Technische Universität München
TUM School of Computation, Information and Technology

Network Design and Resource Management for LiFi-WiFi Heterogeneous Networks

Hansini Vijayaraghavan, M.Sc.

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades einer

Doktorin der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr.-Ing. Antonia Wachter-Zeh

Prüfer der Dissertation:
1. Prof. Dr.-Ing. Wolfgang Kellerer
2. Prof. Dr.-Ing. habil. Falko Dressler

Die Dissertation wurde am 09.09.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 04.11.2024 angenommen.

Network Design and Resource Management for LiFi-WiFi Heterogeneous Networks

Hansini Vijayaraghavan, M.Sc.

MediaTUM

Kurzfassung

Das rasante Wachstum der Anzahl internetfähiger Geräte und der erwartete Anstieg der globalen Datenübertragung stellen erhebliche Herausforderungen für bestehende drahtlose Netze dar. Neue Anwendungen wie Mobile Metaverse, Augmented Reality (AR), Virtual Reality (VR) und High Definition (HD) Streaming belasten diese Netzwerke weiter, da sie Latenzzeiten im Sub-Millisekundenbereich und hohe Bandbreiten erfordern. Diese Arbeit befasst sich mit diesen Herausforderungen, indem sie das Potenzial des gemeinsamen Betriebs von Light-Fidelity (LiFi) und Wireless-Fidelity (WiFi) untersucht, um die Quality of Service (QoS) Anforderungen moderner drahtloser Kommunikationssysteme zu erfüllen.

LiFi arbeitet im sichtbaren Licht- und Infrarotspektrum und bietet ein viel größeres Spektrum im Vergleich zu traditionellem WiFi. Zur Einrichtung eines Hochgeschwindigkeits-Kommunikationssystems für die drahtlose optische Übertragung zwischen mehreren Nutzern nutzt LiFi bestehende Beleuchtung mittels Leuchtdioden (LEDs). Trotz Vorteilen wie hohen Datenraten, Einsatzmöglichkeiten in Umgebungen, die empfindlich auf elektromagnetische Strahlung reagieren, und inhärenter Sicherheit, steht LiFi Herausforderungen wie Lichtkegelüberlagerungen und Schattenwürfen durch Objekte gegenüber.

Um die Vorteile beider Technologien zu nutzen, wurde der gemeinsame Betrieb von LiFi und WiFi als heterogenes Netzwerk vorgeschlagen. Die bestehende Literatur konzentriert sich jedoch hauptsächlich entweder auf das Ressourcenmanagement oder auf Designaspekte von LiFi bzw. WiFi getrennt voneinander. Diese Dissertation argumentiert, dass ein ganzheitlicher Ansatz, der sowohl Designaspekte als auch das Ressourcenmanagement gleichzeitig berücksichtigt, entscheidend ist, um das volle Potenzial heterogener LiFi-WiFi-Netzwerke auszuschöpfen.

Unsere Arbeit beginnt mit der optimalen Platzierung von Access Points (APs) in LiFi-WiFi Netzwerken. Die begrenzte Reichweite von LiFi und die Abhängigkeit von Verbindungen in Sichtlinie (Line-of-Sight (LoS)) stellen spezifische Herausforderungen dar. Mit der Einführung eines Multi-Objective Optimization Frameworks betrachten wir die AP-Platzierung im dreidimensionalen Raum, um Kosten zu minimieren, die Netzleistung zu maximieren und gleichzeitig eine angemessene Beleuchtung zu gewährleisten. Dieser Ansatz wird auf heterogene Netzwerke ausgeweitet, um die Abdeckung, Kosteneffizienz und Nutzeranforderungen durch verschiedene Optimierungsmethoden, einschließlich genetischer Algorithmen und heuristischer Ansätze, auszubalancieren. Das Ergebnis ist eine Netzwerkarchitektur, die nicht nur technische Spezifikationen erfüllt, sondern sich auch an praktische Einschränkungen und Nutzerbedürfnisse anpasst, was einen bedeutenden Fortschritt gegenüber isoliert betrachteten AP-Platzierungsstrategien darstellt.

Neben der Planung untersucht diese Arbeit auch den Echtzeitbetrieb im Hinblick auf ein dynamisches Ressourcenmanagement, um Latenzen zu minimieren und eine hohe Netzstabilität sicherzustellen. Zu diesem Zweck formulieren wir ein Optimierungsproblem zur Minimierung der Netzlatenzen, während gleichzeitig die QoS-Anforderungen erfüllt werden. Für dieses Problem schlagen wir anschließend verschiedene Ressourcenoptimierungsstrategien vor. Unsere Methoden gewährleisten die Stabilität des Systems, indem sie sowohl die Kanaleigenschaften als auch die Bewegungsmuster der Nutzer berücksichtigen. Dies ermöglicht wiederum eine robuste Übergabe bestehender Verbindungen ohne erneute TCP-Verbindung zwischen den Technologien. Darüber hinaus entwickeln wir ein bewegungs- und anforderungsbasiertes Framework, das verschiedene Netzressourcen auf der Grundlage von Echtzeit-Netzeigenschaften und Nutzeranforderungen zuweist. Dieser ganzheitliche Ansatz stellt sicher, dass das Netzwerk sowohl robust als auch effizient bleibt, sich dynamisch an veränderliche Kanaleigenschaften anpasst und gleichzeitig höchste QoS für die Benutzer bietet.

Um die Leistung des Netzwerks langfristig zu verbessern, schlagen wir MobiFi vor, ein proaktives Ressourcenmanagement-Framework, entwickelt um Benutzerbewegungen und Kanaländerungen vorherzusagen. Durch den Einsatz fortschrittlicher Optimierungstechniken, wie etwa Branch-and-Bound-basierter Solver und Evolutionary Game Theory (EGT)-basierter Algorithmen, weist MobiFi Ressourcen effizient zu, mit dem Ziel die Netzleistung über einen langen Zeitraum stabil zu halten. Wir zeigen, durch Simulationen, die verschiedene Vorhersagefehler berücksichtigen, dass MobiFi auch unter sich ändernden Bedingungen zuverlässig bleibt. MobiFi bietet somit eine proaktive Lösung, die Serviceunterbrechungen deutlich verringert und die Ressourcennutzung maximiert.

Zum Abschluss präsentieren wir ComputiFi, ein Framework, dass das Task-Offloading in industriellen Umgebungen in einem LiFi-WiFi-Netzwerk optimiert. Dabei wird der Anteil des Datenstroms bestimmt, der über LiFi und jener, der über WiFi übertragen wird. Zudem wird festgelegt, an welchem Ort die Aufgaben verarbeitet werden – dies kann lokal, an den APs, den Routern oder in der Cloud geschehen. Hierbei kommen verschiedene dynamische Strategien zum Einsatz, die eine effiziente Zuordnung von Rechenressourcen sicherstellen und somit niedrige Latenzzeiten für latenzkritische Anwendungen wie Ultra Reliable Low Latency Communications (URLLC), Fabrikautomatisierung und Künstliche-Intelligenz-Bildklassifikation gewährleisten. Durch die Optimierung des Ortes der Task-Verarbeitung reduziert ComputiFi sowohl die Latenz als auch den Energieverbrauch und wird somit zu einem unverzichtbaren Werkzeug für moderne, hochanspruchsvolle Anwendungen.

Zusammenfassend tragen die vorgestellten Beiträge zu erheblichen Fortschritten im Netzdesign und Ressourcenmanagement von LiFi-WiFi-Netzwerken bei. Durch die Verbindung von Design und Ressourcenmanagement präsentieren wir eine ganzheitliche Lösung, die sowohl die Effizienz des Netzwerks als auch die Qualität in Umgebungen mit hoher Nutzerdichte optimiert. Unsere Ergebnisse bieten praktische Ansätze für die nahtlose Integration von LiFi und WiFi Technologien, positionieren diese als zentrale Bausteine der 6G-Technologie und bereiten den Weg für die nächste Generation drahtloser Kommunikationssysteme in Innenräumen. Diese Erkenntnisse werden die Weiterentwicklung von 6G beschleunigen, die Verbreitung von internetfähigen Geräten unterstützen und den wachsenden Anforderungen datenintensiver Anwendungen gerecht.

Abstract

The rapid growth in the number of internet-connected devices and the anticipated increase in global data transmission pose substantial challenges to existing wireless networks. Emerging applications like mobile metaverse, Augmented Reality (AR), Virtual Reality (VR), and high-definition streaming further strain these networks, requiring sub-millisecond latency and high capacity. This thesis addresses these challenges by exploring the potential of integrating Light-Fidelity (LiFi) and Wireless-Fidelity (WiFi) technologies to meet the Quality of Service (QoS) demands of modern wireless communications.

LiFi, operating in the visible light and infrared spectrum, offers a much larger spectrum compared to traditional WiFi. Utilizing existing lighting infrastructure, LiFi employs Light Emitting Diodes (LEDs) to create a high-speed, multi-user optical wireless communication system. Despite its advantages, such as high data density, suitability for Radio Frequency (RF)-sensitive environments, and inherent security, LiFi faces challenges as a standalone technology due to signal outages in overlapping cell regions and sensitivity to signal blockages.

The integration of LiFi with WiFi has been proposed to form heterogeneous networks, leveraging the strengths of both technologies. However, existing literature has primarily focused on either the resource management or the design aspects of these networks in isolation. This thesis argues that a holistic approach, considering both design and resource management simultaneously, is essential for fully realizing the potential of LiFi-WiFi integration.

This thesis begins with addressing the optimal placement of Access Points (APs) in LiFi-WiFi networks. LiFi's limited range and dependence on Line-of-Sight (LoS) connections present unique challenges. By introducing a multi-objective optimization framework, we consider three-dimensional AP placement to minimize costs, maximize network performance, and ensure adequate illumination. This approach is extended to heterogeneous networks, balancing coverage, cost efficiency, and user requirements through various optimization methods, including genetic algorithms and heuristic approaches. The result is a network design that not only meets technical specifications but also adapts to practical constraints and user needs, offering a significant advancement over isolated AP placement strategies.

Moving from design to real-time operations, our thesis explores dynamic resource management to ensure minimal delay and high network stability. We propose delay-aware resource optimization strategies, formulating an optimization problem to minimize network delays while meeting QoS requirements. Our stability-ensuring methods address transient channel conditions and user mobility, ensuring seamless vertical handovers without requiring TCP re-connections. A mobility-aware resource allocation framework is developed to optimize resource distribution based on real-time network conditions and user demands. This comprehensive approach ensures that the network remains robust and efficient, adapting dynamically to changing channel conditions, and thus providing a superior user experience.

To further enhance long-term network performance, we introduce MobiFi, a proactive resource management framework designed to anticipate user mobility and channel variations. By employing advanced optimization techniques, such as Branch and Bound-based solvers and Evolutionary Game Theory (EGT)-based algorithms, MobiFi effectively allocates resources to maintain long-term network performance. Through simulations incorporating potential prediction errors, we ensure that MobiFi remains reliable even under uncertain conditions, offering a proactive solution that significantly reduces service interruptions and maximizes resource utilization.

Finally, we present ComputiFi, recognizing the potential of integrating LiFi and WiFi networks for optimized task offloading in industrial environments. This task offloading framework optimizes the

distribution of tasks between LiFi and WiFi access technologies, managing data processing across local devices, APs, routers, and cloud servers. ComputiFi's dynamic strategy ensures efficient computational resource allocation, providing guaranteed low latency for latency-sensitive applications like Ultra Reliable Low Latency Communications (URLLC), factory automation, and Artificial Intelligence (AI) image classification. By optimizing the destination for task processing, ComputiFi reduces both latency and energy consumption, making it an invaluable tool for modern, high-demand applications.

In summary, the contributions of this thesis significantly advance the network design and resource management of LiFi-WiFi networks. By integrating design and resource management, we offer a holistic solution that enhances network efficiency and user experience in high-density environments. Our findings provide practical solutions for the integration of LiFi and WiFi technologies, establishing them as cornerstones for 6G technology and paving the way for the next generation of indoor wireless communication systems. These outcomes will drive forward the evolution of 6G connectivity, supporting the growing demands of data-intensive applications and the proliferation of internet-connected devices.

Contents

Abbreviations

AI	Artificial Intelligence	139
AP	Access Point	141
AR	Augmented Reality	1
ARP	Address Resolution Protocol	78
ASF	Achievement Scalarization Function	42
BE	Best Effort	55
CDF	Cumulative Distribution Function	60
CSI	Channel State Information	110
DE	Differential Evolution	117
DGE	Dynamic Grid Explorer	137
DRL	Deep Reinforcement Learning	140
EGT	Evolutionary Game Theory	138
eMBB	enhanced Mobile Broadband	139
FCFS	First Come First Serve	112
FoV	Field of View	26
GA	Genetic Algorithm	135
GAE	Generalized Advantage Estimation	119
HD	High Definition	1
IoT	Internet of Things	136
IR	Infrared	25
LAA	License Assisted Access	86
LED	Light Emitting Diode	87
gLED	Leuchtdiode	v
LiFi	Light-Fidelity	141
LoS	Line-of-Sight	138
LP	Linear Programming	39
LSTM	Long Short-Term Memory	86
LTE	Long Term Evolution	86
MAC	Medium Access Control	136
MCDM	Multi-Criteria Decision Making	29
MEC	Multi-access Edge Computing	106
MINLP	Mixed Integer Nonlinear Programming	106
MOO	Multi-Objective Optimization	134
MPTCP	Multipath Transmission Control Protocol	54

Chapter 1

Introduction

The number of internet-connected devices is projected to reach 125 billion globally by 2030 [Mar17]. Additionally, global data transmission is anticipated to grow at an average annual rate of 50% over the next 15 years [Mar17]. Mobile metaverse Augmented Reality (AR)/Virtual Reality (VR) traffic is forecasted to hit 62 EB per month by 2030 [Nok23]. As VR, AR, 4K and 8K videos, and High Definition (HD) streaming proliferate over wireless links, networks are under pressure to deliver sub-millisecond latency and high capacity per user. This places immense strain on existing wireless networks to meet the Quality of Service (QoS) requirements.

To address this growing demand, the shift is towards higher frequency bands like Terahertz (THz) frequencies. Terahertz communications (0.1 - 10 THz) promise vast transmission bandwidths and high capacity [Ela+20]. However, higher path loss at these frequencies necessitates sophisticated beamforming techniques [Sha+23] to enhance power in the Line-of-Sight (LoS). Consequently, smaller cell sizes are required, leading to increased infrastructure demands.

Continuing this trend towards higher frequencies, Light-Fidelity (LiFi), operating in the 300 THz visible light and infrared spectrum, offers a spectrum size 2600 times that of the 300 GHz radio frequency spectrum [Haa18]. LiFi leverages existing lighting infrastructure, using Light Emitting Diodes (LEDs) prevalent in homes, offices, and other indoor environments, to provide a high-speed, networked, multi-user optical wireless communication system. It employs light sources that emit power directionally, leveraging existing lighting infrastructure to mitigate the challenges faced by Terahertz Radio Frequency (RF) communications.

Key advantages of LiFi include:

- High data density: LiFi can utilize the extensive available bandwidth to provide high capacity. Data rates exceeding 15 Gb/s have been demonstrated using off-the-shelf LEDs [BTH19], compared to the 11 Gb/s theoretical maximum data rate of 802.11ax with a channel bandwidth of 160 MHz. LiFi also offers high data density due to the small cell radius of 1-4 m [Haa+16], wherein the capacity of a cell is shared by fewer users.

- Utilization of existing infrastructure: With Power over Ethernet (PoE) and Power Line Communication (PLC), LiFi can leverage existing data or power infrastructure for connectivity. Streetlights, home, and office lighting can potentially offer wireless internet access.

- Suitability for RF-sensitive environments: LiFi can be safely used in hospitals and airplanes, where RF interference is undesirable.

- Inherent security: The inability of light waves to penetrate walls ensures high physical layer security.

Despite these benefits, LiFi faces challenges as a standalone technology, including complete signal outage in overlapping cell regions and sensitivity to LoS blockages. Therefore, the adoption of LiFi is envisioned in heterogeneous LiFi-RF networks [Ayy+16], which offer the advantage of no interference between RF and LiFi, potentially providing greater aggregate throughput than the standalone

technologies. Since Wireless-Fidelity (WiFi) is a complementary indoor technology to LiFi, this thesis focuses on LiFi-WiFi networks.

The optimal design of such networks involves challenges due to cost differences in placing the Access Points (APs) of the two technologies, rate coverage area discrepancies, and the need for adequate illumination provided by LiFi APs operating on visible light. Existing research on AP placement primarily addresses either WiFi or LiFi networks separately [DBS18; DBS20; VB19; MS23; Yan+20; Gop+22], with no works addressing the integration of both technologies.

The differing characteristics of LiFi and WiFi technologies and their interaction with each other brings forth the need for a global perspective on the design of these networks. The short range as well as the directional nature of LiFi implies a heavy dependence of the channel quality and coverage area on the height of the APs. Therefore, an investigation of the third dimension of placement is necessary to deploy networks with LiFi APs.

Following the design and deployment of a LiFi-WiFi heterogeneous network, users have multiple options of WiFi or LiFi APs to associate to at any time necessitating sophisticated resource management to efficiently manage resources across multiple wireless access technologies. This resource management must ensure network resilience in case of a link failure. There is an existing body of work on load balancing in LiFi-RF networks [WH15; WWH17; WH19a]. However, these works approach the problem from the perspective of load balancing, focusing on maximizing network throughput by using only the channel quality as input.

The potentially huge data rates of LiFi imply that it should serve as a complementary technology rather than just for offloading from a congested RF network. This implies a resource management framework that pools resources from all wireless access technologies, allocating them based on user requirements, rate fairness, and latency goals, not merely network throughput. By considering all these metrics we can fully utilize the potential of a LiFi-WiFi network. Therefore, resource management must take a holistic view in order to utilize the full potential of LiFi as a 6G technology.

A key application area for a heterogeneous LiFi-WiFi network is the factory environment where high data rate transmission and low latency are crucial. LiFi's RF-interference-free communication is ideal for such high user density settings where WiFi struggles with bandwidth saturation and security concerns [Haa+16]. Additionally, LiFi's capability to provide localization services aligns with the needs of a smart manufacturing environments. Such environments demand latency-sensitive applications like Ultra Reliable Low Latency Communications (URLLC), factory automation, and Artificial Intelligence (AI) image classification, requiring efficient computational task offloading [3GP22a; 3GP22b]. Integrating LiFi and WiFi ensures optimal task offloading, balancing speed, coverage, and reliability [Wu+21a]. Thus, designing a task offloading solution in multipath LiFi-WiFi networks presents an opportunity to reduce latency. This involves managing not only wireless resources but also computational resources, ensuring efficient performance for latency-sensitive applications.

This thesis aims to address the design and resource management of LiFi-WiFi networks, establishing them as a cornerstone for the next generation of networks. By developing novel methodologies for optimal AP placement, wireless and computational resource allocation strategies, and comprehensive performance analysis in heterogeneous networks, this research contributes significantly to advancing the capabilities and integration of LiFi and WiFi technologies. The outcomes of this thesis will not only enhance network efficiency and user experience but also pave the way for practical implementation of next-generation indoor wireless communication systems, thereby driving forward the evolution of 6G connectivity.

1.1 Research Questions and Contributions

This section presents the main research questions and challenges addressed in this thesis and explains the contributions made towards resolving these questions.

1.1.1 RQ1: How can access points be effectively placed in a LiFi-WiFi network to optimize cost efficiency, coverage, and illumination quality in various indoor environments?

The core research question explores the optimal placement of APs in a LiFi-WiFi network, focusing on maximizing cost efficiency, coverage, and illumination quality in various indoor environments. This question stems from the unique challenges presented by LiFi technology, characterized by its limited range and dependence on LoS connections. These constraints necessitate ultra-dense AP deployments to ensure effective communication coverage and adequate indoor lighting. However, optimizing the placement of APs in such environments raises issues: How can we minimize costs while maximizing network performance and ensuring adequate illumination? What strategies can effectively balance coverage requirements, avoiding both excessive overlap and coverage gaps? How can the integration of LiFi and WiFi be optimized to leverage the strengths of each technology? Additionally, considering the expected user distribution within these environments is crucial for tailoring AP placement to actual usage patterns.

Addressing these complex questions, our research introduces two major contributions that leverage multi-objective optimization strategies to optimize AP placement in both LiFi-only and heterogeneous LiFi-WiFi environments. These contributions not only aim to enhance the theoretical framework for network design but also offer practical guidelines tailored to the demands of users and network operators in modern indoor wireless networks.

1. **C1a:** Optimized 3D Placement in LiFi Networks: The first contribution addresses the three-dimensional placement of APs in LiFi-only networks through a multi-objective optimization framework. This approach minimizes the number of APs to reduce costs while maximizing the network's sum rate and ensuring both a minimum guaranteed rate and sufficient illumination. The research considers the APs' height as a variable which significantly affects the network's performance metrics. A Genetic Algorithm (GA) is applied to resolve this complex optimization scenario, demonstrating the effectiveness of considering three-dimensional space in network planning.

2. **C1b:** Comprehensive Placement Strategy in LiFi-WiFi Networks: Extending the methodologies from LiFi-only scenarios, the second contribution develops a holistic framework for the placement of APs in heterogeneous LiFi-WiFi networks. This integrated approach is crucial for meeting the diverse requirements of modern indoor environments, combining visible light communication and infrared in LiFi with traditional WiFi. By formulating a multi-objective optimization problem, the framework seeks to significantly improve average network rates and ensure uniform illumination across the covered area. It considers varying costs of APs and technological needs of users, employing a variety of solution methods — ranging from heuristics to meta-heuristics and sophisticated solvers — to achieve optimal outcomes.

1.1.2 RQ2: How can resource management in a LiFi-WiFi network take a holistic view, dynamically adjusting to changing user demands, ensuring minimal delay and high network stability?

Research Question 2 explores the dynamics of resource management in a LiFi-WiFi network, focusing on a holistic approach that dynamically adjusts to changing user demands to ensure minimal delay and high network stability. This question addresses the challenge of effectively managing network resources in environments where user demands fluctuate frequently and require rapid adjustments. The goal is to optimize resource allocation in such a way that the network can handle these variations without significant delays or stability issues, which are crucial for delay-sensitive applications like live video streaming and Voice over Internet Protocol (VoIP).

For a deeper understanding of this challenge, we must consider the unique aspects of both LiFi and WiFi technologies and their integration. LiFi, often constrained by LoS requirements and limited range, and WiFi, with its broader coverage but potentially more congested channels, must be managed

in tandem to exploit their respective strengths without allowing their limitations to degrade overall network performance. Dynamic resource management becomes essential in such settings, where user mobility and varying communication needs can drastically affect the QoS. The critical performance metrics here are the minimization of packet delays and the enhancement of network stability, ensuring that the network remains robust even under instantaneous channel changes, high traffic conditions and frequent changes in user location and service requirements.

Addressing the challenges of dynamic resource management, our research introduces three significant contributions that enhance network performance and stability in LiFi-WiFi hybrid environments.

1. **C2a:** Delay-aware Resource Optimization: The first contribution focuses on optimizing network resource allocation specifically for delay-critical applications in LiFi-WiFi heterogeneous networks. We develop a strategy that not only minimizes network delays but also meets diverse user requirements for data rate and latency. This involves formulating a Mixed Integer Nonlinear Programming (MINLP) problem aimed at minimizing average network packet delay while ensuring quality of service requirements for each user. Our approach expands to accommodate multi-homing user devices, enabling simultaneous resource allocation across both LiFi and WiFi technologies. For precise solutions, we employ a Branch and Bound-based solver and propose a Genetic Algorithm for scenarios requiring faster, near-optimal solutions.

2. **C2b:** Stability Enhancement amidst Channel Challenges: Our second contribution proposes methods to maintain a stable LiFi-WiFi network amidst transient channel conditions caused by light path blockages, instantaneous receiver orientation changes, and user mobility. This includes a system approach to manage unavoidable vertical handovers, ensuring seamless interface switching. This approach has been implemented and tested on hardware in a LiFi-WiFi network setup, with performance evaluated through practical measurements, demonstrating that vertical handover overhead can be significantly minimized, thus maintaining network stability without requiring TCP re-connections.

3. **C2c:** Mobility- and Demand-aware Resource Allocation: Lastly, we introduce a framework for mobility-aware resource allocation in mobile LiFi-WiFi networks. This framework adapts to currently observed network conditions and optimizes resource allocation by solving an optimization problem using the weighted alpha-fair throughput. This method balances individual user needs with overall network efficiency, aiming to achieve a balanced distribution of resources that considers both user satisfaction and network performance.

1.1.3 RQ3: How can predictive technologies be utilized to proactively manage network resources, anticipating congestion and changing channel quality to sustain long-term network performance?

Research Question 3 investigates proactively managing wireless resources in LiFi-WiFi heterogeneous environments, aiming to anticipate congestion and dynamically adjust to changing user channel conditions. This question addresses the essential challenge of maintaining optimal long-term network performance in the face of fluctuating channel conditions and user mobility. The transition from reactive to proactive resource management can significantly enhance network efficiency, reduce service interruptions, and maintain consistent QoS. This approach relies on forecasting user movements and upcoming network conditions, optimizing wireless resource allocation in advance to align with expected channel conditions. Such a strategy not only improves the overall user experience but also maximizes the utilization of network resources. This shift raises several crucial questions: How can a balance between user rate and fairness be optimized for long-term network performance? What strategies can be implemented to adapt resource allocation in real-time based on predicted channel changes? Moreover, what are the implications of potential inaccuracies in user position prediction on overall network performance?

In response to RQ3, we introduce MobiFi, a proactive wireless resource management framework for LiFi-WiFi networks.

1. **C3:** Proactive Wireless Resource Management Framework: MobiFi is designed to enhance long-term network performance by proactively managing wireless resources in anticipation of user mobility and variable channel conditions. The framework establishes an optimization problem that uses an alpha-fairness objective to allocate resources effectively and fairly over time. Advanced solution techniques such as Branch and Bound-based solvers and iterative algorithms based on evolutionary game theory are employed to address the complexity of these optimization challenges. Additionally, MobiFi tests the robustness of its proactive strategies through simulations that incorporate potential errors in predicting user positions, ensuring that the framework remains reliable even when predictions do not perfectly match real-world scenarios.

1.1.4 RQ4: How can task offloading in LiFi-WiFi networks be optimized to minimize latency and energy consumption, while dynamically adapting to changes in network topology and user requirements?

Research Question 4 explores how latency-critical tasks can be processed in LiFi-WiFi networks while minimizing latency and energy consumption. This investigation is crucial for supporting latency-sensitive applications such as URLLC, factory automation, and AI image classification. These applications demand not only minimal latency for effective real-time processing but also high energy efficiency to reduce operational costs and environmental impacts.

The challenge lies in developing a framework that effectively integrates and manages multiple network technologies to meet these stringent requirements. Traditional single-path offloading methods often fail to fully utilize available network resources, resulting in increased latency and higher energy consumption. The integration of LiFi with WiFi in a multipath network architecture offers a unique opportunity to leverage the strengths of both technologies, where LiFi provides a high data rate communication channel and WiFi offers extensive coverage.

In response to RQ4, we introduce ComputiFi, a comprehensive task offloading framework designed for multipath, multihop LiFi-WiFi networks. This framework aims to reduce both latency and energy consumption by optimizing how tasks are offloaded and processed across the network.

1. **C4:** Task Offloading Framework: ComputiFi enables the efficient offloading of latency-sensitive tasks across multipath, multihop LiFi-WiFi networks. It determines the most efficient offloading destinations and manages data distribution between LiFi and WiFi, optimizing both network and user device energy consumption. ComputiFi implements a dynamic strategy for allocating resources across a diverse range of network devices—local devices, LiFi and WiFi APs, routers, and cloud servers. This strategy is tailored to minimize energy consumption at both the network and user device while providing guaranteed latency and minimizing task completion latency.

