

Integrating Auction-Based Mechanisms For Enhanced Resource Management In 5G Slicing

Qasim Abduljabbar Hamad¹, Morteza Valizadeh^{2,*}, Vahid Talavat³

^{1, 2, 3} Department of electrical and computer engineering, Urmia University, Urmia, Iran.

*Corresponding author: Morteza Valizadeh

Abstract:

Fifth-generation computer networks (5G) are designed to accommodate a diverse range of heterogeneous and often incompatible services. Unlike traditional networks, the "one-size-fits-all" approach is not applicable to 5G. To enable a seamless transition to this new generation, several theories have been explored, with network slicing being among the most critical. Network slicing involves creating and managing multiple logical networks, which are then mapped to a physical infrastructure. Each slice is characterized by a unique set of network functions and resources, allowing specific services to be assigned to individual slices. This approach ensures that the network functions and resources of each slice are tailored to meet the specific requirements of the assigned service. Network slicing has attracted significant research interest, but this study focuses primarily on resource management within network slices. Resource management is broadly categorized into two key areas: "slice admission control" and "cross-slice resource allocation." From a decision-making perspective, these areas employ two main approaches: policy-based decisions and auction-based decisions. This research combines both slice admission control and cross-slice resource allocation using an auction-based decision-making framework, which, to the best of our knowledge, is a novel approach. While slice admission control and cross-slice resource allocation may appear distinct, they are closely interconnected, making it essential to study and combine them together with an appropriate technique. This study combines the mentioned methods and examines the impact of dynamic cross-slice resource allocation on slice admission control. Through simulations and analysis, the research demonstrates how these mechanisms collectively enhance resource management in fifth-generation computer networks.

Keywords: Network Slicing, Fifth Generation Cellular Networks, Auction-Based Decision.

1- Introduction

Network slicing enables the implementation and execution of various service requirements of fifth-generation network services in the form of network slices with suitable network resources and functionalities. Network slicing and resource management within 5G networks have been discussed and investigated in various articles and references, some of which will be referenced in the following sections.

Network slicing, both in the Core Network and in the radio access network, has been discussed and investigated in previous researches. One of the earliest definitions of network slicing and the need for its implementation and use in fifth-generation networks was explored in [7]. It introduced and examined the concept of slicing as a Service (SlaaS), which enables the separation of the structure of wireless networks from the services running on them, providing network operators with the ability to lease network slices as services to users and requesters instead of running and managing network services. In this model, service providers can rent a slice from network operators without owning a network infrastructure and offer their desired services to customers. For example, slice admission control and utilization of slices with the same size was focused in [8]. An example of a slice as a service is illustrated in Figure (1). As shown in this figure, network operators lease network slices composed of various network functions and resources to different tenants who then offer these slices to service requesters [9]. Another research [10] addressed the issue of enhancing resource management in network slicing for the radio access network. Available resources in the system, including base stations and

backhaul capacity, are simultaneously examined and optimized in the form of a Bi-convex problem in this reference. Although this source discusses resource allocation among slices, it does not delve into slice admission control based on auctions. In [11], the authors utilized auction mechanisms for resource allocation and network slicing, where the limited resources such as storage, computational resources, and radio resources are considered to allocate a balanced rate to auctions so that the overall network revenue and system efficiency in allocating resources to network slices are maximized. The model used to define network slices and the services allocated to them is illustrated in Fig. (2).

Fig. (1). Slice as a Service (SlaaS) in Fifth-Generation Networks [7].

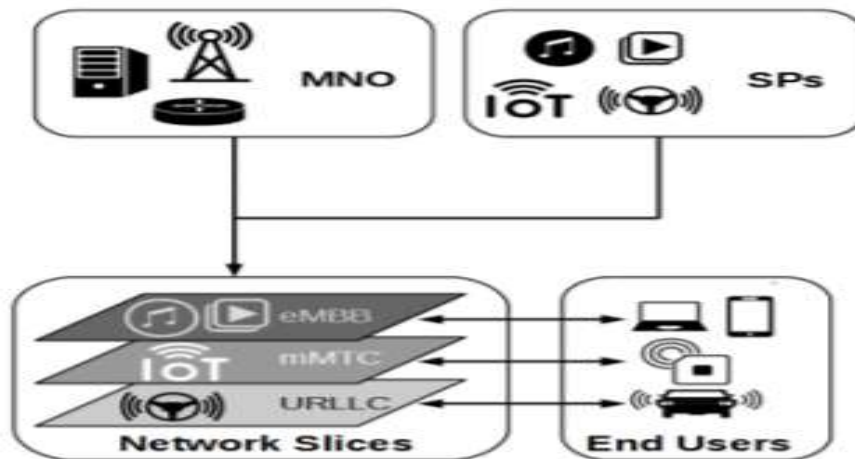
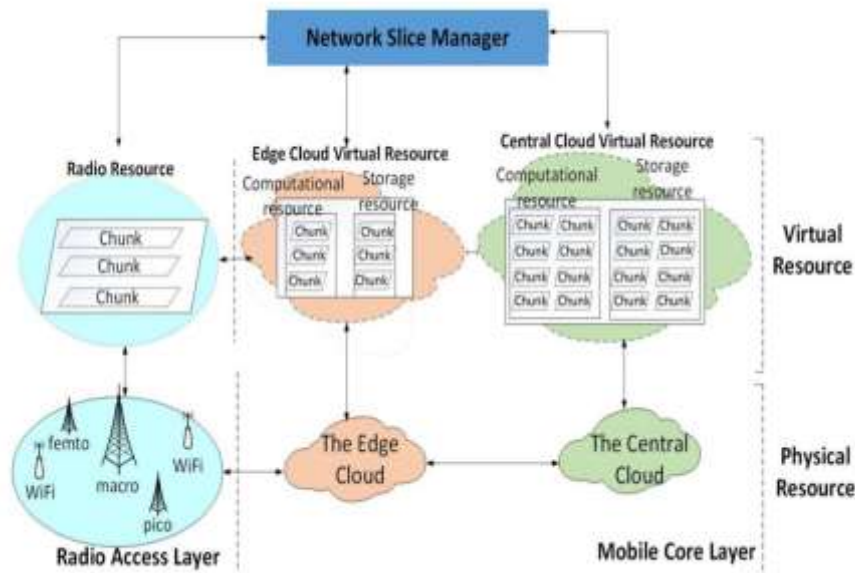


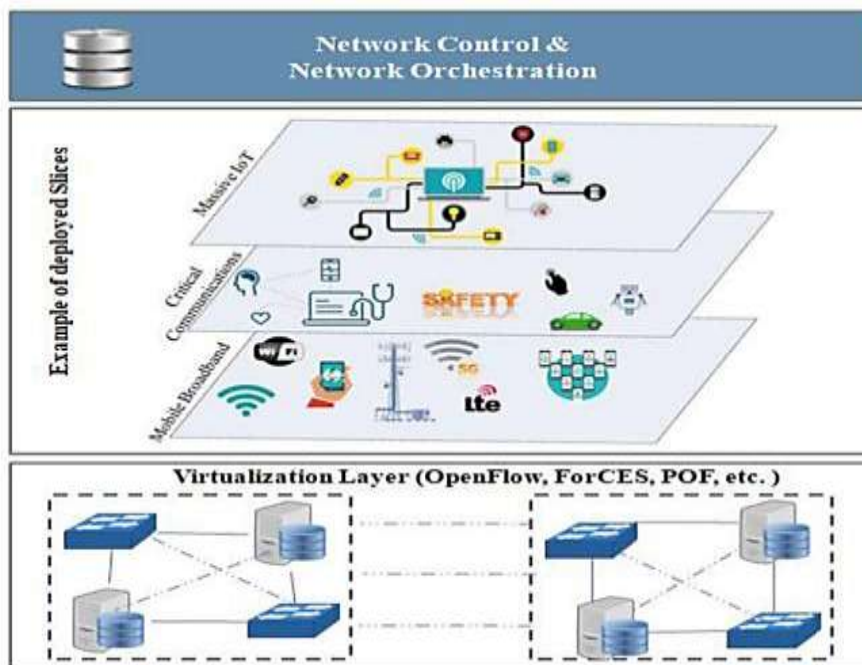
Fig. (2) A sample network slicing model [11].



