IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 1 February 2024, accepted 28 March 2024, date of publication 16 April 2024, date of current version 23 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3384562

RESEARCH ARTICLE

Framework for Optimized Resource Allocation in Multi-User, Multi-Service, Multi-Device Aerial Networks

MUHAMMAD IRFAN MUSHTAQ[1], OMER CHUGHTAI[1], (Member, IEEE), YUDONG ZHANG[2,3], (Senior Member, IEEE), ALI H. ALENEZI[4], (Member, IEEE), AND NAYEF ALQAHTANI[5], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, COMSATS University Islamabad, Wah Campus, Wah Cantt 47040, Pakistan
[2]School of Computing and Mathematical Sciences, University of Leicester, LEI 7RH Leicester, U.K.
[3]School of Computer Science and Engineering, Southeast University, Nanjing 210096, China
[4]Remote Sensing Unit, Electrical Engineering Department, Northern Border University, Arar 76321, Saudi Arabia
[5]Department of Electrical Engineering, College of Engineering, King Faisal University, Al-Hofuf, Al-Ahsa 31982, Saudi Arabia

Corresponding author: Omer Chughtai (omer.chughtai@cuiwah.edu.pk))

This work was supported by the Deanship of Scientific Research, Northern Border University, Arar, Saudi Arabia, through the research work under Project NBU-FFR-2024-2159-02.

ABSTRACT With the increasing prevalence of multi-user, multi-service, and heterogeneous multi-device environments, there is a need to address the imperative for efficient resource allocation in contemporary wireless networks, such as those involving unmanned aerial vehicles (UAVs) or drones. In this regard, this work addresses the challenges within a five-dimensional heterogeneous wireless network model, focusing on diverse services such as Big Data Analytics, Video Rendering, and Computer-Aided Design, and the allocation of resources among heterogeneous devices, including UAVs, tethered balloons, and multi-rotors. The resource allocation is facilitated through multiple interfaces like LTE, Wifi, LoRa, and Sigfox, catering to the diverse needs of users operating in aerial Networks. Additionally, this work introduces a novel Intelligent Relaxation using the Penalty Function (IRPF) approach for resource allocation, treating it as an integer programming problem to balance user needs while ensuring affordability. A comparative analysis is conducted between the proposed approach and the traditional branch-and-bound algorithm. In scenarios requiring resource allocation for numerous services based on user demand and device capabilities, the proposed work presents a penalty-based integrality gap solution adept at managing fractional values. The resulting optimization framework is meticulously designed to minimize activation and operating costs while optimizing utility. Additionally, the computing efficiency of the proposed approach is demonstrated by extensive simulations that prove its superiority over the traditional algorithm. Consequently, this research emphasizes the essential role of the proposed model in navigating the intricate challenges of resource allocation in modern drone-centric wireless networks.

INDEX TERMS Resource allocation, aerial networks, multi-user environments, heterogeneous devices, intelligent relaxation, penalty function, utility minimization, service diversity.

I. INTRODUCTION

Heterogeneous network (HetNet) [1] is a term used for modern mobile communication networks that incorporate diverse and varied elements, such as network nodes, technologies, protocols, and components, to provide effective and efficient connectivity. Unlike traditional homogeneous networks, which typically use uniform infrastructure and technology, HetNets are designed to integrate various types of base stations, frequency bands, and communication protocols to enhance the performance and efficiency of wireless communication networks. HetNet deployment is common

The associate editor coordinating the review of this manuscript and approving it for publication was Irfan Ahmed.

in modern wireless systems, particularly in 4G LTE (Long-Term Evolution) and 5G networks, enhancing performance, boosting capacity, and improving service quality [2] in different scenarios, including urban and rural environments. These networks exhibit characteristics of multi-tiered [3], or multi-dimensional [4] architectures, ensuring seamless communication across a variety of interconnected devices.

The evolution of new technologies, responding to the diversity of HetNets, results in a growing number of users engaging in flexible services such as video streaming [5], augmented reality (AR) [6], virtual reality (VR), and e-gaming [7]. To support these trends, robust computing devices are necessary. These computing devices can be strategically placed on aerial platforms, such as multi-rotors [8] and tethered balloons [9], as well as fixed infrastructure, like base stations [10]. The goal is to enhance connectivity and communication in diverse environments, following the demand for the applications. However, satisfying the demands of different users and applications in HetNets simultaneously is challenging, which emphasizes the necessity for effective resource management. Therefore, an effective method is required to optimize resource utilization to improve network performance by considering several criteria and the combination of different communication protocols and technologies. Effective techniques for multi-criteria decision-making in [11] and compute-intensive application placement strategies in [12], may provide valuable insights to tackle challenges related to efficient resource management for applications that require high computation.

The contemporary computing paradigm in wireless networks [13] integrates cloud, fog, edge, and mist computing models to create a versatile architecture, optimize resource allocation (the process of allocating resources efficiently), and provide scalable, real-time solutions for diverse applications. Nevertheless, compute-intensive applications [14] may not always perform optimally in the cloud, particularly for time-sensitive tasks. Exploring fog and edge computing offers an alternative to managing substantial data bandwidth for end devices [15]. However, offloading tasks to devices integrating cloud, fog, and edge computing introduces challenges such as latency, resource constraints, and interoperability issues, which arise when application and user demands are overlooked. Resource allocation in these environments, involving tasks like load balancing, provisioning, and scheduling, becomes a critical consideration. Addressing these concerns is vital for optimizing task offloading and system performance within modern wireless networks. Effective resource allocation to satisfy varied user needs and the computing capacities of different devices is made more challenging by the dynamically changing nature of multi-user, multi-service, and heterogeneous multi-device platforms.

Addressing the escalating demand for services like big data analytics, video rendering, and computer-aided design, this paper focuses on resource allocation across diverse devices such as unmanned aerial vehicles, tethered balloons, and multi-rotors. It makes notable contributions by tackling challenges in a five-dimensional heterogeneous wireless network model, organizing users to services, and linking services to computing devices for efficient resource accommodation. Introducing Intelligent Relaxation using the Penalty Function (IRPF) Method, this research treats resource allocation as an intricate integer programming problem. Through a comparative analysis with the traditional branch and bound algorithm, the study evaluates the efficacy of the novel approach. IRPF serves as an integrality gap penalty-based approach, demonstrating flexibility for fractional values in resource distribution for various services based on user demand and device capabilities. The resulting multi-tiered framework is meticulously designed to minimize activation and operating costs while optimizing utility. The major contributions of this paper are outlined as follows:

- Proposed a specialized network model for efficient resource allocation in diverse wireless environments to address the optimization problem associated with activation and operational costs.
- Emphasized relevance to modern networks in addressing challenges of resource allocation in contemporary wireless networks with multi-user, multi-service, and heterogeneous environments.
- Implemented advanced optimization technique for precise resource allocation in the complex five-dimensional heterogeneous wireless network model.
- Conducted comparative analysis with the traditional branch-and-bound algorithm, showcasing the efficiency of the proposed IRPF approach in diverse resource allocation scenarios.

The remainder of the paper includes Section II, which provides a concise literature review. Section III introduces the system model proposed in this paper. Section IV presents simulation results and discussions. Finally, Section V concludes the paper and outlines directions for future research.

II. LITERATURE REVIEW

Next-generation heterogeneous aerial networks have experienced incredible growth in recent years, as a result of the rapid growth of wireless networks and the increasing demand for seamless connectivity. Efficient resource allocation is crucial in the multi-tiered heterogeneous wireless network, with various users demanding flexible services [5]. High computational services require computing devices [8] with high processing power. Each computing device has its interfaces [16], and each interface has its heterogeneous resources [17]. Challenges include efficient spectrum utilization, interference management, and latency reduction [18].

To address these challenges more effectively, researchers have investigated multi-dimensional, multi-criterion utility optimization techniques in HetNets. Specifically, the authors in [19] have addressed optimized resource allocation to achieve optimal power management and tier assignment within a multi-tier HetNet. This optimization aims to maximize user association and increase the average number of connected users and devices in a specific tier. However, aggressively maximizing user association might lead to higher power consumption, potentially causing challenges in achieving optimal power management goals. The author of [20] addressed resource allocation and financial challenges in Mobile Edge Computing (MEC), emphasizing user satisfaction in Quality of Service (QoS). Formulating the Edge Resource Allocation Problem (ERAP) as a Mixed-Integer Linear Program (MILP) was innovative, but scalability concerns arose due to its NP-hard nature. However, the potential challenge with this technique lies in its computational complexity, particularly in large-scale MEC systems.