1.2 Thesis Organization

An overview of the structure of the thesis is given in Figure 1.1. The forthcoming chapters are structured as follows:

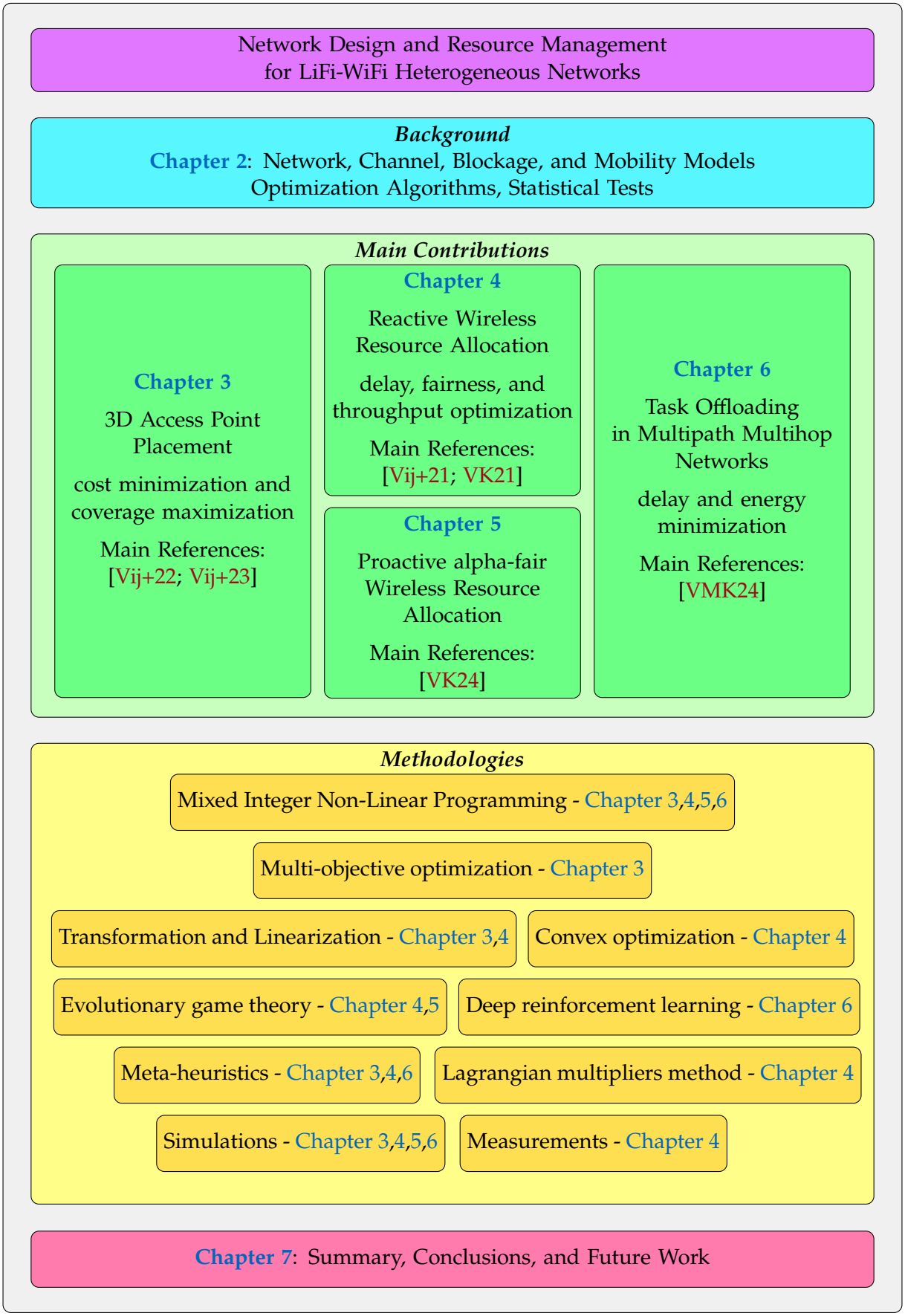


Figure 1.1 Thesis outline, including the main contributions, methodologies, and references of each chapter

Chapter 2: Essential Concepts and Models for LiFi-WiFi Network Design and Management provides a comprehensive background for understanding the thesis. It covers the fundamental models and concepts of LiFi-WiFi heterogeneous networks, detailing the network model, channel models, and blockage and illuminance models specific to LiFi. Additionally, it describes mobility models, commonly used optimization techniques, and methods for statistically evaluating results.

Chapter 3: Designing LiFi-WiFi Networks with a focus on Access Point Placement introduces two major contributions using multi-objective optimization to place APs in LiFi-only and heterogeneous LiFi-WiFi environments. The first contribution focuses on optimizing 3D placement of APs in LiFi networks to minimize costs and maximize network performance, using a Genetic Algorithm to handle the complexity. The second contribution extends this approach to heterogeneous networks, developing a comprehensive framework that balances average network rates, illumination, costs, and technological needs through various optimization methods. This chapter is based on the following publications:

- Hansini Vijayaraghavan et al. "Optimized 3D Placement of LiFi Access Points towards maximizing Wireless Network Performance". In: *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*. 2022, pp. 1278–1283. DOI: 10.1109/GLOBECOM48099.2022.10000893
- Hansini Vijayaraghavan et al. "PlaciFi: Orchestrating Optimal 3D Access Point Placement for LiFi-WiFi Heterogeneous Networks". In: *IEEE Access* 11 (2023), pp. 115415–115429. DOI: 10.1109/ACCESS.2023.3325097

Chapter 4: Wireless Resource Allocation in LiFi-WiFi Networks introduces three significant contributions to enhance network performance and stability in LiFi-WiFi heterogeneous environments. The first contribution optimizes resource allocation for delay-critical applications, employing strategies like Mixed Integer Nonlinear Programming solvers and a Genetic Algorithm to minimize delays and meet QoS requirements. The second contribution focuses on maintaining network stability amidst challenges such as light path blockages and user mobility, implementing methods to ensure seamless vertical handovers. The third contribution develops a mobility-aware weighted alpha-fair resource allocation framework that adapts to network conditions and balances user needs with overall network performance. This chapter is based on the following publications:

- Hansini Vijayaraghavan and Wolfgang Kellerer. "Delay-aware Wireless Resource Allocation and User Association in LiFi-WiFi Heterogeneous Networks". In: *2021 IEEE Global Communications Conference (GLOBECOM)*. 2021, pp. 01–06. DOI: 10.1109/GLOBECOM46510.2021.9685276
- Hansini Vijayaraghavan et al. "Algorithmic and System Approaches for a Stable LiFi-RF HetNet Under Transient Channel Conditions". In: *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. 2021, pp. 1048–1054. DOI: 10.1109/PIMRC50174.2021.9569271

Chapter 5: Mobility-aware Proactive Wireless Resource Allocation in LiFi-WiFi Networks introduces MobiFi, a proactive wireless resource management framework for LiFi-WiFi networks. MobiFi enhances long-term network performance by anticipating user trajectory and future channel conditions, employing advanced optimization techniques like Branch and Bound-based solvers and Evolutionary Game Theory algorithms. The framework's robustness is tested through simulations accounting for potential prediction errors, ensuring reliable performance in real-world scenarios. This chapter is based on the following paper:

- Hansini Vijayaraghavan and Wolfgang Kellerer. "MobiFi: Mobility-Aware Reactive and Proactive Wireless Resource Management in LiFi-WiFi Networks". In: *IEEE Transactions on Network and Service Management* 21.6 (2024), pp. 6597–6613. DOI: 10.1109/TNSM.2024.3455105

Chapter 6: Task Offloading in Multipath Multihop LiFi-WiFi Networks introduces ComputiFi, a comprehensive task offloading framework for multipath, multihop LiFi-WiFi networks. ComputiFi optimizes task offloading to reduce latency and energy consumption, determining efficient offloading

destinations and managing data distribution between LiFi and WiFi. The framework employs a Deep Reinforcement Learning-based algorithm to allocate computational resources across various network devices, ensuring minimal network and user device energy consumption with guaranteed low latency for task completion. This chapter is based on the following publication:

- Hansini Vijayaraghavan, Jörg von Mankowski, and Wolfgang Kellerer. "ComputiFi: Latency-Optimized Task Offloading in Multipath Multihop LiFi-WiFi Networks". In: *IEEE Open Journal of the Communications Society* 5 (2024), pp. 4444–4461. DOI: 10.1109/OJCOMS.2024.3426278

Finally, Chapter 7: Conclusions and Future Research summarizes the key findings and contributions of this thesis, highlighting the advancements in LiFi-WiFi heterogeneous network design, resource management, and task offloading. Additionally, the chapter outlines potential areas for future research, such as exploring new optimization techniques, integrating emerging technologies, and further improving the adaptability and robustness of LiFi-WiFi networks in dynamic environments.

Chapter 2

Essential Concepts and Models for LiFi-WiFi Network Design and Management

This chapter provides the necessary background to understand the subsequent chapters in this thesis. It begins by introducing the fundamental technologies and concepts relevant to the design and management of Light-Fidelity (LiFi)-Wireless-Fidelity (WiFi) heterogeneous networks. The network model, channel models for both LiFi and WiFi technologies, and blockage and illuminance models for LiFi are discussed in detail. Additionally, this chapter explores mobility models, some commonly used optimization techniques, and methods to statistically evaluate the results presented.

2.1 Overview of LiFi Technology

This section provides a concise overview of the LiFi technology, focusing on concepts to aid the understanding of the rest of the thesis. While not intended as a comprehensive tutorial, it covers key aspects necessary to understand the integration and operation of LiFi within heterogeneous networks. We also discuss the current state of LiFi hardware, highlighting its capabilities and limitations. For a detailed picture of the history and functioning of the technology, we point the readers to [DH15].

LiFi, is an emerging wireless communication technology that utilizes the visible light or infrared spectrum [CBH18] to transmit data. It operates by modulating the intensity of light emitted by light sources such as Light Emitting Diode (LED) bulbs to encode data, which is then received by photodetectors and converted back into electronic form. LiFi's operation within the visible light spectrum, enables its use in environments such as museums and conference rooms where both lighting and data transmission are essential. This dual-use capability of LiFi in visible light settings adds a layer of utility, as it serves both functional and data transmission roles.

In designing LiFi networks, particularly in indoor settings like museums or conference rooms, the placement of Access Points (APs) is crucial for achieving optimal coverage and efficient resource management. The challenge of ensuring adequate illumination while maintaining reliable data transmission capabilities requires strategic placement of these APs. This is a challenge that is tackled in this thesis as a step towards designing and planning such networks before their deployment.

Unlike traditional Radio Frequency (RF) communication, LiFi leverages the light spectrum, which offers several unique advantages. LiFi can achieve exceptionally high data transfer rates, surpassing those of conventional WiFi in many scenarios [BTH19]. The vast bandwidth available in the light spectrum allows for the transmission of large amounts of data [Haa18]. In industrial environments, such as factories, the high data rates provided by LiFi enable efficient task offloading requiring large data transfers and real-time data handling, crucial for automation and smart manufacturing processes. LiFi's inherent limitation of signal penetration through solid objects translates to enhanced security. Data transmission is confined to the illuminated area, preventing unauthorized access from outside the coverage zone. Another key advantage is reduced interference. LiFi operates on a different spectrum than radio-based technologies, eliminating interference with sensitive RF equipment commonly used in hospitals and airplanes.

However, LiFi also faces limitations that our thesis aims to address. LiFi signals require a clear path between the transmitter and receiver. Blockages by objects or people can significantly attenuate the signal strength, impacting data transmission. Our thesis models these blockages to evaluate the robustness of our proposed resource allocation strategies. Furthermore, our chapter on proactive resource allocation anticipating potential obstructions or channel quality issues that could impact connectivity is crucial in addressing this challenge.

The current state of LiFi hardware has seen significant advancements, yet it also faces certain limitations. A significant limitation in the field is the lack of open-source hardware solutions for LiFi that provide access to the Medium Access Control (MAC) layer, which is essential for implementing and testing the resource allocation algorithms that we propose in this thesis. This issue is well-documented in the literature [Wu+21b]. Despite this hurdle, our LiFi testbed, based on available market technology, validates aspects of our simulation models used in this thesis. We are optimistic that the recent release of the IEEE 802.11bb standard [IEE23] will catalyze the development of new open-source hardware. This development will enable future implementations and proof of concept for our resource allocation strategies, facilitating more robust testing and refinement of our algorithms in practical scenarios.

Our testbed provides a controlled environment to experiment with and evaluate the performance of LiFi technology in real-world scenarios. The setup includes several key components such as multiple LiFi APs, user equipment, and a network management system. Each AP is equipped with an optical front end operating on infrared for data transmission and photodetectors from [Gmb] for receiving signals. The user equipment comprises devices with integrated photodetectors and infrared LEDs to communicate with the LiFi APs. These devices are plugged in to PCs running a Linux system for performance measurements.

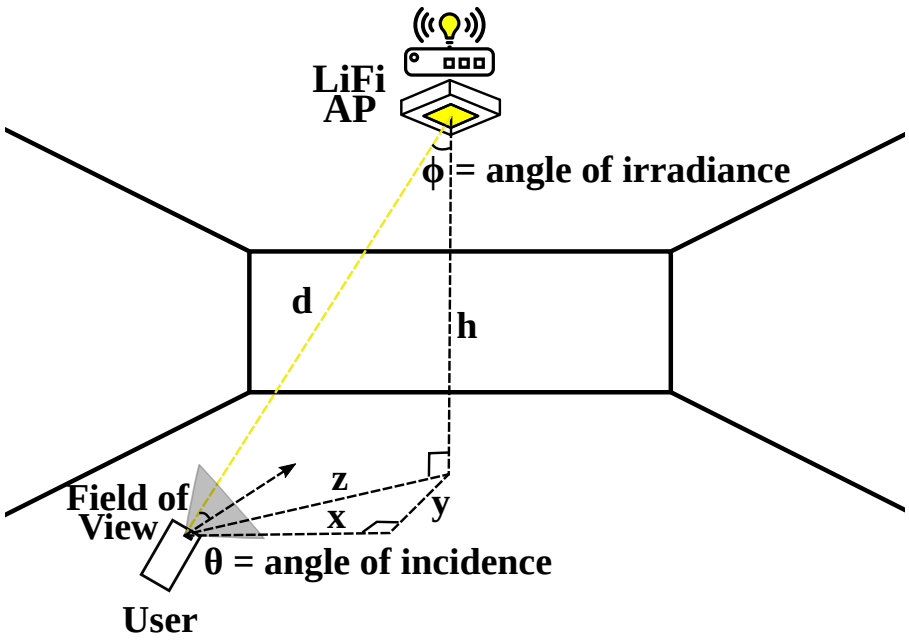


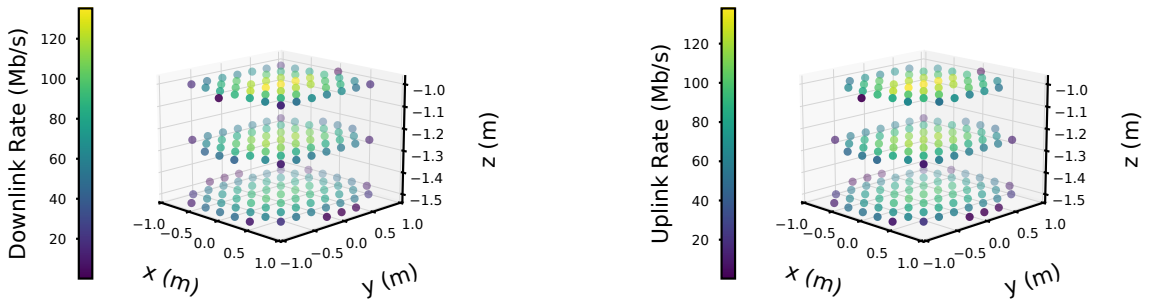
Figure 2.1 Architecture of a LiFi system showing the dependence of the signal quality on the distance, device orientation, Field of View, and beam angle of the AP

Figure 2.1 shows an example architecture of a LiFi system with one AP and one user. As can be observed, the signal between the two communicating devices depends on the distance between the two, the orientation of the user device which affects the angle of incidence, the beam angle of the transmitter which affects the angle of irradiance, and the Field of View of the receiver. To measure the

practical rate coverage affected by all these components, we conduct a series of experiments within the testbed environment.

The testbed is set up in a room with controlled lighting conditions to minimize interference from external light sources. Multiple data collection points within the coverage area are defined to measure the data rate at various locations. The methodology for measuring rate coverage involves initial calibration of the APs and user devices to start bidirectional iperf data streams. Static measurements are taken at each data collection point while the user remains stationary, providing a baseline for coverage performance. The measurements were taken in three slices along the z-axis. The orientation of the user device was flat, with infrared being used in both uplink and downlink, representing just one type of device available in the market [Gmb], though there are more with varying capabilities. The collected data is then analyzed to evaluate the rate coverage across the testbed area and identify the range at which the LiFi channel falls to zero due to the Field of View of the receiver device.

The results of the rate coverage measurements are summarized in Figure 2.2. The z-axis represents the vertical distance from the AP, with the AP located at coordinates (0,0,0). This summary shows the average data rate at each collection point in the uplink and downlink.



(a) Downlink rate coverage

(b) Uplink rate coverage

Figure 2.2 LiFi Rate Coverage as experienced in the testbed

The results indicate symmetric uplink and downlink rates, with the highest rates occurring closest to the AP. As the distance increases, the data rate decreases, eventually cutting off to zero after a diameter of 2 m along the x and y directions due to the Field of View of the receiver. These results are obtained with a fixed transmission power. Increasing the power saturates the receiver at vertical distances around 1 m, leading to lower rates at this range. However, increased power does result in higher rates at larger vertical distances. Despite the power adjustment, the x and y range remains the same due to angle cutoffs, while the z range increases because of the enhanced power.

To address the limitations of LiFi and enhance network robustness, integrating LiFi with WiFi has been proposed [Ayy+16]. This hybrid approach allows the network to leverage the high data rates and low latency of LiFi while utilizing the seamless coverage of WiFi. Such integration is essential in creating multipath transmission scenarios, which are useful in industrial settings where task offloading and real-time data processing are critical.

2.2 Hybrid LiFi-WiFi Network Model

This thesis focuses on deploying and managing a LiFi-WiFi heterogeneous network, integrating these wireless technologies for enhanced indoor communication. The network consists of a maximum of M^L LiFi APs and a maximum of M^W WiFi APs strategically placed throughout the environment. Throughout this thesis, the APs of both technologies are assumed to be non-colocated and operate independently. Physical infrastructure constraints, such as the need for LiFi APs to be integrated

with lighting systems, may necessitate their placement in locations that are not the same as WiFi APs. Additionally, the distinct architectures of the two technologies in most commercially available devices further justify their separate placement.

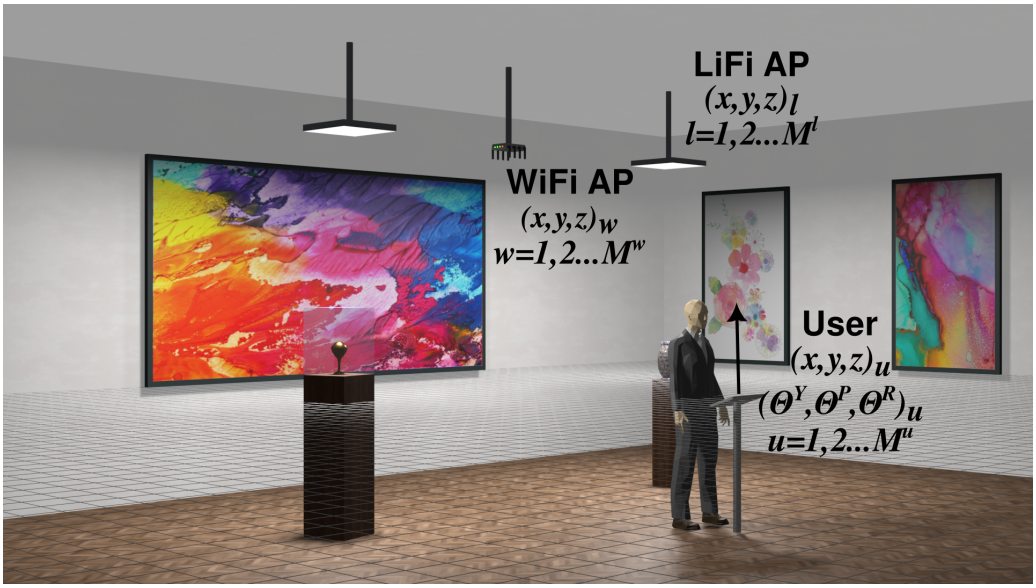


Figure 2.3 Example architecture of a LiFi-WiFi communication and illumination network in a Museum, one among several application scenarios

LiFi APs operate on the visible light or infrared spectrum and are generally installed above the user plane, oriented downwards. In our case, this is depicted by the white LED panels shown in Figure 2.3. When using visible light, they offer both illumination and data transmission capabilities. They are also equipped with photodiodes to receive uplink signals. The coordinates for each LiFi and WiFi AP are represented by $\boldsymbol{c}_l = (x, y, z)_l$ and $\boldsymbol{c}_w = (x, y, z)_w$, respectively. Since all LiFi APs share the same frequency channel, co-channel interference occurs where their coverage areas overlap, unlike the WiFi network.

The system supports a total of M^U users, each equipped with both LiFi and WiFi transmitters and receivers for both uplink and downlink traffic. Each user position $u \in \mathcal{U}$ is given by the coordinates $\boldsymbol{c}_u = (x, y, z)_u$. The orientation of the receivers is indicated by $\boldsymbol{\Theta_u} = (\Theta^Y, \Theta^P, \Theta^R)_u$, which corresponds to the device's yaw, pitch, and roll angles. A value of $\boldsymbol{\Theta_u} = (0, 0, 0)$ signifies that the user device is parallel to the floor and facing the ceiling.

2.3 LiFi Channel Model

Our work adapts the LiFi channel model detailed in [WWH17]. This model accounts for the three-dimensional positions of APs and users, as well as the orientation of user devices. The three-dimensional distance between a user u and an AP l is defined as follows:

$$\mathrm{d}_{u,l} = \|\boldsymbol{c}_l - \boldsymbol{c}_u\|_2 \tag{2.1}$$

The cosines of the angles of irradiance ($\phi_{u,l}$) at the AP and the angle of incidence ($\theta_{u,l}$) at the user are defined as follows:

$$\cos\phi_{u,l} = \frac{z_l - z_u}{d_{u,l}} \tag{2.2}$$

$$\cos\theta_{u,l} = \frac{(x_l - x_u)\hat{n}_{u,x} + (y_l - y_u)\hat{n}_{u,y} + (z_l - z_u)\hat{n}_{u,z}}{d_{u,l}} \tag{2.3}$$

where $\hat{\boldsymbol{n}}_u$ is the normal vector of the rotated user device. Equations (2.1), (2.2), and (2.3) combine to determine the Line-of-Sight (LoS) gain, HLoS:

$$\mathrm{HLoS}_{u,l} = \begin{cases} \frac{H_0}{d_{u,l}^2} \cdot \cos\phi_{u,l}^m \cdot \cos\theta_{u,l} & \text{if } \theta_{u,l} \leq \Theta_f \\ & \text{and } \phi_{u,l} \leq \Phi_f \\ 0 & \text{elsewhere} \end{cases} \tag{2.4}$$

where H_0 is a constant given by $((m+1)A_p\chi^2 T_s)/2\pi$. Here, m is the Lambertian order of the AP, A_p is the area of the photodiode receiver, χ is the refractive index, and T_s is the gain of the optical filter. The gain exists only when both the uplink and downlink signals are within the Field of View (FoV) of the transmitter (Φ_f) and receiver (Θ_f). The received signal power is calculated as:

$$P_{u,l} = (\mathrm{HLoS}_{u,l} \cdot P_l \cdot k)^2 \tag{2.5}$$

where P_l is the optical transmission power of the AP in Watts and k is the optical to electrical conversion efficiency. If we assume that the user connects to the AP offering the highest signal strength, then the signal power is given by

$$S_u^L = \max_l P_{u,l}. \tag{2.6}$$

Therefore, the Signal to Interference and Noise Ratio (SINR) at the receiver is given by:

$$\mathrm{SINR}_u^L = \frac{S_u^L}{\sum_l P_{u,l} - S_u^L + \mathrm{noise}} \tag{2.7}$$

The link data rate between a user u and a LiFi AP l is calculated using the modified Shannon formula [WH20]:

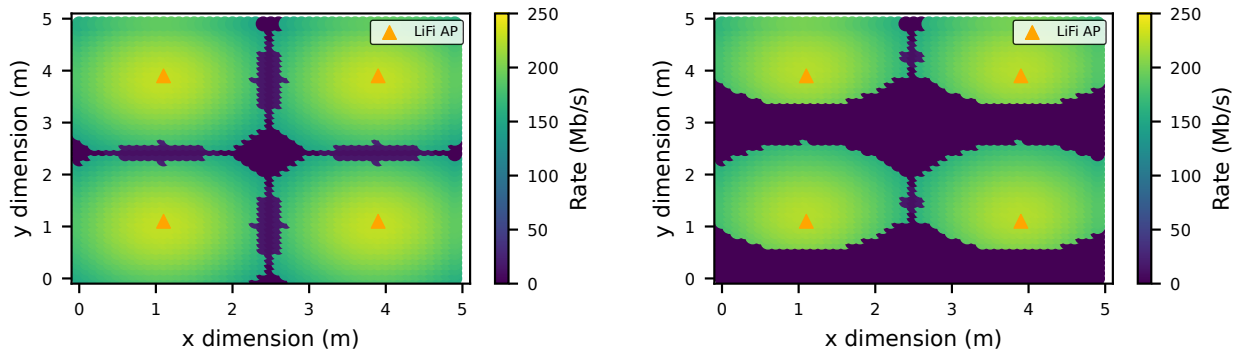
$$R_u^L = \min\left(B^L \cdot \log_2\left(1 + \frac{e}{2\pi} \cdot \mathrm{SINR}_u^L\right), R_{\max}^L\right) \tag{2.8}$$

where B^L denotes the LiFi modulation bandwidth of an LED. We assume a maximum data rate of 250 Mb/s for a LiFi AP, denoted by $R_{\max}^L$. We also introduce the normalized rate $\tilde{R}_u^L$:

$$\tilde{R}_u^L = \frac{R_u^L}{R_{\max}^L}. \tag{2.9}$$

where $\tilde{R}_u^L$ represents the ratio of the achieved rate to the technology's maximum capacity.

Figure 2.4 illustrates the LiFi data rate distribution on a user plane located at a height of 1.4 m, with four LED panels arranged in a lattice grid at a height of 3 m.



(a) Flat device orientation

(b) Walking device orientation

Figure 2.4 LiFi Rate Coverage in a 5 x 5 m room

The Figure 2.4a shows the data rate coverage for a user device positioned flat, parallel to the ceiling, like a laptop. This results in a circular coverage area, typical for APs. In contrast, Figure 2.4b displays the data rate coverage for a user device tilted with a pitch angle of approximately 28°, simulating the use of a mobile phone while walking [Sol+19]. In this case, the coverage area is no longer circular. Understanding how device orientation affects LiFi coverage is vital for optimizing network performance and ensuring reliable communication.

2.4 Blockage Models for LiFi Networks

One of the significant challenges in LiFi systems is their susceptibility to blockages. Understanding and modeling these blockages is crucial for designing robust and efficient LiFi networks. LiFi relies heavily on LoS communication between the transmitter (light source) and receiver (photodetector). Any obstruction in the path can lead to significant signal degradation or complete loss of communication. In typical indoor environments, where LiFi is deployed, objects and people move constantly, creating dynamic and unpredictable blockages. Additionally, changes in the orientation of the receiver, such as a user moving with a mobile device, can easily disrupt the light path. Accurately modeling these dynamics is essential to ensure consistent and reliable performance.

This section describes three blockage models used in our simulations: the instantaneous blockage model, the correlated blockage model, and the geometric blockage model.

2.4.1 Instantaneous Blockage Model

The instantaneous blockage model assumes that blockages occur suddenly and without prior indication. This model is characterized by abrupt transitions between blocked and unblocked states. It is useful for scenarios where obstacles appear and disappear quickly, such as people moving in and out of a LoS path. This model can also be used to capture the effect of fast changes in user device orientation. The instantaneous blockage model is based on a Bernoulli random variable, representing blockage events as probabilistic occurrences. The blockage event, B_{inst}, is defined as:

$$B_{\text{inst}} \sim \text{Bernoulli}(p) \tag{2.10}$$

where p is the probability of a human body obstructing the LoS path at any given instant and is set at 0.1.

