engines. The last stage is physical resource allocation, where resources are distributed to the slices and users. Simulations on MATLAB software indicated this model to be effective in enhancing resource allocation.

Keywords— Network slicing, resource allocation, AI model, algorithm, communication network, deep reinforcement learning, radio resource management

I. INTRODUCTION

The current trends in information technology are changing the paradigms and approaches deployed in networking. The recent milestones in information technology include the development of fifth-generation technology, the emergence of diverse services, the explosive growth of data-intensive applications, and the adoption of the Internet of Things, which have changed the paradigm shift in how communication networks are designed and managed (Debbabi 162748). The existing communication networks are prompted to deal with the resulting wide variety of quality-of-service demands. Network slicing has emerged to meet the diverse demands of terminal devices using the existing communication networks. Network slicing technology mainly focuses on dividing the traditional physical network into several dedicated, virtualized, customed, and isolated logical networks, such that each portion is dedicated to fulfilling the requirements of a specific service (Khan et al. 36010). According to Wang et al., the major scenarios for network slicing include enhanced mobile broadband, ultra-reliable, low-latency communication, and massive machine-type communication services (p. (12505). Each scenario puts different pressure on network resources, determining the services they can provide. For instance, enhanced mobile broadband services can adequately serve high-data-rate applications. The efficiency of each network slice largely depends on the resources allocated. Network slicing relies on network function virtualization and software-defined networking to manage and allocate resources (Rakkiannan et al., 640). Integration of artificial intelligence is proposed to have the potential to optimize network slicing by facilitating efficient resource allocation. Although network slicing has unleashed a new era of flexibility and performance in communication networking, the full potential of this technology can be realized by deploying AI-driven resource allocation (Shahjalal et al. 345). Through advanced machine learning techniques, AI can analyze real-time network data, predict future demands, and dynamically allocate resources, thereby leveraging the adaptability and responsiveness of communication infrastructure. This research delves into integrating AI-driven approaches to optimize resource allocation within network slicing.

II. PROPOSED METHODOLOGY BLOCK DIAGRAM

Resource allocation in network slicing is a pivotal element for the success of network slicing. Depending on their demands and requirements, various processes allocate resources to different network slices. The following diagram highlights the major steps in AI-driven resource allocation in network slicing.

