World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(Research Article)

Development of improved disjoint resource allocation for cost-effective and scalable 5G network slicing

Comfort Damilola Oluwatimilehin [1, *], Sagir Lawan [1], Ashraf Adam Ahmad [1], Abdulrazaq Abdulaziz [2] and Chinedum Martins Harbor [1]

[1] *Department of Electrical/Electronic Engineering, Nigerian Defence Academy, Kaduna, Nigeria.*
[2] *Department of Electrical and Electronics Engineering, Usmanu Danfodiyo University Sokoto, Nigeria.*

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(03), 1649-1660

Publication history: Received on 27 April 2025; revised on 08 June 2025; accepted on 11 June 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.15.3.0976

Abstract

Efficient resource allocation and cost management for network slicing have become critical challenges for communication service providers with the rapid proliferation of 5G networks. This paper presents an Improved Disjoint Resource Allocation (IDRA) framework designed to enhance resource management and cost efficiency in 5G network slicing. The IDRA method independently addresses access control and resource allocation sub-problems, leading to optimized utilization of network resources while maintaining quality of service. Comparative analyses against conventional Disjoint Resource Allocation (DRA) and Joint Resource Allocation (JRA) approaches demonstrate that IDRA achieves significantly lower Communication Service Provider (CSP) costs, with stable values around 2400 from the eighth to sixteenth slice requests, outperforming peak costs of 5715 and 2725 recorded by DRA and JRA, respectively. Bandwidth consumption cost under IDRA is markedly reduced and predictable, with a peak of 1,986,000 compared to higher and more oscillatory costs in alternative methods, ensuring better budget management. The framework also maintains manageable increases in power consumption and computing capacity while demonstrating reasonable execution times, confirming its scalability and computational feasibility. The results validate IDRA as a robust, scalable, and cost-effective solution for resource allocation in 5G network slicing, with future work suggested to address dynamic tenant behavior and heterogeneous service demands.

Keywords: 5G Network Slicing; Resource Allocation; IDRA; Communication Service Cost; Bandwidth Consumption; Access Control; Network Optimization

1. Introduction

The rapid evolution of telecommunications has ushered in the era of Fifth Generation (5G) and beyond 5G (B5G) wireless communication systems. These next-generation networks are designed to deliver ultra-high data rates, extremely low latency, and significantly increased capacity, enabling a broad spectrum of innovative applications and services [1]. By building upon the limitations of earlier generations, particularly 4G LTE, 5G introduces transformative capabilities such as seamless high-definition media streaming, real-time video conferencing, and accelerated data transfers [2]. These advancements hold the potential to revolutionize numerous domains, from consumer entertainment to critical industrial systems.

To achieve these capabilities, 5G and B5G networks leverage a suite of enabling technologies. These include millimeter wave (mmWave) frequencies for high-throughput data transmission, dense deployment of small cells to enhance coverage, and massive multiple-input multiple-output (MIMO) systems to boost spectral efficiency and network capacity [3]. Complementary innovations such as Network Function Virtualization (NFV), Software-Defined Networking

* Corresponding author: Comfort Damilola Oluwatimilehin.

(SDN), edge computing, artificial intelligence (AI), and network slicing collectively enhance the flexibility, scalability, and responsiveness of modern communication infrastructures. As depicted in Figure 1.1, the 5G architecture incorporates these elements in a modular and layered structure, supporting diverse application requirements [4].

Among these technologies, network slicing has emerged as a critical enabler of service differentiation and efficient resource management. It involves partitioning a shared physical network into multiple virtualized and logically isolated slices, each customized to meet specific application demands, including bandwidth, latency, and security constraints [5]. This approach allows operators to tailor service delivery for industries such as healthcare, autonomous transport, and smart cities while maintaining strict Quality of Service (QoS) levels across slices [6]. Network slicing thus facilitates coexisting services with heterogeneous performance needs, optimizing resource utilization and system flexibility—key objectives of 5G and B5G systems.

However, the increasing heterogeneity of services and user demands poses significant challenges in resource allocation and management. Traditional approaches often fall short in addressing the dynamic, multi-tenant nature of 5G networks [7–10]. In this context, efficient disjoint resource allocation techniques are essential to ensure optimal performance and QoS provisioning across slices. Current methods suffer from scalability issues and limited adaptability to network fluctuations [11, 12]. This study aims to develop an Improved Disjoint Resource Allocation (IDRA) technique that addresses these limitations, enhances slice management, and outperforms existing solutions such as Joint Resource Allocation (JRA) and conventional Dynamic Resource Allocation (DRA). Through design, implementation, and performance evaluation, this research contributes practical insights for optimizing network slicing in next-generation wireless systems.

The remainder of this paper is structured as follows: Section 2 presents a review of recent related works, focusing on existing resource allocation strategies in 5G network slicing and their associated challenges. Section 3 outlines the methodology adopted in this study, detailing the design and implementation of the proposed Improved Disjoint Resource Allocation (IDRA) technique, along with the system model and performance evaluation parameters. Section 4 discusses the simulation results, comparative performance analysis, and key insights derived from the evaluation of the proposed technique against existing methods. Finally, Section 5 concludes the paper by summarizing the major findings and contributions, while also highlighting potential directions for future research in efficient 5G network slicing and resource management.

2. Review of recent related works

Recent advances in 5G network slicing have explored diverse techniques to optimize resource allocation and guarantee Quality of Service (QoS) across various use cases. For instance, the work in [13] addressed resource coordination in Radio Access Network (RAN) slicing by formulating a biconvex optimization problem to minimize both slice load and backhaul delay. Despite proposing effective minimization algorithms that converged to near-optimal solutions in small-scale scenarios, the complexity arising from interdependent slice coordination remained a limitation. Similarly, [14] focused on dynamic RAN slicing under high-density environments, such as aircraft cabins, formulating an optimization framework that ensured resource isolation and QoS stability. However, the scope was restricted to a software-modeled, domain-specific scenario.