As mentioned, the issue of resource management in network slicing is examined in two aspects: resource allocation among slices and slice admission control. One of the fundamental challenges in implementing network slicing is efficiently sharing network

resources. Statically sharing resources among network slices based on their allocations is a known method to solve this challenge [12]. In this method, resources are exclusively assigned to slices, preventing resource sharing between slices. Since system loads may vary at different times, this approach restricts overall system efficiency. Ideally, it is better to Elastically allocate resources to slices based on network load and allow each slice tenant to independently manage their slice resources to meet their customer needs effectively. The resource sharing in network slicing has been discussed and examined by various researchers. Some researchers utilized a static approach to model resource sharing, with their common focus being on slice admission control, making decisions based on policies or auctions. Some others [13] employ heuristic methods, while others like [14] use Integer Linear Programming for implementing the model. The employed architecture and an example of network slicing are illustrated in Fig. (3) [13]. In [22], the authors address the improvement of slice admission control by employing a reinforced learning method and decision-making based on policies. The authors claimed that the utilized Q-Learning method can relatively discover the best slice admission policy in the network, maximizing operator revenue to the fullest extent possible. It is worth mentioning that compared to value iteration methods that reliably converge to the optimal solution, this approach achieves optimal network revenue with only a slight difference. However, on the other hand, the computational cost of Q-Learning is significantly lower than numerical methods, which is considered as one of the advantages of the mentioned method. In [23], authors utilized reinforced learning methods to enhance resource allocation among slices. Operators use a specific framework to assess slice requests for obtaining additional resources. Based on the priority of slices and the available resources in the system, slice requests are either accepted or rejected for resource allocation consideration.

Fig. (3): An example of network slicing architecture [13].



Resource allocation in such problems is often likened to a fishermen market scenario. In other words, buyers in the fishermen market are synonymous with network slices that

have limited and defined budgets. Based on the priority of these buyers, auctions are announced, and according to these auctions, network resources are allocated. Resource allocation in such scenarios is often likened to a fisher market problem. In other words, buyers in the fisher market represent network slices with limited and defined budgets. Based on these buyers' priorities, auctions are announced, and resources are allocated to them according to these auctions. Analyzing the fisher market reveals that if the auction's impact on product prices is negligible, achieving equilibrium is feasible, and there exist different distributed allocation algorithms that can converge to this state [15]. Since the auction's effect on prices is usually not negligible, the fisher market cannot fully and comprehensively model resource allocation in this research problem. One other resource allocation model relevant to this subject is the general Kelly's Mechanism [16]. In which, based on conducted auctions, existing network resources are allocated to buyers or network slices. However, unlike the fishermen market, the auction's impact on prices is taken into account. Analyzing and scrutinizing the efficiency of Kelly's Mechanism have examined in [17], , and [18] its scalability was enhanced through careful consideration of additional parameters in problem-solving. In these methods, buyers simply react to the auction amount, given that in this research problem, network slices have a limited share of the total network resources. Models that do not incorporate this parameter cannot be adequate models for this investigation. In any case, network slicing has emerged as one of the crucial mechanisms and expected features of fifth-generation networks [19], and 3GPP has also defined its criteria and measures for network slicing [20]. This section refers to researches focusing on resource allocation among slices. In [21], user admission was examined and modeled using policy-based decision-making. In scenarios where the service quality of a slice falls below expectations under specific system conditions, one or multiple users are evicted based on another corresponding policy of the system. The investigated problem was modeled using game theory and it was discussed that by employing these two acceptance and eviction policies, the system reaches an equilibrium state. The resource allocation through auctions among users was investigated, but dynamic allocation among slices was not explored. In [24], improving resource allocation among slices was also addressed, but it used a more general model to solve the problem compared to other methods. It considers all slices requesting that resource simultaneously, instead of a binary perspective, when allocating a resource to a slice. This perspective introduces complexity and nonlinearity to the problem, leading the authors to employ artificial neural network methods to address this challenge. Another research has utilized artificial neural networks to solve the mentioned complex and nonlinear problem [25]. In this research, a three-dimensional convolutional neural network was used to accurately predict the required capacity of a service using data-centric methods. This prediction aids operators in making decisions for accepting a slice and achieving optimal resource allocation among slices. Also, the reinforcement learning technique was employed to recursively update the auction rates of each resource [26]. To solve this issue, auctions are modeled using game theory, and the authors claimed that this learning method converges to an equilibrium state, aiming to enhance resource allocation among slices. The resource allocation problem was tackled among slices using reinforcement learning, implementing a policy to maximize network revenue while minimizing the cost of each slice [27]. By modeling the problem with game theory, they argued that their learning approach also converges to a non-Nash equilibrium state. It is noteworthy that the focus of this reference lies in improving resource allocation among slices. In the existing literature, various approaches have been explored to address resource management challenges in network slicing. Mo and Walrand [28] introduced a fair congestion control mechanism for end-to-end networks, emphasizing window-based strategies. The Topology Zoo [29] has been a significant resource for simulating network

topologies, offering valuable insights into network slicing structures. The 3GPP specifications [30] outlined foundational physical layer procedures crucial for resource allocation in 5G. Thantharate and Beard [31] proposed ADAPTIVE6G, an adaptive resource management framework tailored for current 5G and emerging 6G networks. Singh et al. [32] investigated the use of machine learning techniques for efficient network slicing in 5G, highlighting the potential of intelligent systems in managing complex networks. Alotaibi et al. [33] demonstrated the effectiveness of SDN-based technologies for traffic management in 5G network slicing. Saibharath et al. [34] addressed joint QoS and energy-efficient resource allocation in 5G, introducing methods to balance performance and energy consumption. Additionally, Kim and Lim [35] employed multi-agent reinforcement learning for dynamic resource management, presenting a novel approach for end-to-end network slicing optimization. These studies collectively form the foundation for advancing resource management in network slicing architectures. Following this, Table (1) is presented to summarize recent references.

Table (1): Summarizing the references discussed in Section Two.

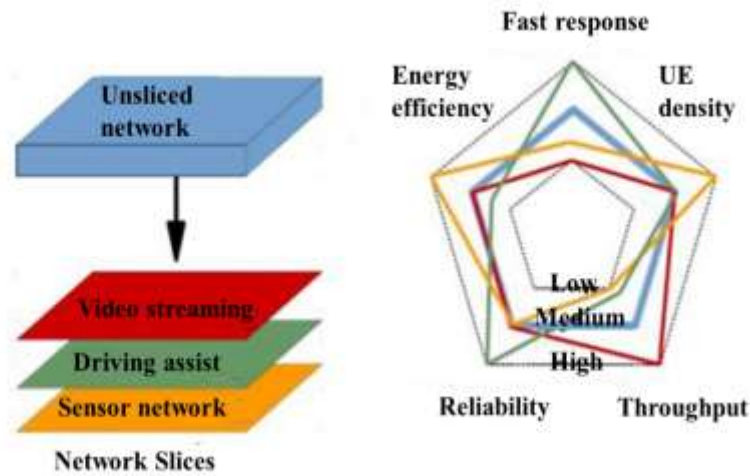
Ref.	Objective	Service Quality	The mechanism	Approach
[8]	Increasing Network Throughput	Not examined	Resource Allocation Mechanism between Slices	Decision-making based on auction
[10]	Resource Management Optimization	Not examined	Resource Allocation between Slices	Decision-making based on policy
[11]	Enhancing network throughput and resource management optimization.	Not examined	Resource Allocation between Slices	Decision-making based on auction
[12]	Network revenue and system throughput improvement.	Examined	Slice Admission Control	Decision-making based on policy
[13]	Increasing user base and enhancing service rates.	Examined	Slice admission control	Decision-making based on policy
[14]	Boosting user numbers and improving network throughput.	Examined	Resource allocation among slices	Decision-making based on auction
[17]	Optimizing resource management and network throughput.	Not examined	Resource allocation among slices	Decision-making based on auction

[18]	Enhancing throughput and network scalability for network acceptance.	Examined	Resource allocation among slices	Decision-making based on auction
[21]	Reduction in the Churn Rate of New Users in the Network	Examined	Slice admission control	Decision-making based on policy
[22]	Increase in Network Revenue	Examined	Slice admission control	Decision-making based on policy
[23]	Improvement in Resource Management and Network Throughput.	Examined	Resource allocation among slices	Decision-making based on policy
[24]	Increasing Network Throughput	Examined	Resource allocation among slices	Decision-making based on policy
[25]	Increasing Network Revenue.	Examined	Resource allocation among slices	Decision-making based on policy
[26]	Increasing Network Throughput.	Examined	Resource allocation among slices	Decision-making based on auction
[27]	Increase in Network Revenue and Throughput	Examined	Resource allocation among slices	Decision-making based on policy

2- Problem statement

Network slicing refers to the creation and maintenance of multiple independent logical networks that are mapped onto a common physical infrastructure. Each of these logical networks represents a slice of the network where network functions and resources to be executed are explicitly defined [4]. Physically, logically, in terms of resources, and so forth, network slices can be entirely independent, and required services in the network can be executed across all or part of these slices. Given that 5th generation services may have heterogeneous requirements such as throughput, access capacity, or expected latency, network slicing enables the definition of network resources and functions to address the specific needs of a service. An example of network slicing with several different services (video media, autonomous vehicle control, and sensor networks) is illustrated in Fig. (4) and (5).

Fig. (4) and (5). Network Slicing with the Capability to Create Different Services in Fig. (4) and (5).