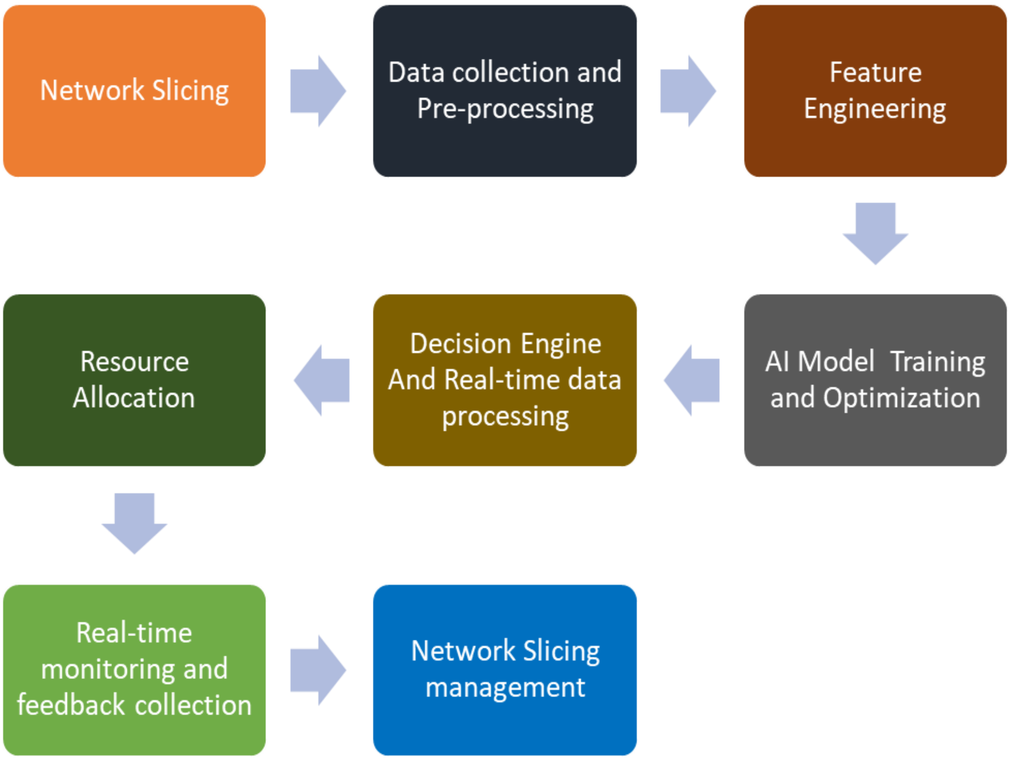

Figure 1: Proposed methodology block diagram

A. DESCRIPTION

The first step entails subdividing the physical communication network into multiple and isolated network slices. In this step, the major network slices developed from the traditional communication network include enhanced mobile broadband, ultra-reliable, low-latency communication, and massive machine-type communication services (Nadeem et al. 37596). Resources will be allocated to these slices depending on the requirements and demands of the network users. The second step entails the collection of data and preprocessing. Essential network information is gathered in this step, including traffic patterns, user demands, and resource utilization. This information is crucial in developing a suitable algorithm for the resource allocation to different network slices (Shen et al. 50). Pre-processing entails cleaning the raw data to remove noise and outliers. In this stage, data is normalized to ensure consistency and comparability. Also, essential features for the input into the AI model are extracted during the preprocessing stage. Feature engineering will focus on the development of new features or transformation of the existing ones to leverage the AI model's ability to learn and make accurate predictions. In this stage, the most important factors to consider include latency requirements, bandwidth needs, and application characteristics (Long et al 168). Feature engineering marks the final stage of network slice customization, ensuring that they are ready for the integration of the AI model for resource allocation.

AI model training step focuses on model customization to facilitate efficient resource allocation in network slicing. Machine learning algorithms, including deep learning and reinforcement learning, are used to train the AI model. The training uses historical data to establish patterns and relationships between variables (Sarker 160). The model is then optimized for accuracy and efficiency. The next step is implementing the AI decision engine, which receives real-time data from the network and processes it through the trained model. The decision engine facilitates resource allocation by assessing demands, network conditions, and application requirements (Ramezanpour). Based on the decisions from the AI decision engine, the resources are allocated dynamically. Resource allocation entails adjusting bandwidth, computing power, and other resources to align with the needs identified by the AI model (Fu and Ye 27). Following resource allocation, the AI model will continuously monitor the network performance and resource allocation. Feedback is collected to evaluate the effectiveness of resource allocation decisions and ensure continuous learning and improvement of the AI model. The last aspect, network slicing management, entails managing and orchestrating network slices based on the allocated resources. This step ensures that all network slices meet their specified requirements and objectives.

B. ALGORITHM

The algorithm developed for resource allocation in network slicing mainly focuses on implementing the proposed methodology steps, identifying potential problems, and developing effective and real-time solutions. In this research, the algorithm will focus on resource management and allocating network resources depending on requirements and user demands. The entire process of resource allocation will be segmented into three stages: dynamic resource reservation, autonomous radio resource management, and customized physical resource allocation.

C. RESOURCE RESERVATION

Resource reservation schemes are critical for mitigating the signaling burden when the demand of users in each base station for a slice increases. Reservation is determined and facilitated by the controller, which has a universal network view (Wu et al. 98). The controller is responsible for determining the minimum resource requirement for each slice. The following Algorithm will be implemented for dynamic resource reservation.

```
Initialization
Set the allocated V= zeros(m , n ) for all slices on all BSs, the reserved E= zeros(m , n ) for all slices on all BSs, the unused F= [1,1, ... ,1] for the BSs;
Set T←0
Iteration
Calculate the weights Islicem,n of BSs to each slice;
If(T==0)
Collect the minimum resource requirements of each user of the slice and sum up them to find the minimum requirement of the slice;
Calculate the reservation weight IBSm,n and the initial slice resource reservation Em,n;
Else if
Collect the updated resource allocation V¯m,n from all slices;
End if
Update the unused resource Fn on each BS;
If (traffic statistics changed)
Update the reservation weight IBSm,n and Islicem,n for all slices in all BSs;
Update the slice resource reservation Em,n;
End
Output vm and em to DQN;
T←T+1 ;
End
```

Figure 2: Dynamic resource reservation (Sun 45764).