2.4.2 Correlated Blockage Model

The correlated blockage model incorporates temporal correlation. This model assumes that blockages are not purely random but have some temporal dependency. For instance, if a blockage is detected, the likelihood of the blockage persisting in subsequent time intervals is higher. This approach provides a more realistic representation of environments where obstacles, such as furniture or stationary objects, influence the blockage pattern over time. The correlated blockage model incorporates memory, where the probability of blockage depends on the previous state:

$$P_{\text{block}}(t) = \begin{cases} \text{prob}_{\text{one}} & \text{if } B_{\text{prev}}(t-1) = 0 \\ \text{prob}_{\text{two}} & \text{if } B_{\text{prev}}(t-1) = 1 \end{cases} \tag{2.11}$$

where $B_{\text{prev}}(t-1)$ indicates the blockage state at the previous time step. prob_{one} is the probability of blockage if previously unblocked, and prob_{two} (with $\text{prob}_{\text{two}} > \text{prob}_{\text{one}}$) is the probability of continued blockage if already blocked. In our thesis, $\text{prob}_{\text{one}} = 0.1$ and $\text{prob}_{\text{two}} = 0.7$.

2.4.3 Geometric Blockage Model

The geometric model represents users as cylindrical obstructions, impacting the LoS paths between ceiling-mounted APs and user devices, based on the work in [Fir+21]. This cylindrical model is chosen for its accuracy in reflecting the height and width of a typical user, making it effective in evaluating signal blockage. This approach is applicable not only to humans but also to robots or other machines. The focus is on modeling users as blockages due to their dynamic nature, which presents a complex challenge for maintaining reliable connections. In contrast, stationary obstacles are not modeled, as they can be strategically positioned or removed through careful environment planning, thereby minimizing their impact on signal transmissions.

Each cylinder is defined by a height of 1.8 m and a radius of 0.2 m, with the height aligning with the vertical axis between the user device and the AP. The position of the cylinder's base is represented by coordinates (x_c, y_c) on the horizontal plane. A LoS path is blocked when the line segment connecting an AP and a user device intersects the cylindrical volume of a user.

Mathematically, the LoS blockage is formulated by considering the geometric relationship between the AP, user device, and cylindrical obstruction. Given an AP positioned at (x_a, y_a, z_a) and a user device at (x_u, y_u, z_u), the LoS path is the line segment joining these points. The blockage condition occurs when this line segment intersects the cylindrical volume of a human body, determined by the cylinder's height h, radius r, and base coordinates (x_c, y_c).

In all these blockage models, when a user's connection to an AP is blocked, the user's rate offered by that AP is 0. Blockage models are integral to the channel model, simulating blockages in the LoS of LiFi links. It influences channel quality and, consequently, the wireless LiFi data rate experienced by the user during transmission or reception. By incorporating the blockage model, we aim to simulate a realistic channel for users, ensuring an accurate representation of the impact on data rates.

2.5 Illumination Model

When LiFi operates using Visible Light Communication (VLC), the APs also provide illumination for the indoor area. Similar to the user grid, the room is divided into an illuminance grid. Each illuminance grid point v is positioned at $\boldsymbol{c}_v = (x, y, z)_v$, with a total of M^V grid positions. The illuminance at grid position v from a single LiFi AP l is defined as:

$$I_{v,l} = I_0 \frac{1}{d_{v,l}^2} \cos \phi_{v,l}^m \cdot \cos \theta_{v,l} \tag{2.12}$$

where I_0 is the luminous efficacy of the LED in lumens/Watt. The total illuminance at grid position v from all LiFi APs is the sum of the individual contributions

$$I_v = \sum_l I_{v,l}. \tag{2.13}$$

The total illuminance achieved using this model is shown in Figure 2.5

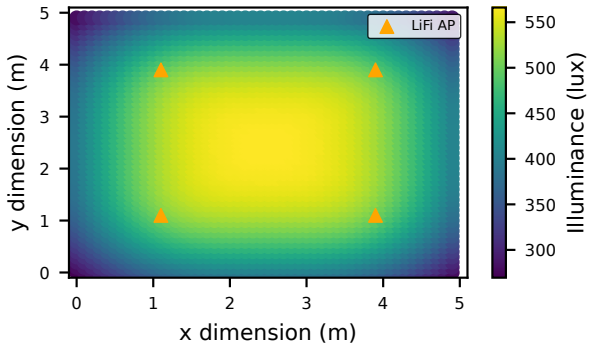


Figure 2.5 Illuminance in lux achieved in a 5 m x 5 m x 3 m room

Furthermore, the illumination uniformity (I) is given by the ratio of the minimum and the average illumination intensity [IAY20]:

$$I = \frac{\min_v I_v}{\frac{\sum_v I_v}{M^V}}. \tag{2.14}$$

The summary of parameters related to the LiFi channel used in the rest of the thesis is given in Table 2.1.

Table 2.1 LiFi Channel Parameters

Parameter	Notation	Value
Optical Power of a LiFi AP	P_l	5 W
Half power beam width	$\theta_{1/2}$	60°
Physical area of the receiver	A_p	1 cm^2
FoV of the receiver	Θ_f	90°
FoV of the transmitter	Φ_f	90°
Optical filter gain	T_s	1
Refractive index	χ	1
Lambertian order	m	1
Modulation bandwidth of LED	B^L	20 MHz
Noise power spectral density	noise	10^{-21} A^2/Hz
Maximum capacity LiFi	$R^L_{\max}$	250 Mb/s

2.6 WiFi Channel Model

We model the WiFi network according to the IEEE 802.11n standard with a channel bandwidth of 40 MHz and a total capacity of 160 Mb/s per AP, denoted by $R^W_{\max}$. Similar to the LiFi channel model, the three-dimensional distance between user u and AP w is defined as:

$$\mathrm{d}_{u,w} = \|\boldsymbol{c}_w - \boldsymbol{c}_u\|_2 \tag{2.15}$$

The channel gain is adapted from [WWH17], rewritten here in linear terms for ease of comparison with the LiFi model. The gain is expressed as: Hence, the gain is given by

$$H_{u,w} = \frac{1}{d^2_{u,w}} \cdot \frac{1}{f^2_W} \cdot 10^{14.45} \cdot h_r \tag{2.16}$$

where h_r is the small scale fading gain, with an average power of 2.46 dB, and f_W is the carrier frequency of transmission. The received signal power is calculated as:

$$P_{u,w} = (H_{u,w} \cdot P_w)^2 \tag{2.17}$$

where P_w is the transmission power of the AP in Watts. Similar to the LiFi model, assuming that the user connects to the AP offering the highest signal strength, the received signal is:

$$S_u^W = \max_w P_{u,w}. \tag{2.18}$$

Assuming frequency reuse and no interference between WiFi APs, the Signal to Noise Ratio (SNR) at the receiver is given by:

$$\mathrm{SNR}_u^W = \frac{S_u^W}{\mathrm{noise}} \tag{2.19}$$

The link data rate between a user u and a WiFi AP w is calculated using the Shannon formula:

$$R_u^W = \min\left(B^W \cdot \log_2\left(1 + \mathrm{SNR}_u^W\right), R_{\max}^W\right) \tag{2.20}$$

where B^W denotes the transmission bandwidth. As with LiFi, the normalized rate for WiFi is introduced as:

$$\tilde{R}_u^W = \frac{R_u^W}{R_{\max}^W}. \tag{2.21}$$

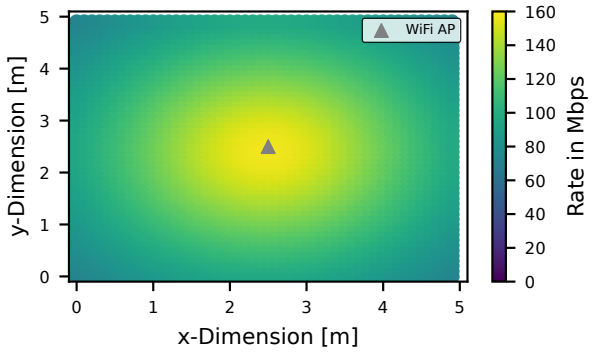


Figure 2.6 WiFi rate coverage achieved in a 5 m x 5 m x 3 m room

Figure 2.6 illustrates the WiFi data rate distribution on a user plane located at a height of 1.4 m, with one WiFi AP placed at the center of the ceiling at a height of 3 m.

The summary of parameters related to the WiFi channel used in the rest of the thesis is given in Table 2.2.

Table 2.2 WiFi Channel Parameters

Parameter	Notation	Value
Small-scale fading gain	h_r	2.46 dB
Transmission Power of a WiFi AP	P_w	0.1 W
Bandwidth of WiFi	B^W	40 MHz
Noise power spectral density	noise	10^{-15} A^2/Hz
Frequency of WiFi	f_W	2.45 GHz
Maximum capacity WiFi	$R_{\max}^W$	160 Mb/s

2.7 Mobility Models

Simulating user movements is essential for accurately assessing the performance of LiFi-WiFi networks, especially when evaluating resource allocation methods. In real-world environments, user

mobility significantly influences network dynamics, affecting connectivity, signal quality, and overall system performance. Properly modeling mobility in simulations enables us to predict and optimize network behavior, leading to more reliable and efficient resource allocation strategies.

User mobility plays a critical role in the handover processes between LiFi and WiFi networks. In hybrid LiFi-WiFi systems, users frequently switch between networks to maintain optimal connectivity [Wu+20]. Effective mobility models simulate these transitions, allowing us to evaluate and optimize handover mechanisms to minimize latency and packet loss, which is crucial for maintaining seamless resource allocation. Mobility models are also vital for studying network coverage and capacity. By simulating various movement patterns, we can identify areas with poor coverage or high congestion and develop strategies to improve resource allocation. This is particularly important in environments like offices, campuses, and public spaces, where user density and movement can vary significantly.

Given these considerations, this section discusses four mobility models used in our simulations to assess the performance of resource allocation methods: Random Waypoint (RWP) with and without pauses, Truncated Levy Walk [Rhe+11], Self-Similar Least Walk (SLAW) [Lee+09], and Reference Point Group (RPG) [Hon+99].

2.7.1 Random Waypoint Model

The RWP model is one of the most widely used mobility models due to its simplicity and suitability for indoor scenarios. In this model, users move randomly from one waypoint to another within the simulation area. Each user selects a random destination and moves towards it with the chosen speed. Once the destination is reached, a new waypoint is selected, and the process repeats. The RWP model can be implemented in two versions: with and without pauses.

In the version with pauses, after reaching a waypoint, the user pauses for a random duration before selecting the next waypoint. This pause time introduces variability in the movement patterns, simulating realistic scenarios where users stop occasionally. In the version without pauses, users continuously move from one waypoint to another without any pause. The RWP model with pausing is particularly useful for simulating mobile users in wireless communication scenarios. This model introduces moments of stationary activity, making it more realistic for environments where users might pause, such as indoors or at specific points of interest. The pausing condition is typically determined by factors such as user behavior, environment characteristics, or application requirements. The distribution governing the pausing duration can be chosen based on the desired characteristics of the model.

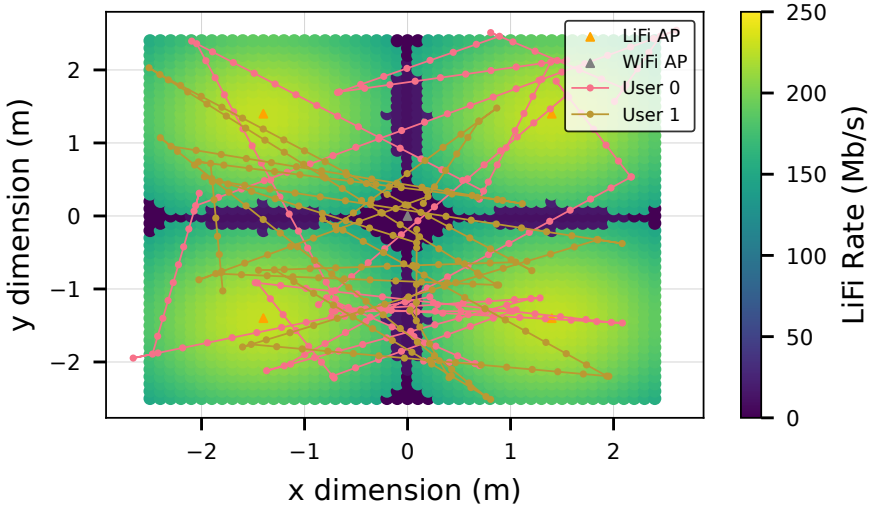


Figure 2.7 Trajectory of two exemplary users following the RWP Mobility model overlaid on the rate coverage of 4 LiFi APs along with one WiFi AP positioned at the center of the room

Figure 2.7 illustrates the trajectories of two users simulated using the RWP model without pausing. The users' movements are depicted over time, showcasing their paths in the LiFi-WiFi network environment. AP positions are overlaid on the trajectory, highlighting cases where users move into interference regions of LiFi APs, necessitating support from the WiFi AP for reliable connectivity. In the rest of this thesis the term RWP refers to the model without pausing unless otherwise specified.

2.7.2 Truncated Levy Walk Mobility Model

The Truncated Levy Walk model captures more complex and realistic human movement patterns compared to the RWP model. This model is based on the observation that human movements often consist of many short trips interspersed with occasional long-distance travels. The model uses a step length distribution that follows a power-law for small distances, transitioning to an exponential decay for larger distances. This truncation prevents excessively long steps that are unrealistic in natural settings. Each step in the walk is followed by a change in direction, which is uniformly distributed. This truncated Levy distribution helps maintain a balance between long and short steps, effectively capturing the intermittent nature of human mobility. This model is particularly useful for simulating scenarios where user movements exhibit heavy-tailed distributions, such as in urban environments.

Let the user's position at time t be denoted as $(x(t), y(t))$ in a 2D space. The levy walk model describes the movement of a user by defining random steps of varying lengths and directions. The position update is given by:

$$\begin{aligned} \Delta s &\sim P(s) \\ \Delta\theta &\sim U(0, 2\pi) \\ x(t+1) &= x(t) + \Delta s \cos(\Delta\theta) \\ y(t+1) &= y(t) + \Delta s \sin(\Delta\theta) \end{aligned}$$

where Δs is the step length, drawn from a Power-law distribution $P(s) \sim s^{-\mu} \exp\left(-\frac{s}{s_{\max}}\right)$ where μ is the exponent characterizing the power-law regime, and $s_{\max}$ is the truncation length. $\Delta\theta$ is the step direction, drawn from a uniform distribution $U(0, 2\pi)$.

2.7.3 Self-similar Least Action Walk (SLAW) Model

The SLAW model incorporates both fractal-like (self-similar) properties and least-action principles observed in human mobility. This model is based on the observation that human movements exhibit self-similarity over different time scales. For example, the way people move around a city on a daily basis is similar to how they might move around within a neighborhood over a shorter period. The SLAW model replicates this fractal-like behavior, ensuring that the movement patterns remain consistent whether observed over short or long duration. Humans tend to minimize their effort when moving, choosing the shortest or easiest paths to reach their destinations. This is the least-action principle. The SLAW model captures this by generating movement paths that favor shorter, more direct routes, while still allowing for occasional longer trips. Other than these aspects, the model incorporates aspects of Levy walks, where the step sizes (distances between consecutive waypoints) follow a heavy-tailed power-law distribution. This property allows the model to efficiently cover large areas with a minimal number of long jumps, interspersed with many shorter, more localized movements.

It is particularly effective in simulating realistic human mobility patterns where users frequently visit a set of preferred locations, such as in urban and campus environments. These preferred locations as represented by waypoints that are clustered together following a Hurst parameter. The waypoints are generated such that the density of waypoints decreases with distance, emulating the fractal nature of human mobility. A simplistic description of the working of the SLAW model is given in Algorithm 1.

Algorithm 1 Self-Similar Least Walk (SLAW) Model

Input: Number of waypoints N, Hurst parameter H, Power-law parameter μ
Output: Sequence of waypoints and paths
Initialize starting location x_0
Generate N waypoints $\{x_1, x_2, \ldots, x_N\}$ using fractal distribution with parameter H
for $i \leftarrow 1$ to N **do**
    Generate flight length $s_i \sim s^{-\mu}$
    Choose direction randomly
    Set $x_i \leftarrow x_{i-1} + s_i$
end for
Compute initial action $S \leftarrow 0$
for $i \leftarrow 1$ to $N - 1$ **do**
    Calculate cost $c(x_i, x_{i+1})$ considering total distance to travel
    Update action $S \leftarrow S + c(x_i, x_{i+1})$
end for
Optimize path to minimize total action S

2.7.4 Reference Point Group (RPG) Mobility Model

The RPG model simulates scenarios where users move in groups, such as at conferences, events, or guided tours at museums. Each group has a logical center or reference point that moves according to the RWP model or another mobility model. Individual users within the group move around this reference point, maintaining cohesion with the group while exhibiting individual variability. The RPG model effectively captures the dynamics of group mobility, where users show correlated movements but also maintain some level of individual movement.

Let $G = \{G_1, G_2, ..., G_N\}$ be the set of reference point groups, where N is the total number of groups. Each group G_i consists of M_i reference points, denoted as $P_{i,j}$ for $j = 1, 2, ..., M_i$. The position of a reference point $P_{i,j}$ at time t is defined as $(x_{i,j}(t), y_{i,j}(t))$, representing its coordinates in a 2D space. The movement of reference points is determined by:

$$x_{i,j}(t+1) = x_{i,j}(t) + \Delta x_{i,j}(t)$$
$$y_{i,j}(t+1) = y_{i,j}(t) + \Delta y_{i,j}(t)$$

where $\Delta x_{i,j}(t)$ and $\Delta y_{i,j}(t)$ are the displacements of $P_{i,j}$ in the x and y directions, respectively. These displacements can be modeled using various methods, such as random walks, correlated movements, or predefined trajectories.

The overall movement of the group G_i is influenced by the movements of its constituent reference points. The center of mass of the group, denoted as $(X_{c,i}(t), Y_{c,i}(t))$, is calculated as:

$$X_{c,i}(t) = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{i,j}(t)$$

$$Y_{c,i}(t) = \frac{1}{M_i} \sum_{j=1}^{M_i} y_{i,j}(t)$$

This center of mass represents the collective movement of the group.

For all these mobility models, the height of the user device over time follows a Gaussian distribution, with a mean height of 1.4 m for mobile users and 0.8 m for stationary users, and a standard deviation of 5 cm. The orientation generation for these devices is consistent across all models. After determining the x and y positions using the respective mobility model, the direction of movement is calculated to determine the yaw angle of the user's orientation, as described in [Sol+19]. The pitch angle follows

a truncated Laplace distribution with a mean of 28° for moving users and 0° for stationary or sitting users. The roll angle also follows a truncated Laplace distribution, with a mean of -1.35°. The overall user orientation evolves over time as a correlated random process, as described in [Sol+19].

2.8 Optimization Algorithms

In our research on LiFi-WiFi networks, we frame optimization problems to enhance network performance, particularly in the placement of APs and resource allocation. We employ a variety of optimization techniques to find optimal or near-optimal solutions and compare these tools to find the best approach to solving these problems. This chapter provides an overview of the common algorithms used in our work, focusing on their fundamental concepts and their relevance to our specific research context. Detailed problem-specific tuning of these algorithms will be discussed in the respective chapters. The techniques covered include Meta-heuristics, Black-box optimizers, Evolutionary Game Theory, Lagrangian Multipliers Method, Branch and Bound, and the deep reinforcement learning-based Proximal Policy Optimization.

Branch and Bound is an algorithmic method for solving integer programming and combinatorial optimization problems. It systematically explores the solution space by dividing it into smaller subproblems and using bounds to eliminate those that cannot yield better solutions than the current best solution. Branch and Bound is particularly effective for solving large-scale optimization problems with discrete variables, such as the optimal placement of APs in a network.

The Lagrangian Multipliers Method is a powerful technique for solving constrained optimization problems. By transforming a constrained problem into an unconstrained one, this method allows us to incorporate various network constraints into the optimization process. This technique is particularly useful for finding optimal solutions that satisfy all necessary constraints while maximizing network performance.

Evolutionary Game Theory (EGT) applies the principles of game theory to evolving populations, providing a framework for modeling and solving optimization problems where multiple agents interact and adapt over time. In our research, EGT is used to model competitive and cooperative interactions among and users, enabling us to develop strategies that model users competing for shared bandwidth resources at APs.

Meta-heuristics are high-level procedures designed to generate good-enough sub-optimal solutions to optimization problems, especially those that are too complex for exact methods. In our research, the primary focus within meta-heuristics is on Genetic Algorithms (GAs). Inspired by the process of natural selection, GAs are particularly effective for solving complex optimization problems involving large search spaces. In the context of AP placement and resource allocation, GAs help in exploring various configurations by simulating the evolutionary processes of selection, crossover, and mutation. This iterative approach enables us to identify high-quality solutions that balance network performance and efficiency.

Proximal Policy Optimization (PPO) is a reinforcement learning algorithm that has gained significant popularity due to its robustness and ease of implementation. PPO falls under the class of policy gradient methods, which optimize the policy directly by adjusting the policy parameters to maximize the expected return. Unlike traditional policy gradient methods that can be sensitive to step sizes, PPO uses clipping mechanisms that constrain the policy updates, preventing them from deviating too far from the current policy. PPO has been effectively used in various domains, including robotics, game playing, and autonomous driving, demonstrating strong performance and sample efficiency. Its balance between complexity and performance makes PPO a widely adopted choice for reinforcement learning tasks. PPO is especially effective in high-dimensional, continuous action spaces, making it suitable for dynamic resource allocation in LiFi-WiFi networks. By learning optimal policies through interactions with the environment, PPO helps in adapting to changing network conditions and user behaviors.

Black-box optimizers are algorithms that do not require explicit knowledge of the underlying objective function's structure or derivatives, making them ideal for optimizing complex and non-differentiable functions. In our research, we focus on Bayesian Optimization, Random Forests, and Extra Trees. Bayesian Optimization is particularly useful for optimizing expensive-to-evaluate functions. It builds a probabilistic model of the objective function and uses this model to select the most promising points for evaluation. In the context of AP placement, Bayesian Optimization helps efficiently explore the configuration space, reducing the number of evaluations needed to find optimal solutions. Random Forests and Extra Trees are ensemble learning techniques used to model the relationship between input parameters and network performance metrics. These methods provide insights into the importance of different parameters and help identify optimal configurations. Their ability to handle large datasets and complex interactions makes them valuable tools in our optimization framework.

2.9 Statistical testing

In our research, statistical testing is fundamental to validate our findings. This section outlines the statistical methods employed across all evaluations presented in this thesis, focusing on the Mann-Whitney U test [MW47] and the Benjamini-Hochberg procedure [BH95]. The motivation behind these testing methods is to find out whether observed differences in performance metrics are statistically significant or merely due to random variation. To compare two independent samples and determine whether they come from the same distribution, we employed the Mann-Whitney U test. This non-parametric test is particularly suitable for our data, which may not adhere to the normal distribution assumptions required by parametric tests such as the t-test.

The Mann-Whitney U test operates under the null hypothesis that there is no difference in the distributions of the two parameters being compared. This test ranks all the values from both groups together, then compares the sum of ranks between the groups. A significant p-value indicates that one group tends to have higher or lower values than the other. The significance level applied in this thesis is 5% or 0.05.

The results of the Mann-Whitney U test are annotated in the corresponding result figures with symbols based on the p-values, using the following star notation:

$$\begin{aligned} \text{ns} &: p > .05 \\ * &: .01 < p <= .05 \\ ** &: .001 < p <= .01 \\ *** &: .0001 < p <= .001 \\ **** &: p <= .0001 \end{aligned}$$

When conducting multiple statistical tests on the same dataset, there is an increased risk of Type I errors (false positives). To address this, we applied the Benjamini-Hochberg procedure to control the false discovery rate (FDR). The Benjamini-Hochberg procedure adjusts the p-values obtained from multiple hypothesis tests to account for the number of tests being performed. By applying this method, we ensure that the overall rate of false discoveries remains controlled, thereby enhancing the reliability of our statistical inferences.

2.10 Summary

This background section provides a comprehensive foundation for the advanced concepts discussed in this thesis, covering essential topics such as LiFi technology, LiFi-WiFi heterogeneous networks, channel models, blockage models, illuminance models, mobility models, optimization algorithms, and statistical testing of results. We began with an introduction to LiFi technology, highlighting its

high data transfer rates, enhanced security, and abundant bandwidth. The integration of LiFi with existing RF-based networks to create heterogeneous networks was then explored, emphasizing the benefits of seamless connectivity and optimized performance.

We discussed the specifics of channel modeling for both LiFi and WiFi, which is crucial for predicting signal behavior and optimizing network performance. The discussion on blockage models for LiFi focused on the impact of physical obstructions and user movements, presenting various models to simulate real-world scenarios. The illuminance model for LiFi addressed the dual function of providing adequate lighting while maintaining high data transmission rates. Mobility models were introduced to simulate user movements and assess resource allocation methods. We explored optimization algorithms, such as Meta-heuristics, Black-box optimizers, Evolutionary Game Theory, Lagrangian Multipliers Method, Branch and Bound, and Proximal Policy Optimization, which are crucial for solving complex optimization problems in LiFi-WiFi networks. Lastly, statistical testing methods like the Mann-Whitney U test and Benjamini-Hochberg procedure are described to ensure the accuracy and reliability of our findings

Building on this foundational knowledge, the rest of the thesis will present strategies to design and manage LiFi-WiFi networks. By leveraging the concepts and methods discussed in this background section, we aim to offer solutions that enhance the integration and operation of LiFi within heterogeneous networks, ultimately leading to reliable, efficient, and high-performing heterogeneous systems.

Chapter 3

Designing LiFi-WiFi Networks with a focus on Access Point Placement

Wireless communication technologies have seen significant advancements driven by the increasing demand for fast and reliable connectivity. While Wireless-Fidelity (WiFi) networks have long been the dominant communication means, the emergence of Visible Light Communication (VLC) and Infrared (IR) communication networks, such as Light-Fidelity (LiFi), presents a promising alternative. LiFi offers advantages like higher bandwidth and enhanced security compared to WiFi.

Integrating multiple wireless technologies becomes essential as the demand for high data rates, low latency, and ubiquitous coverage grows. One emerging combination is the LiFi-WiFi heterogeneous network, which efficiently utilizes available resources, enhances coverage and capacity, and provides flexibility to accommodate diverse user demands. Designing such networks, however, is challenging due to cost constraints, rate coverage requirements, and the need for adequate illumination.

A critical aspect of both LiFi-only and LiFi-WiFi heterogeneous network design is the placement of Access Points (APs). The LiFi channel's rapid degradation with distance confines communication to short ranges, necessitating ultra-dense deployment of LiFis cells. Optimal placement of LiFi APs is essential for providing optimal communication coverage and proper illumination indoors. In heterogeneous networks, the complexity increases as the distinct characteristics of LiFi and WiFi technologies must be considered.

This chapter integrates insights from a conference [Vij+22] and journal publication [Vij+23]. These combined insights address complementary aspects of the same challenge: placement of APs to achieve optimal network performance and infrastructure cost-efficiency in next-generation indoor wireless networks. The first section of the chapter focuses on AP placement in LiFi-only networks, based on [Vij+22]. This section lays the groundwork by addressing the fundamental issues and strategies for optimizing the placement of LiFi APs, operating on visible light, to ensure robust communication coverage and adequate illumination. The second section builds on this foundation by addressing the more complex scenario of AP placement in LiFi-WiFi heterogeneous networks, as published in [Vij+23]. Here, the integration of WiFi introduces additional variables and challenges due to the differing properties of LiFi and WiFi channels, necessitating more sophisticated optimization techniques.

Each section presents a unique contribution, beginning with an introduction and motivation, followed by a state-of-the-art review for the specific contribution. Subsequently, methodologies and results are discussed, ending in a concluding summary.

3.1 Optimized 3D Placement of LiFi Access Points towards Maximizing Wireless Network Performance

The unique characteristics of LiFi, such as its short range where the channel degrades rapidly with distance and sensitivity to Line-of-Sight (LoS) conditions, pose challenges for effective network deployment. Therefore, LiFi cells are typically deployed ultra-densely. In such networks, the placement

of LiFi APs or Light Emitting Diodes (LEDs) is crucial for providing optimal communication coverage and indoor illumination. In this work, the terms LiFi AP and LED are used interchangeably.