In general, network slicing is extensively researched and utilized in the wireless network domain. Examples include network architecture design, radio access network virtualization, network resource management, mobility management, and cybersecurity establishment. However, in our research focus, the emphasis will be on resource management in network slicing, which will be further elaborated on the various aspects of this research.

2-1- Resource Management in Network Slicing

Resource management in network slicing is examined in the following two aspects [5]:

- Slice Admission Control
- Resource Allocation Among Slices

Considering that network operators have a specific and limited amount of resources at their disposal, sometimes it may not be feasible to provide resources for allocating to a new slice within the system. Thus, issues within slice admission control domain deal with policies that generally evaluate accepting or rejecting a request to create a new slice in the system. Therefore, it is expected that various slice admission control strategies will directly influence system efficiency and performance.

Upon allocating resources and network functions required for a service to a slice based on its needs, the service commences operation on the network. Due to the fluctuating traffic load of a slice over time, operators may, at times, assign more than the capacity of a slice to users for enhancing their resource utilization and increasing the revenue. Mechanisms focusing on resource allocation among slices must promptly and dynamically address the shared resources among slices to boost overall system efficiency and maintain expected service quality within the slice.

On the other hand, decision-making mechanisms, whether in slice admission control or resource allocation among slices, are divided into the following two categories:

- **Policy-based Decision-making:** In this approach, network operators make decisions regarding network events based on system conditions and a policy that is uniform across all customers. For instance, a system that makes decisions regarding slice admission control based on policy considers factors such as slice pricing, system state comprising available resources, active slices in the system, queue length of current requests, and so forth when deciding to accept or reject a request for creating a new slice in the system.

- **Auction-based Decision-making:** In contrast to the previous method, operators do not consider uniform prices for slices and resources within the system but adjust these prices based on proposals and auctions for various requests. Essentially, requests for different resources and slices are collected periodically by operators, and auction winners receive their desired resources. A crucial aspect of this method involves determining a minimum threshold for an auction to maximize revenue from imposed network costs.

2-2- Problem Statement

Based on the aforementioned foundations, the problem of enhancing resource management in network slicing can be broadly categorized into slice admission control and resource allocation among slices. Each of these approaches can also be examined from two decision-making perspectives: policy-based decision-making or auction-based decision-making. This paper aims to simultaneously improve resource management in slice admission control and resource allocation among slices for auction-based decision-making systems, ensuring that system efficiency is maximized and the required quality of service for requests is met.

3- Proposed Method

This section introduces the system model and the proposed method in this research. The system model presentation begins with an explanation of dynamically allocating resources among slices through auctions, followed by outlining slice admission control.

3-1- Dynamic Resource Allocation Among Slices through Auctions

As mentioned in the first section, resource allocation among slices refers to network resource management after creating a slice and generally occurs in two forms: static and dynamic. In the static mode, a resource is exclusively assigned to a slice, while in the dynamic mode, a resource may be shared among one or multiple network slices.

This research focuses on the dynamic mode, defining a wireless network where system resources are denoted by B , network slices by S , and customers by U . By utilizing indexing, U^s refers to users and devices utilizing slice s , while U_b refers to users and devices present at base station b .

Each slice possesses a specific share of the total network resources denoted by p_s . The relationship among the shares of all slices is defined as follows:

$$\sum_{s \in S} p_s = m \quad (1)$$

Where m represents the total network resources. On the other hand, each slice can freely distribute its share of network resources among its users. If each user's share is denoted as w_u , the following relationship exists for the users' share of resources within a slice:

$$\sum_{u \in U^s} w_u = p_s \quad (2)$$

Therefore, users will receive a portion of the resources allocated to each base station according to the weight assigned to them within the slice. Since decision-making is to be based on auctions, slice tenants must adjust these weights such that users willing to pay more will receive a larger share of the slice.

Similar to the simplification done in [14], the relationship between each user's rate and the assigned weight can be expressed as:

$$r_u = \frac{w_u}{\sum_{u \in U_b} w_u} C_u = \frac{w_u}{l_b} C_u \quad (3)$$

In the given expressions, r_u denotes the rate of user u , l_b represents the total share of existing users on base station b from the resources, and C_u is the maximum achievable rate when only one user is present at a base station.

To prioritize users, the parameter Φ_u is employed as follows:

$$\sum_{u \in U} \Phi_u = 1 \quad (4)$$

Two points are pertinent regarding this relationship. Firstly, to prioritize users, the parameter Φ_u is employed as follows:

is a non-negative parameter ($\Phi_u \geq 0$), and secondly, the total priority of users within a slice will sum up to one. To apply auctions in resource allocation among slices, users willing to pay more will receive a greater Φ_u .

Based on the aforementioned, the utility function for a slice is derived as follows [14]:

$$U^s(w) = \sum_{u \in U} \Phi_u f_u(r_u(w)) \quad (5)$$

In this scenario, w represents the vector of resource shares of active users within a slice, f_u signifies the utility function of user u , and among the diverse utility functions discussed [15], the following relationship can be calculated:

$$f_u(r_u) = \frac{r_u^{(1-\alpha)}}{1-\alpha} = r_u \quad (6)$$

Essentially, this is an α -fair utility function with $\alpha = 0$ (linear sum utility). In previous researches, dynamically resource allocation among slices, similar to research, has been performed, where [15] albeit investigating α values greater than one and utilizing a heuristic model to tackle the issue. In this paper, a numerical linear programming method is employed, and as further elucidated, the topic of slice admission control will be concurrently examined alongside resource allocation among slices. Therefore, the forefront problem is to maximize the sum of utility functions for slices, a task that, as per [14], falls under the np (Nondeterministic Polynomial) category. In this study, the issue is soluble and investigable using a mixed-integer linear programming (MILP) methodology [31].

3-2- Solution of the MILP Problem

The MILP problem is generally formulated as follows:

$$\max_x = \{f^T x\} \quad (7)$$

Subject to:

$$Ax \leq b \quad (8)$$

$$A_{eq}x = b_{eq} \quad (9)$$

$$lb \leq x \leq ub \quad (10)$$

$$x \in \mathbb{Z} \quad (11)$$

For the specific problem addressed in this research, the formulation is transformed as:

$$\min_w (-\left\{ \sum_{s \in S} P_s \sum_{u \in U^s} \Phi_u \frac{w_u}{l_b} C_u \right\}) \quad (13)$$

Subject to:

$$w_u \in \mathbb{Z} \quad (14)$$

Considering that in equation (12) the values of P_s , Φ_u , l_b , and C_u are specified, this equation can easily be translated into equation (7). Furthermore, equation (13) can also be directly transformed into (9), making the MILP problem solvable in this manner. By solving this MILP problem, the vector w representing the users' shares of slice resources is obtained. Equation (12), based on this derived share, maximizes the users' data rate or in other words, maximizes the network throughput.

3-3- Slice Admission Control

Slice admission control deals with the decision to accept or reject a request to create a new slice in the network, occurring when a user requests a new service from the network that cannot be supported by existing network slices.

In the virtualized network structure, each slice comprises several Forwarding Elements interconnected according to a specific topology, along with several servers connected to these elements. Therefore, each slice consists of a set of virtual machines, nodes, and links known collectively as network resources. Each of these network resources is responsible for performing network Tasks, estimating metrics such as latency, reliability, and accessibility, on average, and representing them for each resource as follows [13]:

$$B_j: \begin{pmatrix} \text{Reliability} = R_j \\ \text{Availability} = A_j \\ \text{Latency} = L_j \\ \text{Load} = O_j \end{pmatrix} \quad (15)$$

On the other hand, each slice also requires a minimum of these metrics determined by the Service Level Agreement (SLA) contract. Therefore, for any given slice, one can express:

$$S_j: \begin{pmatrix} \text{Reliability} = R_i \\ \text{Availability} = A_i \\ \text{Latency} = L_i \\ \text{Load} = O_i \end{pmatrix} \quad (16)$$

From the definition of w_u in the previous section, which was the allocation share to a user of the resources of a slice, the load on a resource and slice is computable. Therefore, a resource equals the sum of shares of users that have utilized the mentioned resource.

$$O_j = \sum_{u \in U_b} w_u \quad (17)$$

On the other hand, the load on a slice equals the sum of weights of users present in the slice, or in other words:

$$O_i = \sum_{u \in U^s} w_u = p_s \quad (18)$$

In this study, the delay value L_j is obtained from the research [29] without losing the generality of the problem, the reliability of network resources R_j is assumed to be a random value with the distribution [99.990, 100], and the accessibility of resources A_j is also estimated to be a random value with distribution [99.990, 100] [13].

Furthermore, the delay values L_i , reliability R_i , and accessibility A_i expected for a slice may vary depending on the type of service to be executed on it, as derived from the table (2) provided at the end of the section.