To provide a comprehensive overview of the various methodologies and algorithms discussed in the literature review, we present Table 1. This table summarizes key aspects such as Resource Management, Objective Function, Challenges, and Limitations associated with each approach. In [26], the authors introduced a hybrid algorithm that combined gray wolf and genetic algorithms to tackle heterogeneous resource allocation problems. The algorithm aimed to mitigate local optimization issues, leading to reduced energy consumption and latency. However, potential challenges emerged concerning algorithm scalability and adaptability to dynamic environments. In [3], the authors introduced an optimization framework for fog nodes, operators, and subscribers. They utilized the Stackelberg game algorithm for DSS resource allocation analysis. However, the challenge with this algorithm include computational complexity and potential suboptimality. The investigation in [33] optimized RAN for energy efficiency, reducing interference, and addressing power, bandwidth, and cache concerns.

The authors of [34] introduced a cloud and edge computing-based framework incorporating collaborative computation offloading and resource allocation to optimize system profit while adhering to response time limits. Challenges include potential issues in the resource allocation technique for virtual machines, impacting processing times. To address the resource allocation problem in cloud computing, [29] proposed a genetic algorithm. However, unequal resource allocation for virtual machines could lead to extended processing times. In [35], rising resource prices complicate infrastructure-as-a-service (IaaS) network management. The problem here is that users striving to maximize utility during cloud resource acquisition may face complexities due to budget and resource needs.

To address resource allocation challenges, [11] explored multi-criteria methods for cost-based drone selection, focusing on optimizing selection criteria based on cost. In [12], the emphasis shifted to deploying fog applications for optimizing the quality of experience (QoE), utilizing a multi-criteria approach that considered both time and cost. Proposing BEHAVE in [30], the authors targeted efficient utility optimization in multi-tier wireless networks, specifically allocating edge resources to heterogeneous IoT devices. Meanwhile, in [36], innovative scheduling algorithms were introduced to optimize resource allocation within the frequency domain, addressing challenges in wireless networks. The challenges in Next Generation Wireless Networks (NGWN), such as limited radio resources and unreliable terminals, were thoroughly discussed in [37].

In [38], a two-phase task offloading technique was introduced to limit job outages, addressing challenges in service continuity. Simultaneously, [31] proposed algorithms to minimize energy consumption during task allocation, focusing on efficient resource use. However, challenges in the execution efficiency of the proposed binary computation offloading technique were encountered. Magnetic RAM (MRAM), discussed in [4], aimed to free up IoT device resources but faced challenges in ensuring efficient fog-based services, posing potential hurdles in service optimization. Moreover, fog computing, proposed for low-latency data services in [4], faced challenges in service efficiency. In [25], a genetic algorithm addressed challenges in efficient resource allocation for HetNets, striving to optimize resource utilization despite inherent complexities.

Efficient access to the spectrum poses multiple challenges, complicating wireless network setups [18]. Optimizing the spread spectrum within cognitive spaces is crucial for realizing cognitive radio benefits. However, challenges in interference minimization and data rate maximization persist. Cloud computing offers advantages, but network congestion can increase latency. To address this, fog computing deploys resources at end-user edges. The Enhanced Dynamic Resource Allocation Method (EDRAM) proposed by [32] addresses load balancing using particle swarm optimization (PSO). EDRAM minimizes task waiting time, latency, and network bandwidth consumption, enhancing QoE. Video streaming, a high-computation application, targets user satisfaction and utility improvement.

Given the challenges identified in the literature review, addressing the increasing demand for services like big data analytics, video rendering, and computer-aided design necessitates an efficient resource allocation strategy within a five-dimensional heterogeneous wireless network model. There is a need for a model that organizes users into services and establishes connections between services and computing devices for optimal resource accommodation. The effectiveness of this resource allocation relies on the seamless integration of computing devices with high processing power, considering their interfaces and heterogeneous resources to meet diverse requirements.

III. PROPOSED INTELLIGENT RELAXATION METHOD USING THE PENALTY FUNCTION

In this section, we present the system model and corresponding problem formulation for optimizing multi-dimensional, multi-criterion utility in next-generation heterogeneous aerial networks based on the proposed intelligent relaxation using the penalty function (IRPF). Users in the system, requesting various services such as video streaming, podcasting, and augmented reality, drive the demand for computing resources.

TABLE 1. Summary of key aspects in reviewed algorithms for efficient resource management in multi-user, multi-service, and heterogeneous multi-device networks.

Ref No.	Algorithm	Resource Management Aspect	Objective Function	Challenges	Limitations
[21]	Deep Reinforcement Learning (DRL) -based twin actor Deep Deterministic Policy Gradient (DDPG) algorithm	Resource Allocation (Communication, computing, and caching)	Maximize the utility and to ensure Quality-of -Service (QoS)	Effective integration of network slicing and multi-access edge computing is a challenging task	DDPG and Lyapunov optimization (LO) do not estimate the hidden nodes states assuming that all nodes are reliable [22]
[23]	Branch-and-bound Algorithm	Scheduling (multi-energy system scheduling)	To reduce the Cost and Computational complexity	MEC scheduling problem with uncertainties due to renewable generation and demand	Limited processing power and storage capacity
[24]	Game theory algorithm with and without transmission power	Spectrum Allocation	(1) Fair spectrum allocation (2) Maximize spectrum utilization (3) Priority among sensor data (4) Reduce spectrum handoff	The problem of the centralized spectrum allocations in WSN of moderate size	In centralized radio networks, the failure of a single node can significantly disrupt the entire network
[25]	Genetic algorithm	Resource Allocation	To maximize the spectrum efficiency for the 5G mobile networks	To mitigate the Inter-cell interference for 5G HetNets	A tradeoff between energy efficiency and total achieved throughput exists
[26]	Hybrid algorithm (gray wolf and genetic algorithms)	Resource Allocation (Devices and fog nodes)	To reduce energy consumption and latency	Getting stuck in local optimizations	The energy required for convergence is increased by increasing the number of devices. Genetic Algorithm is a robust optimization technique and possesses large-scale computational applications [27]. Premature convergence and a tendency to stagnation in local optima in GWO [28]
[29]	Genetic Algorithm	Resource Allocation (CPU and memory)	Load balancing and reduces Cloud's processing time	Unequal allocation of resources results in longer processing times	Prediction of accuracy and unexpected tasks. Genetic algorithms possess large-scale computational applications
[30]	Resource management framework named BEHAVE	Resource Management (memory, CPU, and bandwidth)	BEHAVE addresses the management technical barriers by (1) Modeling and assessing the BRD of IoT devices (2) RFTA model that binds the devices' BRD and resource allocation to achieve fair allocation	The management of multidimensional resources efficiently by the edge is a challenging task	Low processing power and storage capacity
[31]	Approximation algorithm	Task Allocation	Minimize total energy consumption	(1) Power consumption models may have different complexity and accuracy. (2) Designing energy-efficient MEC systems	
[18]	Hybrid optimization algorithm with effective decision-making mechanism	Resource Allocation	Maximizing the capacity and the data rate of the network and minimizing interference		
[32]	PSO-based Enhanced Dynamic Resource Allocation Method (EDRAM)	Resource Allocation (RAM)	Reduces task waiting time, latency, and network bandwidth consumption and enhances the Quality of Experience (QoE)		PSO has a low convergence rate in the iterative process. DRA results in slightly higher latency for tasks, as resources may need to be allocated or released dynamically

To address these user and application demands, diverse computing devices are employed. Each computing device is equipped with interfaces such as Wi-Fi, Zigbee, Bluetooth, and Cellular, each possessing specific properties like data rate, bandwidth, and speed. These interfaces act as conduits connecting users to computing devices. Importantly, each computing device is linked to distinct hardware resources, such as RAM and CPU, ensuring efficient processing and storage capabilities. Cloud-based resources are also leveraged to enhance the functionality of the network.

The system model, depicted in Figure 1, encapsulates the complex interactions of users, services, computing devices, interfaces, and resources within the system. It illustrates the dynamics of a network with diverse users seeking high-computational services such as video streaming, online gaming, AR, and VR. To address these computationally demanding services, high computation is necessary, facilitated by computing devices such as multi-rotors, base stations, and tethered balloons. These devices are equipped with interfaces and essential resources required for efficient processing. So, multi-dimensional multi-criterion utility optimization can help achieve this by considering multiple criteria simultaneously and optimizing them based on their relative importance.

In the proposed system model, variables are strategically defined to capture the intricacies of the heterogeneous wireless network. These variables include K for users, S for services, M for devices, I for interfaces, and R for resources.