While several studies analyze LiFi cell coverage, they often overlook optimizing the number of APs. Deploying more LEDs can ensure adequate illumination but can lead to network interference in overlapping coverage areas and increase the number of handovers. Conversely, using fewer LEDs can create coverage gaps and fail to meet user and illumination requirements. When optimizing AP positions, the height of each access point from the user plane significantly affects communication range and Signal to Interference and Noise Ratio (SINR) at user devices, impacting overall network performance. Additionally, the expected user distribution in an indoor environment must be considered to optimize placement effectively. Although some studies optimize LiFi AP placement for objectives like minimizing the number of APs or maximizing network throughput, they do not address both objectives simultaneously. Considering both objectives simultaneously like as in Multi-Objective Optimization (MOO) is essential because these objectives are contradictory and must be balanced for optimal placement.

This section focuses on the optimized 3D placement of LiFi APs to maximize network performance while ensuring adequate illumination. Unlike traditional 2D placement approaches, this research considers the height of the APs as a critical factor, formulating the placement problem as a multi-objective optimization. By addressing the trade-offs between minimizing the number of APs, maximizing data rates, and providing sufficient illumination, this work aims to enhance the overall performance and efficiency of LiFi networks.

3.1.1 State-of-the-art Analysis

In [Yan+20], the authors address an LED array with a fixed number of LEDs, deriving the optimal x,y positions to minimize power consumption while meeting data rate and illumination constraints. In [VB19], the authors consider a variable number of LEDs, optimizing their 2D placement along the x and y coordinates to minimize the number of LEDs. However, these approaches did not consider the height of the AP as a variable.

While [VB19] accounts for expected user distribution in a room, it does not optimize for network performance in terms of throughput. Reference [DBS20] focuses on maximizing average throughput, incorporating the stationary distribution of users based on the Random Waypoint (RWP) mobility model. However, this model may not accurately represent many indoor scenarios. Our analysis includes multiple indoor scenarios with varying user distributions. Additionally, the authors in [DBS20] assume dedicated APs for specific regions, which raises issues in interference-prone areas and ignores user device orientation, although orientation is crucial in LiFi communication, as receivers have a Field of View (FoV) outside of which signals are not received. The works mentioned do not simultaneously optimize the number and position of APs nor consider height as a variable.

In contrast, works like [Pan+19] and [UUC19] propose solutions for the 3D placement of Unmanned Aerial Vehicles (UAVs) for wireless communication resource allocation, assuming known user positions. This assumption is invalid for our problem since LED placement optimization should occur during the network planning stage before deployment in a user environment.

3.1.2 Key Contributions

This section focuses on optimizing the 3D placement of APs in LiFi-only networks. Our contribution builds on foundational works in network planning by framing a complex multi-objective optimization problem tailored to the unique characteristics of LiFi communication and illumination in indoor environments. The primary objectives of this research include minimizing the number of APs to reduce costs, while maximizing the sum rate and ensuring a minimum guaranteed achievable rate and adequate illumination. To achieve these goals, we incorporate several approaches and methodologies, as detailed below:

1. Multi-Objective Optimization: This research analyzes the planning of a LiFi communication and illumination network in an indoor environment, framing an optimization problem with multiple objectives: minimizing the number of APs (thereby reducing costs), and maximizing the sum rate while ensuring a minimum guaranteed achievable rate and minimum illumination level. Additionally, we consider the expected user distribution to calculate expected rates.
2. 3D Placement Framework: Among the optimization variables is the height of each AP, which we allow to be either freely placed (3D free-height) or constrained such that all APs share the same height (3D fixed-height). This approach highlights the impact of AP height on network performance.
3. Optimization Algorithm: To solve this 3D placement problem, we propose using a genetic MOO algorithm. We evaluate the solutions for varying system parameters and scenarios, demonstrating the validity of our optimization framework.

3.1.3 System Model

Table 3.1 summarizes the notation used throughout this section. We use bold lowercase letters for vectors and cursive capital letters for sets.

Table 3.1 List of Notations used in the 3D Placement of LiFi APs

Notation	Description
l, M^L	Index, size of LiFi APs
$u, M^U, \mathcal{U}$	Index, size, set of all user positions
v, M^V	Index of, total positions on the illuminance grid
$\boldsymbol{c} = (x, y, z)$	Vector of 3D coordinates
$\boldsymbol{\Theta_u} = (\Theta^Y, \Theta^P, \Theta^R)_u$	Vector of yaw, pitch and roll of user device
p_u^L	Probability of occurrence of LiFi
$R_u, \tilde{R}_u$	Achievable and normalized achievable rate
I_v	Illuminance at the grid position
I_0	Luminous efficacy of the LED
α	Binary existence variable for AP
$\tilde{R}_{\text{thresh}}$	Normalized Rate requirement
I_{thresh}	Illumination requirement

This work considers an indoor LiFi network designed to provide data communication and illumination. The network consists of a maximum of M^L APs, which can be positioned in three dimensions, with the coordinates $\boldsymbol{c}_l = (x, y, z)_l$. The horizontal coordinates are confined to the dimensions of the room, and the vertical coordinate ranges from a minimum height z_{mindim}, at least one meter above the user plane to prevent receiver saturation, to a maximum height z_{maxdim}, which corresponds to the ceiling height. This flexibility in height allows the optimization framework to exploit the vertical dimension to minimize interference and enhance signal coverage.

The user plane, where users are expected to be, is quantized into a grid with a resolution of 0.25 m, and each grid position is represented by $\boldsymbol{c}_u = (x, y, z)_u$, totaling M^U positions. The orientation of each user at a grid position is given by angles $\boldsymbol{\Theta_u} = (\Theta^Y, \Theta^P, \Theta^R)_u$ representing Yaw, Pitch, and Roll. Since the exact positions of users are unknown and can vary dynamically, we model the user location with a probabilistic distribution. Each grid position u is associated with a weight p_u^L representing the likelihood of a user being present at that location. Users connect to the AP offering the highest SINR which is dependent on where the APs are placed, adding flexibility to our model as there are no dedicated APs pre-associated to the users. The 3D placement of LiFi APs is optimized during the network planning phase, after which the LiFi-enabled LEDs are deployed indoors.

The LiFi channel model is detailed in Section 2.3. The LiFi LEDs, operating on the visible light spectrum, also provide environmental illumination. The area where we expect a certain illumination level is quantized into the illuminance grid which is not necessarily identical to the user grid. One quantized grid position is denoted by $(x, y, z)_v$ with M^V such positions. The illuminance I_v at grid position v is described in Section 2.5. The simulation parameters for the LiFi channel are consistent with those listed in Table 2.1, in addition to the luminous efficacy of the LED, I_0, set at 525 lm/Watt.

3.1.4 3D Placement Problem Formulation

This section outlines the formulation of the 3D placement optimization problem. The objective is to determine the optimal 3D placement of LED APs to minimize the number of APs (and thus the cost) and maximize the sum rate on the user plane, weighted by the user occurrence probability P_u. The optimization problem is constrained by a minimum guaranteed rate requirement $\tilde{R}_{\text{thresh}}$ at all user grid positions, ensuring reliable coverage across the user plane. Since the exact number of users is unknown at the planning stage, wireless resource sharing is not considered; instead, the focus is on overall user plane performance. To ensure the model's applicability to future technology generations with higher capacities, rate constraints are expressed as a ratio of the maximum supported rate rather than absolute values. Additionally, LEDs must provide a minimum required illumination level I_{thresh}. This illumination constraint can apply to different planes, such as floor level or desk height, and is separate from the user plane.

The optimization problem can be formulated as follows

$$\min_{\boldsymbol{c}_l, \alpha_l} \quad \sum_{l=1}^{M^L} \alpha_l \tag{3.1}$$

$$\max_{\boldsymbol{c}_l, \alpha_l} \quad \sum_{u=1}^{M^U} \tilde{R}_u^L p_u^L \tag{3.2}$$

$$\text{s.t.} \quad \tilde{R}_u^L \geq \tilde{R}_{\text{thresh}} \qquad \forall u = 1, 2, \ldots, M^U \tag{3.3}$$

$$I_v \geq I_{\text{thresh}} \qquad \forall v = 1, 2, \ldots, M^V \tag{3.4}$$

The objective function in (3.1) focuses on minimizing the cost, while the function in (3.2) aims to maximize the sum rate for the 3D placement problem. The constraint in (3.3) ensures that every position on the user grid meets a minimum guaranteed rate, expressed as a ratio of the maximum supported rate. The constraint in (3.4) mandates a minimum illumination level at each position on the illumination grid.

The optimization variable α_l represents the presence of an AP, a binary variable set to 1 if the AP is placed and 0 otherwise. To highlight the impact of AP height on network performance, we consider two different height models in the optimization:

3D free-height: This model allows the height of each AP to be independently selected within specified bounds. The AP position variables are:

$$\begin{aligned} &x_l \in [0, x_{\text{dim}}] \qquad y_l \in [0, y_{\text{dim}}], && l = 1, 2, \ldots, M^L \\ &z_l \in [z_{\text{mindim}}, z_{\text{maxdim}}], && l = 1, 2, \ldots, M^L \end{aligned}$$

3D fixed-height: This model constrains all AP heights to be the same, but this value is freely selected within the bounds. The AP position variables are:

$$\begin{aligned} &x_l \in [0, x_{\text{dim}}] \qquad y_l \in [0, y_{\text{dim}}], && l = 1, 2, \ldots, M^L \\ &z_l = z^L && \forall l = 1, 2, \ldots, M^L, \\ &z^L \in [z_{\text{mindim}}, z_{\text{maxdim}}]. \end{aligned}$$

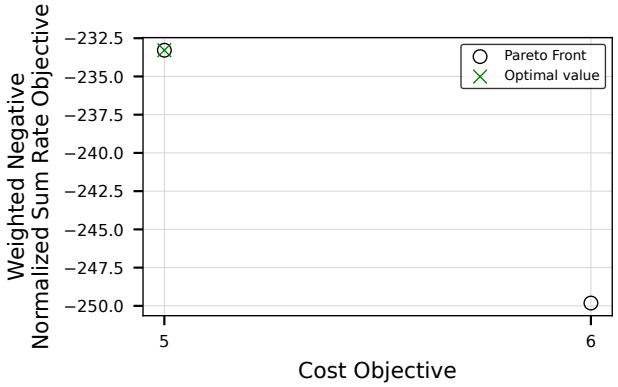


Figure 3.1 An exemplary Pareto-front for the 3D Placement Optimization

The described optimization problem is a MOO problem. MOO problems involve multiple conflicting objective functions, resulting in a set of solutions that offer the best trade-offs between objectives, known as Pareto-optimal solutions [MA03]. Unlike single-objective optimization, where the value of the objective function determines the superiority of a candidate solution, in MOO problems, a solution's superiority is judged by its dominance. All non-dominated solutions within the feasible space form the Pareto-optimal front. Our optimization problem is an Mixed Integer Nonlinear Programming (MINLP) problem involving integer variables α_l and real variables $(x, y, z)_l$. Typically, MINLP problems are mathematically challenging and often intractable.

3.1.5 Method to solve the 3D Placement Problem

To solve this MINLP, a genetic algorithm is employed. Genetic Algorithms (GAs), or evolutionary algorithms, are meta-heuristic algorithms that operate on a set of candidate solutions. They select the fittest candidates from each generation, which are then reproduced to create candidates for the next generation. A standard and powerful MOO algorithm based on GA is the Non-dominated Sorting Genetic Algorithm (NSGA-II) [Deb+02], which classifies solutions into multiple non-dominated sets.

The proposed solution method consists of the following components.

1. Population: The population comprises all possible solutions, where each individual represents a set of AP placements with their 3D coordinates. The algorithm starts with an initial population randomly sampled from within the bounds described by the optimization problem. The initial population size is set to 100 candidate solutions.

2. Selection: Individuals are grouped into fronts by their fitness values, determined by evaluating the objective functions. The fronts are ranked by their level of non-domination. The best individuals are selected by comparing their rank and a second metric called the crowding distance, which is determined by the density of solutions around each candidate. Individuals that violate constraints are made undesirable, as their objective function values (and hence their fitness) are penalized.

3. Crossover: Selected individuals are combined using the Simulated Binary Crossover [DA+95] operator to produce offspring that will form part of the next generation.

The algorithm converges when the constraints are satisfied, and the solutions belong to the Pareto-optimal front. The optimization terminates upon convergence.

An exemplary run with two solutions in the Pareto-front is shown in Figure 3.1. Two objective functions define the objective space: minimizing the number of APs and maximizing the weighted normalized sum rate (which can also be formulated as minimizing the negative sum rate). A Multi-Criteria Decision Making (MCDM) method is employed to select a unique solution from the Pareto

front. The two objectives are weighted according to their importance. In this case, we set the weights to 0.8 for the cost minimization objective and 0.2 for the rate maximization objective, prioritizing the number of access points used. Each solution in the Pareto front is assigned a pseudo weight corresponding to its normalized distance to the worst solution of each objective function. The solution with pseudo weights closest to the objective weights is considered optimal.

3.1.6 Evaluation Methods

To evaluate the proposed 3D placement optimization framework, we consider various application areas that reflect common indoor environments. Each scenario is designed to test the framework under different user distributions and layout configurations, as illustrated in Figure 3.2.

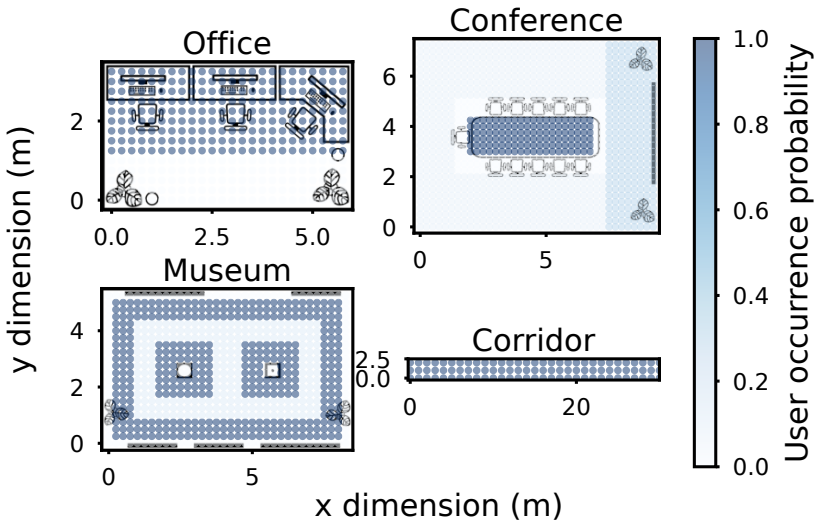


Figure 3.2 Floor plans of indoor scenarios with LiFi user occurrence probability

The selected scenarios include an office, conference room, museum, and corridor, each with unique characteristics and requirements as summarized in Table 3.2. These include two distinct illumination requirements: one for work desks and another for other areas. The pitch angle Θ^P of the user is set to either 0° for seated users or 28° for standing users. The pitch angle of 28° is based on typical measurements for standing users [Sol+19].

Table 3.2 Scenario parameters to evaluate the proposed LiFi AP Placement

Scenario	Size	$\tilde{R}_{\text{thresh}}$	I_{thresh}	$(z_{\text{mindim}}, z_{\text{maxdim}})$
Regular	5.0 m x 5.0 m	0.01	300 lux	(2.5 m, 3.5 m)
Office	6.0 m x 3.5 m	0.01	500, 200 lux	(2.5 m, 3.5 m)
Museum	8.5 m x 5.5 m	0.01	200 lux	(2.5 m, 3.5 m)
Conference	9.5 m x 7.5 m	0.01	500, 200 lux	(2.5 m, 3.5 m)
Corridor	30 m x 2.5 m	0.01	300 lux	(2.5 m, 3.5 m)

All results presented are based on 1000 simulation runs. To validate our claims, we performed hypothesis testing using the Mann–Whitney U test [MW47]. The alternative hypothesis states that

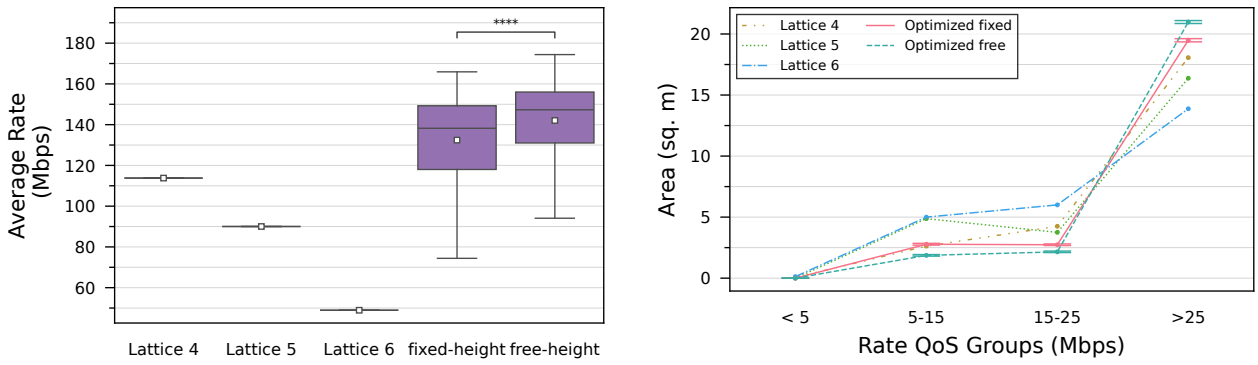
the distribution underlying the left box plot is stochastically less than that of the right box plot. The test results are annotated on the relevant figures [Cha+22], following the notation provided.

$$\begin{aligned} \text{ns} &: p > .05 \\ * &: .01 < p <= .05 \\ ** &: .001 < p <= .01 \\ *** &: .0001 < p <= .001 \\ **** &: p <= .0001 \end{aligned}$$

Since we perform multiple tests on the dataset, we also apply the Benjamini–Hochberg procedure to control the false discovery rate [BH95].

3.1.7 Results and Comparative Analysis

To validate the necessity of our optimization approach, we compare it against several baseline scenarios where LEDs are placed in a deterministic pattern on the ceiling. One such pattern is a lattice grid with four LEDs, as illustrated in Figure 2.4. We also compare lattice grid patterns with 5 and 6 LEDs. The optimization is conducted for a standard 5 m x 5 m room scenario, with a rate ratio requirement of 0.01 and an illumination requirement of 300 lux. The average rate in the room after optimization is presented in Figure 3.3a, alongside the results from the deterministic models.



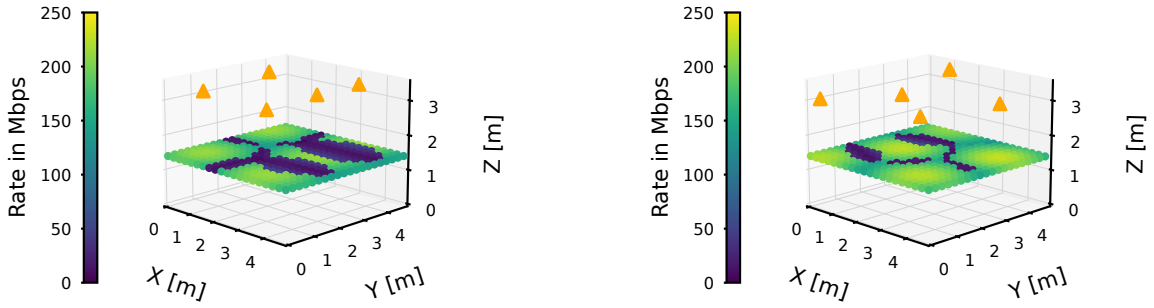
(a) Average rate achievable in Mb/s

(b) Coverage area per rate Quality of Service (QoS) level

Figure 3.3 Results for varying placement models in a Regular room showing the need for optimization

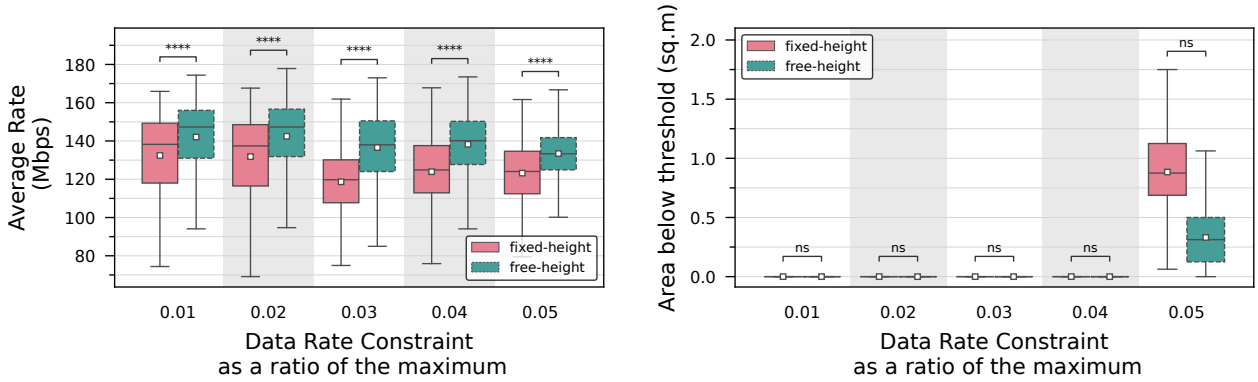
In the deterministic model, increasing the number of access points leads to lower rates due to increased interference. This demonstrates the clear need for optimization, as the average achievable rate is significantly higher than the deterministic models. To further analyze these results, we divide the data into four rate groups and plot the area of the room covered by each rate range. This analysis for the different strategies is shown in Figure 3.3b. The optimized placement results in substantially larger areas achieving data rates higher than 25 Mb/s.

Examining the results for the free and fixed height models in the regular scenario, as shown in Figure 3.4, we observe that the fixed-height model places all APs near the ceiling to maximize the rate. In contrast, the free-height model utilizes varying heights for the APs, which helps minimize interference regions and achieve a higher overall rate.



(a) fixed-height Model **(b)** free-height Model

Figure 3.4 3D Positions of APs and Rate coverage in a regular room



(a) Average rate achievable in Mb/s **(b)** Outage Area in sq.m

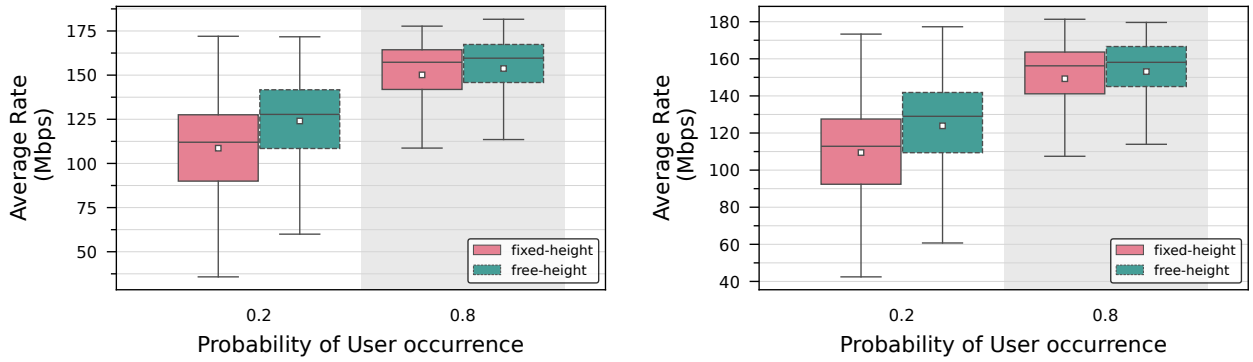
Figure 3.5 Results for varying minimum rate requirements in a regular room

Figure 3.5a shows the average rate achieved in Mb/s for various minimum rate requirements in a regular room. In all cases, the free-height model outperforms the fixed-height model. This finding is statistically confirmed with hypothesis tests, annotated in the figure.

Figure 3.5b illustrates that the free-height model consistently provides coverage as the minimum rate requirement increases, resulting in a negligible outage area. When the minimum requirement is set to 12.5 Mb/s (0.05 rate ratio), both height models fail to meet the constraint. However, the free-height model achieves a significantly lower outage area and a higher average rate of 132 Mb/s. These results highlight the critical role that AP height plays in enhancing network performance.

Next, we analyze the effect of user occurrence probabilities in an indoor environment. For this purpose, the regular room is modified so that half of the user plane has a low probability of 0.2 while the other half has a higher probability of 0.8. To ensure robustness in the results, we also examine a scenario where the user plane is flipped, starting with a probability of 0.8 and ending with 0.2. The average rate achieved in each probability group is shown in Figure 3.6.

Our proposed 3D placement framework consistently maximizes the rate in areas with higher user occurrence probabilities. This demonstrates the importance of incorporating user occurrence probabilities in optimizing LED placement, ensuring optimal performance in regions where users are most likely present.

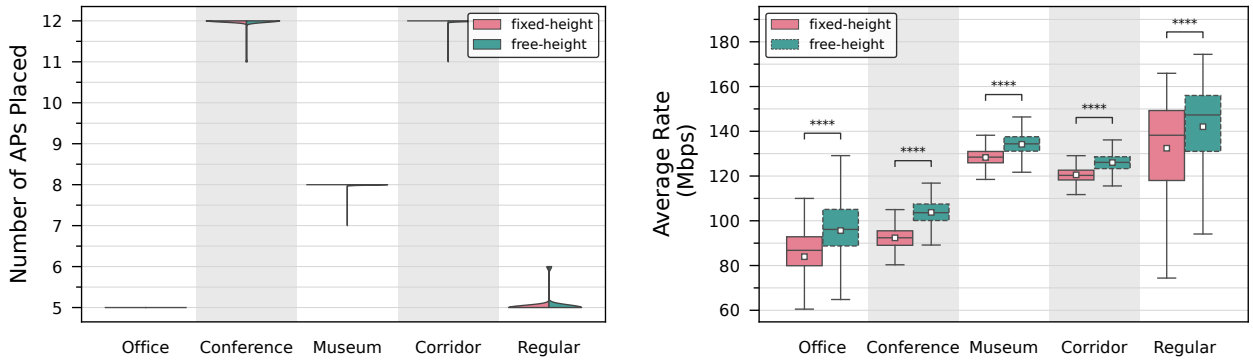


(a) User Occurrence (0.2, 0.8)

(b) User Occurrence (0.8, 0.2)

Figure 3.6 Rate achieved per user probability group in regular rooms where the first half of the user plane is weighted with the first probability value in the tuple and the second half with the second value of the tuple

To understand the planning required for various indoor scenarios, we optimize the 3D placement for each scenario and plot the number of APs needed and the average achievable rate in such rooms. The parameters for these scenarios are detailed in Table 3.2, and the results are presented in Figure 3.7.



(a) Violin plot of the number of APs placed

(b) Average Rate achievable in Mb/s

Figure 3.7 Comparing the various application scenarios optimized with our proposed 3D placement

These results indicate that the number of APs placed is very similar for both the free-height and fixed-height models. However, the free-height model consistently achieves a higher average rate, sometimes with even fewer APs than the fixed-height model. This demonstrates the significance of considering AP height in the optimization process. Therefore, the proposed optimized 3D placement framework effectively plans LiFi networks in various indoor scenarios with different user distribution patterns.

3.1.8 Summary

In this section, we addressed the LiFi AP 3D placement problem with the dual objectives of minimizing the number of APs and maximizing the user occurrence probability-weighted sum rate. Both free-height and fixed-height models were analyzed, with constraints on data rates and illumination. Using an NSGA-II-based algorithm, we demonstrated that allowing free selection of AP heights results in better network performance, and considering user distribution enhances coverage in high user likelihood areas.

Building on the insights gained from optimizing LiFi-only networks, the next section extends these principles to the more complex scenario of LiFi-WiFi heterogeneous networks. While the current

section focused on minimizing the number of LiFi APs and maximizing the sum rate with constraints on data rate and illumination, the upcoming section expands the scope to WiFi APs and introduces cost considerations.

Both sections employ MOO techniques and consider user distribution; however, the next section incorporates distinct user occurrence probabilities for LiFi and WiFi. This allows for a more tailored placement strategy that accounts for the specific requirements and characteristics of each technology.