Other researchers have investigated data-driven and learning-based slicing frameworks. In [15], resource sharing efficiency in metropolitan-scale networks was analyzed through empirical traffic data, uncovering how various performance demands impact infrastructure utilization. Meanwhile, [16] proposed a reinforcement learning approach for eMBB and V2X service slices, employing offline Q-learning with softmax decision policies to enhance exploration and exploitation trade-offs. Although promising, their method relied on heuristic rules rather than fully algorithmic designs and lacked scalability across broader multi-traffic scenarios.

End-to-end slicing frameworks have also been proposed to unify RAN and core resource management. For instance, [17] introduced the 5GIIK test bed, integrating multiple open-source tools for managing slices across 4G/5G networks. While impactful for experimentation, the implementation remained limited to testbed validation. Similarly, [18] proposed a Slice-as-a-Service (SlaaS) model using integer linear programming to minimize power and bandwidth costs, with a joint resource and admission management mechanism. The inclusion of an admission control method addressed feasibility issues under resource constraints, demonstrating practicality for cloud service providers managing multiple tenants.

Further innovation has emerged with the integration of AI techniques. In [19], an end-to-end deep Q-learning-based slicing model was proposed to dynamically allocate resources, significantly improving access rates under varying service constraints. Additionally, [20] offered a QoS-centric slicing architecture using SDN/NFV, with tailored algorithms for eMBB, mMTC, and URLLC services. Finally, [21] introduced SliceNetVSwitch, a software-defined datapath enabling real-time multi-tenant slicing with strong performance isolation, achieving up to 8,192 concurrent slices at 10 Gbps. Despite these efforts, most works remain constrained by context-specific scenarios, limited scalability, or lack of integration across all network domains, thereby motivating the need for improved, adaptive, and scalable network slicing mechanisms in 5G and beyond.

Despite the advancements in 5G network slicing, significant research gaps remain, particularly in achieving strict slice isolation, ensuring scalability, and enhancing energy efficiency. Existing approaches often fail to prevent resource interference between slices, leading to performance degradation, especially in multi-tenant or high-density user environments. Moreover, many current solutions are not robust enough to scale effectively with growing user demands and dynamic service requirements. Energy efficiency, though critical for sustainable network operations, is frequently overlooked or insufficiently addressed. Therefore, there is a pressing need for a more comprehensive slicing framework—such as an enhanced Disjoint Resource Allocation (DRA) strategy—that can guarantee slice isolation, scale with user and service growth, and optimize power consumption to support the evolving demands of next-generation 5G networks.

3. Methodology

3.1. System Architecture

The system architecture diagram illustrated in Figure 1. shows a network architecture for end-to-end (E2E) orchestration in a Software-Defined Networking (SDN) environment adapted from [20].

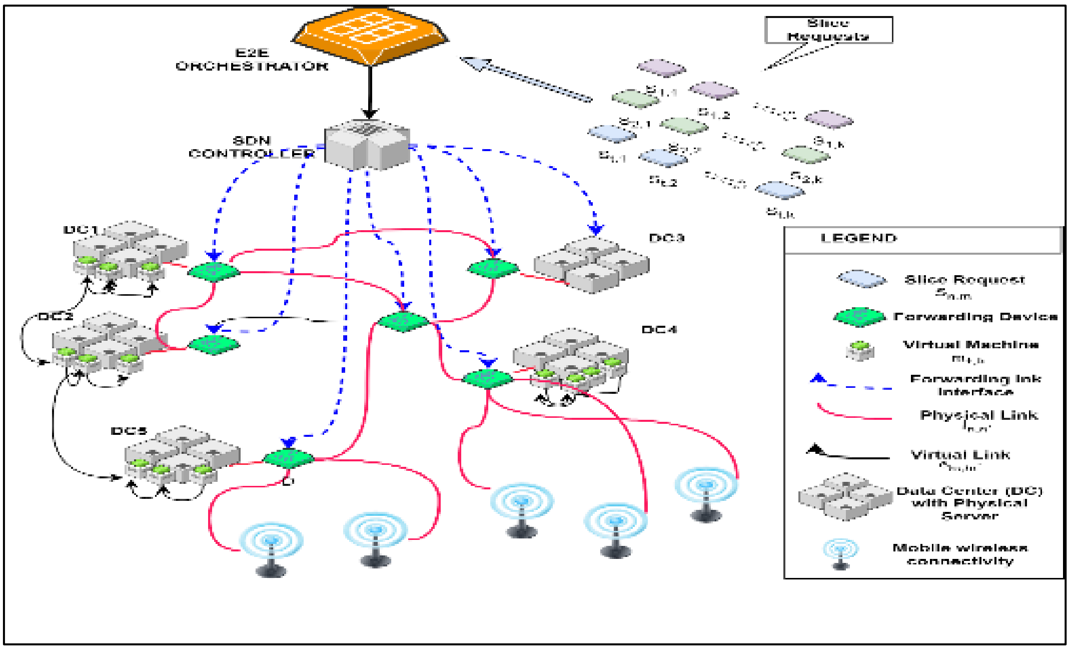

Figure 1 Overall System Architecture

Figure 1 illustrates a comprehensive network slicing architecture within a Cloud Service Provider (CSP) environment, highlighting the interaction of key components across multiple layers. At the top layer, the End-to-End (E2E) Orchestrator oversees the overall orchestration of network resources, while the Software-Defined Networking (SDN) Controller manages the control plane, regulating data flow across the infrastructure. The network infrastructure includes forwarding devices (depicted as circular icons), physical servers (represented by stacked rectangles hosting virtual machines), and both physical and virtual links, shown as solid and dashed lines, respectively. These components are distributed across several interconnected data centers (DC1 to DC5), each containing multiple physical servers supporting virtualized environments. At the bottom of the architecture, antenna symbols signify edge connectivity, enabling wireless and mobile access to the network. On the right side of the diagram, various "Slice Requests" are illustrated, categorized into types (e.g., S1,1; S1,2; S2,1) and levels (e.g., 1st, 2nd Slice Requests), reflecting the concept of network slicing—where the physical network is partitioned into isolated virtual slices tailored to specific service requirements. A legend further clarifies the symbols used, including the number of forwarding devices (n) and the notation for virtual machines (m{i,j}). This architecture enables dynamic and programmable resource management, with the E2E Orchestrator and SDN Controller collaboratively allocating virtual network slices to tenants based on slice requests that specify desired virtual resources, delay tolerance, and data rates. The CSP evaluates these requests against available physical resources, approving or modifying them as needed to ensure efficient and reliable service delivery.