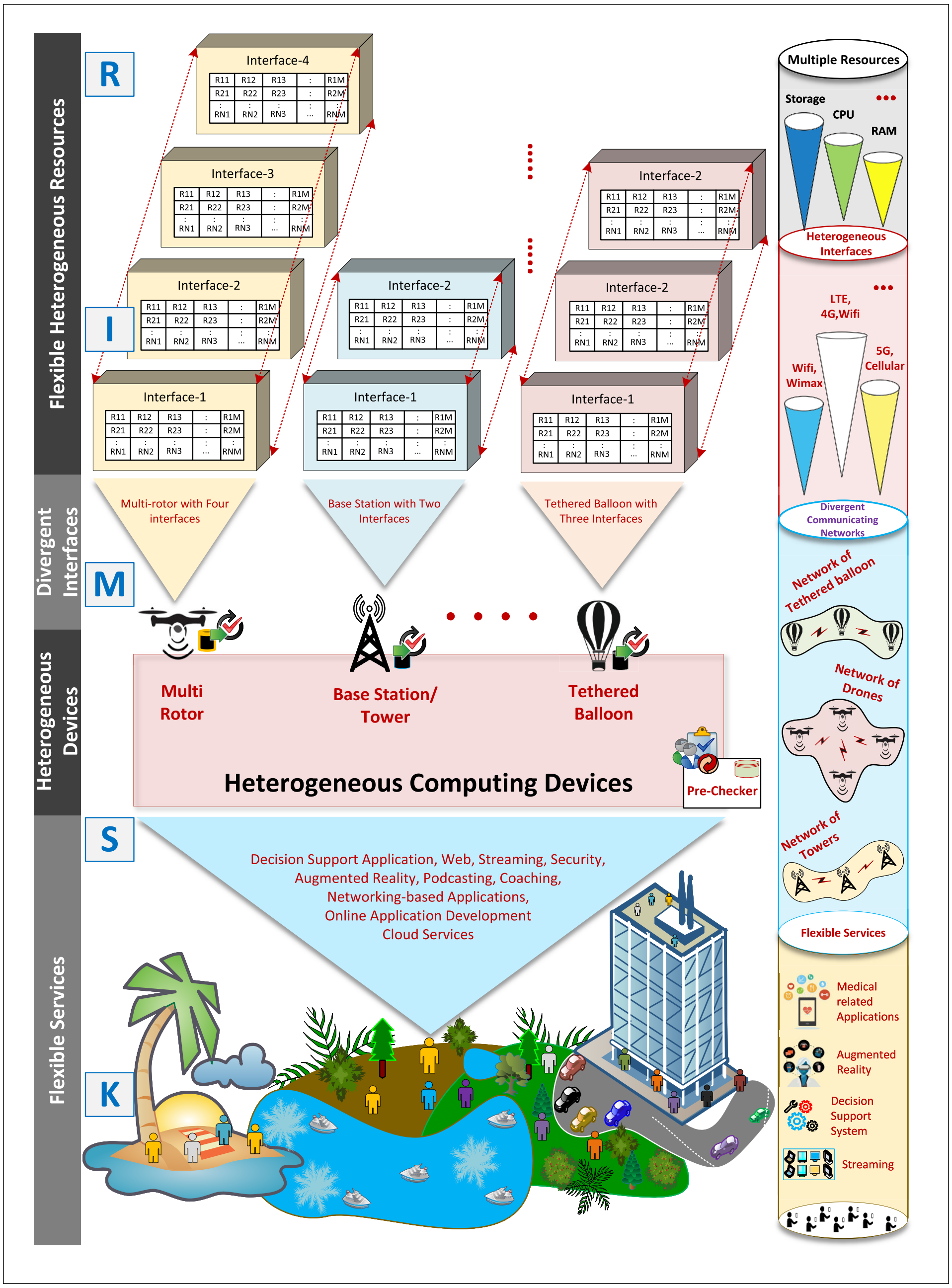

FIGURE 1. Framework for multi-dimensional multi-criterion utility optimization in heterogeneous aerial network.

The multi-dimensional nature of the network is reflected in the product of these dimensions, shaping its diverse characteristics. Additionally, the formulation involves user-specific parameters such as operational costs, activation costs, and user priorities. The representation of these variables within the system model establishes a foundation for effective optimization, enabling the network to dynamically adapt to user demands and operational requirements. In the following equation, $y_m^{k,s}$ is a binary variable indicating whether the k_{th} user's s_{th} service is fulfilled by the m_{th} computing device. It takes the value 1 if the service is fulfilled and 0 otherwise.

$$y_m^{k,s} = \begin{cases} 1: & \text{if } k_{th} \text{ user, } s_{th} \text{ service is fulfilled} \\ & \text{by the } m_{th} \text{ computing device/machine} \\ 0: & \text{otherwise} \end{cases}$$

The binary variable $x_{i,r,s}^{k,m}$ represents whether the k_{th} user requests the computation of the s_{th} service on the m_{th} machine of the i_{th} interface using the r_{th} resource. When $x_{i,r,s}^{k,m} = 1$, it indicates an active computation request; otherwise, it is 0, signifying no such request. This variable is crucial for

TABLE 2. Notations and descriptions.

Notation	Description
K	Number of Users/IoT Devices.
S	Number of Flexible Services.
M	Number of Heterogeneous Computing Devices.
I	Number of Interfaces supported by the mth devices Heterogeneous Computing Device.
R	Number of flexible resources assimilated at the i_{th} interface of m_{th} Heterogeneous Computing Device.
x	Resource allocation
y	Service assignment
$x_{m,i,r}^{k,s}$	k_{th} user, s_{th} service is computed by m_{th} computing device and i_{th} interface uses r_{th} resource
$y_m^{k,s}$	The k_{th} user, s_{th} service is fulfilled by the m_{th} computing device.
C_s^k	Service capacity for each user-service combination
$d_{m,i,r}^{k,s}$	Demand of k_{th} user to fulfill s_{th} service to allocate r_{th} resource through i_{th} interface of m_{th} computing device.
C_r	Capacity limits for each resource
C_{rem}^r	Remaining capacity for each resource

optimizing resource allocation in the heterogeneous wireless network.

$$x_{i,r,s}^{k,m} = \begin{cases} 1: & \text{If the } k_{th} \text{ user requests the computation of the} \\ & s_{th} \text{ service on the } m_{th} \text{ machine of the} \\ & i_{th} \text{ interface utilizing the } r_{th} \text{ resource.} \\ 0: & \text{otherwise} \end{cases}$$

The problem formulation for multi-dimensional multi-criterion utility optimization in next-generation heterogeneous wireless networks is given below. Table 2 shows the notations and their descriptions used in the problem formulation.

A. PROBLEM FORMULATION

Optimizing resource allocation in the dynamic nature of wireless networks becomes particularly challenging within the intricate framework of five-dimensional hierarchical multi-tier structures. The intricate interactions among diverse users, services, computing devices, interfaces, and resources contribute to these challenges. Balancing activation costs, operational costs, and user priority, while addressing the complex nature of resource allocation through a penalty term, becomes a critical focus requiring effective solutions. The objective is to minimize the total cost in (1), composed of operational cost, activation cost, user priority, and a penalty term. Operational cost represents resource usage expenses; activation cost accounts for device activation costs; and user priority is subtracted to prioritize user preferences. Additionally, the penalty term, introduced with a coefficient γ, addresses the optimization challenges by penalizing infeasible solutions. The goal is to find an efficient resource allocation strategy that minimizes total costs, considering the dynamic and complex nature of the heterogeneous wireless network.

$$\text{Minimize: Total Cost} = \text{Operational Cost} + \text{Activation Cost} - \text{User Priority} + \text{Penalty Term} \quad (1)$$

The objective is to minimize the total cost represented in (2), including operational and activation costs, user priority, as well as a penalty term, while meeting constraints ranging from C1 to C10. The penalty coefficient (γ) is added to the objective function to enforce a penalty when $x_{m,i,r}^{k,s}$ deviates from being binary. Adjusting the γ allows controlling the trade-off between optimizing the objective and penalizing non-binary values. Here, the objective is to find values for the decision variables that jointly minimize the total cost, considering operational, activation, and priority costs, while adhering to the penalty term to encourage binary decisions.