3.2 PlaciFi: Orchestrating Optimal 3D Access Point Placement for LiFi-WiFi Heterogeneous Networks

In this section, we extend the methodologies and insights from optimizing LiFi-only networks to encompass the integration of WiFi with LiFi, presenting a more complex and holistic optimization problem for designing a heterogeneous network. Designing networks in indoor heterogeneous wireless environments is crucial for achieving optimal performance and user satisfaction. A key factor in designing a LiFi-WiFi heterogeneous network is the strategic placement of APs. However, positioning APs in such a network is challenging due to the differing characteristics of LiFi and WiFi technologies.

Current research on AP placement mainly addresses either WiFi or LiFi networks individually [DBS18; DBS20; VB19; MS23; Yan+20; Gop+22], with little focus on the integration of both technologies. Consequently, there is a need for a systematic and efficient methodology to determine the optimal AP placement in a LiFi-WiFi heterogeneous network, taking into account the specific requirements and interactions of both systems. This section introduces PlaciFi, a comprehensive framework for the optimal 3D placement of APs in LiFi-WiFi heterogeneous networks. By addressing the unique challenges of such networks, PlaciFi aims to optimize AP placement to enhance network performance and cost efficiency.

3.2.1 State-of-the-art Analysis

In this section, we provide an overview of relevant literature related to the placement of APs in LiFi and Radio Frequency (RF) networks, highlighting gaps that motivate our research.

Placement in LiFi Networks

In [DBS18], the authors explore optimal AP placement in LiFi networks by considering the stationary distribution of user mobility. They analyze the stationary distribution of users following a Random Waypoint model in an indoor environment. This initial work demonstrates feasibility using a small Fixed-cell Single-user environment with four APs.

In [DBS20], the initial work is extended to optimize average throughput in an indoor environment by considering the stationary distribution of users, placing APs in a 2D space using an adaptive gradient projection algorithm. Unlike these works, our approach does not assume the actual positions of users. Instead, it considers coverage on a user plane and various user occurrence patterns, which is advantageous when user positions or numbers are unknown during the planning stage. Furthermore, these studies do not address the optimal number of APs to be placed.

The work in [VB19] focuses on minimizing the number of LEDs based on the expected user distribution in a room. However, it does not optimize for network performance in terms of throughput. Instead, it aims to maximize the number of users served or minimize the number of APs placed. Their approach uses an exhaustive search method, impractical for our problem.

The previously mentioned works ensure only a minimum illumination level in the room, which could lead to uneven illumination since the positions or powers of the APs are optimized for maximizing the rate. Our objective is to ensure a minimum uniformity of illumination.

In [MS23], the authors aim to improve the arrangement of a fixed number of LEDs in LiFi systems by minimizing outage probability, addressing 2D placement using a heuristic approach. Similarly, [Yan+20] also addresses the limitations of exhaustive search by solving the 2D placement problem for an LED array through successive convex sub-problems. They propose a power-efficient LED

placement algorithm for indoor VLC networks, aiming to reduce power consumption while ensuring reliable communication. However, their approach is limited to 2D placement with a fixed number of LEDs. Both studies also consider achieving uniform illumination.

In contrast, our research considerably advances the existing literature by incorporating the distinct characteristics and constraints of both LiFi and WiFi technologies. We introduce an optimization framework that positions APs in a 3D space and accounts for various user occurrence patterns.

Placement in WiFi and Heterogeneous RF Networks

In the field of RF placement research, [Gop+22] presents a modified vector quantization method for small-cell WiFi networks. Their optimization strategy focuses on minimizing interference and optimizing the placement of APs within the network. This evaluation is conducted for single-user networks and includes a discussion on the limitations of their approach in small-scale networks.

In a different work, [Pas+17] tackles the issue of colocated and non-colocated node placement in Long Term Evolution (LTE)-WiFi aggregation networks by formulating it as an MINLP problem. This work is particularly relevant to our work as it examines heterogeneous networks.

Similarly, [EAE19] proposes an optimal deployment strategy for heterogeneous wireless nodes in integrated LTE-WiFi networks. They also frame the problem as an MINLP and introduce a genetic algorithm to achieve near-optimal solutions. We also employ a genetic algorithm to demonstrate its effectiveness in solving placement problems.

While these works focus on optimizing AP placement in WiFi or LTE-WiFi networks, our research addresses the unique challenges of integrating LiFi with WiFi in a heterogeneous network. This integration is particularly complex due to the orientation dependency of the LiFi channel, which complicates formulating a mathematically convex optimization problem.

3D Placement

Previous works have primarily focused on the 2D placement of APs. To address 3D placement optimization, we extend our review to include the placement of UAVs.

In their survey, [MM22] examine various methods for optimizing UAV placement and flight path design. One section of their work investigates different placement optimization strategies. Their focus on UAV placement is relevant to our work due to the 3D spatial considerations, though it differs as it involves actual users during optimization.

Another work by [Gha+18] introduces an efficient strategy for 3D aerial base station placement, taking user mobility into account through reinforcement learning. This work marks the beginning of more recent research utilizing machine learning.

The authors in [WZZ18] explore joint trajectory and communication design for multi-UAV enabled wireless networks. They tackle a non-convex problem by solving approximate convex sub-problems, providing an interesting mathematical formulation.

An AP placement approach for UAV-terrestrial small-cell networks is proposed in [GRV21]. Their work aims to optimize AP placement by considering UAVs as small cells while minimizing interference, using a vector quantization approach similar to [Gop+22].

Previous research has examined various aspects of AP placement in homogeneous and heterogeneous networks. While some studies have concentrated on WiFi-based networks and others on LiFi-based networks, there has been limited focus on integrating these technologies and optimizing AP placement in LiFi-WiFi heterogeneous networks. Our work addresses this gap by tackling the unique challenges of LiFi-WiFi heterogeneous networks. We introduce novel optimization techniques and adopt a comprehensive approach considering multiple objectives and the placement of three-dimensional AP.

3.2.2 Key Contributions

We introduce a novel framework for the optimal three-dimensional placement of APs in a LiFi-WiFi heterogeneous network, addressing key factors such as cost, rate coverage, user distribution, and

illumination. This approach provides a comprehensive network plan considering the indoor layout and strategically positions the APs.

1. Holistic Network Optimization: Our framework offers a comprehensive strategy for the optimal placement of APs in a LiFi-WiFi heterogeneous network, incorporating both VLC and IR communication for LiFi.

2. Multi-Objective Optimization: We tackle the challenge of optimal AP placement in three dimensions by formulating a MOO problem, which achieves significantly higher average rates than state-of-the-art 2D power optimization models. Our primary objective is to minimize the cost of placing APs while maximizing network rate coverage. The results demonstrate the effectiveness of MOO techniques in addressing multiple objectives simultaneously. Additionally, we incorporate constraints to ensure uniform and adequate illumination throughout the coverage area. We refine our objective to minimize placement costs by accounting for the different expenses associated with APs for LiFi and WiFi, thereby optimizing the network economically. To meet the diverse needs of users, we include distinct user technology occurrence probabilities for each technology (LiFi and WiFi), allowing us to tailor AP placement to the specific requirements of each user group.

3. Versatile Solution Methods: We explore a range of solution methods, including heuristics, metaheuristics, and off-the-shelf solvers, consistently outperforming the baseline random approach. Additionally, we examine various strategies to combine multiple objectives to obtain a single optimal solution. This flexibility in solution approaches allows network planners to select the most appropriate method based on the specific requirements of their deployment scenario.

3.2.3 System Model

Table 3.3 List of Notations used in the 3D Placement of LiFi and WiFi APs

Notation	Description
l, M^L	Index, size of LiFi APs
w, M^W	Index, size of WiFi APs
u, M^U, $\mathcal{U}$	Index, size, set of all user positions
v, M^V	Index of, total positions on the illuminance grid
$\boldsymbol{c} = (x, y, z)$	Vector of 3D coordinates
$\Delta\boldsymbol{c}$	Grid spacing in 3 dimensions
α	Binary existence variable for AP
C	Cost of placing an AP
$\boldsymbol{\Theta_u} = (\Theta^Y, \Theta^P, \Theta^R)_u$	Vector of yaw, pitch and roll of user device
$\hat{\boldsymbol{n}}_u$	Normal vector of the rotated user device
p_u^L, p_u^W	Probability of occurrence of LiFi, WiFi user
$\boldsymbol{d}_{u,l}$, $\boldsymbol{d}_{u,w}$	3D Euclidean distance between user and AP
ϕ	Angle of irradiance of LiFi signal
θ	Angle of incidence of LiFi signal
P_l, P_w	Transmission power of the APs
P_u	Received power at the user
R_u, $\tilde{R}_u$	Achievable and normalized achievable rate
I_v	Illuminance at the grid position
I	Illumination uniformity
$\tilde{R}_{\text{thresh}}$	Normalized Rate requirement
I_{thresh}	Illumination uniformity requirement

Table 3.3 summarizes the notation used throughout this section. Specifically, we use bold lowercase letters for vectors, cursive capital letters for sets, and $\mathbb{1}$ to represent indicator functions.

This work focuses on deploying a LiFi-WiFi heterogeneous network, integrating wireless technologies for indoor communication. The network comprises up to M^L LiFi APs and up to M^W WiFi APs, strategically positioned throughout the environment.

The LiFi APs operate on the visible light or Infrared spectrum and are mounted above the user plane facing downwards. When operating on visible light, they provide both illumination and data transmission. The three-dimensional coordinates of each LiFi and WiFi AP are denoted by $\boldsymbol{c}_l = (x, y, z)_l$ and $\boldsymbol{c}_w = (x, y, z)_w$, respectively. The maximum height of the APs is indicated by z_{maxdim}, corresponding to the ceiling height in the indoor environment. The minimum height, denoted by z_{mindim}, is assumed to be at least one meter above the user plane to prevent receiver saturation.

The user plane, where users are expected to be located, is represented by a grid with a spacing of $\Delta\boldsymbol{c}_u = 0.25$ m at a height of 1.4 m. Each position $u \in \mathcal{U}$ on this grid is specified by the coordinates $\boldsymbol{c}_u = (x, y, z)_u$, with M^U such positions. The orientation of these receivers is indicated by $\boldsymbol{\Theta_u} = (\Theta^Y, \Theta^P, \Theta^R)_u$, representing the device's Yaw, Pitch, and Roll angles. A $(0, 0, 0)$ value signifies that the user device is parallel to the floor and facing the ceiling. Given the dynamic and unknown exact user locations, each grid position u is assigned a weight p_u^L or p_u^W, corresponding to the expected probability of a LiFi or WiFi user being in that position. Users connect to the AP offering the strongest signal strength within the same technology, providing flexibility in the model as there are no dedicated APs for specific users. If users can access both technologies, they choose the one with the highest link rate.

The LiFi channel model is described in Section 2.3, while the WiFi channel model is detailed in Section 2.6. In this work, we aim to ensure a minimum level of illumination uniformity. Building upon the illumination model discussed in Section 2.5, illumination uniformity is quantified by the ratio of the minimum to the average illumination intensity, as outlined in [IAY20].

$$I = \frac{\min_v I_v}{\frac{\sum_v I_v}{M^V}}. \tag{3.5}$$

The simulation parameters for the LiFi channel are the same as in the previous section, except for the luminous efficacy of the LED, I_0, set at 200 lm/Watt. The illumination uniformity threshold is defined as $I_{\text{thresh}} = 0.7$. The specific simulation parameters for the WiFi channel are detailed in Table 2.2.

3.2.4 3D Placement Problem Formulation

The primary focus of our 3D placement model is to determine the optimal number and positions of LiFi and WiFi APs within a given environment. This optimization aims to minimize the deployment cost of the APs while maximizing the network coverage, as reflected in the sum rate over the user grid. This sum rate is weighted by the user occurrence probabilities for LiFi and WiFi, denoted as p_u^L and p_u^W, respectively. The cost associated with the placement is calculated based on the number of APs deployed and their respective unit costs, C^L for LiFi and C^W for WiFi.

To ensure reliable coverage across the entire user plane, we impose a constraint on the minimum guaranteed rate, defined in terms of a normalized rate threshold $\tilde{R}_{\text{thresh}}$. This constraint varies depending on the technology used by the user. Specifically, if a user position u has a zero probability of a LiFi user ($p_u^L = 0$), the minimum rate must be provided by WiFi. Conversely, if both technologies have a non-zero probability at position u, at least one must meet the minimum rate requirement.

Furthermore, we differentiate between the use of VLC and IR for LiFi. For the VLC model, we introduce an additional constraint to ensure minimum illumination uniformity, denoted by I_{thresh}, across the illumination grid.

The decision variables in this 3D placement problem include the presence and positions of both LiFi and WiFi APs. The presence of a LiFi AP $\alpha_l \in \{0, 1\}$ is a binary variable indicating whether AP

l is deployed. Similarly, $\alpha_w \in \{0,1\}$ indicates the presence of a WiFi AP. The positions of these APs are defined within specific bounds as follows:

For WiFi APs:

$$x_w \in [0, x_{\text{dim}}], \quad y_w \in [0, y_{\text{dim}}], \qquad w = 1, 2, \ldots, M^W$$
$$z_w \in [z_{\text{mindim}}, z_{\text{maxdim}}], \qquad w = 1, 2, \ldots, M^W$$

For LiFi APs:

$$x_l \in [0, x_{\text{dim}}], \quad y_l \in [0, y_{\text{dim}}], \qquad l = 1, 2, \ldots, M^L$$
$$z_l \in [z_{\text{mindim}}, z_{\text{maxdim}}], \qquad l = 1, 2, \ldots, M^L$$

The room size constrains the x and y dimensions, while the z dimension is bounded by a minimum of 2.5 m and a maximum at the ceiling height of 3.5 m.

The optimization problem can be mathematically formulated as follows:

$$\min_{\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l} \quad C^W \sum_{w=1}^{M^W} \alpha_w + C^L \sum_{l=1}^{M^L} \alpha_l \tag{3.6}$$

$$\max_{\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l} \quad \sum_u \left(\tilde{R}_u^W p_u^W + \tilde{R}_u^L p_u^L \right) \tag{3.7}$$

$$\text{s.t.} \quad \max \left(\tilde{R}_u^L \lceil p_u^L \rceil, \tilde{R}_u^W \lceil p_u^W \rceil \right) \geq \tilde{R}_{\text{thresh}} \qquad \forall u \in \{ \mathcal{U} \mid p_u^L > 0 \vee p_u^W > 0 \} \tag{3.8}$$

$$\frac{\min_v I_v}{\frac{\sum_v I_v}{M^V}} \geq I_{\text{thresh}} \tag{3.9}$$

The first objective function (3.6) aims to minimize the deployment cost, while the second objective (3.7) seeks to maximize the sum rate. The constraint in (3.8) ensures a minimum guaranteed rate at each user grid position, and the constraint in (3.9) enforces illumination uniformity in the VLC mode.

This problem is categorized as an MINLP problem due to both integer and continuous variables, and it is non-convex due to the nature of the second objective function. Consequently, it is mathematically intractable and computationally complex. Additionally, this is a MOO problem involving multiple conflicting objectives. In such scenarios, the optimal solution comprises a set of Pareto-optimal solutions, representing the best trade-offs between the objectives [MA03].

In addition to the 3D Placement formulation, we also consider the "2D Pow" formulation. This problem shares the same optimization framework, but instead of optimizing the height of the APs, it optimizes their transmission power. This formulation aligns more closely with existing state-of-the-art methods. The literature on LiFi-WiFi heterogeneous networks is scarce, so we draw comparisons with LiFi-only studies. For instance, [Yan+20] addresses a similar problem in the "2D Pow" context, focusing on power minimization rather than our multi-objective approach. By providing this alternative, we highlight the advantages and limitations of our model in relation to current state-of-the-art methodologies.

3.2.5 Methods to solve the 3D Placement Problem

To address the MINLP and MOO problem outlined in Section 3.2.4, we have developed a comprehensive orchestration framework featuring various solution methods. These methods include a random optimizer as a baseline, an analytical approach suitable for off-the-shelf solvers, and a novel heuristic called Dynamic Grid Explorer (DGE). Additionally, we adapt existing meta-heuristic and black-box optimization algorithms to solve our specific problem.

3.2.5.1 Random Baseline

To benchmark and validate the effectiveness of our optimization framework for AP placement in LiFi-WiFi heterogeneous networks, we implement a random optimizer as a baseline. This involves generating random candidate solutions within the defined 3D search space of potential AP locations. These candidate solutions represent different configurations for the network's AP placements. This approach helps establish a baseline for comparing the effectiveness of our optimization framework can be measured.

3.2.5.2 Optimizer

To tackle the MINLP nature of the 3D placement problem, the Gurobi optimization software [Gur23] is employed. Gurobi efficiently handles Linear Programming (LP) relaxations for MINLP problems.

The nonlinear aspects of the problem are approximated through transformations to make them solvable by Gurobi. The first objective function is a linear sum of binary variables representing the presence of APs. The second objective, involving user rate calculations, is a non-convex, nonlinear function of the optimization variables.

To calculate the rate for LiFi, we first determine the distance using auxiliary variables and the convex Euclidean norm function. Next, we calculate the cosine of the irradiance and incidence angles using auxiliary variables and bilinear transformations for the multiplications and divisions. These terms are combined with additional bilinear constraints to obtain the gain HLoS. Recalling (2.4), the gain is non-zero only within the FoV of the receiver. This results in adding indicator constraints. The indicator variables for the incidence angles at the receiver and transmitter are expressed as

$$\mathbb{1}_{0\leq\theta_{u,l}\leq\Theta_f} \tag{3.10}$$

$$\mathbb{1}_{0\leq\phi_{u,l}\leq\Phi_f} \tag{3.11}$$

where $\mathbb{1}$ is the indicator variable that takes the value of 1 if the condition in the subscript is satisfied and is 0 otherwise. Thus, the gain can be formulated as:

$$\text{HLoS}_{u,l} = \frac{H_0}{d_{u,l}^2} \cos\phi_{u,l}^m \cos\theta_{u,l} \mathbb{1}_{0\leq\theta_{u,l}\leq\Theta_f} \mathbb{1}_{0\leq\phi_{u,l}\leq\Phi_f} \tag{3.12}$$

The SINR for the user is calculated by considering the AP offering the maximum signal strength as the connected AP. The max function is replaced with binary variables, which have the value of 1 for the AP connected to the user. An additional constraint is included such that the sum of these binary variables equals one, ensuring each user connects to only one AP. The signal at the receiver is given by,

$$S_u^L = \sum_l g_{u,l} P_{u,l}. \tag{3.13}$$

where $g_{u,l}$ is the binary association variable, constrained by

$$\sum_l g_{u,l} = 1 \qquad \forall u = 1, 2, \ldots, M^U \tag{3.14}$$

The rate calculation involves converting the division in the SINR formula to subtraction using the $\log_2$ function, which converts the SINR to the rate as in (2.8). In contrast to the equation, we do not limit the rate to the maximum capacity of the technology since that would involve a Piecewise function. This transformation is applied to both LiFi and WiFi rates. The second objective function (3.7) is then computed.

We set a lower bound on an auxiliary rate variable to ensure that users achieve the necessary data rates. This constraint applies when a user can connect to only one technology. However, when a user can utilize both LiFi and WiFi, we require that at least one technology meets the minimum guaranteed

rate. To implement this, we introduce binary variables to choose between the two technologies, ensuring their sum equals one for each user.

To address illumination uniformity constraint within the VLC model, we need to calculate the minimum illuminance in the room, recalling (3.9). We introduce an auxiliary variable, $I_{\min}$, representing the minimum illuminance. We then add a constraint ensuring this variable is less than or equal to the illuminance at every grid position:

$$I_{\min} \leq I_v \qquad \forall v = 1, 2, \ldots, M^V. \tag{3.15}$$

This approach leverages the lower bound on the uniformity constraint to maximize $I_{\min}$. In contrast, the above constraint (3.15) ensures $I_{\min}$ does not exceed any grid position's illuminance value I_v.

By introducing these auxiliary variables and constraints, we simplify the problem, though the non-convexity persists due to the interactions of the optimization variables in calculating the LiFi gain.

Single Objective Optimization: Weighted Sum: The optimization problem can be simplified by combining the two primary objectives — cost minimization and rate maximization — into a single objective function using the weighted sum method. Here, the cost objective is given a weight of 60%, and the rate maximization objective is given a weight of 40%. This results in the following combined objective function:

$$F = 0.6 \cdot f_1(\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l) + 0.4 \cdot f_2(\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l).$$

Multi-objective Optimization: Epsilon-constraint: To handle MOO problems, the epsilon-constraint method is employed. This approach converts the MOO problem into a series of single-objective sub-problems by introducing an epsilon constraint on one of the objectives. The epsilon constraint sets a threshold, ensuring the first objective does not exceed a specified value. By varying the epsilon value, we generate a set of solutions that represent trade-offs between the objectives.

$$\begin{aligned} \text{Maximize:} \quad & f_2(\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l) \\ \text{Subject to:} \quad & f_1(\boldsymbol{c}_w, \boldsymbol{c}_l, \alpha_w, \alpha_l) \leq \epsilon \end{aligned}$$

Solving these sub-problems individually results in Pareto-optimal solutions for the different epsilon values, enabling us to choose the solution that most effectively balances the combined objectives.

Single Objective Optimization: Simplified: While Gurobi is effective for finding precise optimal solutions in smaller, less complex instances, it faces limitations in larger, more complex scenarios typical in real-world LiFi-WiFi network deployments. To simplify the problem, the 3D AP space is quantized into a grid with half the spacing of the user grid in the x and y dimensions ($\Delta x = \Delta x_u/2$ and $\Delta y = \Delta y_u/2$) and one-fourth the spacing in the z dimension ($\Delta z = \Delta z_u/4$).

Signal powers $P_{u,\tilde{l}}, P_{u,\tilde{w}}$ between each AP grid point $(\tilde{l}, \tilde{w})$ and user grid point (u) are pre-calculated. During optimization, binary variables $\alpha_{\tilde{l}}$ and $\alpha_{\tilde{w}}$ are used to select which grid points should be occupied by APs.

$$\alpha_{\tilde{l}} = \begin{cases} 1 & \text{if LiFi AP is placed at grid point } \tilde{l} \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{\tilde{w}} = \begin{cases} 1 & \text{if WiFi AP is placed at grid point } \tilde{w} \\ 0 & \text{otherwise} \end{cases}$$

The modified objective functions become

$$\begin{aligned} \min_{\alpha_{\tilde{w}}, \alpha_{\tilde{l}}} \quad & C^W \sum_{\tilde{w}} \alpha_{\tilde{w}} + C^L \sum_{\tilde{l}} \alpha_{\tilde{l}} \\ \max_{\alpha_{\tilde{w}}, \alpha_{\tilde{l}}} \quad & \sum_u \left(\tilde{R}_u^W p_u^W + \tilde{R}_u^L p_u^L \right). \end{aligned}$$

While this grid-based approach reduces computational complexity, it still requires significant memory due to the number of potential AP positions. However, it performs well in smaller scenarios, such as rooms up to 5 x 5 m with a grid spacing of Δc_u =1 m.

3.2.5.3 Dynamic Grid Explorer

The DGE algorithm is introduced to efficiently explore the 3D placement space. It begins by fixing the number of APs, similar to the epsilon-constraint method. Once the number of APs is set, let N represent the total number of APs available for placement, and let $C = \{c_0, \ldots, c_N\}$ denote the set of all coordinates. These APs are initially randomly placed within the search space.

In the next step, the set of positions is updated according to predefined movement directions. Let Δ be the set of possible movement directions for an AP

$$\Delta = \{(\pm d, 0, 0), (0, \pm d, 0), (0, 0, \pm d), (0, 0, 0)\} \tag{3.16}$$

Here, d represents a multiple of the grid size; for instance, $d = 2$ indicates two steps on the grid. The possible movement directions are illustrated in Figure 3.8.

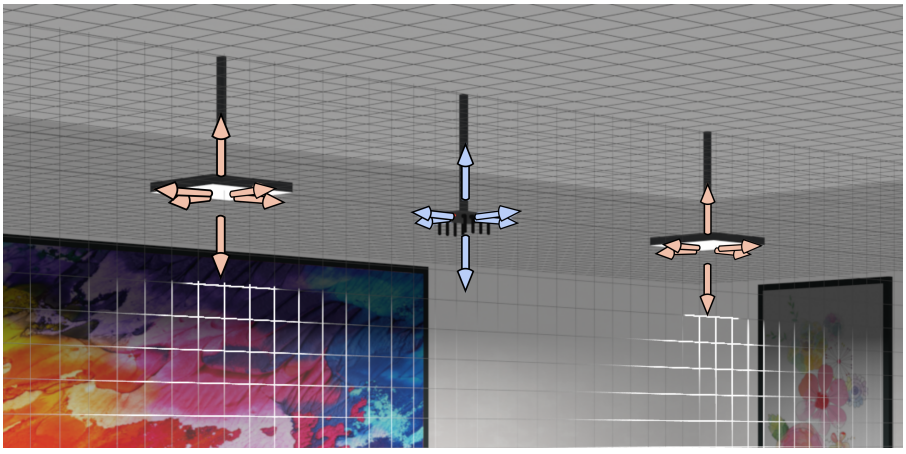


Figure 3.8 Possible movement directions for LiFi and WiFi APs in the DGE algorithm

The coordinates of a single AP are then updated as follows:

$$c = c + \delta \mid \delta \in \Delta. \tag{3.17}$$

Each AP is moved in all possible directions, and the value of the objective function is calculated for each of these movements. Let $f_{\text{dge}} : f(C) \rightarrow \mathbb{R}$ denote a function that maps a set of coordinates to a real value representing our objective

$$f_{\text{dge}} = \sum_u \left(\tilde{R}_u^W p_u^W + \tilde{R}_u^L p_u^L \right). \tag{3.18}$$

The objective function is penalized for any constraint violations to ensure compliance with both rate and illumination constraints. Specifically, if any user position violates the rate constraint, the objective is penalized by the worst violation:

$$f_{\text{dge}}^{\text{corr}} = \begin{cases} f_{\text{dge}}(1 - \min_u(\tilde{R}_{\text{thresh}} - \tilde{R}_u)) & \text{if } \exists u : \tilde{R}_u < \tilde{R}_{\text{thresh}} \\ f_{\text{dge}} & \text{otherwise} \end{cases} \tag{3.19}$$

A similar penalty is applied for violations of the illumination uniformity constraint

$$f_{\text{dge}}^{\text{corr}} = \begin{cases} f_{\text{dge}} \left(1 - \frac{I_{\text{thresh}} - I}{2}\right) & \text{if } I < I_{\text{thresh}} \\ f_{\text{dge}} & \text{otherwise} \end{cases} \tag{3.20}$$

Here, I represents the achieved illumination uniformity.

Based on the objective value, the best combination of coordinates is selected, and the APs are moved to this new position. This process is repeated until the objective value no longer increases, indicating that either a global optimum has been found or the algorithm is stuck in a local optimum. To avoid local optima, the step size is increased incrementally as $d \leftarrow d + 1$, and the search continues until no further movements are possible and the entire grid has been explored. Once a solution for a fixed number of APs is found, the process is repeated for the next higher number of APs. Finally, the optimal solution is selected using a weighted sum of the two objectives, considering only solutions with the least constraint violations.