3.2. Disjoint Resource Allocation Technique

Disjoint resource allocation in the context of 5G network slicing involves allocating resources in a way that ensures the isolation of resources between different slices. This is crucial to prevent interference and guarantee the quality of service (QoS) for each slice. Disjoint resource allocation ensures that resources allocated to one slice are not simultaneously used by another, enhancing the overall reliability and performance of the network. First, based on the work presented by [18], the system model for Disjoint Resource Allocation scheme is formulated as follows.

Let's denote a binary variable Y_{usr} to indicate whether user u in slice s is allocated resource r in a disjoint manner

$$Y_{usr}\square\{0,1\} \qquad \text{............ (3.1)}$$

The disjoint resource allocation constraint can be modified as

$$Y_{usr}+X_{usr}\square\ \square\ \square\ \square\ \square\ \square\ \square\ 1 \quad {}_{u}U, \qquad {}_{s}S, \qquad {}_{r}R \text{.......... (3.2)}$$

This constraint ensures that if a resource is allocated to a user in a slice $(X_{usr}=1)$, it cannot be simultaneously allocated to another user in the same slice or any other slice $(Y_{usr}=1)$

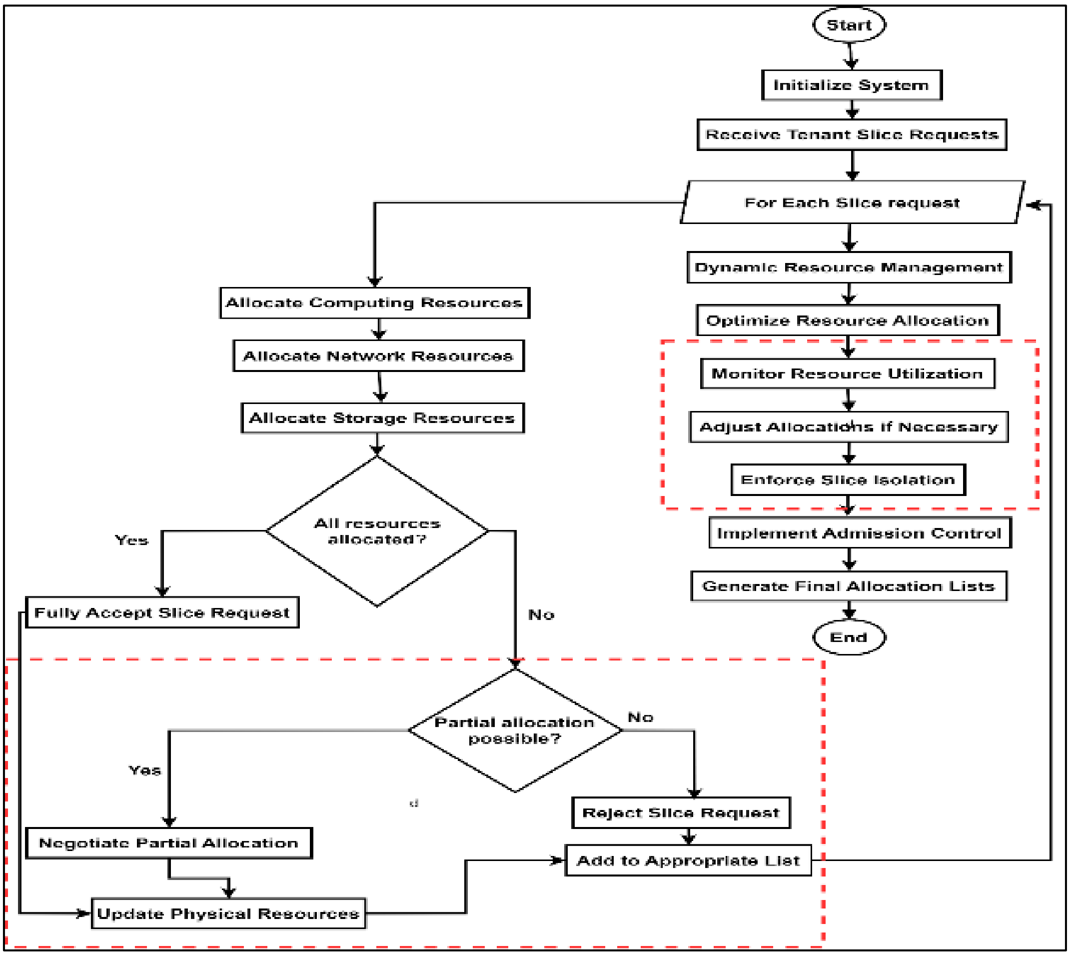

Figure 2 Proposed Improved Disjoint Resource Allocation Scheme flowchart

Disjoint resource allocation is vital for 5G network slicing as it ensures that the resources assigned to one slice do not overlap with those of another slice. This isolation enhances the reliability and performance of each slice, especially in scenarios where stringent QoS requirements must be met.

Implementing disjoint resource allocation ensures that the 5G network slices operate independently, minimizing interference and maximizing the overall efficiency and effectiveness of the network. Figure 2 indicate the areas of improvement such as the resource utilization monitoring, partial allocation and slice isolation policy introduced to ensure slices are allocated in a disjoint fashion to properly meet the QoS and ensure optimized utilization.