$$OP: \min_{x_{m,i,r}^{k,s}} \Bigg(\overbrace{\sum_k \sum_s \sum_m \sum_i \sum_r \underbrace{\text{Cost}_{m,i,r}^{k,s}}_{\text{Per Unit Cost}} \cdot \underbrace{x_{m,i,r}^{k,s}}_{\text{integer}}}^{\text{Operational Cost}}$$

$$+ \overbrace{\sum_m \sum_i \sum_r \text{ActCost}_{m,i,r} \cdot y_m^{k,s}}^{\text{Activation Cost}}$$

$$- \overbrace{\sum_k \sum_s \sum_m \text{P}_k \cdot y_m^{k,s}}^{\text{User Priority}}$$

$$+ \overbrace{\gamma \sum_m \sum_i \sum_r x_{m,i,r}^{k,s} \cdot (1 - x_{m,i,r}^{k,s})}^{\text{Penalty Term}} \Bigg)$$

subject to

Resource allocation constraint:

$$C1: \sum_m \sum_i x_{m,i,r}^{k,s} \leq 1 \forall k, s, r$$

Interface allocation constraint:

$$C2: \sum_m y_m^{k,s} \leq 1 \quad \forall k, s$$

Resource and interface compatibility constraints:

$$C3: x_{m,i,r}^{k,s} \leq y_m^{k,s} \forall k, s, m, i, r$$

$$C4: \sum_k \sum_s x_{m,i,r}^{k,s} d_{m,i,r}^{k,s} \leq C_{m,i,r}, \quad \forall m, i, r$$

Service capacity constraint:

$$C5: \sum_m \sum_s y_m^{k,s} \leq C_s^k, \quad \forall k$$

User-service pair and machine allocation constraint:

$$C6: \sum_m x_{m,i,r}^{k,s} \leq 1 \forall k, s, i, r$$

Interface allocation constraint:

$$C7: \sum_i x_{m,i,r}^{k,s} \leq 1 \forall k, s, m, r$$

Resource Capacity Constraint:

$$C8: \sum_r x_{m,i,r}^{k,s} \leq C_r \forall k, s, m, i$$

Resource allocation and demand constraint:

$$C9: \sum_{r} x_{m,i,r}^{k,s} \cdot d_{m,i,r}^{k,s} \leq C_{rem}^{r} \forall k, s, m, I$$

Resource Allocation and Service Assignment Constraint:

$$C10: \sum_{r} \sum_{i} x_{m,i,r}^{k,s} = y_{m}^{k,s} \forall k, m, s \quad (2)$$

The constraints from $C1$ to $C10$ restrict how resources are allocated, how services are assigned, and how interfaces are used in the framework of an optimization problem. They are essential for preserving an appropriate balance between the use of resources and meeting compatibility and capacity requirements. Each constraint plays a specific role in governing resource allocation and service assignment. The following is an exploration of these constraints to understand their contributions to the entire system.

Here, C_1 ensures exclusive resource allocation by enforcing that each user-service pair (k, s) and resource (r) can be assigned to at most one machine (m) through a designated interface (i). This constraint establishes a one-to-one relationship between resources and machine-interface pairs for each user-service pair. In contrast, C_2 focuses on the interface allocation constraint, limiting the assignment of services to a single interface on a machine for each user-service pair (k, s). It ensures that the k_{th} user demanding the s_{th} service can only be admitted by one machine, guaranteeing exclusive use of a particular interface on the same machine simultaneously. This constraint is crucial for efficient interface utilization. On the other hand, C_3 represents the resource and interface compatibility constraint, maintaining compatibility between resource allocation (x) and service assignment (y) decisions. It stipulates that a resource can only be allocated if the corresponding service is assigned to the machine interface, ensuring that allocated resources correspond to service assignments for each user-service pair on a machine.

Addressing both resource and interface capacity, C_4 imposes restrictions on resource allocation to prevent exceeding defined capacity limits for a machine's interface, ensuring adherence to capacity restrictions. Specifically, it prevents user (k) from exceeding available resources on interface (i) of machine (m) while utilizing service (s), which requires resource (r). Simultaneously, C_5 introduces the service capacity constraint, setting an upper limit on total services assignable to a machine for each user (k) as a safeguard against overloading services on a machine, ensuring that the assignment does not surpass the designated service capacity (C_s^k) for each user-service combination. Transitioning to C_6, the user-service pair and machine allocation constraint, enforces exclusivity in resource allocation for each user-service pair on a machine through a specific interface, preventing simultaneous allocation to multiple pairs and avoiding duplication. Concurrently, C_7 maintains exclusivity at the interface level, restricting resource allocation to one user-service pair on a machine through a specific interface, ensuring that each interface is used by a single pair at any given time.

The resource capacity constraint is introduced in C_8 that limits the total allocation of resources for a user-service pair on a machine through various interfaces. This constraint ensures that resource allocation does not exceed the capacity limits (C_r) defined for each resource, controlling resource utilization within their capacity constraints. C_9 is the resource allocation and demand constraint, managing resource allocation while considering resource demand (d). It enforces that resource allocations do not exceed the remaining capacity (C_{rem}^r) for each resource, ensuring alignment between resource allocation, demand, and capacity constraints. Lastly, C_{10} guarantees consistency between resource allocation (x) and service assignment (y). It requires that resource allocations match service assignments for each user (k) and service (s) on a machine (m), fostering a close integration between these critical components of the optimization problem.

The iterative optimization process considers convergence criteria based on changes in the objective function and constraint violations. The dynamic adjustment of the penalty factor is a key feature, contributing to the adaptability of the optimization approach. The overall performance of this approach is evaluated in terms of total cost minimization, convergence speed, and the feasibility of solutions. Comparative analyses with alternative optimization approaches provide insights into the effectiveness of this methodology for resource allocation in a heterogeneous wireless network. Sensitivity analysis explores the trade-off between penalty strength and solution quality, shedding light on the impact of penalty factor variations. The interpretation of results involves identifying the strengths and limitations of the approach and providing valuable insights for practical implementation in real-world scenarios.

B. PROPOSED METHODOLOGY

In addressing the dynamic resource allocation challenges within wireless networks, particularly in the context of five-dimensional hierarchical multi-tier structures, the first step involves a comprehensive problem formulation. This involves identifying and articulating the complexities associated with balancing activation costs, operational costs, and user priority. The subsequent phase focuses on formulating constraints to guide the optimization process. These constraints encompass considerations like computational capacity $(C_{i,r}^m)$, user demands $(d_{m,i,r}^{k,s})$, resource capacity (C_r), and remaining capacity (C_{rem}). The optimization objective is then developed to minimize the total cost, with operational costs, activation costs, and user priority. A penalty term, represented by $\gamma \sum_m \sum_i \sum_r x_{m,i,r}^{k,s} \cdot (1 - x_{m,i,r}^{k,s})$, is integrated to address the inherent complexities of resource allocation. Efficiently allocating network resources and bandwidth to various devices and applications is difficult for optimal network utilization and QoS provision. However, effectively managing heterogeneity in multi-tiered or multi-dimensional

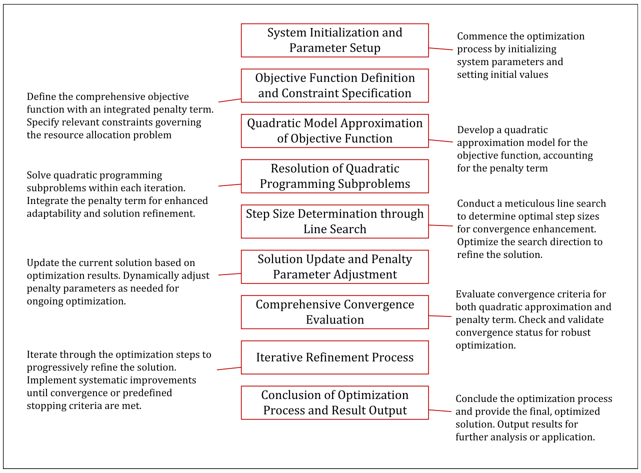

FIGURE 2. Intelligent relaxation using penalty function (IRPF) optimization process.

wireless networks is essential and still challenging for seamless connectivity, access, fair distribution of resources, and an improved user experience. The problem of inefficient and unfair resource allocation can lead to a decline in network performance in terms of utility.

Figure 2, shows the methodological structure as flowchart for the proposed IRPF to generate optimized results tailored for resource allocation in a heterogeneous drone network. The initialization phase begins with setting an initial feasible solution, denoted as $x^{(0)}$, and defining important parameters such as the penalty factor γ. Following this, the objective function $f(x)$ and constraints are formulated to encapsulate resource allocation requirements. The objective function is expressed in (2). The quadratic approximation phase involves developing a quadratic approximation $Q_k(x)$)of the objective function and constraints around the current solution $x^{(k)}$, utilizing the gradient and Hessian matrix. The subsequent step encompasses solving quadratic programming subproblems with the inclusion of a penalty term. A line search is then conducted to determine the optimal step size α for the current iteration. The updated solution $x^{(k+1)}$ is obtained by incorporating the results of the line search. The process iterates through the quadratic approximation, quadratic programming, line search, and update steps until convergence is achieved. The convergence check involves evaluating specific criteria to ascertain the proximity to an optimal solution. This iterative methodology continues until the convergence criteria are met, leading to the final resource allocation solution.