Now, the problem of accepting or rejecting a request to create a new slice, considering the loads on the slices and characteristics of the new service, can be examined. For this purpose, the following utility function is utilized [13]:

$$U(x) = 1 - e^{-ax} \quad (19)$$

where X is the determining vector and a is the associated weight with this vector. The value of a ranges between zero and one, and the closer it is to one, the more the impact coefficient of vector X in calculating the utility function increases. Currently, using the equations (20), the score of a network resource is calculated, and if the scores of individual resources present on a slice are greater than zero, that slice is considered a candidate for running the desired service for the user:

$$\begin{aligned} \text{Score}(R_j) = & (1 - e^{-a_A A_j})(A_i \otimes A_j) \times \left(1 - e^{-\frac{a_L}{L_j}}\right)(L_i \otimes L_j) \\ & \times (1 - e^{-a_R R_j})(R_i \otimes R_j) \times (1 - e^{-a_O O_j})(O_i \otimes O_j) \end{aligned} \quad (20)$$

In the above equations:

$$A_i \otimes A_j = \begin{cases} 1 & \text{if } A_j \geq A_i \\ 0 & \text{otherwise} \end{cases} \quad (21)$$

$$R_i \otimes R_j = \begin{cases} 1 & \text{if } R_j \geq R_i \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

$$L_i \otimes L_j = \begin{cases} 1 & \text{if } L_j \geq L_i \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

$$O_i \otimes O_j = \begin{cases} 1 & \text{if } O_j \geq O_i \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

In relation to Equation (20), a_A , a_L , a_R , and a_O respectively represent the coefficients of accessibility, delay magnitude, reliability, and load on the resource, in calculating the score of that resource. Equations (21) to (24) further elucidate that, for instance, if the accessibility value of a resource R_j is greater than the expected service value to be executed on a slice, R_i , the value $R_i \otimes R_j$ is calculated as one; otherwise, it is considered zero.

Thus, following Equation (20), the score of each resource on a slice is calculated based on the service slated for execution on it. If none of the resource scores equate to zero, the slice is introduced as a candidate for service execution. If multiple candidate slices exist, the slice with the highest sum of resource scores is chosen. Conversely, if no candidate slice for service execution is available, a new slice is generated from resources that have scores greater than zero, if feasible.

3-4 Assumptions

- The simulated traffic is considered elastic and may vary over time.
- Implementing Equation (3) necessitates that each slice lessor is aware of the total load on individual base stations and informs the infrastructure of the allocated share to each user.
- Since this issue is examined in the form of MILP, the number of network nodes must be limited. Otherwise, problem-solving time increases significantly.
- Slice lessors are inherently non-malicious, and the exchanged information between them and the infrastructure will be truthful.
- The minimum values of delay, reliability, and accessibility metrics for various applications in fifth-generation networks are taken from reference [13].
- The problem is analyzed assuming no failures and is modeled accordingly.
- The existing slides in the network can receive their required resources from all network entities.
- The algorithms described in this section run in a centralized manner on the network.

the system model consists of two parts, which include resource allocation among slices and slice admission control. In the resource allocation among slices section, the MILP problem investigates the potential for dynamic resource sharing among slices and optimizes the data rate of users. This results in obtaining the weight array, w , representing users' shares of slice resources. On the other hand, in the slice admission control section, considering the requirements of new services and the obtained resource scores in the network, the problem of admitting or rejecting the creation of new slices in the network is examined. It is worth mentioning that the resource allocation algorithms among slices and slice admission control algorithms are executed in case of an event

occurrence in the network. For example, if a user disconnects from the network, resource allocation among slices is rerun to recalculate users' shares of network resources. Additionally, if a new user connects to the network and requests a new service, slice admission control is first examined. If feasible, running the service on one of the network slices is considered, followed by a reassessment of resource allocation among slices based on the newly added user. The next section will focus on simulation results and their analysis.

4- Results and Analysis

4-1 Initiation

As mentioned in the previous section, the problem is divided into two main sections: resource allocation among slices and slice admission control. The first section focuses on the dynamic allocation of resources among slices, and by solving the corresponding MILP problem based on Equation (12), the optimal state of resource sharing among slices is obtained. To simulate this part, the utility function of network slices in this paper is compared under the same network conditions with the utility function in the reference material [14], which addresses static resource allocation among slices, and the results are examined in the next section. Subsequently, by obtaining the vector w , or in other words, the user weight vector, the slice admission control problem is investigated, and the number of supported users in the proposed method is compared with the reference [13], and the results are compiled in the following section.

4-2 Simulation

For simulating the proposed method, the software MATLAB R2017b has been utilized. Table (2) and (3) demonstrate the list of parameters, their values, and corresponding relationships used in this simulation [13].

Table 2: The minimum values of delay, reliability, and accessibility indices necessary for various applications in fifth-generation networks.

Use Case Family	Example of use cases	Requirements ,Latency,Reliability and Availability	Latency _{max}	Service Availability Level (SAL)	Availability _{min}	Reliability _{min}
Enhanced Mobile Broadband (eMBB)	Pervasive, video Dense Urban ,Society Video streaming in stadium Ultra-low cost network High Speed	There is no accuracy need for the reliability and availability for this use case family. For most use cases, the latency should be	Between 10 and 5 ms	:SAL 3 Availabl e in normal conditio ns only	95%	95%

	,train Moving Hot .spots, etc	very low				
Massive IoT	Smart wearables Sensor , networks	This family is relatively tolerant to the delay and it requires a normal to high reliability and availability	ms 10			
Critical Communi- cations	Ultra high reliability and ultra low latency use :cases This category includes applicati ons such as automate d traffic control, autonom ous driving, collabora tive robotics, and remote object manipula tion. These use cases demand extremel y low latency,	Use cases of this category are characterize d by a very low latency, a very high reliability and a high .availability	ms 1	:SAL 2 Availabl e in normal and overloa ded conditio ns	99.9%	99.999%

	exceptionally high reliability , and robust availability.					
	Ultra low latency :use cases Examples include the Tactile Internet, where applications require minimal latency alongside high reliability and strong availability.	Use cases in this category are defined by extremely low latency, along with high reliability and availability .	ms 1	:SAL 2 Available in normal and overloaded conditions	99.9%	99.9%
	Ultra high reliability and availability use :case)Applications like eHealth (critical for life-threatening scenarios) and public safety fall into this category. They are characterized by	Use cases in this category are distinguished by exceptionally high reliability and availability , combined with low latency.,	5ms	:SAL 1 Always available	99.999%	99.999%

	extremely high reliability, exceptional availability, and low latency.					
--	--	--	--	--	--	--

Table (3): Simulation Parameters

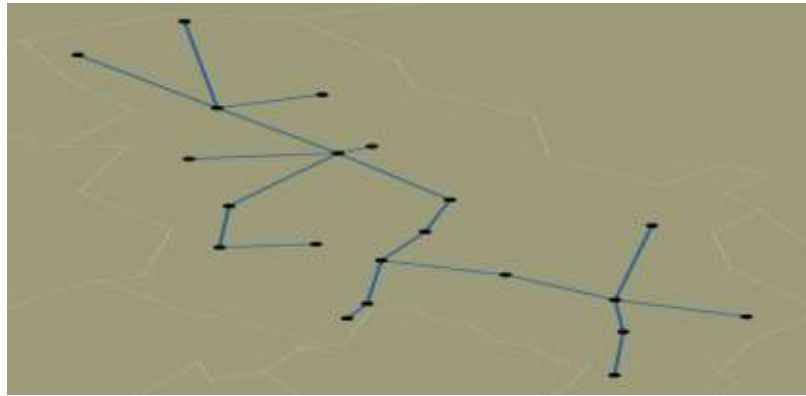
Explanation	Value/Relation	Parameter
the share of each section in network resources	$\sum_{s \in S} p_s = m$	p_s
Each user's share in a segment	$\sum_{u \in U} w_u = p_s$	w_u
Each user's rate in a segment	$r_u = \frac{w_u}{l_b} C_u$	r_u
Hypothetical average rate, obtained from 30.	10Gbit/s	c_u
Load on the base station	$l_b = \sum_{u \in U_b} w_u$	l_b
Auctioneer parameter	$\sum_{u \in U_h} \Phi_u = 1$	Φ_u
User utility function relationship	$f_u(r_u) = r_u$	f_u
Linear utility function	0	α
Network density parameter, taken from [14]	57	$ B $
Load on a resource	$Load_{NRJ} = \sum_{u \in S_j} w_u$	$Load_{NRJ}$
Load on a segment	p_s	$Load_{SCRi}$
Impact coefficient of accessibility level in calculating the resource score	0.25	a_A
Impact coefficient of delay level in calculating the resource score	0.25	a_L
Impact coefficient of reliability level in calculating the resource score	0.25	a_R

Impact coefficient of load on the resource in calculating the resource score	0.25	a_0
--	------	-------

4-3- Analysis Results in Amres

Fig. (6) illustrates the network topology of Amres derived from the Internet Topology Zoo [29]. Subsequent details of the analysis are explained further.