3.2.5.4 Meta-heuristics

The NSGA-II [Deb+02] is a widely used MOO technique that employs a genetic algorithm framework to efficiently discover Pareto-optimal solutions. In MOO problems, scalarization functions [Chu20; Wie82] can be used to break down the objectives and convert them into a single-objective problem that can be addressed using standard optimization algorithms. The scalarization functions utilized in this approach include:

- Weighted Sum: This method linearly combines multiple objectives with predefined weights, transforming the multi-objective problem into a single-objective problem. For n objectives $f_i(x)$ with weights w_i, the weighted sum objective $F(x)$ is represented as:

$$F(x) = \sum_{i=1}^{n} w_i \cdot f_i(x)$$

- Tchebicheff: This function evaluates the objectives by considering the maximum weighted deviation from an ideal point, focusing on the worst-case scenario. For n objectives $f_i(x)$ with weights w_i and an ideal point ideal(x), Tchebicheff is represented as:

$$T(x) = \max_i \{w_i \cdot |f_i(x) - \text{ideal}_i(x)|\}$$

- Achievement Scalarization Function (ASF): ASF aims to minimize the weighted sum of deviations from predefined aspiration levels for each objective. For n objectives $f_i(x)$ with weights w_i and reference points $\text{ref}_i(x)$, ASF is represented as:

$$\text{ASF}(x) = \sum_{i=1}^{n} w_i \cdot |f_i(x) - \text{ref}_i(x)|$$

- Penalty-Based Boundary Intersection (PBI): PBI uses a penalty function to combine the objectives, encouraging convergence to the Pareto front while penalizing deviations. For n objectives $f_i(x)$ with weights w_i, and a penalty parameter ρ, PBI is represented as:

$$\text{PBI}(x) = \sum_{i=1}^{n} w_i \cdot f_i(x) + \rho \cdot \sqrt{\sum_{i=1}^{n} (f_i(x))^2}$$

Applying these scalarization functions allows the GA to handle the multi-objective AP placement problem effectively. These decomposition techniques also serve as MCDM tools to select a single optimum from a Pareto set. Additionally, Pseudo-Weights [BD20] can be used as an alternative scalarization method. All meta-heuristics in this work are implemented using the Pymoo framework [BD20], which has been adapted for our specific purposes.

3.2.5.5 Black-box

We investigate the use of black-box optimization techniques, such as Bayesian optimization [Moc89], Random Forest [Bre01], and Extra Trees[GEW06], to solve the AP placement problem. These methods do not require explicit mathematical formulations of the objective functions or constraints. Instead, they treat the optimization problem as a black box, relying on objective function evaluations without needing gradient information or access to the underlying problem structure. This approach is particularly useful for handling complex, nonlinear optimization problems where traditional methods may be less effective.

3.2.6 Evaluation Methods

Our proposed PlaciFi framework for optimal placement of APs in LiFi-WiFi heterogeneous networks is evaluated through extensive simulations using a custom simulation environment implemented in Python.

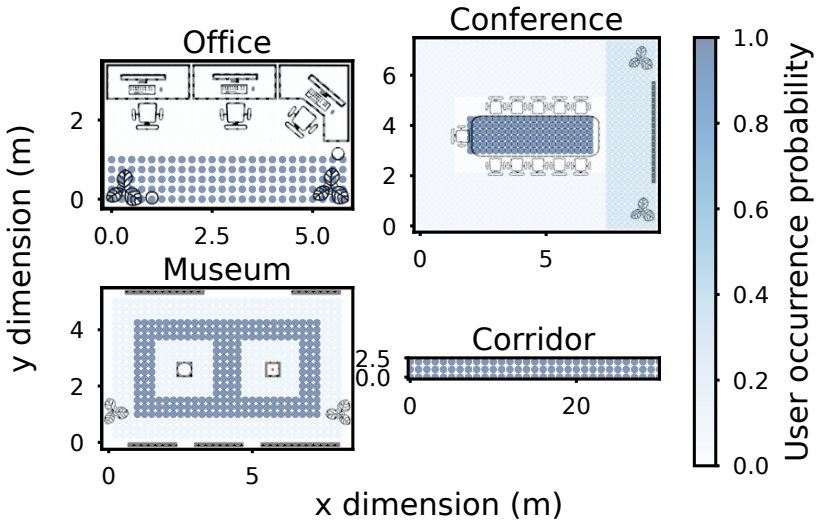


Figure 3.9 Floor plans of indoor scenarios with WiFi user occurrence probability. The pattern for LiFi probability is the same as WiFi in Conference and Corridor, while it is the opposite for Office and Museum.

To assess the performance of our proposed 3D placement algorithm, we simulate various typical indoor environments, visualized in Figure 3.9. This figure also shows the user occurrence probability for WiFi users. It is important to note that the LiFi user probability is also considered and may differ from the WiFi user probability. For instance, the patterns for LiFi and WiFi are identical in Conference and Corridor scenarios but opposite in Office and Museum scenarios.

- Office: This scenario represents a typical office setting where users need reliable, high-speed wireless connectivity. The environment includes multiple workstations where users primarily use LiFi for communication, while WiFi is valuable in areas with more movement.

- Museum: Museums frequently use wireless technologies to enhance visitor experiences. In this scenario, LiFi APs are strategically placed to provide information and interactive content, with WiFi APs supplementing coverage for mobile users.

- Conference: Conferences and large events require robust wireless connectivity to support numerous participants. This scenario involves deploying a dense network of LiFi and WiFi APs to ensure high-capacity coverage and seamless connectivity for attendees.

- Corridor: Corridors serve as critical pathways in buildings. This scenario considers deploying LiFi and WiFi APs along corridors to provide continuous wireless connectivity for users moving through these areas

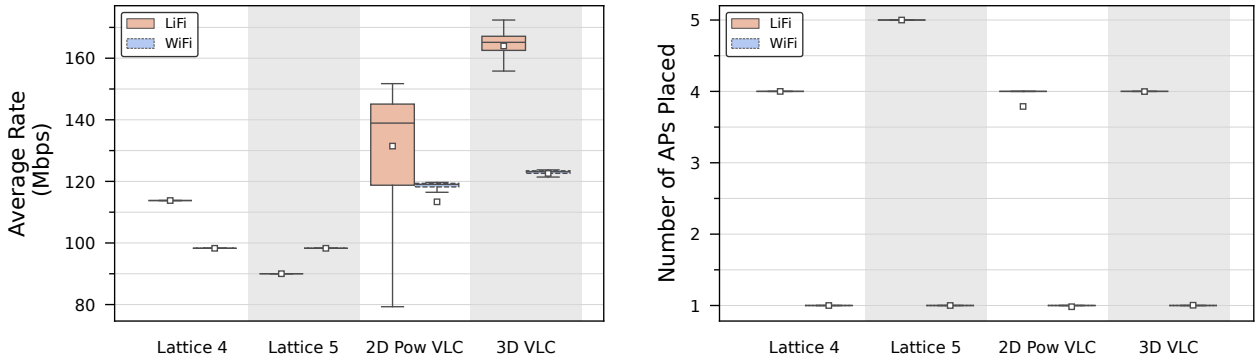
While we do not focus on a minimum illumination level, we ensure uniform illumination across all scenarios. Apart from these application-specific scenarios, we evaluate a standard 5 x 5 m room with a uniform user occurrence probability throughout. The additional simulation parameters used to generate results are listed in Table 3.4.

Table 3.4 Simulation Parameters used to evaluate our implementation of PlaciFi

Parameter	Notation	Value
Normalized Rate threshold	$\tilde{R}_{\text{thresh}}$	1%
Illumination Uniformity threshold	I_{thresh}	0.7
Cost of placing WiFi AP	C^W	10 cost units
Cost of placing LiFi AP	C^L	5 cost units

All collected results are based on 1000 simulation runs to ensure robustness. To validate our findings, we employ hypothesis testing.

3.2.7 Results and Comparative Analysis



(a) Average rate achievable in Mb/s

(b) Number of APs placed

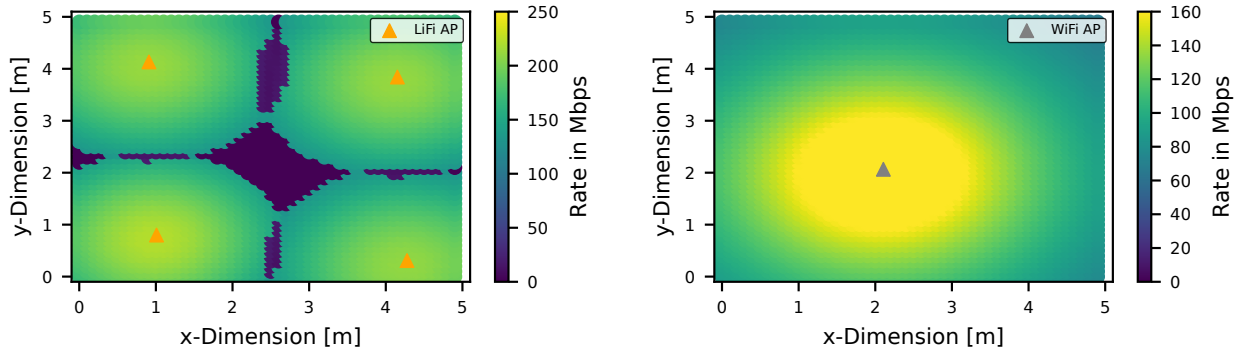
Figure 3.10 Deterministic and optimized placements for the VLC model showing the need for optimization

In the following analysis, we primarily focus on results obtained using meta-heuristic methods unless otherwise specified. We begin by examining the regular room scenario to evaluate the algorithm's performance independently of any real-world scenario-specific effects.

Figure 3.10a presents a comparison of the average data rates achieved in a regular room using LiFi and WiFi technologies with three different AP placement techniques: deterministic placement, 2D power optimization (2D Pow), and our proposed 3D optimized placement. The deterministic placement involves using a fixed lattice grid configuration with either 4 or 5 LiFi APs and one WiFi AP placed at the center of the room, all positioned at a height of 3.5 m on the ceiling. In this scenario, APs are deployed based on predefined positions without any optimization strategy, resulting in sub-optimal data rate coverage and areas with poor signal quality and lower average data rates.

Compared to the deterministic placement, the 2D Pow optimized placement approach enhances average overall data rate coverage. Deploying 3 or 4 LiFi APs results in a wider rate spread than other methods. In contrast, our proposed 3D optimized placement technique consistently utilizes four LiFi APs, as seen in Figure 3.10b. This consistency is due to the illumination uniformity constraint,

which would be violated if fewer APs were used. The 3D technique also manages interference more effectively than other strategies. By fully utilizing the three-dimensional space for AP placement, this strategy outperforms the deterministic and 2D Pow methods in terms of overall performance.

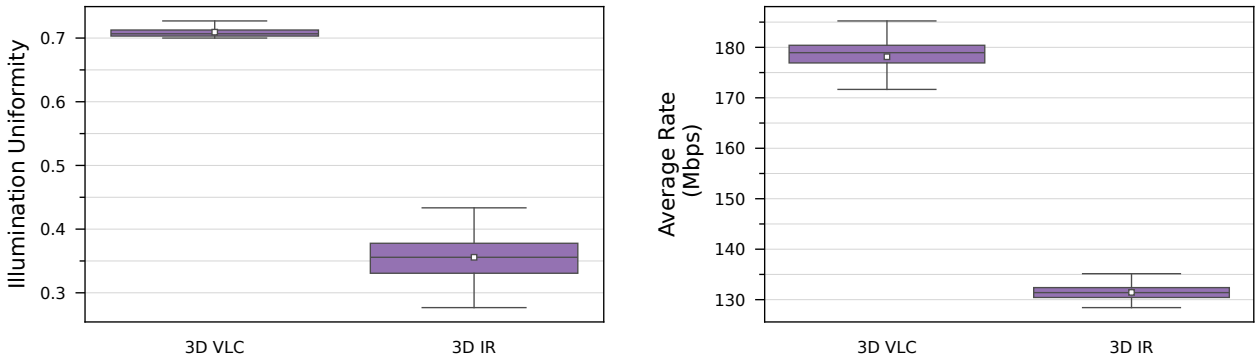


(a) LiFi Rate

(b) WiFi Rate

Figure 3.11 Rate Coverage achieved by the meta-heuristic NSGA-II for the VLC model with our proposed optimized placement

The superior performance of our proposed 3D optimization technique is illustrated in Figure 3.11, which shows a representative result from the simulation. The APs utilize height differences to achieve varied coverage areas, effectively minimizing regions of interference and leading to significantly improved network performance.



(a) Illumination uniformity

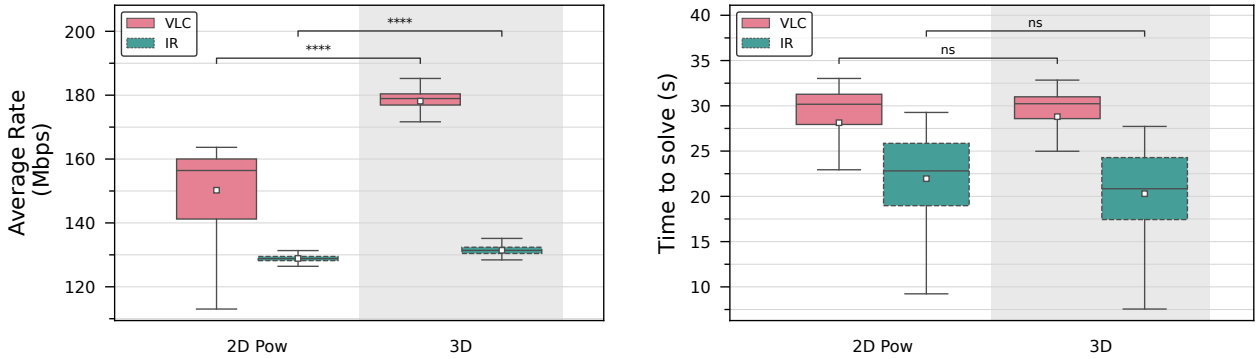
(b) Average rate achievable in Mb/s

Figure 3.12 Performance achieved using 3D VLC and 3D IR models showing the need to constrain the uniformity when using visible light.

Figure 3.12a compares illumination uniformity achieved using two techniques: 3D IR and 3D VLC. The key difference between these methods is the inclusion of the illumination uniformity constraint in the 3D VLC approach. The results indicate that the 3D VLC technique provides superior illumination uniformity, effectively distributing light evenly across the coverage area. This demonstrates the necessity of including the illumination uniformity constraint, as optimizing solely for data rate does not ensure uniform lighting.

In contrast, the 3D IR model, which lacks this constraint, is still appropriate for scenarios where LiFi relies on IR communication and the APs do not contribute to illumination. Thus, the choice between the 3D VLC and 3D IR techniques should be based on the specific requirements of the LiFi system.

Furthermore, as shown in Figure 3.12b, the 3D IR model achieves a lower average data rate because it reduces the number of APs placed while still optimizing the overall objective function. Despite this reduction, the data rate remains significantly above the minimum guaranteed rate.

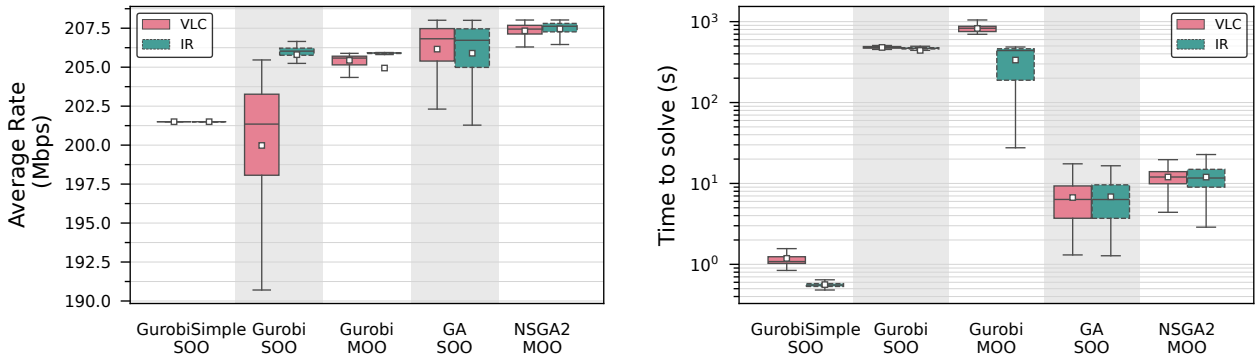


(a) Average rate achievable in Mb/s

(b) Time to solve in seconds

Figure 3.13 Comparing the proposed 3D model and extended State of Art 2D Pow model for both VLC and IR models of LiFi. The figure is annotated with the significance levels of p-values achieved using hypothesis testing, indicating the better performance of the proposed 3D optimization.

Understanding the distinctions between the two LiFi models, we compare their performance when utilizing our proposed 3D optimization against the extended state-of-the-art 2D Pow Model. As illustrated in Figure 3.13a, our 3D optimization method consistently achieves a significantly higher average data rate than the 2D Pow model for both LiFi models. The results of the significance tests are indicated by star notation on the figure. This improvement in data rate does not come at the expense of solving time. As shown in Figure 3.13b, the difference in solving time between the models is negligible, indicating that our 3D optimization method is efficient and effective.



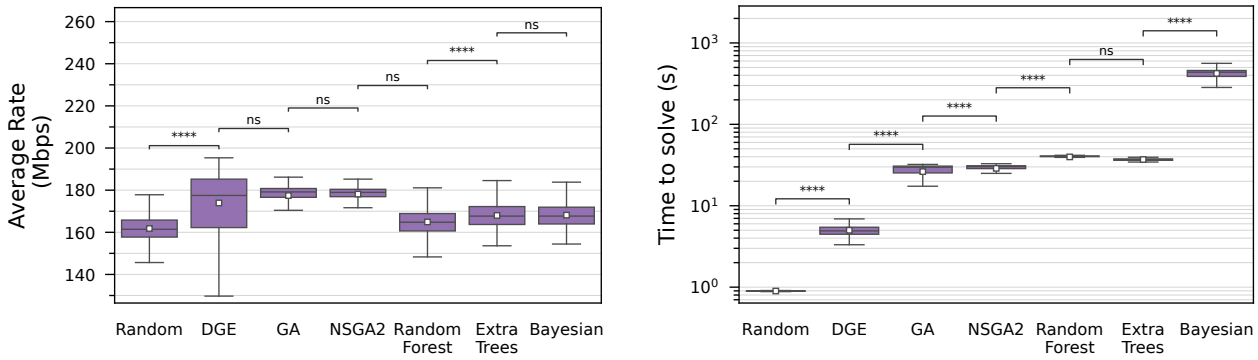
(a) Average rate achievable in Mb/s

(b) Time to solve in seconds

Figure 3.14 Comparing the solutions obtained with the Gurobi optimizer and meta-heuristic approaches for smaller rooms of size 3 x 2 m. The figure displays close rate values indicating the competitive performance of the meta-heuristic solver. The complexity of the full Gurobi models is also evident.

With the efficacy of our proposed 3D optimization approach established, we now evaluate the performance of various solvers within our optimization framework. We first compare the optimality of the proposed meta-heuristic techniques against solutions produced by different implementations using the Gurobi solver. These simulations are conducted in a smaller version of the regular room, measuring 3 x 2 m. The Figure 3.14a shows that the average data rates achieved by the meta-heuristic techniques and the Gurobi solver differ by a maximum of 7 Mb/s. This narrow range of differences

is encouraging, indicating that both approaches can deliver competitive results for smaller scenarios. Upon closer inspection, the meta-heuristic results are slightly higher than those obtained with the Gurobi solver. This can be attributed to minor variations in combining the two objectives during optimization. Despite these slight differences, the overall performance of our proposed meta-heuristic techniques remains exceptionally close to that of the Gurobi solver, demonstrating the optimality of our results for both the VLC and IR models. Figure 3.14b highlights the complexity of the MINLP solution methods. Although the GurobiSimple method shows low solve time, it does not scale well to larger scenarios.



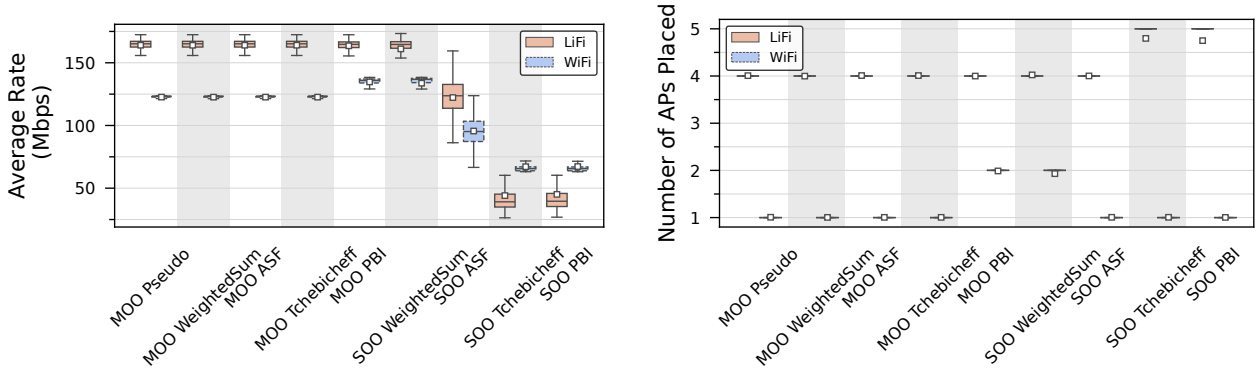
(a) Average rate achievable in Mb/s

(b) Time to solve in seconds

Figure 3.15 Comparing all proposed solvers of PlaciFi alongside a random solver as baseline focusing on the VLC model of LiFi. The figure is annotated with the significance levels of p-values achieved using hypothesis testing, indicating the higher achieved values of each solver than the previous.

Encouraged by these results, we compared the average data rates obtained using our proposed heuristic, meta-heuristic, and black-box solvers against a random solver as a baseline. Our focus is on VLC, given the increased challenge introduced by the illumination constraint. The rate results, depicted in Figure 3.15a, show the performance of each solver. The plot demonstrates that all the proposed solvers outperform the baseline random solver on average. Among the solvers, the DGE, NSGA-II, and GA stand out, achieving the highest average rates. While the black-box optimization performs slightly worse than the meta-heuristics, it still yields competitive results. Significance test results are annotated on the plot to validate these findings.

Additionally, we conducted a time-to-solve comparison for all the solvers, as shown in Figure 3.15b. The significance test results effectively represent the comparison. The DGE exhibits very competitive performance. However, it is essential to note that DGE has relatively high memory requirements, which grow with factorial complexity proportional to the number of APs, making it less scalable for larger scenarios. In summary, the DGE is a highly effective solution for small scenarios, delivering outstanding results. For larger scenarios, meta-heuristics like NSGA-II and GA are the preferred choices, providing optimal performance with competitive solving times. The combined evaluation highlights the strengths and limitations of each solver, enabling network planners to select the most suitable approach based on their specific deployment requirements and constraints.

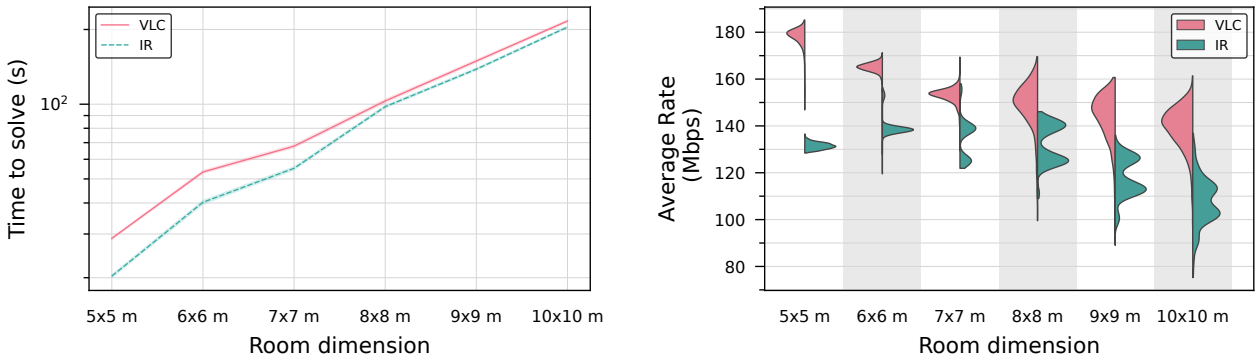


(a) Average rate achievable in Mb/s

(b) Number of APs placed

Figure 3.16 Comparing single and multi-objective solution methods with the meta-heuristic solver focusing on the VLC model of LiFi. The figure depicts the superior performance of the multi-objective methods.

Examining meta-heuristics further, we compare the average rates of LiFi and WiFi achieved by different MOO and Single-Objective Optimization (SOO) techniques for the VLC model. These solution methods employ various decomposition and MCDM approaches to handle the multiple objectives. The results, as illustrated in Figure 3.16, demonstrate that, overall, the MOO methods outperform the SOO methods. The SOO Weighted Sum method, while performing close to the MOO techniques, highlights the difficulty of combining multiple objectives into a single one, as it tends to place an extra WiFi AP. In summary, MOO methods demonstrate superior capability in handling multiple objectives. However, the SOO Weighted Sum method also provides a compelling alternative with competitive performance.



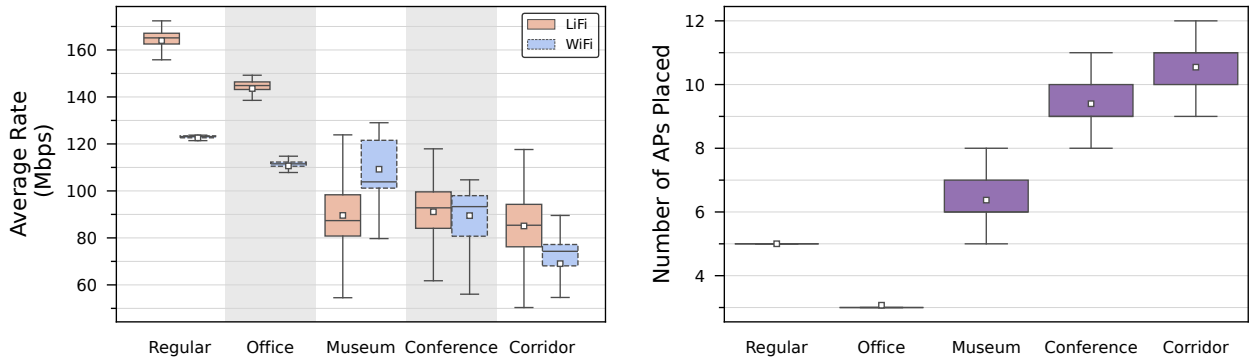
(a) Time to solve with 95% confidence interval band

(b) Average rate achievable in Mb/s

Figure 3.17 Scalability of the meta-heuristic solvers for increasing room size

Finally, we assess the scalability of our meta-heuristic solution by analyzing the time required to solve for increasing room sizes in a standard scenario. Both the VLC and IR models are considered to identify any performance differences. The results of this analysis are displayed in Figure 3.17. As anticipated, the average time to solve shows a linear trend, indicating a steady rise in computation time with increasing room size. However, breakpoints in the slope are observed, attributed to changes in various variables, such as the number of user positions and the maximum number of APs that can be placed. The VLC and IR models exhibit similar trends, suggesting that the choice of communication technology does not significantly impact the scalability of the meta-heuristic solution. To measure uncertainty, the figure also includes a 95% confidence interval band. Although the band is not clearly visible due to its narrow width, it widens as room size increases, reflecting greater variability in the time-to-solve for larger scenarios.

When examining the average rate achieved, as shown in Figure 3.17b, we observe a decrease with increasing room sizes. This trend is due to the objective of placing as few APs as necessary while still ensuring adequate rate coverage.

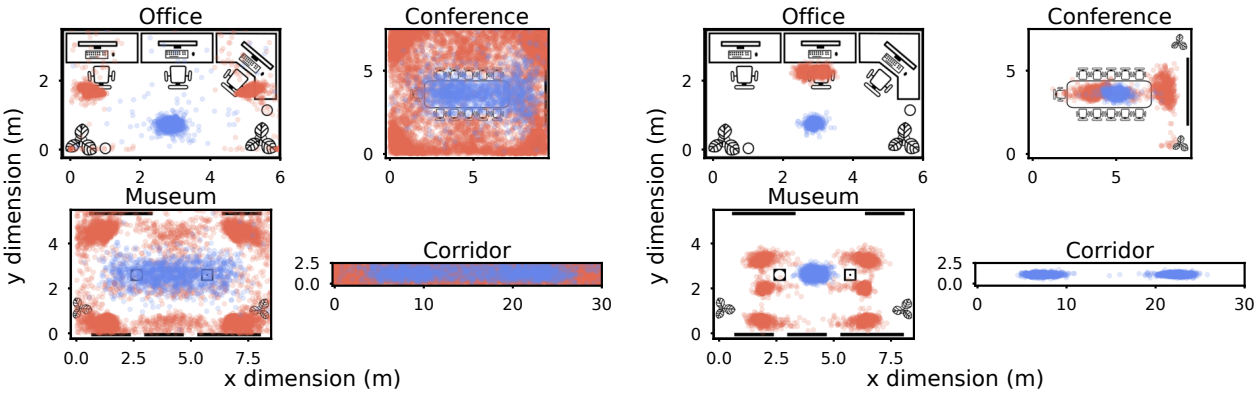


(a) Average rate achievable in Mb/s

(b) Number of APs placed

Figure 3.18 Performance that can be achieved with the NSGA2 solver for the VLC model of LiFi for various application scenarios

Having demonstrated the effectiveness of our optimization framework in a regular room, we now evaluate its practical applicability in real-world use cases by conducting simulations across various indoor scenarios. Each scenario incorporates different user occurrence distributions for both LiFi and WiFi as detailed in Figure 3.9. Figure 3.18 highlights the challenges posed by the corridor scenario, where achieving uniform coverage and higher data rates is more complicated. Consequently, this scenario requires a higher number of APs compared to other use cases, resulting in a wider range of data rates. In contrast, the office layout allows for more strategic placement of a smaller number of APs, ensuring adequate coverage for users with fewer resources.