The optimization approach to be adopted for this research is the Branch and Bound approach [22], which is generally described as a deterministic global optimization technique. The branch and bound technique are a comprehensive term that encompasses mathematical optimization problems featuring both continuous and discrete/integer variables, as well as linear constraints and an objective function [23]. The branch and bound technique are widely utilized across diverse industries for modeling optimization problems involving integer requirements and linear relationships. Hence, it serves as an adept optimization candidate for the approach delineated in this investigation of network slicing disjoint resource allocation [24]. As it requires less computation power and can be solved disjointedly.

To define the requirements, equations (2.4) to (2.12), describes the disjoint integer, the linear constraints and the objective function.

The overall cost function is defined as [25]

$$C_{\text{Total}}(\boldsymbol{\pi}, \gamma, \boldsymbol{\xi}) = \zeta\beta + \Upsilon \sum_{n\in\mathcal{N}} P_n \quad \text{.............. (3.3)}$$

$\epsilon^{m,' n,}$ respectively. And ζ and Υ are scaling where $\boldsymbol{\pi}, \boldsymbol{\gamma}, \boldsymbol{\xi}$ are the sets of all $\pi p_{nb,n'}, \gamma_{n,} \xi_m$ factors for translating bandwidth and power consumption into cost.

The linear constraints are [26]

$$\text{Cl: } \sum^{n}_{i=1} \gamma_{ij} \leq B_j \quad \text{................. (3.4)}$$

$$\text{C2: } \sum^{n}_{i=1} \gamma_{ij} \geq D_j \quad \text{................. (3.5)}$$

$$\text{C3: } \sum^{n}_{i=1} \xi_{ij} = C_j \quad \text{.................} \quad \text{(3.6)}$$

$$\text{C4: } \sum^{n}_{i=1} \pi_{ij} \leq E_j \quad \text{.................} \quad \text{(3.7)}$$

$$\text{C5: } \sum^{n}_{i=1} \pi_{ij} \geq F_j \quad \text{................. (3.8)}$$

$$\text{C6: } \sum_{i=1}^{n} Aij\ \xi ij + \sum ni=1\ \ \pi ij \leq Gj \quad \text{................. (3.9)}$$

$$\text{C7: } \sum_{i=1}^{n} \xi_{ij} + \sum_{i=1}^{n} \pi_{ij} = H_j \quad \text{................. (3.10)}$$

These constraints are solved for each slice disjointedly to aggregate the maximizing of the overall objective function. And the overall objective function is described as;

$$C\text{Total} = \max \sum_{i=1}^{n}(Iij\ \gamma ij - Jij\ \xi ij - Kij\ \pi i) \quad \text{................. (3.11)}$$

Where;

$\pi_p{}^{b}{}_{n,'}{}^{\epsilon}{}_{m,m'}$: represents a decision variable in the optimization problem. This variable is associated with the routing of virtual links between nodes n and n' through path b in the network. The exponent $\epsilon_{m,'}$ indicates additional characteristics or attributes related to the link.

: The variable γ with subscript n represents a binary decision variable associated with node n in the optimization problem. It is likely used to indicate the activation or inactivation status of node n.

: The variable ξ with subscripts m and n represents a binary decision variable associated with the placement of a virtual machine m at node n in the network. It is used to specify whether a virtual machine is placed at a particular node.

3.3. Access Control Mechanism

The proposed access control mechanism decides which incoming slice requests to accept or reject based on whether they would cause resource infeasibility in the network. The control mechanism ensures the fair allocation of available resource before granting request and passing it to the overall Disjoint resource allocation mechanism. The flowchart presented in Figure 3 provides a detailed implementation procedure.