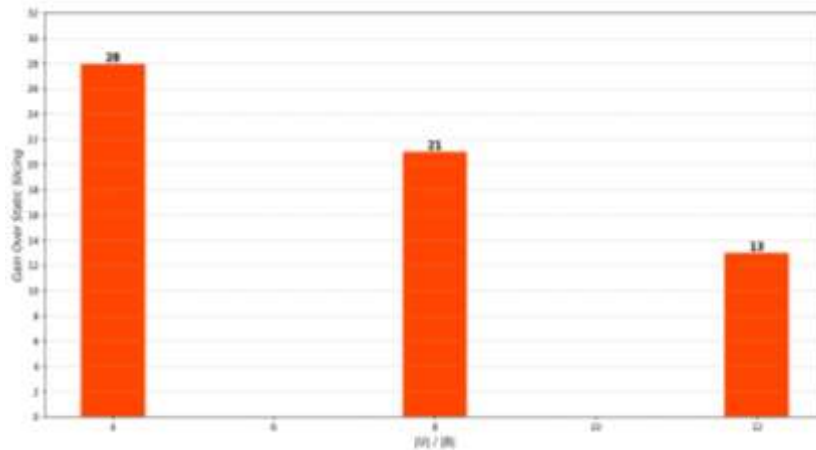
Fig. (6): The network topology, nodes, and links present in Amres [29].



4-3-1- Analysis Details

The Fig. (1) displays the level of resource allocation improvement between slices in the proposed method compared to the static resource allocation mentioned in [14]. The horizontal axis in this graph represents the ratio of users to resources in the network, while the vertical axis indicates the percentage of improvement in the resource allocation method of the proposed approach, obtained by comparing the utility function of equation (12) with the utility function value of the previous research [14] for the same network conditions. It is worth noting that the mentioned utility functions are both calculated towards optimizing network throughput and hence serve as the basis for comparison in this study.

Fig. (7): The percentage improvement in inter-slice resource allocation in Amres compared to the previous research [14].



As evident from the Fig. (7), in general, dynamic inter-slice resource allocation outperforms static allocation, although with an increase in user density, the degree of improvement in dynamic resource allocation compared to static becomes less pronounced. This is due to network resource saturation, which limits the sharing of resources among different slices. Fig. (8) illustrates the net value of the utility function obtained by the proposed method in this example, showing that as long as the network has not reached saturation, this value relatively increases with rising user density.

Fig. (8): The net value of the utility function relationship (12) in Amres.

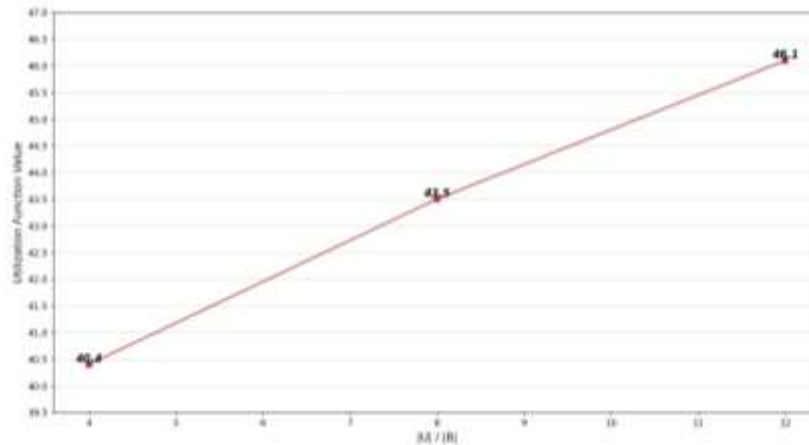
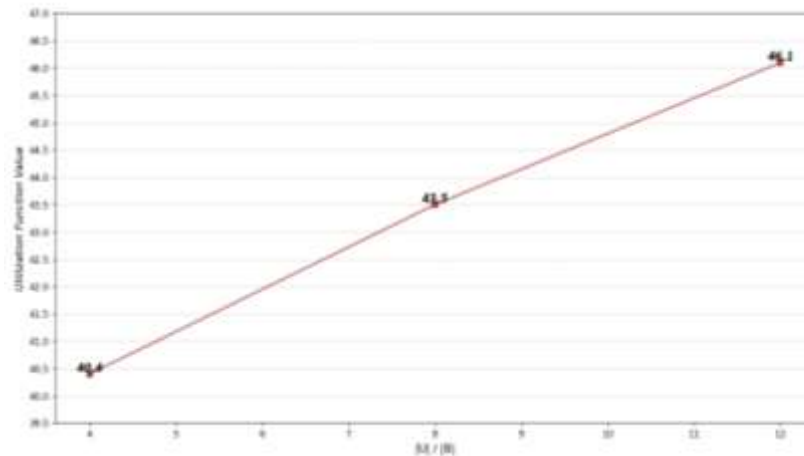


Fig. (9): comparison of the supported users in Amres with [13].

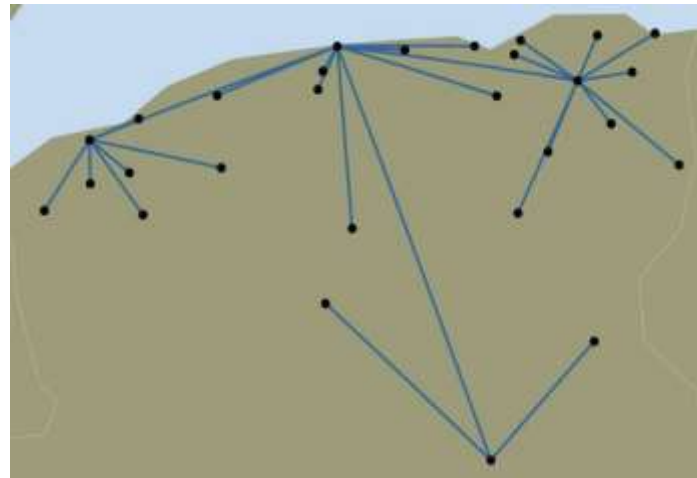


Subsequently, the problem of acceptance control for cuts, in each of the three cases $|U|/|B| = 4.8, 12$, is examined, and the supported user count of the proposed method is compared with the research presented in [13] and depicted in Fig. (9). As evident from this chart, the proposed method relatively outperforms the previous presented method under comparison, attributed to the dynamic resource sharing effect enabling support for a greater number of users in the network.

4-4 - Analytical results in Arn

Fig. (10) displays the network topology of Arn derived from [Internet Topology Zoo [29]. Further details of the analysis are explained subsequently.

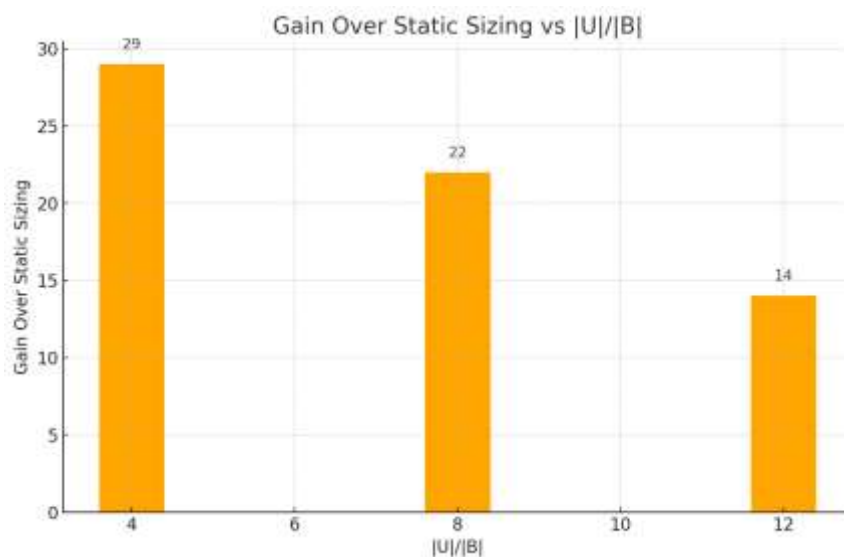
Fig. (10): The network topology, nodes, and links present in Arn [29].



4-4-1- Analysis Details

Fig. (11) depicts the improvement in resource allocation efficiency across segments in the proposed method compared to static resource allocation as highlighted in reference [14]. The horizontal axis in this chart represents the ratio of users to available resources in the network, while the vertical axis indicates the percentage improvement in the resource allocation method of the proposed approach in equation (12) compared to the reference method, obtained under identical network conditions. It is worth noting that the utility functions mentioned are both calculated to optimize network throughput and hence serve as the benchmark for comparison in this study.

Fig. (11): the percentage improvement in inter-slice resource allocation in Arn compared to reference [14].