(a) All scenarios with VLC model

(b) All scenarios with IR model

Figure 3.19 Distribution of LiFi (red) and WiFi (blue) AP positions obtained with the NSGA-II solver over 1000 runs for various application scenarios

Figure 3.19 offers a detailed view of the distribution of LiFi and WiFi APs over the 1000 simulation runs. The results reflect the users' probability patterns. In the Office scenario, the APs are predominantly placed where the occurrence of users for that technology is the highest. The LiFi APs are positioned closer to the room's outer edges to ensure uniform illumination for the VLC model. The Conference scenario shows a broader distribution of APs due to the identical probability patterns for both technologies. The Museum scenario highlights the differences between the VLC and IR models. While both models place the APs near high-probability user areas, the placement patterns differ to

achieve illumination uniformity. The LiFi APs are positioned close to artifacts, whereas WiFi APs are placed nearer to areas with higher user mobility. Overall, our framework's performance across these diverse indoor scenarios demonstrates its versatility and effectiveness in tailoring AP placement to accommodate various user occurrence distributions for both LiFi and WiFi.

These results highlight the strengths and limitations of the proposed optimization framework, providing valuable insights for network planners and designers when considering deployment scenarios of different sizes.

3.3 Summary and Conclusions

This chapter has addressed the problem of optimizing 3D placement of APs in both LiFi-only and LiFi-WiFi heterogeneous networks. We tackled the LiFi AP 3D placement problem with dual objectives: minimizing the number of APs and maximizing the user occurrence probability-weighted sum rate. Both free-height and fixed-height models were analyzed. The placement problem was formulated as a multi-objective optimization problem, incorporating constraints on minimum data rates and illumination levels. Key findings from this section include:

1. Height Variability in LiFi Networks: The research on LiFi AP placement highlights the pivotal role of AP height in optimizing network performance. Allowing flexible height adjustments significantly enhances data rates and coverage compared to fixed-height placements. This flexibility helps in minimizing interference and maximizing the effective coverage area.

2. User Distribution Considerations: Incorporating expected user distribution into the optimization process ensures that network performance is maximized in areas with high user likelihood.

Building on the principles established in the LiFi-only network optimization, we introduced PlaciFi, a framework for the optimal 3D placement of APs in LiFi-WiFi heterogeneous networks. This framework addresses the critical challenge of maximizing coverage and capacity while minimizing deployment costs through strategic AP placement. Key findings from this section include:

1. Holistic Network Optimization: PlaciFi incorporates both VLC and IR communication for LiFi, and considers the distinct costs associated with LiFi and WiFi APs. This comprehensive approach ensures an economically viable network design.

2. User Distribution and Technology Integration: PlaciFi's approach to considering user technology occurrence probabilities is a significant advancement in heterogeneous network planning. By tailoring AP placement based on the likelihood of users utilizing LiFi or WiFi, the network can more effectively meet diverse user needs and improve overall performance.

3. Multi-Objective Optimization: Both studies employ MOO to balance competing goals such as minimizing AP count and maximizing data rates. This approach is crucial for achieving an optimal trade-off between infrastructure costs and network performance.

4. Solution Methods: The exploration of various solution methods, including genetic algorithms and heuristics, demonstrates the suitability of each optimization technique in network planning. These methods provide network planners with tools to address complex AP placement challenges effectively.

The work of this chapter not only advances the state-of-the-art in AP placement strategies but also provides practical tools and methodologies for network planners to design efficient, reliable, and cost-effective LiFi-WiFi indoor networks. With the network design phase completed, the next logical step is to deploy these optimized networks and focus on their effective management post-deployment to ensure that the designed networks not only meet initial performance expectations but also adapt to changing conditions and continue to serve evolving user needs. The upcoming chapters will

transition from the design aspects to addressing the challenges and strategies involved in managing deployed networks.

We will explore resource management techniques, focusing on both wireless and computational resources. This includes dynamic resource allocation and load balancing to optimize the utilization of wireless resources with mobility and environment awareness. Additionally, we will analyze methods for managing computational resources, such as in edge computing to support the growing demand for data processing needs in advanced wireless networks.

Chapter 4

Wireless Resource Allocation in LiFi-WiFi Networks

Wireless resource management is a critical aspect of modern communication networks, particularly in the context of heterogeneous networks combining Light-Fidelity (LiFi) and Wireless-Fidelity (WiFi) technologies. As these networks evolve, they are expected to support a diverse range of applications, from bandwidth-intensive streaming services to latency-sensitive real-time communications. Effective resource management strategies are essential to meet the stringent Quality of Service (QoS) requirements of these applications while ensuring optimal utilization of available network resources [WO22].

In LiFi-WiFi heterogeneous networks, the unique characteristics of each technology can be leveraged to enhance overall network performance. LiFi, with its high data rates and low latency capabilities, is well-suited for scenarios requiring rapid data transmission and minimal delay [Sol+23]. WiFi, on the other hand, offers broader coverage and robust connectivity, making it indispensable for maintaining seamless communication in diverse environments [Wu+21b]. Integrating these technologies presents both opportunities and challenges in wireless resource management.

This chapter explores the complexities of wireless resource management in LiFi-WiFi networks, presenting novel contributions aimed at optimizing network performance. The first contribution focuses on delay minimization optimization. Given the critical importance of minimizing packet delay for applications such as live video streaming and Voice over Internet Protocol (VoIP) [Est+24], this section explores strategies to allocate resources efficiently, ensuring that delay-sensitive traffic is prioritized without compromising overall network throughput. This work, based on our conference publication [VK21], addresses these gaps by proposing methods that account for network delay and data rate requirements, considering both singlepath and multipath scenarios.

The second contribution addresses the stability of LiFi-WiFi networks. Transient channel conditions, such as light path blockages and changes in receiver orientation, pose significant challenges to maintaining stable connections [ASS23]. This section introduces methods to enhance network stability, ensuring consistent performance even in the face of dynamic and unpredictable environmental factors. The enhancements in network stability introduced by our handover management strategies have been documented in a published conference paper [Vij+21].

The final contribution examines the weighted alpha-fair utility maximization. Fairness in resource allocation is crucial to prevent scenarios where certain users monopolize bandwidth, leading to degraded service for others [Lan+10]. By implementing a weighted alpha-fairness approach, this section aims to balance the trade-offs between efficiency and equity, ensuring that all users receive a fair share of the network resources while maximizing overall utility. Our network optimization through alpha-fair utility maximization is partly detailed in a journal manuscript [VK24].

Through these contributions, this chapter provides a comprehensive framework for wireless resource management in LiFi-WiFi networks. By addressing delay minimization, network stability, and fairness in resource allocation, the proposed strategies aim to enhance the performance and reliability of heterogeneous networks, paving the way for the future of wireless communications.

4.1 Delay-aware Wireless Resource Allocation and User Association in LiFi-WiFi Networks

To optimally leverage the heterogeneity of access technologies in a LiFi-WiFi network, it is essential to develop intelligent methods for user association and resource allocation. While a significant body of work [WSH17; WWH17; WWH16] has addressed resource management in LiFi-Radio Frequency (RF) heterogeneous networks, these works primarily focus on optimizing resource allocation for maximum throughput or data rate while ensuring proportional fairness among users. However, most of these works overlook the critical aspect of network delay requirements.

Future wireless networks are expected to support a wide range of communication applications with varying QoS requirements. Applications such as live video streaming and VoIP are highly delay-sensitive and require a minimum guaranteed bandwidth. Therefore, it is crucial to optimize resource allocation with network packet delay as a key performance metric for QoS flows.

Additionally, most existing works assume that users in a heterogeneous network can be served by only one technology at a time, reflecting the limitations of conventional user equipment that lacks support for Multi-Homing. However, with the growing adoption of multipath transport protocols like Multipath Transmission Control Protocol (MPTCP), it is becoming increasingly important to consider Multi-Homing user devices, which can be served by multiple wireless access technologies simultaneously. This capability enables the aggregation of wireless resources and better utilization of the diverse technologies available.

In this work, which is based on a conference publication [VK21], we address these gaps by proposing methods that account for network delay and data rate requirements, considering both singlepath and multipath scenarios. Our approach not only ensures efficient resource allocation and user association but also enhances the overall network performance by accommodating the unique requirements of delay-sensitive applications and leveraging the potential of Multi-Homing devices.

4.1.1 State-of-the-art Analysis

In their research, [WWH17], the authors utilize the data rate QoS metric as a critical element in the optimization objective for resource allocation within a LiFi-RF heterogeneous network, employing an Evolutionary Game Theory (EGT)-based algorithm for resolution. Despite this advancement, their work overlooks delay-related metrics, which are pivotal for applications requiring timely data transmission.

The concept of effective capacity, as introduced by [WN03], offers a sophisticated method to model the channel with respect to various QoS metrics, specifically accommodating the analysis of delay constraints alongside throughput. Building upon this, [Jin+17] effectively employ effective capacity to maximize resource utilization within a homogeneous RF network. This approach was further explored by [Jin+16] within the context of a homogeneous LiFi network and subsequently extended to a LiFi-RF network in [JZH15]. These studies collectively advance the understanding of effective capacity as a tool for enhancing network performance under QoS constraints. Despite these developments, there remains a noticeable gap in the literature concerning user association in heterogeneous networks with explicit consideration of delay constraints. Our work seeks to bridge this gap by not only aiming to reduce overall network latency but also ensuring that each user meets specific delay requirements.

In parallel research, [Luo17] focus on minimizing network packet delays in an RF-only heterogeneous network. Their strategy involves a distributed algorithm that addresses network efficiency without explicitly imposing constraints on delay or bandwidth. However, their model does not accommodate devices with more than one network connection (Multi-Homing), which can significantly influence network dynamics and user experience by allowing devices to maintain multiple network connections simultaneously.

Our work extends these concepts by incorporating a comprehensive analysis of network packet delays, specifically focusing on heterogeneous network environments where both LiFi and RF technologies coexist. We develop an algorithmic approach that not only optimizes for low latency and meets stringent delay requirements but also effectively integrates the capabilities of Multi-Homing devices, thus offering a solution for contemporary network challenges.

At the time of publication of this work, the previously mentioned works were the state-of-the-art. Currently another work [WO22] must be mentioned due to its similarity in goal. The authors aim to optimize the resource allocation to minimize packet losses and delay as QoS metrics for single- and multi-homing devices. However, the difference to our work is that they do not provide any latency or data rate guarantees as we do.

4.1.2 Key Contributions

In this work, we address the optimization of network resource allocation in LiFi-WiFi heterogeneous networks with a focus on delay-critical applications. We develop a comprehensive strategy that not only minimizes network delays but also meets diverse user requirements for data rate and latency. Our contributions are summarized as follows:

1. Formulation of a Mixed Integer Nonlinear Programming (MINLP) Problem: We introduce a MINLP problem aimed at minimizing average network packet delay while ensuring delay and data rate QoS requirements for each user.

2. Extension to multipath networks: Our model is expanded to accommodate Multi-Homing user devices, allowing for simultaneous resource allocation across both LiFi and WiFi technologies.

3. Solution Methods: We employ a Branch and Bound-based solver for precise solutions and propose a genetic algorithm for scenarios requiring faster, near-optimal solutions.

4. Simulation-Based Evaluation: All proposed solution methods are rigorously evaluated through simulations and results indicate that our delay-optimized algorithm consistently achieves an average packet delay of less than 1 ms even in high traffic scenarios.

4.1.3 System Model

This work explores delay-aware resource management in a singlepath and multipath LiFi-WiFi heterogeneous network with a total of M^A Access Points (APs), consisting of both LiFi and WiFi. The LiFi APs, which are Light Emitting Diodes (LEDs) mounted at the ceiling, operate at the same frequency, leading to co-channel interference in overlapping areas. Conversely, the WiFi APs operate on distinct, non-overlapping frequencies. This AP network serves a total of M^U users, each equipped with both LiFi photodiode and WiFi receivers for downlink communication. These users are served by one access technology in the singlepath configuration of the network and simultaneously served by both technologies in the multipath configuration of the network.

Each user generates its own QoS traffic that must be served in the downlink. The inter-arrival times of these data packets follow an exponential distribution with a mean of $1/\lambda_u$ seconds, independently for each user. Similarly, packet lengths follow an exponential distribution with a mean of L_u bits. In addition to these QoS flows, we assume the presence of Best Effort (BE) traffic or delay-tolerant traffic, which primarily has bandwidth constraints. A representation of the network architecture is depicted in Figure 4.1.

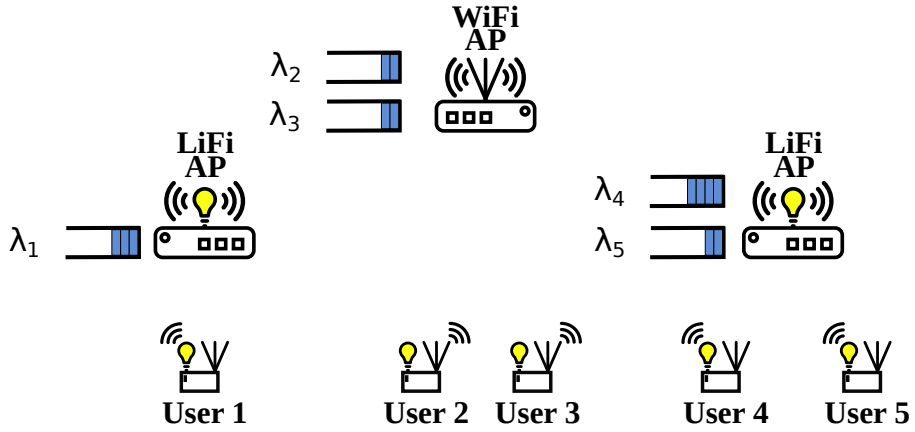


Figure 4.1 Architecture of a LiFi-WiFi heterogeneous network with QoS traffic

APs are assumed to be always active to serve continuous BE traffic. To manage this traffic, each AP a allocates a resource proportion $y_{u,a}$ to each user u. The resulting data rate is the product of this resource proportion and the link data rate $R_{u,a}$. Consequently, the service time for QoS packets to each user is exponentially distributed with a mean of $1/(y_{u,a}\frac{R_{u,a}}{L_u})$. Thus, the traffic to each user can be modeled as an M/M/1 queue.

For resource allocation, all LiFi and WiFi APs are connected to a central controller. This controller has comprehensive wireless channel state information for all users in the network, enabling centralized resource allocation. This allocation must be updated regularly to reflect changing channel conditions. The process involves determining both the resource proportions and the appropriate AP for each user to associate with in a singlepath network. In the next section, the problem formulation for the resource allocation minimizing network delay for a LiFi-WiFi network is described for both single and multipath configurations. For this purpose we use the average network packet delay as the metric to minimize.

The channel models for LiFi and WiFi are elaborated in Section 2.3 and Section 2.6. To analyze the performance of the system under Line-of-Sight (LoS) blockages to the LiFi signal, we use several different blockage models like transient blockages, correlated blockages and the geometric blockage models as explained in Section 2.4. To evaluate the performance of our resource allocation strategy with mobile users, we include the Random Waypoint (RWP) mobility model. Further details about the mobility model are provided in Section 2.7. Table 4.1 provides a summary of the notations used in this section.

Table 4.1 List of Notations used in Delay-aware Resource Allocation

Notation	Description
a, M^A, $\mathcal{A}$	Index, number, and set of total APs
u, M^U, $\mathcal{U}$	Index, number, and set of users
λ_u	Arrival rate of QoS packets of user u in packets per second
L_u	Size of a QoS packet of user u in bits
$\tau_{u,a}$	Network packet delay
$\tau_{\text{thresh},u}$	Delay requirement for user u
$R_{u,a}$	Achievable rate of user u connected to AP a
$R_{\text{thresh},u}$	Rate requirement for u
$x_{u,a}$	Optimization variable indicating association
$y_{u,a}$	Optimization variable indicating resource proportion
$\mathcal{A}_u$	Set of the best APs one of each technology for user u

4.1.4 Delay Minimization Problem Formulation for Singlepath Networks

In a singlepath network, each user can connect to only one AP at any given moment. Therefore, the resource allocation process must determine both the user's association with an AP, represented by $x_{u,a}$, and the proportion of resources allocated by that AP, denoted as $y_{u,a}$. Here, $x_{u,a}$ is a binary variable equal to 1 when user u is connected to AP a.

The network packet delay in a singlepath network, $\tau_{u,a}$, is described by:

$$\tau_{u,a} = \frac{1}{y_{u,a}\frac{R_{u,a}}{L_u} - x_{u,a}\lambda_u} \tag{4.1}$$

where $y_{u,a}R_{u,a}/L_u$ is the service rate for the user u.

The objective of the resource allocation is to minimize the average network packet delay while ensuring that each user's data flow meets the maximum delay and minimum data rate requirements. The optimization problem can be formulated as

$$\min_{x_{u,a},y_{u,a}} \quad \frac{1}{\sum_{u=1}^{M^U}\lambda_u}\sum_{a=1}^{M^A}\sum_{u=1}^{M^U} x_{u,a}\lambda_u\tau_{u,a} \tag{4.2}$$

$$\text{subject to} \quad \sum_a x_{u,a} = 1 \qquad \forall u = 1,2,\dots M^U \tag{4.3}$$

$$\sum_u x_{u,a}y_{u,a} \le 1 \qquad \forall a = 1,2,\dots M^A \tag{4.4}$$

$$\sum_a x_{u,a}\tau_{u,a} \le \tau_{\text{thresh},u} \qquad \forall u = 1,2,\dots M^U \tag{4.5}$$

$$\sum_a x_{u,a}y_{u,a}R_{u,a} \ge R_{\text{thresh},u} \qquad \forall u = 1,2,\dots M^U \tag{4.6}$$

$$x_{u,a} \in \{0,1\} \qquad \forall u,a \tag{4.7}$$

$$0 < y_{u,a} \le 1 \qquad \forall u,a \tag{4.8}$$

The equality constraint in (4.3) ensures that each user can only be associated with a single AP. The constraint in (4.4) sets a limit on the maximum capacity of each AP. The constraints in (4.5) and (4.6) specify the requirements for delay and data rate, respectively. Additionally, we consider that besides the QoS packets, there is other traffic directed to the user that contributes to fulfilling each user's data rate requirement.

4.1.5 Delay Minimization Problem Formulation for Multipath Networks

In a multipath network, each user can be served simultaneously by two different technologies, with one AP per technology. Consequently, the resource allocation process must determine the resource proportions $y_{u,a}$ allocated by the APs of both technologies. The optimal APs are pre-determined based on the highest Signal to Interference and Noise Ratio (SINR) they offer. For each user u, the set $\mathcal{A}_u$ includes the LiFi AP with the highest SINR among all LiFi APs and the best of the WiFi APs. The network packet delay for user u, denoted as τ_u, is then described by:

$$\tau_u = \frac{1}{\sum_{a\in\mathcal{A}_u} y_{u,a}\frac{R_{u,a}}{L_u} - \lambda_u} \tag{4.9}$$

where $\sum_{a\in\mathcal{A}_u} y_{u,a}\frac{R_{u,a}}{L_u}$ represents the total service rate provided to user u by both technologies combined.

The objective is the same as in the singlepath network which is to minimize the average network packet delay while satisfying specific delay and data rate requirements for each user. Therefore, the optimization problem is formulated as follows:

$$\min_{y_{u,a}} \quad \frac{1}{\sum_{u=1}^{M^U} \lambda_u} \sum_{u=1}^{M^U} \lambda_{u,a} \tau_u \tag{4.10}$$

$$\text{subject to} \quad \sum_{u} y_{u,a} \leq 1 \qquad \forall a = 1, 2, \ldots M^A \tag{4.11}$$

$$\tau_u \leq \tau_{\text{thresh},u} \qquad \forall u = 1, 2, \ldots M^U \tag{4.12}$$

$$\sum_{a \in \mathcal{A}_u} y_{u,a} R_{u,a} \geq R_{\text{thresh},u} \qquad \forall u = 1, 2, \ldots M^U \tag{4.13}$$

$$0 < y_{u,a} \leq 1 \qquad \forall u, a \tag{4.14}$$

The constraint in (4.11) specifies the maximum capacity limit for each AP. Meanwhile, the constraints in (4.12) and (4.13) set the requirements for maximum allowable delay and minimum data rate, respectively.

4.1.6 Resource Allocation schemes for delay minimization

The singlepath network optimization problem outlined in the previous section is a MINLP problem, as it involves both integer variables $x_{u,a}$ and real valued $y_{u,a}$, with the multipath problem exhibiting non-linearity. The constraints for both problems define their feasible regions. Generally, MINLP problems are mathematically complex and challenging to solve. Branch and bound algorithms are widely employed to address mixed integer problems. These algorithms recursively decompose the optimization problem into smaller sub-problems until they become manageable. To avoid exhaustively exploring all variable combinations, the algorithm efficiently prunes the search space. In this research, we utilize the Gurobi solver [Gur23], which solves MINLP problems using a spatial branch and bound algorithm. For comparative purposes, we also apply this solver to the multipath problem.

In addition to the branch and bound-based solver, we propose a Genetic Algorithm (GA) metaheuristic [Gol89] to solve the problem with reduced complexity while maintaining solution quality. Genetic algorithms are inspired by the theory of natural selection, where the fittest individuals are chosen for reproduction to form the next generation's population. A standard genetic algorithm typically involves the following stages

1. Initial Population: The initial population comprises a set of potential solutions to the optimization problem, randomly generated to include both integer and real-valued variables.

2. Fitness Evaluation: The fitness function, or objective function, assigns a fitness score to each individual solution. This score determines the suitability of each individual for reproduction. The fitness functions we utilize are the objective functions described in (4.2) and (4.10).

3. Selection: In this stage, individuals with the highest fitness scores are selected for reproduction, ensuring that the best solutions are carried forward to the next generation.

4. Crossover: During crossover, selected parent individuals are combined to produce offspring, which will form part of the next generation. The generation of offspring adheres to the variable bounds specified in (4.7), (4.8), and (4.14). For this process, we employ the Simulated Binary Crossover operator [DA+95], to handle both integer and real variables.

5. Mutation: To maintain genetic diversity and prevent premature convergence to a local optimum, some offspring undergo mutation with a certain probability. The mutation process follows the same probability distribution as the crossover.

6. Termination: The algorithm terminates when it converges to a solution, ensuring that the variable solutions are feasible and optimal or near-optimal.

To ensure the feasibility of the solution, constraints must be incorporated into the model. Typically, meta-heuristic algorithms are designed for unconstrained problems, so adding constraints requires manipulation of the objective or fitness function. This is often achieved through a penalty function, where any violation of a constraint penalizes the objective function value, making the corresponding solution less desirable. In our approach, the genetic algorithm incorporates these constraints by modifying the fitness function to include penalties for constraint violations. As a result, the algorithm ranks all solutions based on their feasibility, prioritizing feasible solutions over those that do not meet the constraints. The algorithm converges when it identifies solutions that satisfy all constraints while minimizing the objective function. Figure 4.2 illustrates the average convergence time of the genetic algorithm used in this work. The algorithm is observed to search around in the population space until it finds the set of feasible solutions. Once it finds the feasible solution it then continues in this feasible space to minimize the objective function until convergence.

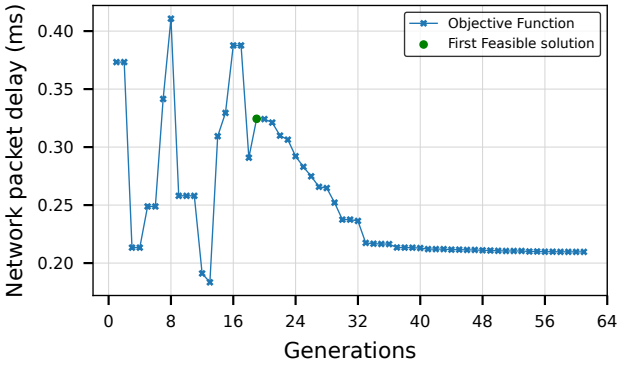


Figure 4.2 Convergence of the genetic algorithm

The meta-heuristic algorithm used in this work starts with an initial population of 100 and is designed to run through a maximum of 200 generations. However, the algorithm may end earlier if convergence is reached before the maximum number of generations and this is evidenced by the termination of the algorithm in 61 generations as seen in Figure 4.2.

4.1.7 Evaluation Methods

Max-SNR Method for Benchmarking:
We develop a baseline model to benchmark the performance of our optimization framework for minimizing delay in LiFi-WiFi networks. User association with APs is determined by selecting the AP that provides the highest Signal to Noise Ratio (SNR) since this is the association strategy that is currently implemented in market devices. The proposed resource allocation is evaluated using a simulation framework implemented in Python 3.10.6 with Gurobi v11.0.0.

Table 4.2 Network Topology under evaluation for delay minimization

Dimension	Users	LiFi APs	Position of LiFi APs	WiFi APs	Position of WiFi APs
5 x 5 x 3 m	10	4	(-1.4, 1.4, 3.0), (1.4,-1.4, 3.0) (-1.4,-1.4, 3.0), (1.4, 1.4, 3.0)	1	(0, 0, 3.0)

To evaluate the effectiveness of the proposed resource allocation approach, we examine their performance in an indoor network topology which includes 1 WiFi AP and 4 LiFi APs positioned at

a height of 3 m. The specific parameters of this topology are provided in Table 4.2. The proposed framework is evaluated under varying LiFi channel parameters like blockage models and device orientation.

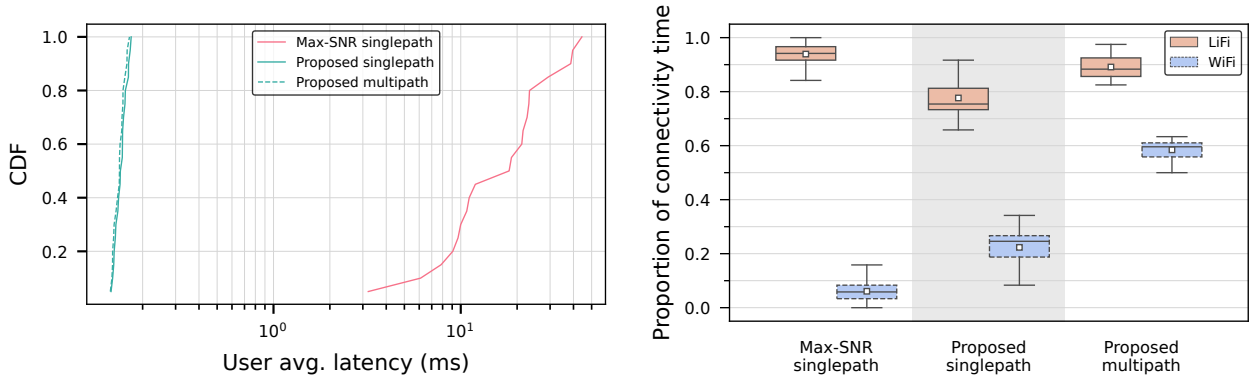
Performance evaluations were performed using an 11th Generation Intel® Core™ i7-11700 16-Core Processor. Each set of results was derived from 20 independent simulation runs, with each run consisting of 120 time steps. To validate our findings comparing the solution methods proposed, we use the Mann-Whitney U test [MW47], operating under the null hypothesis that there is no difference in the distributions of the two parameters being compared. The outcomes of these tests are annotated in the corresponding figures with symbols based on the p-values [Cha+22], using the star notation.

$$\begin{aligned} \text{ns} &: p > .05 \\ * &: .01 < p <= .05 \\ ** &: .001 < p <= .01 \\ *** &: .0001 < p <= .001 \\ **** &: p <= .0001 \end{aligned}$$

Furthermore, due to multiple hypothesis testing on the same dataset, we applied the Benjamini-Hochberg procedure [BH95] to adjust for the false discovery rate, ensuring the reliability of our statistical inferences.