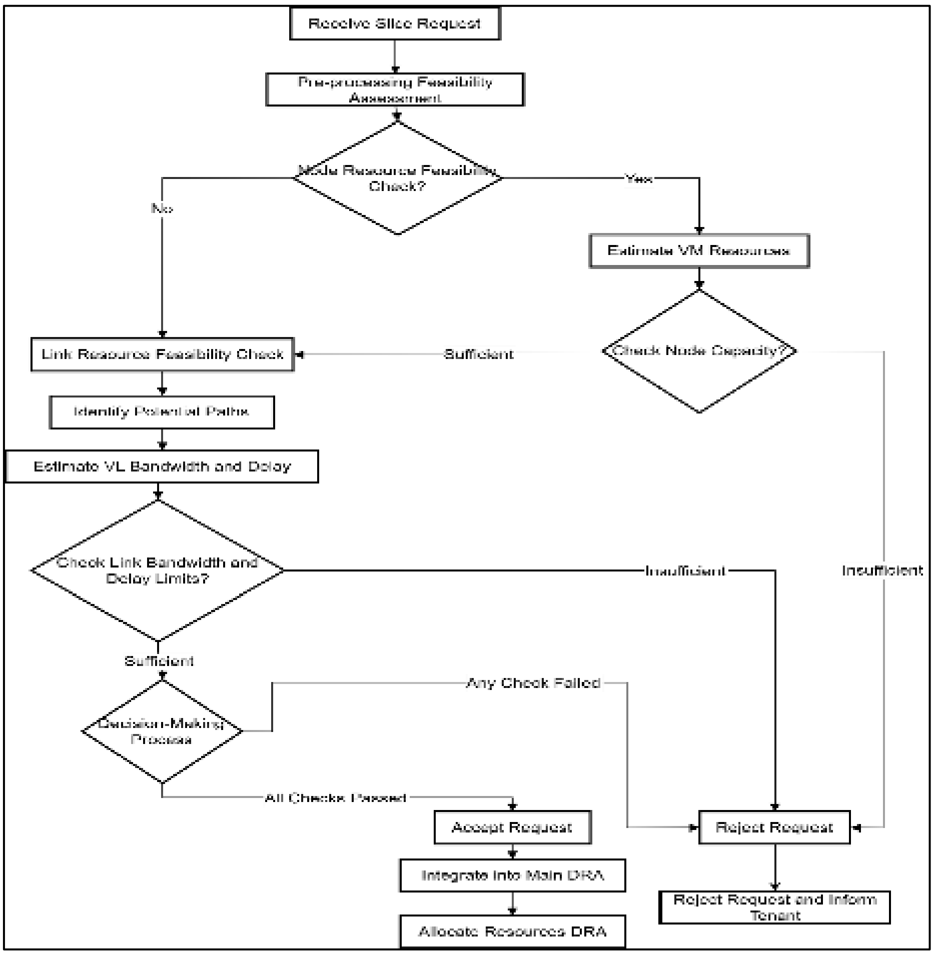

Figure 3 Flowchart of developed Access control integration to Main DRA

A simplified access control mechanism for handling incoming slice requests within a cloud infrastructure. Upon receiving a new slice request specifying virtual machines (VMs), virtual links (VLs), and associated resource demands, the system initiates a preprocessing feasibility assessment to determine whether sufficient resources are available. The first stage involves a Node Resource Feasibility Check, where the required computational, memory, and storage resources for each requested VM are estimated. Specifically, the system calculates the total demands and verifies if the sum of the remaining node resources — computational ($rCom\_n$), memory ($rMem\_n$), and storage ($rSto\_n$) — are adequate to accommodate the new slice without exceeding node capacity constraints. This condition can be expressed as:

$$rCom_n \geq \sum_{m \in \text{VMs}} c \;, \quad rMem_n \geq \sum_{m \in \text{VMs}} m\,em_m \;, \quad rSto_n \geq \sum_{m \in \text{VMs}} s\,to_m \qquad \text{……….(3.12)}$$

where com_m, mem_m, and sto_m represent the computational, memory, and storage requirements of VM m, respectively. If any node fails to meet these conditions, the Node Resource Feasibility Check is deemed unsuccessful.

If the node resources are sufficient, the system proceeds to the **Link Resource Feasibility Check** for each requested VL between VMs m and m'. Here, potential physical paths between nodes are identified based on the network graph (N, L), where

N denotes the set of nodes and L the set of links. For each path, the required bandwidth $e_{m,'}$ is estimated, and the system checks whether each physical link along the path has sufficient residual bandwidth:

$$BW_{ln,'} \geq e_{m,m'} \qquad \text{.............. (3.13)}$$

where $BW_{ln,n'}$ represents the available bandwidth on the physical link between nodes n and n'. Additionally, the propagation delay $\tau_{ln,'}$ along the chosen paths must satisfy the delay constraint:

$$\tau ln,' \leq \tau \text{max}e m,m' \qquad \text{............... (3.14)}$$

where $\tau^{\text{max}em,m'}$ denotes the maximum tolerable delay for the virtual link between VMs m and m'. If either the bandwidth or delay constraint cannot be satisfied, the Link Resource Feasibility Check fails.

Following these feasibility checks, a decision-making process is triggered. If both the Node Resource Feasibility and Link Resource Feasibility Checks pass, the system accepts the slice request and proceeds to integrate it into the main Disjoint Resource Allocation (DRA) solver. The DRA solver then undertakes the allocation of physical nodes and links to the requested VMs and VLs while ensuring resource disjointness and optimizing for cost minimization, in accordance with the Joint Resource Allocation (JRA) model. Conversely, if any of the checks fail, the system rejects the request and informs the tenant accordingly.

This pre-processing step acts as a vital admission control mechanism, preventing overutilization of resources and preserving Quality of Service (QoS) requirements. By rigorously applying resource constraints at an early stage, the system ensures that only feasible and sustainable slice requests are admitted, maintaining the overall efficiency and reliability of the cloud infrastructure. Once the stage is completed successfully, it is then integrated into the main DRA mechanism.

3.4. Performance comparison key indicators and Model Parameters

The performance indicators considered in this research includes several key metrics. Resource Utilization (RU) measures the percentage of allocated resources actually used by the slices. Network Throughput (NT) represents the total amount of data successfully transmitted across the network. Latency (L) assesses the delay experienced during data transmission between users on different slices, which is critical for time-sensitive applications. Power Consumption evaluates the energy consumed during resource allocation among slices, helping to understand the efficiency of the allocation process. Computing Capacity refers to the system's ability to handle and process slice requests within a specific time frame. Bandwidth Improvement (BI) measures the gain in bandwidth allocation relative to a predefined baseline, highlighting the optimization effectiveness. Lastly, the Average Execution Time provides an insight into the processing efficiency and consistency of the algorithm by averaging the total execution time across multiple instances or runs. These performance metrics are applied to assess the proposed Improved Disjoint Resource Allocation (IDRA) technique, alongside Joint Resource Allocation (JRA) and the Conventional Disjoint Resource Allocation (DRA) methods. The detailed comparative results of these techniques based on the identified metrics are presented in Section 4.

To ensure accurate validation of the developed scheme, the simulation and performance evaluation model parameters were adapted from the reference in [18]. The parameter values, notations, and their corresponding descriptions are detailed in Table 3.1. The simulation environment was developed using MATLAB®, where various resource allocation and optimization tasks were implemented. The simulation configuration includes a medium-scale cloud environment with four cloud nodes, each having 7000 MHz computation capacity, 64 GB memory, and 120 GB storage. Each tenant slice requests a fixed number of three virtual machines (VMs), with each VM requiring 1000 MHz computation, 800 MB memory, and 2000 MB storage—parameters selected to reflect real-world application deployment scenarios. Intra-node and inter-node bandwidths were set within variable ranges to simulate diverse network configurations, and the cost and delay associated with data transmission were also parameterized to reflect realistic constraints.

Virtual link parameters, including bandwidth demand and maximum tolerable delay between VMs, were varied to test different 5G application scenarios, ranging from ultra-reliable low latency communications (URLLC) to less delay-sensitive services. The simulation also incorporated cost function weights for balancing power consumption ($\zeta = 9 \times 10^{-5}$) and bandwidth usage ($\Upsilon = 1$), providing a basis for evaluating optimization trade-offs. The scalability of the system was tested with up to 16 tenants, each having a single slice, allowing a focused evaluation of inter-tenant resource dynamics. The simulation was executed on a high-performance Dell Precision 7720 Workstation with 32 GB RAM, Xeon 5 processor, and 2TB SSD storage, running MATLAB® 2023a on Windows® 11, ensuring the computational capacity to handle complex and large-scale optimization problems efficiently.