In a diverse urban setting, a mixed wireless network supports various user needs. Individuals seeking services like video streaming, podcasting, augmented reality (AR), and high-performance gaming connect to this network, which includes base stations, drones, and balloons. These devices host different interfaces such as Zigbee, Bluetooth, WiFi, and cellular, each with its own resources like CPU, RAM, and storage. Users access these resources based on their preferences and requirements, with a system pre-checker containing the optimal resource allocation that is assigned by the solver. For instance, a user looking for video streaming might connect through WiFi, enabling the network to allocate cloud-based resources for smooth streaming. Similarly, someone engaging in AR activities might connect via Bluetooth to a nearby drone, utilizing its fog layer processors for real-time processing and an enhanced user experience. This dynamic allocation of resources ensures efficient and tailored support for diverse user demands. The optimization aims to minimize the total cost, encompassing operational costs, activation costs, and user priority. Constraints are formulated based on the system model, incorporating considerations such as resource capacity and demand fulfillment. Additionally, a penalty term is introduced to handle binary decision variables and ensure the exploration of feasible solution spaces.

The incorporation of a penalty function becomes crucial in guiding the optimization process. This penalty function penalizes infeasible solutions and promotes the exploration of feasible regions. A regularization parameter is introduced to control the impact of the penalty term. This ensures a balance between exploring the solution space and penalizing constraint violations. An intelligent relaxation strategy is employed by dynamically adjusting the penalty factor during optimization. Starting with a relatively low penalty factor encourages exploration, and the factor is adaptively increased based on optimization progress. This strategy helps prevent premature convergence to suboptimal solutions, promoting a more thorough solution for space exploration.

IV. SIMULATION RESULTS AND DISCUSSION

A comparative analysis is carried out to evaluate the performance of Branch and Bound (an optimized algorithm) with the proposed IRPF for resource allocation in a heterogeneous wireless network. To address the complexities of resource allocation further, the proposed IRPF that utilizes the Sequential Quadratic Programming (SQP) algorithm is used, which incorporates a penalty term into the optimization. This penalty term introduces a regularization factor, enhancing the solution's adaptability to intricate resource allocation scenarios. The comparative analysis involves evaluating the efficiency and computational performance of Branch & Bound and proposed IRPF algorithms. Key metrics, including total cost and computational time, are assessed to determine the most effective resource allocation strategy in the context of a heterogeneous wireless network. This methodology assists in determining an optimal resource allocation strategy for practical applications as a five-dimensional heterogeneous wireless network model, organizing users to services, and linking services to computing devices for efficient resource accommodation.

In Figure 3, Utility vs. Number of Users has been analyzed, and the consistent outperformance of IRPF is attributed to its exceptional adaptability to an increasing number of users. The algorithm dynamically adjusts operational costs, activation costs, user priorities, and penalty terms,

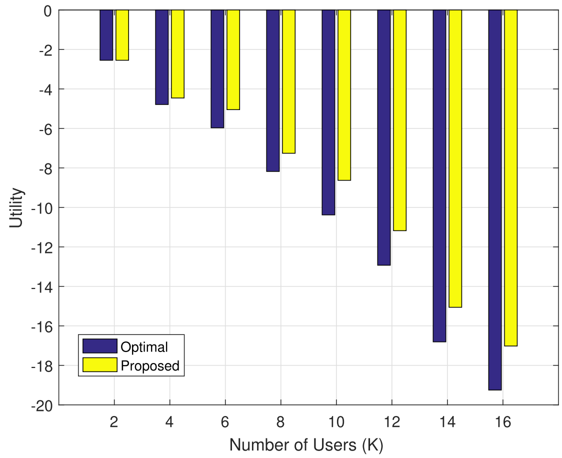

FIGURE 3. Utility vs. Users: Analyzing resource allocation for four services on three machines with a single interface.

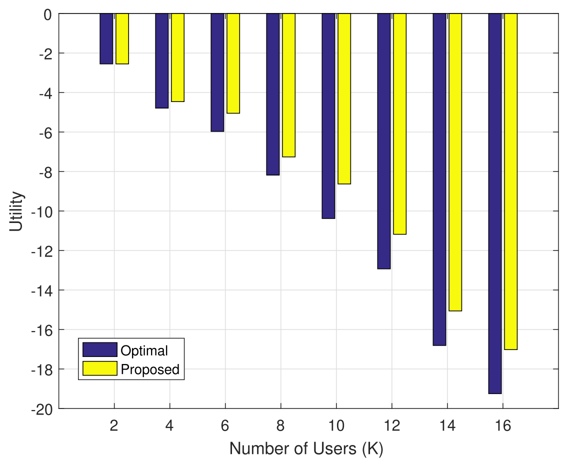

FIGURE 4. Utility vs. Users: Evaluating resource allocation for four services on five machines with three interfaces and three resources.

effectively managing the complexities associated with a growing user base. This adaptability ensures optimal resource allocation to meet evolving demands efficiently. IRPF excels at simultaneously minimizing operational and activation costs, as shown in Figure 4. Its exceptional performance is largely due to its ability to manage cost factors while considering user priorities and penalty modifications. The IRPF achieves a more efficient resource allocation strategy, lowering overall costs. Similarly, the impact of the penalty term on utility is represented in Figure 5. This figure highlights the adaptability of IRPF through penalty adjustments. The ability of the algorithm to dynamically modify the penalty term and optimize resource allocation through efficient fractional value management is crucial in scenarios with changing services, user demands, and device capabilities. This adaptability provides a competitive edge over the branch-and-bound approach. The effectiveness of the proposed IRPF technique in resource allocation is showcased in various scenarios:

Scenario 1 is depicted in Figure 3, which shows the comparison between utility and users for four services on

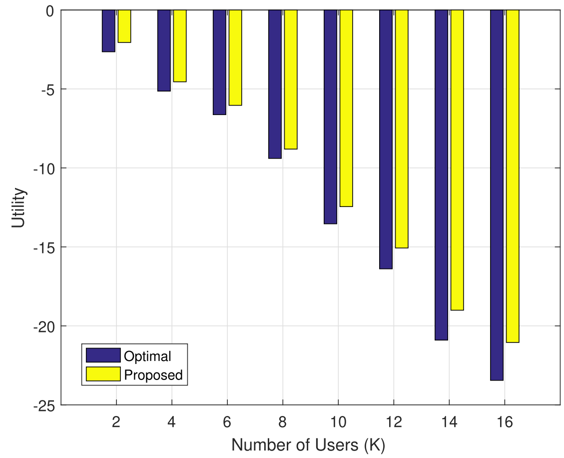

FIGURE 5. Utility vs. Users: Assessing resource allocation for four services on nine machines with two interfaces.

three machines with a single interface. Similarly, Scenario 2 depicted in Figure 4 addresses the resource allocation for four services on five machines with three interfaces and three resources, and Scenario 3 depicted in Figure 5 reflects the resource allocation for four services on nine machines with two interfaces. In a case with two users, three machines, and two services, the IRPF technique successfully acknowledges all four requests, demonstrating its ability to handle relatively smaller problem instances efficiently. The outcomes indicate how effective IRPF is at obtaining near-optimal solutions for certain configurations compared to the optimal solution found by the branch-and-bound algorithm.

For a more extensive scenario with 16 users, 2 services, and 9 machines (each having one interface), the IRPF technique continues to perform well. It acknowledges 22 out of 32 possible services, showcasing its ability to scale effectively to larger problem instances, as shown in Figure 6. Importantly, the proposed technique achieves results comparable to those of the optimized branch-and-bound algorithm but with significantly less computational complexity. This is a notable advantage, especially in scenarios where the exponential complexity of branch-and-bound becomes a limiting factor. Therefore, the proposed IRPF technique demonstrates effectiveness in diverse scenarios.