As evident from Fig. (11), generally, dynamic resource allocation outperforms static allocation. However, as user density increases, the degree of improvement in dynamic resource allocation compared to static allocation diminishes, partly due to network resource saturation, reducing the ability to share resources among different slices. Fig. (12) illustrates the net value of the utility function obtained in the proposed method in

this example, showing that as long as the network remains unsaturated, this value increases relatively with increasing user density. Subsequently, the admission control problem for cuts is investigated in each of the three cases $|U|/|B| = 4.8, 12$, and the supported user count of the proposed method is compared with the reference [13] and illustrated in Fig. (13). As evident from this chart, the proposed method in this instance behaves similarly to the reference method under comparison. This similarity is due to network resource saturation, which practically prohibits resource sharing between cuts, thereby resulting in supporting an equivalent number of users in the network in both cases.

Fig. (12): The net value of the utility function in relation to equation (3-12) in Arn.

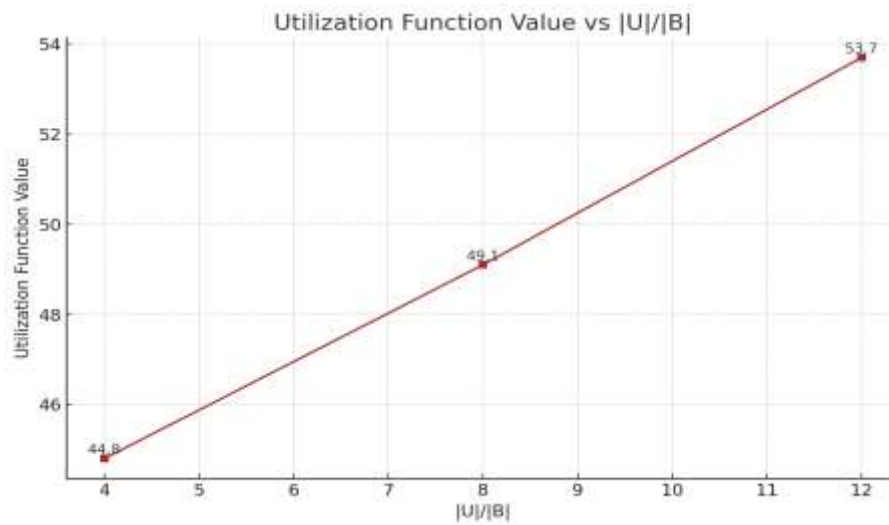
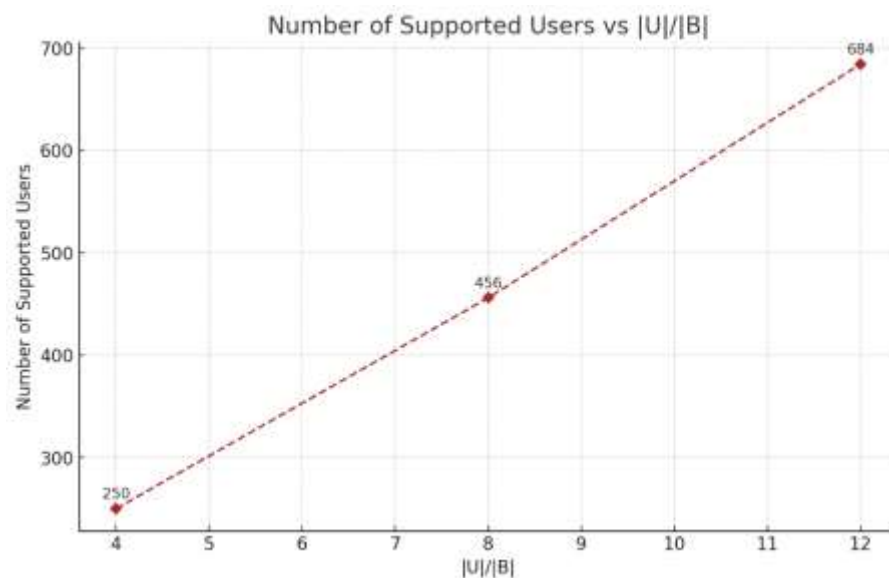


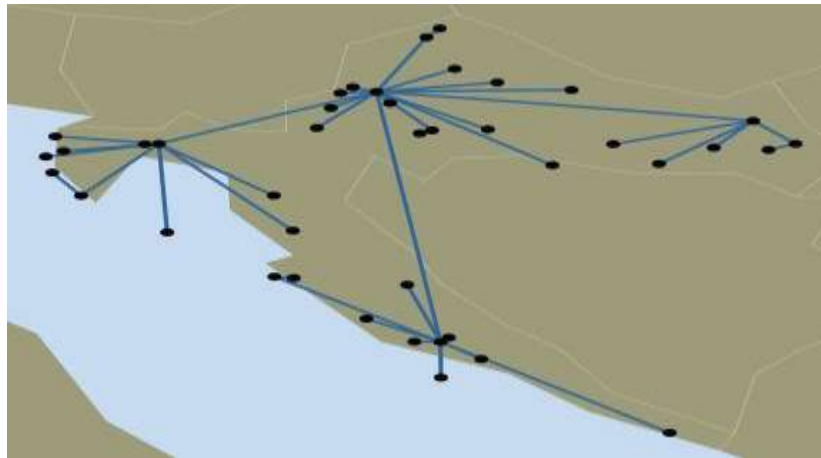
Fig. (13): Comparison of supported users in Arn with [13].



4-5- Analysis Results in Carnet

Fig. (14) displays the network topology of Carnet derived from the Internet Topology Zoo [29]. Further details of the analysis are elucidated.

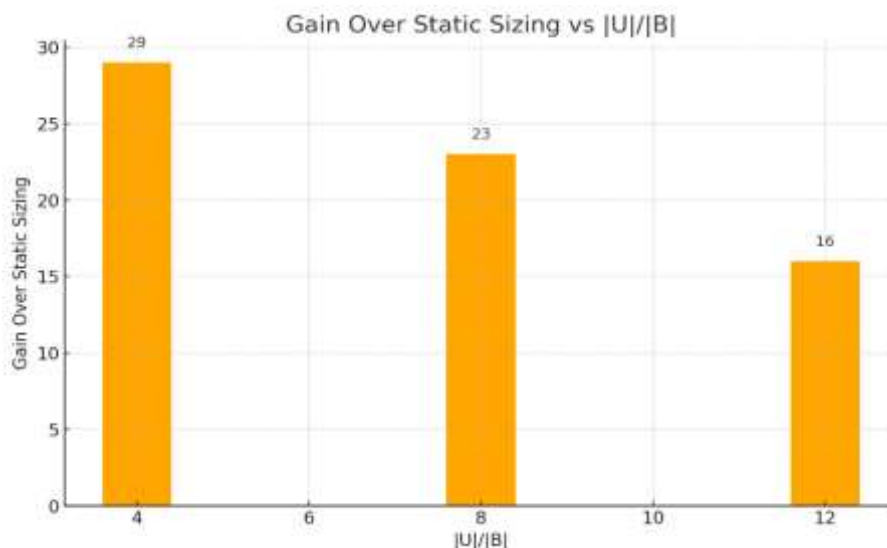
Fig. (14): The network topology, nodes, and links present in Carnet [29].



4-5-1- Analysis Details

Fig. (15) showcases the improvement in resource allocation efficiency across segments in the proposed method compared to the static resource allocation as discussed in reference [14]. The horizontal axis in this chart represents the ratio of users to available resources in the network, while the vertical axis indicates the percentage of improvement in the resource allocation method of the proposed approach. This improvement is obtained through the comparison of the utility function relationship (12) with the utility function value of the reference method under identical network conditions. It is essential to note that the mentioned utility functions are both calculated towards optimizing network throughput, thus serving as the benchmark for comparison in this study.

Fig. (15): The percentage improvement in inter-slice resource allocation in Carnet compared to reference [14].



As evident from Fig. (15), in general, dynamic resource allocation performs better than static allocation in inter-slice resource allocation. However, as user density increases, the degree of improvement in dynamic resource allocation compared to static allocation diminishes, partly due to network resource saturation, which reduces the feasibility of

resource sharing among different slices. Fig. (16) also depicts the net value of the utility function obtained by the proposed method in this example, showing that, until the network reaches saturation, this value relatively increases with increasing user density. Following this, the admission control problem for cuts is examined in each of the three cases $|U|/|B| = 4.8, 12$, and the supported user count of the proposed method is compared with reference [13] and depicted in Fig. (17). As evident from this chart, in this particular scenario, the proposed method behaves similarly to the reference method under comparison. This similarity arises from network resource saturation, which practically precludes resource sharing between cuts, resulting in the support of an equivalent number of users in the network in both cases.

Fig. (16): The net value of the utility function of equation (3-12) in Carnet.