4. Results and discussion

Several simulation scenarios were carried out while developing and evaluating the system model. The scenarios include varied number of tenants (1 to 16), number of slices per tenant is 1, and the virtual machine for each slice is set at 3. The key performance indices for the system are CSP cost, bandwidth consumption and power consumption. The overall performance of the Improved Disjoint Resource Allocation (IDRA) method is evaluated using several key network cost metrics as the number of slice requests increases from 1 to 16. The total Communication Service Provider (CSP) cost peaks at around 2500 when handling 8 slices and then stabilizes at approximately 2400 up to the maximum of 16 slices, indicating efficient and optimized resource utilization as the number of tenants increases. This behavior reflects the effectiveness of IDRA in managing growing traffic demands without proportionally increasing network costs. The bandwidth consumption cost increases predictably with the number of tenants, displaying a nearly linear trend. This consistency highlights the strength of IDRA's access control mechanism, which enables service providers to predict and manage bandwidth usage effectively while maintaining quality of service. In terms of energy efficiency, the total power consumed by cloud nodes rises gradually with an increase in tenants, demonstrating that IDRA scales energy consumption in a controlled manner. A steady power usage phase occurs between certain tenant counts, evidencing optimization in handling resource requests. The computing capacity requirement peaks early at around 14,000 KHz with 5 tenants and remains constant thereafter, thanks to IDRA's strategy of solving resource allocation and access control problems sequentially, which ensures the computing demand does not exceed available resources. Lastly, the average execution time reaches a maximum of about 105 seconds at full tenant load, illustrating the algorithm's computational efficiency in executing resource allocation and admission control tasks across varying levels of network demand.

To better understand the significance of the results obtained from this research effort and show the improvements offered by the developed Improved Disjoint Resource Allocation (IDRA) approach, the results obtained are compared with the JRA and conventional DRA methodology for validation. Several aspects of the performance will be compared and the significance of the result obtained pointed out. The overall CSP cost to the network provider is a key factor consideration in 5G network slicing, and the developed IDRA approach has shown improvement over the JRA and conventional DRA reported in literature as presented in Figure 4.

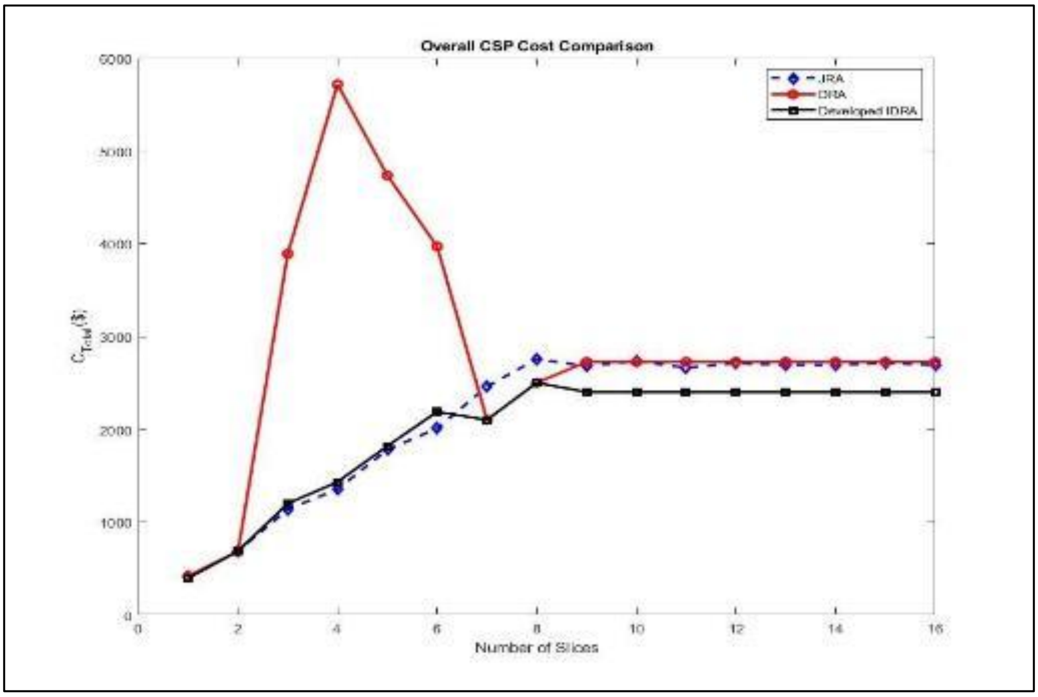

Figure 4 Overall CSP cost comparison for JRA, DRA and IDRA

The most expensive of the CSP cost is the conventional DRA reported in literature as it has a peak cost of around 5715 at 4 slice requests and decreases quickly to match the peak cost of the JRA technique of 2725 at about 9th slice request. Both the JRA and DRA techniques provides and estimated CSP cost of 2725 from slice 9 to 16. The developed IDRA techniques in the research, provides a better and steady CSP cost of 2400 from the 8th slice request to peak request of 16. The highest computed CSP cost under the IDRA approach is 2500, still lower that that recorded by JRA technique. This indicates a better cost consideration and overall CSP cost incurred for transmitting same traffic over the same number of tenant or slice requests.

To compare the bandwidth consumption cost, 3 approaches viz; JRA, DRA and the developed IDRA are compared. The result is presented in figure 5 and 6.