In Figure 7, a visual representation is provided to compare the user accommodation of the proposed IRPF and the optimal Branch and Bound algorithm. The focus is on demonstrating how well the IRPF performs in accommodating users compared to the theoretically optimal solution provided by the Branch and Bound algorithm. The red-highlighted section in the figure draws attention to the convergence point, indicating that the IRPF is capable of accommodating several users equivalent to those of the optimal solution. This suggests that, under certain conditions or scenarios, the IRPF can match the performance of the theoretically optimal algorithm. However, as the complexity of the network configuration increases, which is characterized by a higher number of users, services, machines, interfaces, K = Users | S = Services | M = Machines | I = Interfaces | R = Resources

O = Optimized (branch and bound) technique

P = Proposed Intelligent Relaxation using the Penalty Function (IRPF)

I = 1, R=1

S	K	M=3 O	M=3 P	M=5 O	M=5 P	M=7 O	M=7 P	M=9 O	M=9 P
S=2	2	4	4	4	4	4	4	4	4
	4	7	7	7	6	7	7	7	7
	6	9	9	9	7	9	9	9	9
	8	13	12	13	11	13	13	13	13
	10	15	14	16	14	17	17	17	17
	12	18	17	19	16	20	20	20	20
	14	20	19	23	20	24	24	24	24
	16	23	22	26	23	27	27	27	27
S=3	2	4	3	4	4	4	4	4	4
	4	8	8	7	7	8	8	8	8
	6	10	8	9	9	10	10	10	10
	8	13	11	13	12	14	14	14	14
	10	16	14	19	18	20	20	20	20
	12	19	17	23	22	24	24	24	23
	14	24	22	29	28	30	29	30	28
	16	27	25	33	32	34	32	34	32
S=4	2	4	4	4	4	4	4	4	4
	4	8	8	8	8	8	8	8	8
	6	10	9	10	10	10	10	10	10
	8	14	13	14	14	14	14	14	14
	10	20	19	21	21	21	19	21	20
	12	25	24	26	26	26	24	26	25
	14	31	30	33	32	33	31	33	32
	16	35	34	37	36	37	35	37	36

I = 1, R=3

S	K	M=3 O	M=3 P	M=5 O	M=5 P	M=7 O	M=7 P	M=9 O	M=9 P
S=2	2	4	4	4	4	4	4	4	4
	4	7	7	7	6	7	7	7	7
	6	9	9	9	8	9	8	9	8
	8	13	13	13	12	13	11	13	12
	10	17	15	17	15	17	15	17	16
	12	20	18	20	17	20	17	20	19
	14	24	20	24	21	24	21	24	23
	16	27	23	27	23	27	24	27	26
S=3	2	4	4	4	4	4	4	4	4
	4	8	8	8	6	8	8	8	6
	6	10	10	10	7	10	10	10	8
	8	14	13	14	10	14	14	14	12
	10	19	16	20	15	20	20	20	18
	12	23	18	24	17	24	24	24	22
	14	29	22	30	22	30	29	30	25
	16	33	25	34	24	34	32	34	28
S=4	2	4	4	4	4	4	4	4	3
	4	8	7	8	8	8	7	8	7
	6	10	8	10	9	10	9	10	9
	8	14	12	14	12	14	13	14	13
	10	21	15	21	18	21	20	21	19
	12	26	20	26	23	26	24	26	23
	14	33	27	33	29	33	30	33	29
	16	37	30	37	32	37	33	37	32

I = 2, R=1

S	K	M=3 O	M=3 P	M=5 O	M=5 P	M=7 O	M=7 P	M=9 O	M=9 P
S=2	2	4	4	4	4	4	4	4	4
	4	7	7	7	6	7	7	7	7
	6	9	9	9	8	9	8	9	8
	8	13	13	13	12	13	11	13	12
	10	17	15	17	15	17	15	17	16
	12	20	18	20	17	20	17	20	19
	14	24	20	24	21	24	21	24	23
	16	27	23	27	23	27	24	27	26
S=3	2	4	4	4	4	4	4	4	4
	4	8	8	8	6	8	8	8	6
	6	10	10	10	7	10	10	10	8
	8	14	13	14	10	14	14	14	12
	10	19	16	20	15	20	20	20	18
	12	23	18	24	17	24	24	24	22
	14	29	22	30	22	30	29	30	25
	16	33	25	34	24	34	32	34	28
S=4	2	4	4	4	4	4	4	4	3
	4	8	7	8	8	8	7	8	7
	6	10	8	10	9	10	9	10	9
	8	14	12	14	12	14	13	14	13
	10	21	15	21	18	21	20	21	19
	12	26	20	26	23	26	24	26	23
	14	33	27	33	29	33	30	33	29
	16	37	30	37	32	37	33	37	32

I = 2, R=3

S	K	M=3 O	M=3 P	M=5 O	M=5 P	M=7 O	M=7 P	M=9 O	M=9 P
S=2	2	4	4	4	4	4	4	4	4
	4	7	7	7	7	7	7	7	6
	6	9	9	9	9	9	9	9	8
	8	13	13	13	12	13	13	13	10
	10	17	15	17	15	17	16	17	13
	12	20	18	20	17	20	18	20	16
	14	24	20	24	20	24	22	24	20
	16	27	23	27	23	27	25	27	22
S=3	2	4	4	4	4	4	4	4	4
	4	8	8	8	6	8	8	8	8
	6	10	9	10	8	10	10	10	9
	8	14	12	14	11	14	14	14	13
	10	20	15	20	16	20	20	20	18
	12	24	18	24	20	24	23	24	19
	14	30	22	30	26	30	28	30	24
	16	34	25	34	30	34	31	34	28
S=4	2	4	4	4	4	4	4	4	3
	4	8	7	8	8	8	8	8	5
	6	10	8	10	10	10	9	10	7
	8	14	12	14	13	14	13	14	10
	10	21	18	21	19	21	19	21	15
	12	26	22	26	23	26	23	26	20
	14	33	28	33	30	33	30	33	26
	16	37	32	37	33	37	33	37	30

FIGURE 6. **Performance evaluation of IRPF and branch and bound in diverse network configurations.**

K = Users | S = Services | M = Machines | I = Interfaces | R = Resources

S	K	I = 1, R=1, M=3	I = 1, R=1, M=5	I = 1, R=1, M=7	I = 1, R=1, M=9	I = 1, R=3, M=3	I = 1, R=3, M=5	I = 1, R=3, M=7	I = 1, R=3, M=9	I=3, R=1, M=3	I=3, R=1, M=5	I=3, R=1, M=7	I=3, R=1, M=9	I=3, R=3, M=3	I=3, R=3, M=5	I=3, R=3, M=7	I=3, R=3, M=9
S=2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1
	6	0	2	0	0	0	1	1	1	0	1	1	1	0	0	0	1
	8	1	2	0	0	0	1	2	1	0	1	2	1	0	1	0	3
	10	1	2	0	0	2	2	2	1	2	2	2	1	2	2	1	4
	12	1	3	0	0	2	3	3	1	2	3	3	1	2	3	2	4
	14	1	3	0	0	4	3	3	1	4	3	3	1	4	4	2	4
	16	1	3	0	0	4	4	3	1	4	4	3	1	4	4	2	5
S=3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	2	0	2	0	2	0	2	0	2	0	0
	6	2	0	0	0	0	3	0	2	0	3	0	2	1	2	0	1
	8	2	1	0	0	1	4	0	2	1	4	0	2	2	3	0	1
	10	2	1	0	0	3	5	0	2	3	5	0	2	5	4	0	2
	12	2	1	0	1	5	7	0	2	5	7	0	2	6	4	1	5
	14	2	1	1	2	7	8	1	5	7	8	1	5	8	4	2	6
	16	2	1	2	2	8	10	2	6	8	10	2	6	9	4	3	6
S=4	2	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1
	4	0	0	0	0	1	0	1	1	1	0	1	1	1	0	0	3
	6	1	0	0	0	2	1	1	1	2	1	1	1	2	0	1	3
	8	1	0	0	0	2	2	1	1	2	2	1	1	2	1	1	4
	10	1	0	2	1	6	3	1	2	6	3	1	2	3	2	2	6
	12	1	0	2	1	6	3	2	3	6	3	2	3	4	3	3	6
	14	1	1	2	1	6	4	3	4	6	4	3	4	5	3	3	7
	16	1	1	2	1	7	5	4	5	7	5	4	5	5	4	4	7

FIGURE 7. **Users and services accommodation scenario in heterogeneous environment.**

and resources, a corresponding rise in overall complexity is observed. This heightened complexity has a notable impact on the search space, which refers to the range of possible solutions that the algorithm explores. The trade-off