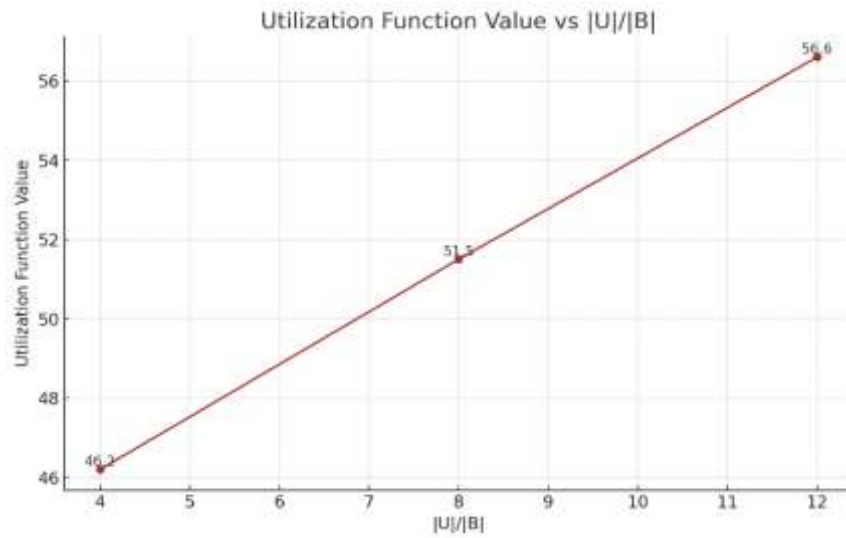
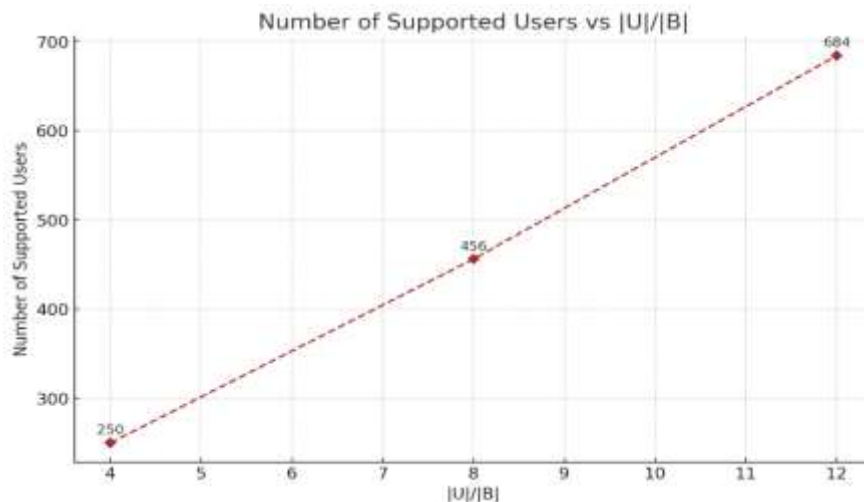


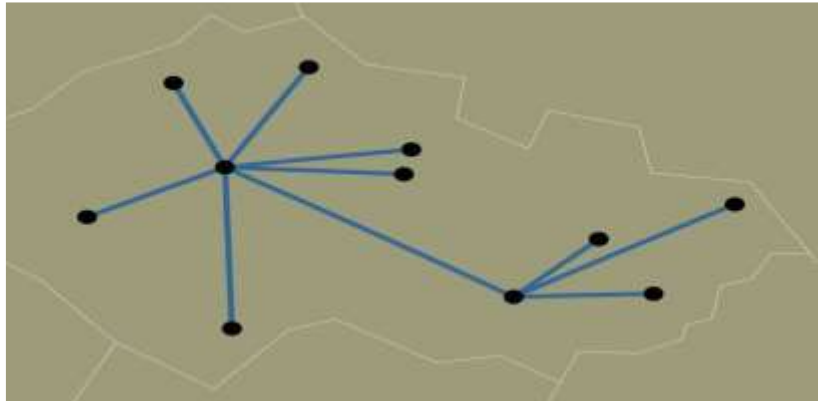
Fig. (17): The supported users in Carnet with [13].



4-6- Analysis Results in Cesnet

Fig. (18) displays the network topology of Cesnet derived from the Internet Topology Zoo [29]. Further details of the analysis are elaborated upon.

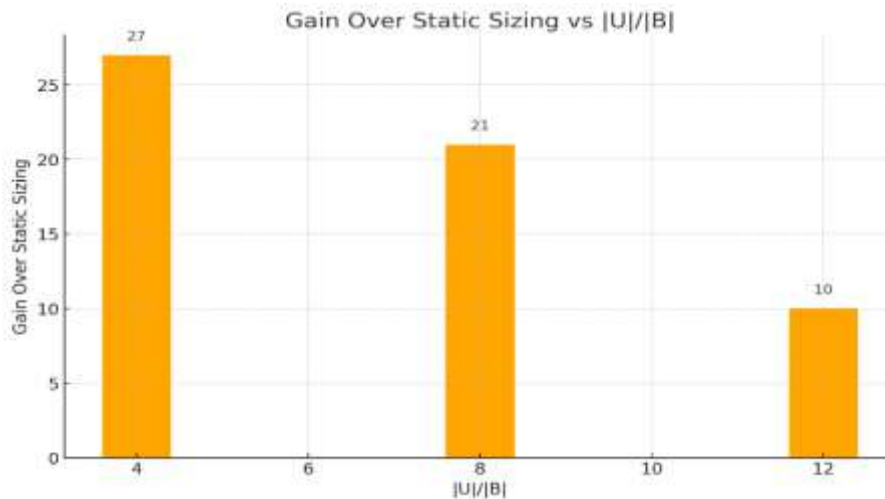
Fig. (18). illustrates the network topology, nodes, and links present in Cesnet [29].



4-6-1- Analysis Details

Fig. (15) displays the level of improvement in inter-slice resource allocation in the proposed method compared to the static resource allocation method presented in reference [14]. The horizontal axis in this chart represents the ratio of users to available resources in the network, while the vertical axis indicates the percentage of improvement in the resource allocation method of the proposed approach. This improvement is derived from comparing the utility function relationship (12) with the utility function value of the reference method under identical network conditions. It is important to note that the mentioned utility functions are both computed to optimize network throughput, hence serving as the benchmark for comparison in this study.

Fig. (19): The percentage improvement in inter-slice resource allocation in Cesnet compared to reference [14].



As illustrated in Fig. (19), resource allocation in the dynamic mode generally outperforms the static mode. However, as user density increases, the improvement achieved by dynamic resource allocation compared to the static mode diminishes. This is primarily due to the saturation of network resources, which limits the possibility of sharing resources among different slices. Fig. (20) further demonstrates the net utility function obtained by the proposed method in this example. It shows that, as long as the network does not reach a state of saturation, this value relatively increases with user density.

Subsequently, the issue of slice admission control was examined for each of the three cases $|U|/|B| = 4, 8, 12$. The number of users supported by the proposed method was compared with the results from reference [13] and presented in Fig. (21). As evident from this figure, the proposed method performs relatively better than the reference method. This improvement is attributed to the effect of dynamic resource sharing, which enables the network to support a larger number of users.

Fig. (20): Net value of the utility function of equation (12) in Carnet

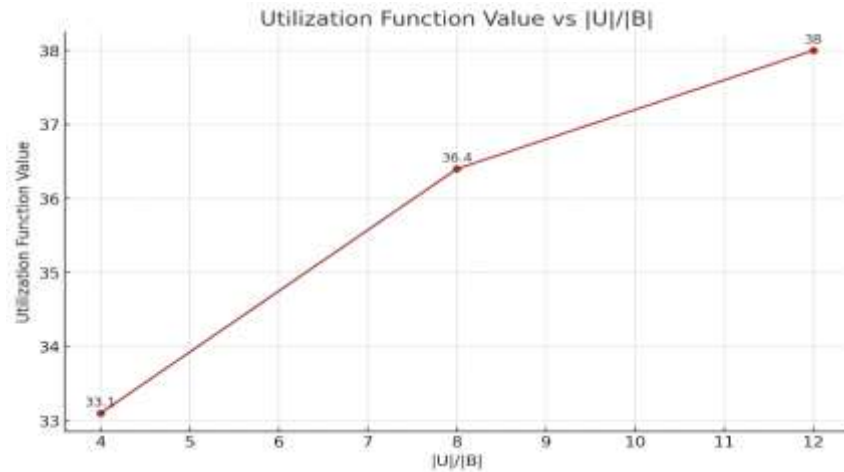
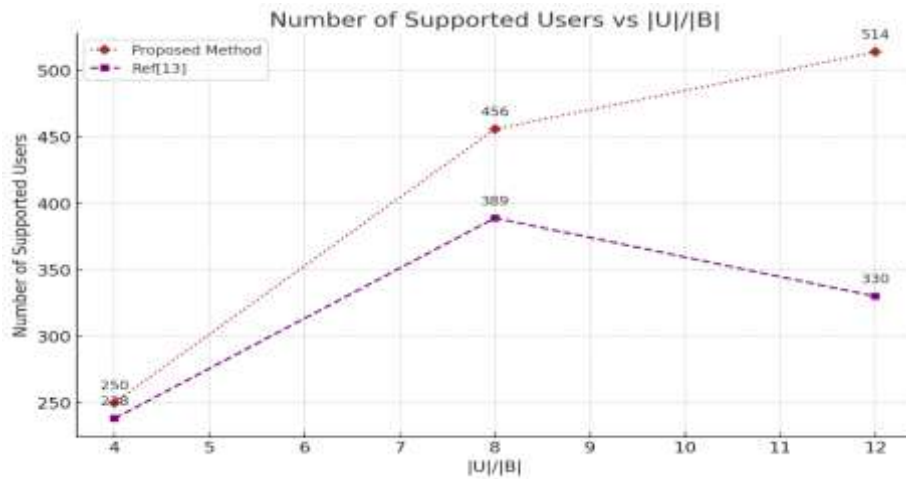