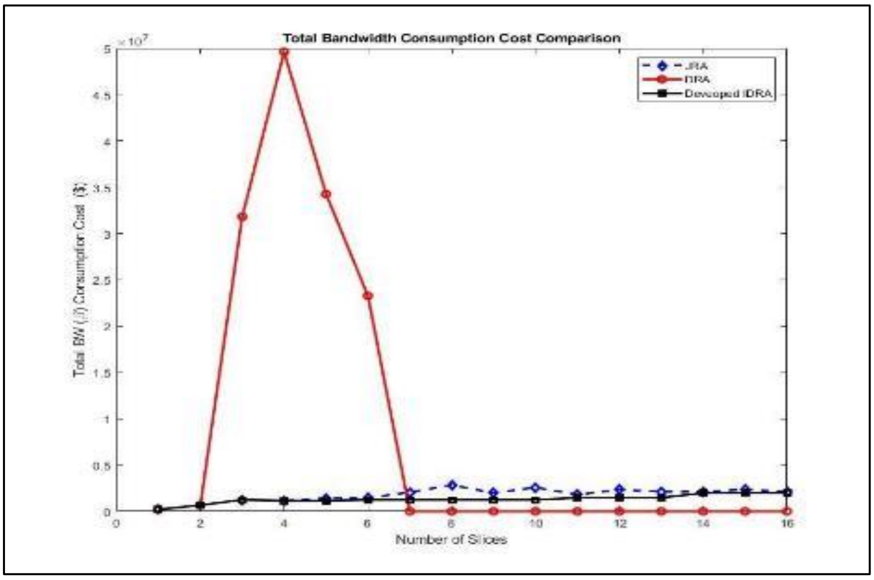

Figure 5 Bandwidth Consumption Cost Comparison for JRA, DRA and IDRA.

The plot in figure 4.7 shows the significant cost incurred by using the conventional DRA method, which records a bandwidth cost of close to 31,823,000 at four slices. This is an accumulation of computational complexity and poor resource utilization experienced while using the conventional DRA approach. The complete fall of the budget cost to 0 from 7th slice in the case of convention DRA, is not an indication of less expense incurred, rather is the failure of request acceptance due to request saturation as reported in literature. This clearly indicates the need to improve the conventional DRA approach, by solving the Access Control and resources allocation sub-problems independently using a better technique. However, it seems from the plot that the bandwidth consumption cost of the JRA and IDRA are almost the same. This is clarified in Figure 6.

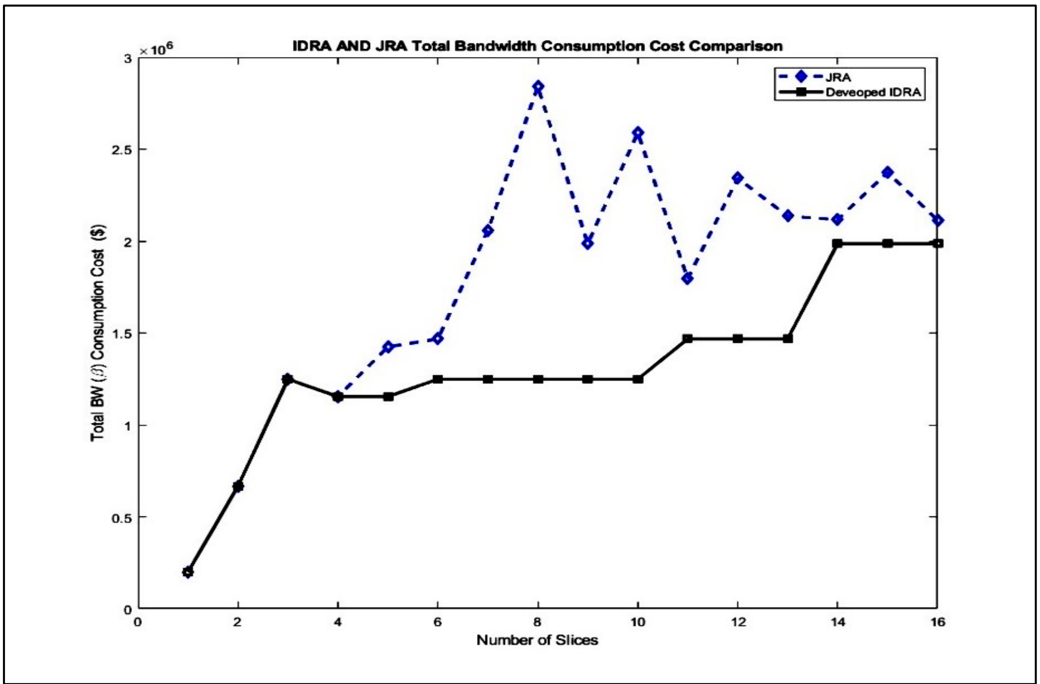

Figure 6 Bandwidth Consumption cost comparison between JRA and IDRA

In Figure 6, the overall superior cost performance of the developed IDRA is evident, as the bandwidth cost under JRA peaks at 2,842,000 while IDRA peaks at 1,986,000 for slice 8 and 14 respectively. This indicates the smooth scalable increase in consumption cost in the case of IDRA which is predictable and therefore budget-able, while the oscillating bandwidth cost occasioned using JRA provides less stability and cost planning for slice request provision.

5. Conclusion

The proposed Improved Disjoint Resource Allocation (IDRA) method demonstrated substantial improvements over conventional Disjoint Resource Allocation (DRA) and Joint Resource Allocation (JRA) techniques in the context of 5G network slicing. The IDRA approach achieved a consistently lower and more stable Communication Service Provider (CSP) cost, maintaining approximately 2400 from the eighth to the sixteenth slice request, compared to peak costs of 5715 and 2725 observed with DRA and JRA, respectively. This indicated enhanced resource utilization and overall cost efficiency. Additionally, the bandwidth consumption cost under IDRA is significantly reduced and more predictable, with a peak of 1,986,000 compared to 2,842,000 for JRA and substantially higher values for conventional DRA, reflecting the effectiveness of solving access control and resource allocation sub-problems independently. The method also exhibited controlled increases in power consumption and computing capacity, demonstrating optimized management of network resources. Furthermore, the average execution time remained reasonable at maximum tenant loads, confirming the computational feasibility of the approach. These findings collectively validate the technical robustness, scalability, and cost-effectiveness of the IDRA framework for resource allocation and admission control in 5G networks. Future research may focus on extending the IDRA model to incorporate dynamic tenant behavior and heterogeneous service requirements to further enhance adaptability and performance in real-world deployment scenarios.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] D. J. Chaithanya and S. M. Dr.Anitha, "Fundamentals of 5g Technology for Next Generation Networks," Interantional Journal of Scientific Research in Engineering and Management, 2022, doi: 10.55041/ijsrem12085.