D. AUTONOMOUS RADIO RESOURCE MANAGEMENT

This stage mainly concerns AI model training, decision engine, and optimization. After the controller completes the initial resource reservation on each base station, the slices can autonomously adjust their resources to leverage the quality-of-service utility and resource utilization among users. Based on the slice requirements, the bandwidth fraction can be physically mapped to the isolated resource blocks (Kasgari 2). Each slice is restricted to the resources allocated by the controller to alleviate performance degradation associated with the mutual interdependence of slices (Esmaeily et al. 730). Reinforcement learning is one of the approaches used in autonomous radio resource management. This approach trains a software agent for decision-making mechanisms by interacting with the environment. Deep-Q reinforcement learning will be implemented in this case, such that the software agent learns by aiming to maximize the rewards by taking appropriate action to change the environment. The following Algorithm will be implemented for reinforcement learning.

1 Set a replay memory size D and mini-batch size D';
2Configure neural network with input to the number of variables in the state and output number of actions;
3 Initialize Q-network with random weights;
4 Choose epsilon value ε ;
5 While (network active)
6 Collect state <v,R,U,e> ;
7 Generate a random number π ;
8 If π<ε
9 Choose an action randomly;
10 Else
11 Input the state to the ANN and choose the action that has maximum Q-value;
12 End If
13 Update resource by (21) and calculate reward by (19);
14 Update the required allocation Vm,n by (23) in all BS;
15 Convert V and E to overall vm and em by (17) and (18);
16 Store <v,R,U,e,a,w> in replay memory;
17 If (activated by timer)
18 Pick a mini-batch of samples from the replay memory;
19 For all of samples
20 Calculate output Q-values by ANN;
21 Update target Q-value of the action by (20);
22 Train ANN by a tuple of (input, output Q-value, and target Q-value) with a loss function as the mean square error of output and target;
23 End for
24 End if
25 End while

Figure 3: Radio resource management algorithm (Sun 45765)

E. CUSTOMIZED PHYSICAL RESOURCE ALLOCATION

At the end of radio resource management, the updated resources of the slice are updated to the base stations and the slice allocates the resources to users. The last stage focuses on the process of allocation of resources to the users. The following algorithm will be implemented for physical resource allocation.

Initialization

1 Receive the data rate, virtual resource fraction, delay constraint, and arrival rates of the users from the slice

2 initialize ρ>0θ(0)km,n=0 , ∀km , n

3 initialize {y(0)km}|Km|km=1 as y(0)km=[0,0,…,0]T

4 initialize {z(0)n}|N|n=1 as z(0)n=[0,0,..,0]T

5 downscale ckm,n=V¯m,nckm,n∀km,n

Iteration

6 Set t← 0

7 For each user

8 Update the resource allocation vector y(t)km in all BSs by (31a) or (33), where τ(t+1)km:=(∑n=1|N|y(t)km,nckm,n−λkm)−1 and μ1≥0 are chosen such that 0<τ(t+1)km≤τmaxkm and μ1(τ(t+1)km−τmaxkm)=0 .

End for

9 For each BS

10 Update the resource allocation for all the served users of the slice at BS n by (31b)

11 End for

12 Update θ(t+1)km,n=θ(t)km,n+ρ(y(t+1)km,n−z(t+1)n) by (31c)

13 t← t + 1;

14 until termination test satisfied

15 return result

Figure 4: Physical Resource Allocation (Sun 45767)