K = Users | S = Services | M = Machines | I = Interfaces | R = Resources

		I = 1								I=3							
		R=1				R=3				R=1				R=3			
	K	M=3	M=5	M=7	M=9	M=3	M=5	M=7	M=9	M=3	M=5	M=7	M=9	M=3	M=5	M=7	M=9
S=2	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	14.3	0.0	0.0	0.0	14.3	0.0	0.0	0.0	14.3	0.0	0.0	0.0	0.0	0.0	14.3
	6	0.0	22.2	0.0	0.0	0.0	11.1	11.1	11.1	0.0	11.1	11.1	11.1	0.0	0.0	0.0	11.1
	8	7.7	15.4	0.0	0.0	0.0	7.7	15.4	7.7	0.0	7.7	15.4	7.7	0.0	7.7	0.0	23.1
	10	6.7	12.5	0.0	0.0	11.8	11.8	11.8	5.9	11.8	11.8	11.8	5.9	11.8	11.8	5.9	23.5
	12	5.6	15.8	0.0	0.0	10.0	15.0	15.0	5.0	10.0	15.0	15.0	5.0	10.0	15.0	10.0	20.0
	14	5.0	13.0	0.0	0.0	16.7	12.5	12.5	4.2	16.7	12.5	12.5	4.2	16.7	16.7	8.3	16.7
	16	4.3	11.5	0.0	0.0	14.8	14.8	11.1	3.7	14.8	14.8	11.1	3.7	14.8	14.8	7.4	18.5
S=3	2	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	0.0	0.0	25.0	0.0	25.0	0.0	25.0	0.0	25.0	0.0	25.0	0.0	0.0
	6	20.0	0.0	0.0	0.0	0.0	30.0	0.0	20.0	0.0	30.0	0.0	20.0	10.0	20.0	0.0	10.0
	8	15.4	7.7	0.0	0.0	7.1	28.6	0.0	14.3	7.1	28.6	0.0	14.3	14.3	21.4	0.0	7.1
	10	12.5	5.3	0.0	0.0	15.8	25.0	0.0	10.0	15.8	25.0	0.0	10.0	25.0	20.0	0.0	10.0
	12	10.5	4.3	0.0	4.2	21.7	29.2	0.0	8.3	21.7	29.2	0.0	8.3	25.0	16.7	4.2	20.8
	14	8.3	3.4	3.3	6.7	24.1	26.7	3.3	16.7	24.1	26.7	3.3	16.7	26.7	13.3	6.7	20.0
	16	7.4	3.0	5.9	5.9	24.2	29.4	5.9	17.6	24.2	29.4	5.9	17.6	26.5	11.8	8.8	17.6
S=4	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.0	0.0	0.0	0.0	25.0	0.0	0.0	0.0	25.0
	4	0.0	0.0	0.0	0.0	12.5	0.0	12.5	12.5	12.5	0.0	12.5	12.5	12.5	0.0	0.0	37.5
	6	10.0	0.0	0.0	0.0	20.0	10.0	10.0	10.0	20.0	10.0	10.0	10.0	20.0	0.0	10.0	30.0
	8	7.1	0.0	0.0	0.0	14.3	14.3	7.1	7.1	14.3	14.3	7.1	7.1	14.3	7.1	7.1	28.6
	10	5.0	0.0	9.5	4.8	28.6	14.3	4.8	9.5	28.6	14.3	4.8	9.5	14.3	9.5	9.5	28.6
	12	4.0	0.0	7.7	3.8	23.1	11.5	7.7	11.5	23.1	11.5	7.7	11.5	15.4	11.5	11.5	23.1
	14	3.2	3.0	6.1	3.0	18.2	12.1	9.1	12.1	18.2	12.1	9.1	12.1	15.2	9.1	9.1	21.2
	16	2.9	2.7	5.4	2.7	18.9	13.5	10.8	13.5	18.9	13.5	10.8	13.5	13.5	10.8	10.8	18.9

FIGURE 8. Percentage increase of user service pair acknowledgment in diverse network configurations.

introduced is between computational efficiency and accuracy in the resource allocation process with the consideration of a five-dimensional heterogeneous aerial network.

Essentially, the more intricate the network becomes, the more challenging it becomes to balance computational efficiency and accuracy, which means providing a solution that closely aligns with the optimal allocation of resources. This dynamic nature of the system underscores the importance of understanding the implications of network complexity on the performance of resource allocation algorithms like IRPF and Branch and Bound. It also emphasizes the need to consider trade-offs and make informed decisions based on the specific requirements and constraints of the given scenario.

Figure 8 presents an analysis of user accommodation, comparing the performance of the Implicit Resource Provisioning Framework (IRPF) and the optimal Branch and Bound algorithm. The IRPF significantly outperforms the Branch and Bound algorithm in accommodating user-service pairs, as indicated by the highlighted red area showing a percentage increase for both algorithms. This suggests that the IRPF can handle a larger number of user-service pairs compared to the optimal Branch and Bound algorithm. The unnoted area signifies a common set of users acknowledged by both techniques, indicating some overlap in their performance. The increasing percentage in the red area reflects a growing search space due to escalating network configuration complexity with additional devices, users, services, and interfaces. The IRPF's notable improvement implies a more efficient strategy for navigating this expanded search space, which is crucial as the network configuration becomes more intricate, requiring the exploration of a larger solution space for optimal resource allocation.

V. CONCLUSION

In the multi-user, multi-service, heterogeneous multi-device scenarios that characterize modern aerial networks, the main goal of this research was to address an essential requirement for efficient resource allocation. The exploration focused on a five-dimensional heterogeneous wireless network model, incorporating diverse services and heterogeneous devices, each with multiple interfaces to cater to varied user needs. In this regard, a novel Intelligent Relaxation using the Penalty Function (IRPF) approach was proposed, treating resource allocation as an integer programming problem to balance between user demands and affordability. In scenarios where resource allocation for multiple services is required based on user demand and device capabilities, the effectiveness of the proposed IRPF approach was demonstrated through a comparative study with the conventional branch-and-bound algorithm. The optimization framework, designed to minimize activation and operating costs while optimizing utility, demonstrated superior computing efficiency through extensive simulations. The findings highlight the important role of the proposed framework in effectively addressing the challenges associated with resource allocation in aerial networks.

REFERENCES

[1] H. F. Alhashimi, M. N. Hindia, K. Dimyati, E. B. Hanafi, N. Safie, F. Qamar, K. Azrin, and Q. N. Nguyen, "A survey on resource management for 6G heterogeneous networks: Current research, future trends, and challenges," *Electronics*, vol. 12, no. 3, p. 647, Jan. 2023.

[2] M. J. Alam, M. R. Hossain, S. Azad, and R. Chugh, "An overview of LTE/LTE-A heterogeneous networks for 5G and beyond," *Trans. Emerg. Telecommun. Technol.*, vol. 34, no. 8, p. e4806, 2023.

[3] H. Zhang, Y. Xiao, S. Bu, D. Niyato, F. R. Yu, and Z. Han, "Computing resource allocation in three-tier IoT fog networks: A joint optimization approach combining Stackelberg game and matching," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1204–1215, Oct. 2017.

[4] Y. Zhao, G. Cheng, C. Liu, and Z. Chen, "Snapshot for IoT: Adaptive measurement for multidimensional QoS resources," in *Proc. IEEE/ACM 29th Int. Symp. Quality Service (IWQOS)*, Jun. 2021, pp. 1–10.

[5] S. Wang, J. Yang, and S. Bi, "Adaptive video streaming in multi-tier computing networks: Joint edge transcoding and client enhancement," *IEEE Trans. Mobile Comput.*, vol. 23, no. 4, pp. 2657–2670, Aug. 2024.

[6] B. Marques, S. Silva, J. Alves, T. Araújo, P. Dias, and B. S. Santos, "A conceptual model and taxonomy for collaborative augmented reality," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 12, pp. 5113–5133, Dec. 2022.

[7] X. Nan, X. Guo, Y. Lu, Y. He, L. Guan, S. Li, and B. Guo, "Delay–rate–distortion optimization for cloud gaming with hybrid streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 12, pp. 2687–2701, Dec. 2017.

[8] A. Bahabry, X. Wan, H. Ghazzai, H. Menouar, G. Vesonder, and Y. Massoud, "Low-altitude navigation for multi-rotor drones in urban areas," *IEEE Access*, vol. 7, pp. 87716–87731, 2019.

[9] S. Wang, M. Chen, C. Yin, W. Saad, C. S. Hong, S. Cui, and H. V. Poor, "Federated learning for task and resource allocation in wireless high-altitude balloon networks," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17460–17475, Dec. 2021.

[10] J. W. Lee, R. R. Mazumdar, and N. B. Shroff, "Joint resource allocation and base-station assignment for the downlink in CDMA networks," *IEEE/ACM Trans. Netw.*, vol. 14, no. 1, pp. 1–14, Feb. 2006.

[11] M. Sohaib Khan, S. I. Ali Shah, A. Javed, N. Mumtaz Qadri, and N. Hussain, "Drone selection using multi-criteria decision-making methods," in *Proc. Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2021, pp. 256–270.