[2] T. I. Băjenescu, "5g and 6g Evolution," Journal of Engineering Science, 2022, doi: 10.52326/jes.utm.2022.29(2).07.

[3] N. Arora, "Issues and Challenges in 5G Mobile Network," International Journal for Research in Applied Science and Engineering Technology, 2022, doi: 10.22214/ijraset.2022.48327.

[4] J. Hartley, "An Introduction to 5G Architecture," (in English), Just Enough Architecture, 2020/10/30/ 2020. [Online].

[5] S. Wijethilaka and M. Liyanage, "Survey on network slicing for Internet of Things realization in 5G networks," IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 957-994, 2021.

[6] F. S. Samidi, R. Nurul Asyikin Mohamed, W. H. Ahmad, A. Fairuz, J. Md. Zaini, and I. Ahmad Fauzi, "5G New Radio: Dynamic Time Division Duplex Radio Resource Management Approaches," Ieee Access, 2021, doi: 10.1109/access.2021.3104277.

[7] M. S. J. Singh, W. S. W. Saleh, A. T. Abed, and M. A. Fauzi, "A Review on Massive MIMO Antennas for 5G Communication Systems on Challenges and Limitations," Jurnal Kejuruteraan, 2023, doi: 10.17576/jkukm-2023-35(1)-09.

[8] A. Mohammed Salah, W. Hua, H. Dongxuan, K. Ziqi, and K. Agnes, "Intelligent Network Slicing in V2X Networks – A Comprehensive Review," Journal of Artificial Intelligence and Technology, 2023, doi: 10.37965/jait.2023.0208.

[9] U. Deepak, "A Review of Network Slicing Based 5G," 2023, doi: 10.1063/5.0154561.

[10] B. Ishan, K. Neeraj, T. Sudhanshu, T. Sudeep and H. Zhu, "A Systematic Review on NOMA Variants for 5G and Beyond," Ieee Access, 2021, doi: 10.1109/access.2021.3081601.

[11] A. Sufyan, K. B. Khan, O. A. Khashan, T. Mir, and U. Mir, "From 5G to Beyond 5G: A Comprehensive Survey of Wireless Network Evolution, Challenges, and Promising Technologies," Electronics, 2023, doi: 10.3390/electronics12102200.

[12] I. Szymon, "Challenges Regarding Network Neutrality for Commercial Deployment of 5G Networks," Nordic and Baltic Journal of Information and Communications Technologies, 2022, doi: 10.13052/nbjict1902-097x.2022.001.

[13] P. L. Vo, M. N. Nguyen, T. A. Le, and N. H. Tran, "Slicing the edge: Resource allocation for RAN network slicing," IEEE Wireless Communications Letters, vol. 7, no. 6, pp. 970-973, 2018.

[14] A. Papa, M. Klugel, L. Goratti, T. Rasheed, and W. Kellerer, "Optimizing dynamic RAN slicing in programmable 5G networks," in ICC 2019-2019 IEEE International Conference on Communications (ICC), 2019: IEEE, pp. 1-7.

[15] M. C. Márquez Colás, M. Gramaglia, M. Fiore, A. Banchs Roca, and X. Costa Pérez, "Resource sharing efficiency in network slicing," 2019.

[16] H. D. R. Albonda and J. Pérez-Romero, "An efficient RAN slicing strategy for a heterogeneous network with eMBB and V2X services," IEEE access, vol. 7, pp. 44771-44782, 2019.

[17] A. Esmaeily, K. Kralevska, and D. Gligoroski, "A cloud-based SDN/NFV testbed for end-to-end network slicing in 4G/5G," in 2020 6th IEEE Conference on Network Softwarization (NetSoft), 2020: IEEE, pp. 29-35.

[18] S. Ebrahimi, A. Zakeri, B. Akbari, and N. Mokari, "Joint resource and admission management for slice-enabled networks," in NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium, 2020: IEEE, pp. 1-7.

[19] T. Li, X. Zhu, and X. Liu, "An end-to-end network slicing algorithm based on deep Q-learning for 5G network," IEEE Access, vol. 8, pp. 122229-122240, 2020.

[20] Z. Shu and T. Taleb, "A novel QoS framework for network slicing in 5G and beyond networks based on SDN and NFV," IEEE Network, vol. 34, no. 3, pp. 256-263, 2020.

[21] A. Matencio-Escolar, Q. Wang, and J. M. A. Calero, "SliceNetVSwitch: Definition, design and implementation of 5G multi-tenant network slicing in software data paths," IEEE Transactions on Network and Service Management, vol. 17, no. 4, pp. 2212-2225, 2020.

[22] S. D. Gupta, B. P. Van Parys, and E. K. Ryu, "Branch-and-bound performance estimation programming: A unified methodology for constructing optimal optimization methods," arXiv preprint arXiv:2203.07305, 2022.

[23] L. Huang, X. Chen, W. Huo, J. Wang, F. Zhang, B. Bai and L. Shi, "Branch and bound in mixed integer linear programming problems: A survey of techniques and trends," arXiv preprint arXiv:2111.06257, 2021.

[24] B. Xiang, J. Elias, F. Martignon, and E. Di Nitto, "Joint planning of network slicing and mobile edge computing: Models and algorithms," IEEE Transactions on Cloud Computing, 2021.

[25] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, and E. C. Sewell, "Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning," Discrete Optimization, vol. 19, pp. 79-102, 2016.

[26] N. Forget, S. L. Gadegaard, K. Klamroth, L. R. Nielsen, and A. Przybylski, "Branch-and-bound and objective branching with three or more objectives," Computers & Operations Research, vol. 148, p. 106012, 2022.