III. FLOW CHART

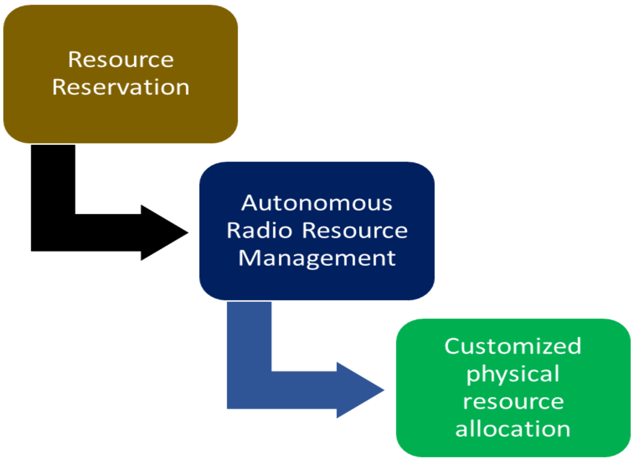

Figure 5: *AI-driven Resource Allocation in the network slicing*

IV. RESULT ANALYSIS

In this research, numerical simulations were conducted using MATLAB software. In the simulation, the parameters are aligned with 5G specifications such that the macro base station is centered within a 500-meter coverage radius and surrounded by 4 small-cell base stations and uniformly distributed users. The objective of the simulation is to shed insights into the AI model's effectiveness in enhancing service quality and optimizing resource utilization. The evaluation focused on the performance of the radio resource management algorithm. The radio resource management algorithm is subjected to high infeasible load to determine their load limitations while periodically increasing users on the slices. Performance is measured as the ratio of satisfied users to the total users. The results indicated a consistent user satisfaction ratio as users increased. This indicates that the AI model effectively allocates resources to slides and enhances user experience.

V. CONCLUSIONS

As network slicing is gaining popularity, there is a growing need to optimize the potential of this technology. The effectiveness of network slicing largely depends on resource allocation. An AI-driven model for resource allocation in network slicing is examined in this research paper. The model's algorithm comprises three stages, each with its specific algorithm. Resource allocation begins with reservation, where the controller reserves minimum resources for each slice. The second stage is autonomous radio resource management, mainly focusing on AI model training and decision engines. The last stage is physical resource allocation, where resources are distributed to the slices and users. Simulations on MATLAB software indicated this model to be effective in enhancing resource allocation.

REFERENCES

[1]. Debbabi, Fadoua, et al. "Algorithmics and Modeling Aspects of Network Slicing in 5G And Beyonds Network: Survey." IEEE Access 8 (2020): 162748-162762.

[2]. Esmaeily, Ali, et al. "Beyond 5G Resource Slicing with Mixed-Numerologies for Mission Critical URLLC and eMBB Coexistence." IEEE Open Journal of the Communications Society 4 (2023): 727-747.

[3]. Fu, Sibao, Fan Yang, and Ye Xiao. "AI Inspired Intelligent Resource Management in Future Wireless Network." IEEE Access 8 (2020): 22425-22433.

[4]. Kasgari, Ali Taleb Zadeh, and Walid Saad. "Stochastic Optimization and Control Framework For 5G Network Slicing with Effective Isolation." 2018 52nd Annual Conference on Information Sciences and Systems (CISS). IEEE, 2018.

[5]. Khan, Latif U., et al. "Network slicing: Recent Advances, Taxonomy, Requirements, And Open Research Challenges." IEEE Access 8 (2020): 36009-36028.

[6]. Long, Wen, Zhichen Lu, and Lingxiao Cui. "Deep Learning-Based Feature Engineering for Stock Price Movement Prediction." Knowledge-Based Systems 164 (2019): 163-173.

[7]. Nadeem, Lubna, et al. "Integration of D2D, Network Slicing, And MEC in 5G Cellular Networks: Survey and Challenges." IEEE Access 9 (2021): 37590-37612.

[8]. Rakkiannan, Thamilselvan, et al. "An Automated Network Slicing at Edge with Software Defined Networking and Network Function Virtualization: A Federated Learning Approach." Wireless Personal Communications 131.1 (2023): 639-658.

[9]. Ramezanpour, Keyvan, and Jithin Jagannath. "Intelligent Zero Trust Architecture For 5G/6G Networks: Principles, Challenges, and the Role of Machine Learning In The Context Of O-RAN." Computer Networks 217 (2022): 109358.
[10]. Sarker, Iqbal H. "Machine Learning: Algorithms, Real-World Applications and Research Directions." SN Computer Science 2.3 (2021): 160.
[11]. Shahjalal, Md, et al. "Enabling Technologies for AI Empowered 6G Massive Radio Access Networks." ICT Express 9.3 (2023): 341-355.
[12]. Shen, Xuemin, et al. "AI-Assisted Network-Slicing Based Next-Generation Wireless Networks." IEEE Open Journal of Vehicular Technology 1 (2020): 45-66.
[13]. Sun, Guolin, et al. "Dynamic Reservation and Deep Reinforcement Learning Based Autonomous Resource Slicing for Virtualized Radio Access Networks." Ieee Access 7 (2019): 45758-45772.
[14]. Wang, Yuxing, et al. "AI-Based Resource Allocation in E2E Network Slicing with Both Public and Non-Public Slices." Applied Sciences 13.22 (2023): 12505.
[15]. Wu, Wen, et al. "AI-Native Network Slicing For 6G Networks." IEEE Wireless Communications 29.1 (2022): 96-103.