[12] V. I. Dankovtsev, N. V. Toutova, A. S. Vorozhtsov, and I. A. Andreev, "Multi-criteria application placement in fog computing," in *Proc. Intell. Technol. Electron. Devices Vehicle Road Transp. Complex (TIRVED)*, Nov. 2022, pp. 1–4.

[13] N. A. Angel, D. Ravindran, P. M. D. R. Vincent, K. Srinivasan, and Y.-C. Hu, "Recent advances in evolving computing paradigms: Cloud, edge, and fog technologies," *Sensors*, vol. 22, no. 1, p. 196, Dec. 2021.

[14] A. Mishra, A. Swain, A. K. Ray, and R. M. Shubair, "HetNet/M2M/D2D communication in 5G technologies," in *5G IoT and Edge Computing for Smart Healthcare*. Amsterdam, The Netherlands: Elsevier, 2022, pp. 45–87.

[15] A. Mijuskovic, A. Chiumento, R. Bemthuis, A. Aldea, and P. Havinga, "Resource management techniques for cloud/fog and edge computing: An evaluation framework and classification," *Sensors*, vol. 21, no. 5, p. 1832, Mar. 2021.

[16] M. Tolani, R. K. Singh, K. Shubham, and R. Kumar, "Two-layer optimized railway monitoring system using Wi-Fi and ZigBee interfaced wireless sensor network," *IEEE Sensors J.*, vol. 17, no. 7, pp. 2241–2248, Apr. 2017.

[17] B. Gul, I. A. Khan, S. Mustafa, O. Khalid, and A. R. Khan, "CPU–RAM-based energy-efficient resource allocation in clouds," *J. Supercomput.*, vol. 75, pp. 7606–7624, Aug. 2019.

[18] Y. Teekaraman, H. Manoharan, A. R. Basha, and A. Manoharan, "Hybrid optimization algorithms for resource allocation in heterogeneous cognitive radio networks," *Neural Process. Lett.*, vol. 55, no. 4, pp. 3813–3826, 2023.

[19] M. Mahbub, B. Barua, and A. G. Alharbi, "Maximizing the probability of user association of a tier of a multi-tier heterogeneous network by optimal resource allocation," in *Proc. Emerg. Technol. Comput., Commun. Electron. (ETCCE)*, Dec. 2021, pp. 1–6.

[20] T. Bahreini, H. Badri, and D. Grosu, "Mechanisms for resource allocation and pricing in mobile edge computing systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 3, pp. 667–682, Mar. 2022.

[21] Z. Wang, Y. Wei, F. R. Yu, and Z. Han, "Utility optimization for resource allocation in multi-access edge network slicing: A twin-actor deep deterministic policy gradient approach," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 5842–5856, Aug. 2022.

[22] A. Asheralieva, D. Niyato, and X. Wei, "Ultra-reliable low-latency slicing in space-air-ground multi-access edge computing networks for next-generation Internet of Things and mobile applications," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 3956–3978, Feb. 2024.

[23] M. Dan, S. Srinivasan, S. Sundaram, A. Easwaran, and L. Glielmo, "A scenario-based branch-and-bound approach for MES scheduling in urban buildings," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7510–7520, Dec. 2020.

[24] S. S. Byun, I. Balasingham, and X. Liang, "Dynamic spectrum allocation in wireless cognitive sensor networks: Improving fairness and energy efficiency," in *Proc. IEEE 68th Veh. Technol. Conf.*, Sep. 2008, pp. 1–5.

[25] X. Qi, S. Khattak, A. Zaib, and I. Khan, "Energy efficient resource allocation for 5G heterogeneous networks using genetic algorithm," *IEEE Access*, vol. 9, pp. 160510–160520, 2021.

[26] Z. Torki and S. M. Matinkhah, "Optimization resource allocation in NOMA-based fog computing with a hybrid algorithm," in *Proc. 11th Int. Conf. Comput. Eng. Knowl. (ICCKE)*, Oct. 2021, pp. 1–6.

[27] S. A. Solangi, D. N. Hakro, I. A. Lashari, K.-ur-R. Khoumbati, Z. A. Bhutto, and and M. Hameed, "Genetic algorithm applications in wireless sensor networks (WSN): A review," *Int. J. Manag. Sci. Bus. Res.*, vol. 1, no. 4, pp. 152–166, 2017.

[28] Z. Miao, X. Yuan, F. Zhou, X. Qiu, Y. Song, and K. Chen, "Grey wolf optimizer with an enhanced hierarchy and its application to the wireless sensor network coverage optimization problem," *Appl. Soft Comput.*, vol. 96, Nov. 2020, Art. no. 106602.

[29] Y.-L. Chen, S.-Y. Huang, Y.-C. Chang, and H.-C. Chao, "Resource allocation based on genetic algorithm for cloud computing," in *Proc. 30th Wireless Opt. Commun. Conf. (WOCC)*, Oct. 2021, pp. 211–212.

[30] I. AlQerm, J. Wang, J. Pan, and Y. Liu, "BEHAVE: Behavior-aware, intelligent and fair resource management for heterogeneous edge-IoT systems," *IEEE Trans. Mobile Comput.*, vol. 21, no. 11, pp. 3852–3865, Nov. 2022.

[31] X. Liu, J. Liu, and H. Wu, "Energy-efficient task allocation of heterogeneous resources in mobile edge computing," *IEEE Access*, vol. 9, pp. 119700–119711, 2021.

[32] D. Baburao, T. Pavankumar, and C. S. R. Prabhu, "Load balancing in the fog nodes using particle swarm optimization-based enhanced dynamic resource allocation method," *Appl. Nanosci.*, vol. 13, no. 2, pp. 1045–1054, 2023.

[33] X. Liu, H. Zhang, K. Long, A. Nallanathan, and V. C. M. Leung, "Energy efficient user association, resource allocation and caching deployment in fog radio access networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1846–1856, Feb. 2022.

[34] H. Yuan and M. Zhou, "Profit-maximized collaborative computation offloading and resource allocation in distributed cloud and edge computing systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 3, pp. 1277–1287, Jul. 2021.

[35] S. Li, J. Huang, and B. Cheng, "Resource pricing and demand allocation for revenue maximization in IaaS clouds: A market-oriented approach," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 3, pp. 3460–3475, Sep. 2021.

[36] S. Sun and S. Moon, "Practical scheduling algorithms with contiguous resource allocation for next-generation wireless systems," *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 725–729, Apr. 2021.

[37] Y. Peng, X. Xue, A. K. Bashir, X. Zhu, Y. D. Al-Otaibi, U. Tariq, and K. Yu, "Securing radio resources allocation with deep reinforcement learning for IoE services in next-generation wireless networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 5, pp. 2991–3003, Sep. 2022.

[38] U. M. Malik, M. A. Javed, J. Frnda, and J. Nedoma, "SMRETO: Stable matching for reliable and efficient task offloading in fog-enabled IoT networks," *IEEE Access*, vol. 10, pp. 111579–111590, 2022.

MUHAMMAD IRFAN MUSHTAQ received bachelor's degree in electronics engineering from the NFC Institute of Engineering and Technology Multan, in 2012, and the master's degree in electrical engineering from COMSATS University Islamabad, Wah Campus, in 2016, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include resource allocation and resource management in heterogeneous wireless networks.

OMER CHUGHTAI (Member, IEEE) received the B.Eng. degree from the University of Engineering and Technology, Taxila, Pakistan, the M.S. degree from COMSATS University Islamabad, Wah Campus, Pakistan, and the Ph.D. degree from Universiti Teknologi PETRONAS, Malaysia. He is currently a tenured Associate Professor with the Department of Electrical and Computer Engineering, COMSATS University Islamabad. He is a professional in computer science and engineering, and brings over two decades of experience. He is an active member of organizations, such as ISOC.

YUDONG ZHANG (Senior Member, IEEE) is currently with the School of Computing and Mathematical Sciences, University of Leicester, Leicester, U.K., and the School of Computer Science and Engineering, Southeast University, Nanjing, China.

ALI H. ALENEZI (Member, IEEE) received the B.S. degree in electrical engineering from King Saud University, Saudi Arabia, the M.S. degree in electrical engineering from the KTH Royal Institute of Technology, Sweden, and the Ph.D. degree in electrical engineering from the New Jersey Institute of Technology, Newark, NJ, USA, in 2018. He is currently an Associate Professor with the Electrical Engineering Department, Northern Border University, Saudi Arabia. His research interests include acoustic communication, wireless communications, and 4G and 5G networks using UAVs.

NAYEF ALQAHTANI (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Oakland University, Rochester, MI, USA, in 2021. He is currently an Assistant Professor with King Faisal University. His research interests include control systems, intelligent systems, digital signal and image processing, smart grids, and cyber-physical systems.