Fig. (21): Comparison of supported users in Cesnet compared to [1]



In this section, the results obtained from the simulation of the proposed method in this study are examined and evaluated. To this end, the outcomes of the simulation for dynamic inter-slice resource allocation were compared with those of the static reference method [14]. Subsequently, the simulations of slice admission control were analyzed, and the number of supported users in the proposed method was compared with the reference method [13]. The results are presented in the corresponding charts. The results indicate that in inter-slice resource allocation, the proposed dynamic method outperforms the static approach under various user densities. However, as user density increases, the efficiency of the dynamic method diminishes. This can be attributed to network resource saturation, which reduces the potential for resource sharing among different slices. On the other hand, a comparison of the number of users supported by the proposed method

and the reference method [13] reveals that, overall, the proposed method accommodates a larger number of users. This improvement is due to the sharing of resources among slices, which allows the network to utilize available resources more efficiently. Consequently, the enhanced resource management translates into a greater number of supported users. The next section will focus on the conclusions of this study and present several potential suggestions related to this paper.

5- Conclusion

The study effectively explores the critical areas of slice admission control and inter-slice resource allocation in the context of 5G network slicing. By integrating these two areas within an auction-based decision-making framework, the research offers a novel approach to resource management. The findings highlight the advantages of dynamic inter-slice resource allocation over static allocation, demonstrating its potential for greater efficiency, particularly in scenarios where resources are shared between slices. This dynamic approach not only improves overall resource utilization but also enhances slice admission control, ensuring more efficient management of user requests for new services. However, the study also reveals that as the number of users and network density increase, the improvements in efficiency become less pronounced, which suggests that additional strategies may be necessary for handling larger-scale deployments. The proposed approach, therefore, contributes significantly to the evolution of resource management in 5G networks, particularly by addressing the interconnectedness of slice admission and resource allocation. Future work can further refine this model, focusing on scalability and optimization as 5G networks expand in size and complexity.

REFERENCES

1. Osseiran++, Afif, Jose F. Monserrat, and Patrick Marsch, eds. 5G mobile and wireless communications technology. Cambridge University Press, 2016.
2. Ict, M. "317669-METIS/D1. Scenarios requirements and KPIs for 5G mobile and wireless system." (2013).
3. International Telecommunications Union Radio (ITU-R), "Requirements related to technical performance for IMT-Advanced radio interface(s)," Report ITUR M.2134, December 2008.
4. Alliance, N. G. M. N. "Description of network slicing concept." NGMN 5G P 1 (2016): 1.
5. Yegnanarayana, Bayya. Artificial neural networks. PHI Learning Pvt. Ltd., 2009.
6. Han, Bin, and Hans D. Schotten. "Machine Learning for Network Slicing Resource Management: A Comprehensive Survey." (2020).
7. Sciancalepore, Vincenzo, Flavio Cirillo, and Xavier Costa-Perez. "Slice as a service (SlaaS) optimal IoT slice resources orchestration." GLOBECOM 2017-2017 IEEE Global Communications Conference. IEEE, 2017.
8. Caballero, Pablo, et al. "Network slicing for guaranteed rate services: Admission control and resource allocation games." IEEE Transactions on Wireless Communications 17.10 (2018): 6419-6432.
9. Kim, Juhoon, and Min Xie. "A study of slice-aware service assurance for network function virtualization." 2019 IEEE Conference on Network Softwarization (NetSoft). IEEE, 2019.
10. Vo, Phuong Luu, et al. "Slicing the edge: Resource allocation for RAN network slicing." IEEE Wireless Communications Letters 7.6 (2018): 970-973.
11. Jiang, Menglan, Massimo Condoluci, and Toktam Mahmoodi. "Network slicing in 5G: An auction-based model." 2017 IEEE International Conference on Communications (ICC). IEEE, 2017.

12. Samdanis, Konstantinos, and Xavier Costa Perez. "Network management system and method for a shared radio access network, RAN, infrastructure." U.S. Patent No. 10,506,471. 10 Dec. 2019.
13. Kammoun, Amal, et al. "End-to-end efficient heuristic algorithm for 5G network slicing." 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). IEEE, 2018.
14. Caballero, Pablo, et al. "Network slicing games: Enabling customization in multi-tenant mobile networks." *IEEE/ACM Transactions on Networking* 27.2 (2019): 662-675.
15. Zhang, Li. "Proportional response dynamics in the Fisher market." *Theoretical Computer Science* 412.24 (2011): 2691-2698.
16. Kelly, Frank P., Aman K. Maulloo, and David KH Tan. "Rate control for communication networks: shadow prices, proportional fairness and stability." *Journal of the Operational Research society* 49.3 (1998): 237-252.
17. Johari, Ramesh, Shie Mannor, and John N. Tsitsiklis. "Efficiency loss in a network resource allocation game: the case of elastic supply." *IEEE Transactions on Automatic Control* 50.11 (2005): 1712-1724.
18. Yang, Sichao, and Bruce Hajek. "VCG-Kelly mechanisms for allocation of divisible goods: Adapting VCG mechanisms to one-dimensional signals." *IEEE Journal on Selected Areas in Com.* 25.6 (2007): 1237-1243.
19. Alliance, N. G. M. N. "Description of network slicing concept." NGMN 5G P 1.1 (2016).
20. 3GPP TR 23.799. "Study on Architecture for Next Generation System." (2016).
21. Caballero, Pablo, et al. "Network slicing for guaranteed rate services: Admission control and resource allocation games." *IEEE Transactions on Wireless Communications* 17.10 (2018): 6419-6432. [22] Bega, Dario, et al. "Optimising 5G infrastructure markets: The business of network slicing." *IEEE INFOCOM 2017-IEEE Conf. on Computer Communications*. IEEE, 2017.
22. Han, Bin, et al. "Admission and congestion control for 5G network slicing." 2018 IEEE Conference on Standards for Communications and Networking (CSCN). IEEE, 2018.
23. Li, Rongpeng, et al. "Deep reinforcement learning for resource management in network slicing." *IEEE Access* 6 (2018): 74429-74441.
24. Bega, Dario, et al. "DeepCog: Cognitive network management in sliced 5G networks with deep learning." *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019.
25. D'Oro, Salvatore, et al. "Auction-based resource allocation in OpenFlow multi-tenant networks." *Computer Networks* 115 (2017): 29-41.
26. Chang, Chia-Yu, Navid Nikaein, and Thrasyvoulos Spyropoulos. "Radio access network resource slicing for flexible service execution." *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2018.
27. Mo, Jeonghoon, and Jean Walrand. "Fair end-to-end window-based congestion control." *IEEE/ACM Transactions on networking* 8.5 (2000): 556-567. Available at <http://www.topology-zoo.org/>
28. 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," TS 36.213, v12.5.0, Rel. 12, Mar. 2015.
29. Thantharate, Anurag, and Cory Beard. "ADAPTIVE6G: Adaptive resource management for network slicing architectures in current 5G and future 6G systems." *Journal of Net. and Systems Management* 31, no. 1 (2023): 9.
30. Singh, Randeep, Abolfazl Mehbodniya, Julian L. Webber, Pankaj Dadheech, G. Pavithra, Mohammed S. Alzaidi, and Reynah Akwafo. "Analysis of network slicing for management of 5G networks using machine learning techniques." *Wireless Communications and Mobile Computing* 2022, no. 1 (2022): 9169568.

31. Alotaibi, Daifallah, Vijey Thayananthan, and Javad Yazdani. "The 5G network slicing using SDN based technology for managing network traffic." *Procedia Computer Science* 194 (2021): 114-121.
32. Saibharath, S., Sudeepta Mishra, and Chittaranjan Hota. "Joint QoS and energy-efficient resource allocation and scheduling in 5G Network Slicing." *Computer Communications* 202 (2023): 110-123.
33. Kim, Yohan, and Hyuk Lim. "Multi-agent reinforcement learning-based resource management for end-to-end network slicing." *IEEE Access* 9 (2021): 56178-56190.