4.1.8 Results and Comparative Analysis

We begin the evaluations by considering a high traffic system with packets arriving at a rate of 1500 pkts/s with a length of 9000 bits. The delay requirement is set to 1 ms and the data rate requirement to 15 Mb/s for all users.



(a) CDF of the user average latency

(b) Proportion of time that a user is connected to a wireless access technology

Figure 4.3 Quality metrics for the Max-SNR benchmark and delay-optimized resource allocation solved with the Expert for a single and multipath network showing the need for our delay-optimized perspective.

Figure 4.3a presents the Cumulative Distribution Function (CDF) of average user packet delay for the baseline Max-SNR user association approach compared to our proposed delay-aware resource allocation using singlepath and multipath technologies. The Max-SNR method, which neglects QoS requirements, results in a wide range of delays due to sub-optimal AP allocations and an inability to manage large traffic volumes. Consequently, it fails to meet the 1 ms latency requirement for all users. In contrast, our delay-optimized algorithms efficiently handle high traffic, achieving an average packet delay of under 1 ms for all users in the network. This emphasizes the benefit of deploying small-range LiFi cells with high data rate density. Additionally, connecting users to both

technologies further improves delay performance, as evidenced by the difference between singlepath and multipath results.

For the same high traffic scenario, Figure 4.3b provides a detailed analysis of the proportion of time each user is connected to a specific access technology. A user is considered connected when the received data rate is greater than zero. The results indicate that users are predominantly connected to LiFi, except during instances of light blockage, which renders LiFi connectivity impossible. In the multi-homing scenario, although users have access to both LiFi and WiFi, they are not always connected to WiFi. This lower proportion of connection time is attributed to the maximum capacity constraints of the WiFi APs with a larger number of users sharing its bandwidth.

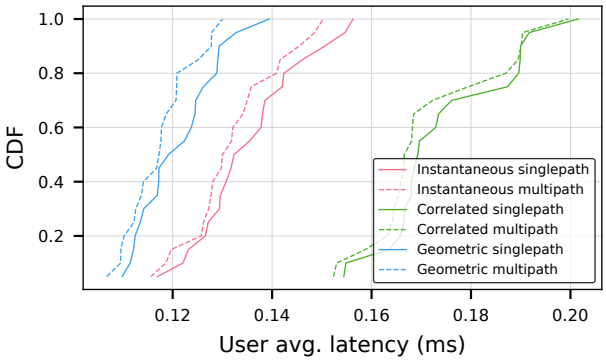


Figure 4.4 CDF of the user average latency for varying LiFi LoS blockage models for our proposed delay-aware resource allocation in a single and multipath network

The behavior of the network is then is evaluated under varying LiFi-specific parameters such as receiver orientation and blockage models, while maintaining constant packet length (9000 bits), arrival rate (20 packets/s), delay requirement (10 ms), and rate requirement (5 Mb/s) for all users. Figure 4.4 illustrates the CDF of average user packet delay across different blockage models which were described in Section 2.4. As the blocking probability increases, the likelihood of connecting to a LiFi AP decreases, leading to more users sharing the resources of a WiFi AP and consequently experiencing increased delays. The correlated blockage model, characterized by continuous periods of blockages, results in the highest delays. Conversely, the geometric blockage model, which considers blockages caused by other users, exhibits the lowest delays. This is because blockages only occur when a user obstructs the direct LoS between another user and an AP , which is rare given that APs are ceiling-mounted. Furthermore, the analysis shows that while multi-homing provides improved delay performance, the improvement is not substantial. This suggests that the primary factor influencing delay is the blocking probability and the resulting resource allocation rather than the mere presence of multiple access technologies.

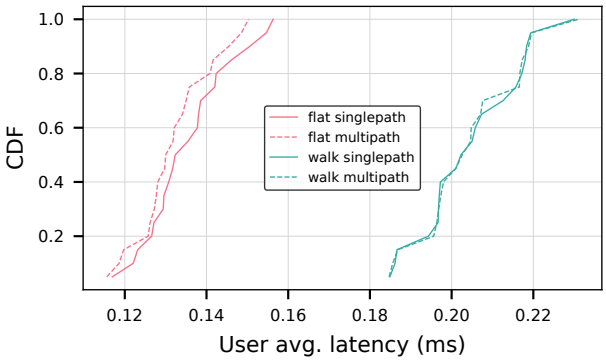


Figure 4.5 CDF of the user average latency for varying LiFi user device orientations for our proposed delay-aware resource allocation in a single and multipath network

Figure 4.5 illustrates the average network packet delay for different user device orientations, a crucial parameter in LiFi systems. The orientation of a user device affects whether the received signal falls within the receiver's Field of View. When the user device is parallel to the ground and facing upwards (elevation angle 0°), it receives the maximum signal from the AP, resulting in minimal delay. This scenario is referred to as the flat device orientation. In contrast, the walk device orientation represents a typical mobile user with an elevation angle of 28°. For this orientation, the singlepath and multipath solutions exhibit similar performance. This similarity arises because the tilted orientation disrupts the longer connection times to LiFi that were previously observed in Figure 4.3b, leading to reduced signal reception and increased reliance on WiFi. This analysis highlights the challenges faced by mobile users with tilted devices, emphasizing the need for adaptive strategies to maintain low delay in varying user contexts.

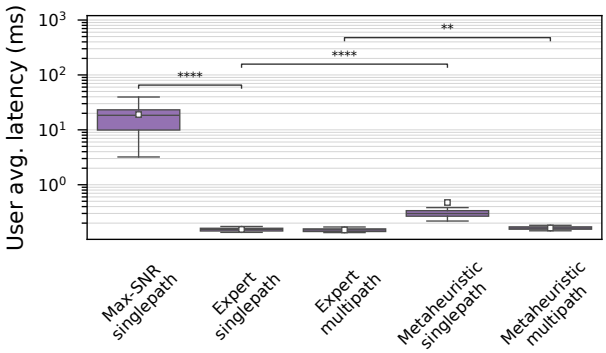


Figure 4.6 User average latency for all proposed solution methods to the delay-aware resource allocation in a single and multipath network comparing them to the Max-SNR approach. Results of statistical tests of significance comparing the Expert and other solution approaches are annotated.

While the Expert solver demonstrates effective delay optimization in a LiFi-WiFi network, its real-time application is limited by the time required to solve the optimization problem. To address this, we investigate an alternative approach using a meta-heuristic GA for high traffic scenarios with a 1 ms delay requirement. The results, depicted in Figure 4.6, show the average user latency for different solution methods. The findings indicate that all proposed solutions outperform the Max-SNR method, consistently maintaining user latency within the 1 ms bound. Although statistical tests reveal significant differences between the Expert solver and the GA, these differences are practically minor. However, in singlepath networks, the GA diverges in performance by relying more on WiFi and less on LiFi compared to the Expert solver. This deviation suggests a sub-optimal tuning of link

usage parameters by the GA. Despite this, the GA still achieves significantly lower latency than the required threshold.

4.1.9 Summary

This work focuses on optimizing wireless resource allocation and user association to minimize average network packet delay, necessary for delay-sensitive applications. We explored this optimization in both singlepath and multipath LiFi-WiFi heterogeneous networks. The problem was formulated to minimize delay while enforcing constraints on maximum allowable delay and achievable data rate for each user. To tackle this problem, we implemented a Branch and Bound algorithm-based solver, Gurobi, and proposed a meta-heuristic GA for comparison. Extensive simulations were carried out to assess the effectiveness of these approaches. Our results indicate that the proposed methods significantly reduce network delay compared to existing max-SNR-based techniques. Additionally, they consistently meet strict delay and data rate requirements, even in scenarios with heavy QoS traffic. Our findings also suggest that LiFi-WiFi networks are highly effective for low-latency applications, achieving sub-millisecond latency, particularly when multi-homing user devices are utilized. This work underscores the potential of advanced optimization strategies in improving the performance of heterogeneous networks for applications requiring stringent delay constraints.

In the next section, we extend this delay-aware perspective to resource allocation to include the metrics of throughput and fairness. By integrating these metrics, we aim to provide a more comprehensive evaluation of resource management strategies in LiFi-WiFi heterogeneous networks. We also address the challenges posed by frequent handovers, which can lead to temporary disruptions in communication, affecting both delay and overall network performance. To mitigate these disruptions, we evaluate various strategies designed to manage resources more effectively in the face of frequent handovers.

4.2 Weighted Alpha-fair Wireless Resource Allocation towards Stability and Utility Maximization

In the previous section of this chapter, we explored initial strategies for managing a heterogeneous LiFi-WiFi network aimed at minimizing latency and ensuring diverse QoS requirements. By leveraging the unique capabilities of this network configuration, we addressed the resource allocation needs of users, with multi-homing capabilities [VK21]. The potential of multi-homing to enhance network reliability and throughput has been clearly shown [WH19a], although its benefits are currently restricted due to the limited multi-homing support in prevalent devices. Most contemporary devices, such as smartphones, do not support multi-homing, which could significantly enhance reliability and throughput by allowing simultaneous connections to multiple networks. With the absence of multi-homing, increased handovers [WSH17], including ping-pong handovers in areas where LiFi and WiFi networks overlap, can occur. Frequent handovers are also caused due to rapid connection changes due to transient blockages of the LiFi signal and changes in the orientation of the user device [Sol+17]. Such frequent handovers not only disrupt service but also degrade the network performance, causing packet losses and delays. These challenges unique in LiFi-WiFi networks necessitates more research to ensure stability compared to handovers in RF networks [LCW16].

In this section, we explore methods to ensure the stability of LiFi-WiFi heterogeneous networks by addressing the challenges posed by handover overheads. We introduce a low-complexity algorithm designed to optimize resource allocation, accounting for the data rate loss due to handover overhead. In the presence of mobility and complete LiFi signal outages, handovers cannot be avoided and soft handover protocols for LiFi-only network have been analyzed in [DEA15] and [VL12]. To mitigate the overhead of unavoidable handovers in LiFi-WiFi heterogeneous networks, we propose a system approach to managing vertical handovers, which we implement and evaluate on a hardware setup. Vertical handovers have the additional challenge of having to manage multiple network interfaces

[ABG14]. Our results indicated that vertical handover overhead can be minimized to negligible levels, thus maintaining network stability without requiring Transmission Control Protocol (TCP) re-connections.

Having established a stable network foundation, we then turn our attention to further optimizing network performance. The objective of this section is to maximize the utility function based on alpha-fairness principles, assigning access points to users and allocating wireless bandwidth resources in an optimal manner. By employing alpha-fair resource allocation, we aim to achieve a balanced distribution of resources that considers both individual user needs and overall network efficiency. This approach not only enhances user satisfaction but also pushes the performance boundaries of the network, ensuring that resources are utilized in the most effective way possible.

The enhancements in network stability introduced by our handover management strategies have been documented in a published conference paper [Vij+21]. Additionally, our network optimization through alpha-fair utility maximization is partly detailed in a journal manuscript currently under revision [VK24]. In the upcoming subsections, we introduce relevant state of art and system model for handover management and resource allocation towards throughput maximization. Then we present the weighted alpha-fair resource allocation problem formulation of which a simplified version is initially used to enhance network stability. This optimization formulation, including considerations for handover losses, is then solved for alpha-fair utility maximization using EGT and Lagrangian optimization techniques. The performance evaluation is split into three parts where the first analyzes the stability evaluations, the second evaluates the alpha-fairness aspect in detail, and the third presents and evaluates a vertical handover protocol that aims to minimize the overhead of unavoidable handovers.

4.2.1 State-of-the-art Analysis

In the analysis presented in [WH19b], user trajectory data is used to strategically skip unnecessary handovers in mobile networks. This approach optimizes network performance by reducing the frequency of handover procedures, which can degrade user experience due to increased latency and potential service interruption. However, the method does not address the needs of stationary users who may experience transient signal blockages, such as those caused by building construction or temporary obstructions.

Expanding on the theme of network optimization, [WSH17] tackle load balancing with a focus on integrating handover management to improve overall network throughput. Their algorithm assigns users to APs and allocates resources to maximize the throughput of the network. This approach, while effective in theory, suffers from high computational costs and necessitates centralized control, which may introduce bottlenecks and scalability issues in larger network deployments.

To address these computational challenges, [WWH17] introduce a novel heuristic algorithm based on EGT. This method reduces processing overhead by simplifying decision-making processes in the face of dynamic network conditions, such as light path blockages caused by intermittent physical barriers. Although it considers the impacts of such blockages, the algorithm stops short of integrating these blockages into its resource allocation.

In another contribution, [WWH16] explore a low-complexity approach using fuzzy logic to manage AP allocation with a focus on the handover processes. Their framework prioritizes reduced computational demands, making it suitable for real-time applications and smaller networks. However, this scheme does not address the optimization of wireless resource allocation.

Extending beyond handover considerations [WWH17] perform alpha-fair resource allocation for load balancing in LiFi-WiFi networks using an EGT-based algorithm which is similar to our approach. However, their model does not account for the heterogeneous resource demands of different users — a gap our research addresses through a weighted alpha-fairness approach.

The principle of alpha-fairness has been thoroughly investigated in RF networks [SHI+14; JY22; Xu+22]. However, it does not address the unique challenges prevalent in LiFi-WiFi heterogeneous networks, such as signal blockages and the density of placement of LiFi APs. Moreover, existing

models often overlook the necessity of tailoring resource allocation to individual user needs and priorities.

In summary, our work differs from the state-of-the-art by specifically tailoring resource allocation strategies to the challenges posed by LiFi-WiFi heterogeneous networks, thereby ensuring network stability and optimizing resource distribution. We adapt our approach to meet the diverse transmission requirements of users by assigning varying priorities or weights based on their specific resource needs. Consequently, our framework promotes efficient and prudent resource allocation, minimizing the potential for resource wastage.

4.2.2 Key Contributions

In this section, we present a comprehensive framework to ensure stability and optimize resource allocation in LiFi-WiFi heterogeneous networks, considering transient channel conditions, mobility, and fairness.

1. We propose methods to maintain a stable LiFi-WiFi network amidst transient channel conditions caused by light path blockages, instantaneous receiver orientation changes, and mobility.

2. We propose a system approach to manage unavoidable vertical handovers, ensuring seamless interface switching. This approach has been implemented and tested on hardware in a LiFi-WiFi network setup, with performance evaluated through practical measurements.

3. We then propose a framework for mobility-aware resource allocation in mobile LiFi-WiFi networks, considering user positions, LoS blockages, and channel quality. This framework adapts to current network conditions, addressing the challenge of optimal resource allocation by formulating an optimization problem that performs AP assignment and wireless resource allocation using the weighted alpha-fairness objective.

4. We propose an EGT-based solution that is benchmarked against a state-of-the-art baseline allocation strategy.

4.2.3 System Model

This work explores resource management in a LiFi-WiFi heterogeneous network comprising M^L LiFi APs and M^W WiFi APs, with a combined total of M^A APs. The LiFi APs, which are LEDs mounted at a height of 3 m, operate at the same frequency, leading to co-channel interference in overlapping areas. Conversely, WiFi APs operate on distinct, non-overlapping frequencies.

The network serves a total of M^U users, each equipped with both LiFi photodiode and WiFi receivers, although connectivity to only one AP is possible at any time. All APs are connected to a central network controller that facilitates instantaneous, error-free communication and monitors the load on each AP by collecting and analyzing the wireless Channel State Information (CSI) from all users. Resource allocation decisions are made by a centralized algorithm that operates at intervals denoted by τ=500 ms. This algorithm uses global network data to allocate resources and communicate these decisions back to the users. Table 4.3 provides a summary of the notations used in this section.

Table 4.3 List of Notations used in Weighted Alpha-fair Resource Allocation

Notation	Description
l, M^L	Index and number of LiFi APs
w, M^W	Index and number of WiFi APs
a, M^A	Index and number of total APs
u, M^U, $\mathcal{U}$	Index, number, and set of users
τ	Length of one time slot
$\eta_{a',a}$	Handover efficiency when switching from AP a' to a
$R_{u,a}$	Achievable rate of user u connected to AP a
T_{loss}	Time loss incurred due to a handover
$\tilde{R}_{u,a}$	Handover-reduced rate for user u to AP a
$f_\alpha()$	Alpha-fairness function
α	Parameter to control the alpha-fairness
w_u	Weights assigned to user for a weighted alpha-fairness
$x_{u,a}$	Optimization variable indicating association
$y_{u,a}$	Optimization variable indicating resource proportion
$U_\alpha(x_{u,a}, y_{u,a})$	Utility function to be maximized
L, λ, ω	Lagrangian function and multipliers
N_a, $\mathcal{U}_a$	Number and set of users associated to AP a
i	Index of iteration of the EGT algorithm
$F_{u,a}$	Payoff for user u to AP a
$\bar{F}^i$	Global average Payoff in i^{th} iteration
p_u^i	Mutation probability of user u

The channel models for LiFi and WiFi are elaborated in Section 2.3 and Section 2.6. To analyze the stability of the system under LoS blockages to the LiFi signal, we use several different blockage models like transient blockages, correlated blockages and the geometric blockage models as explained in Section 2.4. To evaluate the performance of our resource allocation strategy with mobile users, we include four mobility models: RWP, Truncated Levy Walk, Self-Similar Least Walk (SLAW) [Lee+09], and Reference Point Group (RPG) [Hon+99] which also capture the changing orientation of a user device. Further details are provided in Section 2.7.

4.2.4 Handover losses

Due to the limited coverage area of LiFi APs, user mobility can necessitate handovers, which can be either horizontal (within the same technology) or vertical (between different technologies).

In indoor scenarios, handover overhead typically occurs in the range of milliseconds, leading to a reduction in the data rate for users undergoing handovers. The handover efficiency when a user moves from AP a' to a is represented as $\eta_{a',a}$. This efficiency decreases as handover loss increases. Consequently, the maximum achievable link rate $R_{u,a}$ for user u associated with AP a in the current time slot and a' in the previous time slot is reduced to $\tilde{R}_{u,a} = R_{u,a}\eta_{a',a}$. This efficiency can also be viewed as a weighting factor that reduces the data rate. So this weighting factor is defined as a function of the delay incurred due to AP switches. Specifically, $eta_{a',a}$ can be defined as

$$\eta_{a',a} = \begin{cases} 1 - \frac{T_{\text{loss}}}{\tau} & a \neq a' \\ 1 & \text{otherwise} \end{cases} \tag{4.15}$$

where T_{loss} is the time loss incurred due to performing the handover procedure, and τ is the time interval between optimization states.

In this work and all the following chapters, we consider a handover overhead of 200 ms for horizontal handovers and 300 ms for vertical handovers. This results in 60% and 40% handover

efficiency respectively. The vertical handover overhead of 300 ms is validated by experiments on our in-house testbed and will be further discussed in Section 4.2.10.

4.2.5 Optimization Problem Formulation for Weighted Alpha-Fair Resource Allocation

Our resource management framework incorporates the alpha-fairness function [MW00], which provides a method for balancing network efficiency and fairness in resource allocation. We examine cases with $\alpha = 0$, $\alpha = 1$, $\alpha = 2$, and $\alpha = \infty$. For the heterogeneous network stability analysis, only $\alpha = 1$ is analyzed in greater detail.

The weighted alpha-fairness function, $f_\alpha(R_u)$, for a user achieved rate R_u, user weight w_u, and fairness coefficient α is defined as follows:

$$f_\alpha(R_u) = \begin{cases} w_u \log(R_u) & \alpha = 1 \\ w_u \frac{R_u^{1-\alpha}}{1-\alpha} & \alpha \geq 0, \alpha \neq 1 \end{cases} \tag{4.16}$$

The user weights w_u indicate the priority or ordering for these users. These weights can be tuned based on user priorities or users' resource demands. By including these weights into the optimization problem, we offer a flexible framework that differentiates user needs according to these weights. A higher weight indicates a larger user demand or priority.

The implication of the different α values are as follows:

- Maximizing Total Throughput ($\alpha = 0$) The focus is on maximizing the total throughput of the network, often leading to a more efficient but potentially less fair allocation.
- Proportional Fairness ($\alpha = 1$) The utility function becomes the sum of the logarithms of the allocated resources, aiming to achieve proportional fairness, where each user's relative benefit is equalized.
- Delay Fairness ($\alpha = 2$) This represents delay fairness, focusing on minimizing the sum of the inverse of the allocated resources, effectively prioritizing users with higher delays.
- Max-Min Fairness ($\alpha = \infty$) The objective is to maximize the minimum allocated resource, achieving max-min fairness and ensuring the best possible allocation for the worst user.

By varying α, we can customize the resource allocation strategy to align with specific network goals.

Using the alpha-fairness function, the optimization problem is formulated as follows:

$$\max_{x_{u,a}, y_{u,a}} \quad U_\alpha(x_{u,a}, y_{u,a}) \tag{4.17}$$

$$\text{subject to} \quad \sum_a x_{u,a} = 1 \qquad \forall u \tag{4.18}$$

$$\sum_u x_{u,a} y_{u,a} \leq 1 \qquad \forall a \tag{4.19}$$

$$x_{u,a} \in \{0, 1\} \qquad \forall u, a \tag{4.20}$$

$$0 < y_{u,a} \leq 1 \qquad \forall u, a \tag{4.21}$$

where $x_{u,a}$ is a binary variable that indicates whether user u is associated to AP a, and $y_{u,a}$ is a continuous variable representing the proportion of bandwidth allocated from AP a to user u.

The utility U_α for different values of α is expressed as:

$$U_\alpha(x_{u,a}, y_{u,a}) = \begin{cases} \sum_u \sum_a w_u x_{u,a} \log\left(y_{u,a} \tilde{R}_{u,a}\right) & \text{if } \alpha = 1 \\ \sum_u \sum_a w_u x_{u,a} \frac{\left(y_{u,a} \tilde{R}_{u,a}\right)^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1 \end{cases} \tag{4.22}$$

To represent $\alpha = \infty$, we approximate it using $\alpha = 20$. $\tilde{R}_{u,a}$ denotes the achievable link rate from AP a to user u, which is influenced by the user's previous association a'. The handover efficiency factor is applied to adjust the link rate, accounting for the potential reduction in achievable rates resulting from handovers. In the initial network stability analysis we evaluate this optimization problem with and without consideration for this handover efficiency factor.

4.2.6 Proposed Resource Allocation with Lagrangian Optimization

To simplify the MINLP presented in the previous section, we convert the problem into a convex form, by relaxing the binary variable $x_{u,a}$ to take fractional values in the range $[0, 1]$. This relaxed problem is a concave optimization problem for $\alpha > 0$ since the objective is concave and the constraints are affine. The Lagrangian multiplier method can now be used to find the optimal solution to this relaxed problem.

The Lagrangian function is given by:

$$L = \sum_u \sum_a w_u x_{u,a} \frac{\left(y_{u,a}\tilde{R}_{u,a}\right)^{1-\alpha}}{1-\alpha} - \sum_u \lambda_u \left(\sum_a x_{u,a} - 1\right) - \sum_a \omega_a \left(\sum_u x_{u,a} y_{u,a} - 1\right) \tag{4.23}$$

where λ_u is the Lagrangian multiplier for the constraint (4.18) and ω_a the multiplier for (4.19). By re-writing this equation and since $\sum_a x_{u,a} - 1 = 0$ we get

$$L = \sum_u \sum_a x_{u,a} \left(w_u \frac{\left(y_{u,a}\tilde{R}_{u,a}\right)^{1-\alpha}}{1-\alpha} - \omega_a y_{u,a}\right) - \sum_a \omega_a \tag{4.24}$$

The optimal variables can be found by setting the partial derivation of the Lagrangian function L with respect to those variables to 0. Setting the partial derivation with respect to $y_{u,a}$ to 0 and solving for it gives

$$\frac{\partial L}{\partial y_{u,a}} = x_{u,a} \left(w_u \frac{\tilde{R}_{u,a}}{\left(y_{u,a}\tilde{R}_{u,a}\right)^{\alpha}} - \omega_a\right) = 0 \tag{4.25}$$

As $y_{u,a}$ increases from 0 to 1 $\frac{\partial L}{\partial y_{u,a}}$ monotonically decreases since $x_{u,a}$ is non-negative. Therefore, the optimal $y^*_{u,a}$ is calculated as

$$y^*_{u,a} = \begin{cases} 1, & \left.\frac{\partial L}{\partial y_{u,a}}\right|_{y=1} \geq 0 \\ \left(\frac{w_u \tilde{R}_{u,a}}{\omega_a}\right)^{\frac{1}{\alpha}} \frac{1}{\tilde{R}_{u,a}}, & \text{else} \end{cases} \tag{4.26}$$

To obtain the optimal association $x_{u,a}$, we calculate the derivative of the Lagrangian function with respect to it.

$$\frac{\partial L}{\partial x_{u,a}} = w_u \frac{y_{u,a}\tilde{R}_{u,a}}{1-\alpha} - \omega_a y_{u,a} \tag{4.27}$$

By substituting the optimal $y^*_{u,a}$ as found in (4.26) into (4.27) and maximizing it gives the optimal $x^*_{u,a}$ as

$$x^*_{u,a} = \begin{cases} 1, & a = \arg\max_{a'} \frac{\partial L\left(y^*_{u,a'}\right)}{\partial x_{u,a'}} \\ 0, & \text{else} \end{cases} \tag{4.28}$$

To update the Lagrangian multiplier in each iteration, the dual optimum $y^*_{u,a}$ and $x^*_{u,a}$ are substituted back into the Lagrangian function L and the partial derivative with respect to the multiplier ω_a is calculated as

$$\frac{\partial L}{\partial \omega_a} = 1 - \sum_u x^*_{u,a} y^*_{u,a} \tag{4.29}$$

Therefore, the update of the multiplier for the $i+1^{\text{th}}$ iteration is given by:

$$\omega_a^{(i+1)} = \omega_a^i - \epsilon_\omega \left(1 - \sum_u x^*_{u,a} y^*_{u,a}\right) \tag{4.30}$$

where ϵ_ω is the step size in the direction of the negative gradient of ω_a.

Algorithm 2 Lagrangian-based algorithm for alpha-fair wireless resource allocation

1: Initialization: ω_a^i, ϵ_ω, stop_threshold, $i = 0$, $L^{(i+1)} = \infty$, $L^i = 0$.
2: **while** $|L^{(i+1)} - L^i| \geq$ stop_threshold **do**
3: **for each** user u and AP **do**
4: Calculate $y^*_{u,a}$ as per (4.26).
5: Calculate $x^*_{u,a}$ as per (4.28).
6: **end for**
7: **for each** AP a **do**
8: Update Lagrangian multiplier ω_a as per (4.30).
9: **end for**
10: Calculate Lagrangian function $L^{(i+1)}$ by using $y^*_{u,a}$, $x^*_{u,a}$, and $\omega_a^{(i+1)}$ in (4.24)
11: $i \leftarrow i + 1$
12: **end while**
13: Result: $y^*_{u,a}$ and $x^*_{u,a}$

A summary of this algorithm is given in Algorithm 2. This iterative Lagrangian optimization algorithm could take many iterations until convergence and we have also shown that the objective function is only concave when the association variable $x_{u,a}$ is continuous, which could lead to longer solve times. Therefore, in the next section we present another solution approach based on Evolutionary Game Theory (EGT).

4.2.7 Proposed Resource Allocation with Evolutionary Game Theory

EGT is particularly well-suited for wireless resource allocation challenges [WWH17], because it models user interactions as a competitive game for the shared resources. Evolutionary games are advantageous for distributed implementations, requiring minimal signaling in a user-driven system. However, in this work, the algorithm is executed at the central controller. This approach ensures that the controller maintains a global understanding of the entire network and can communicate decisions to users after the algorithm's convergence. The algorithm converges to an evolutionary equilibrium, representing a stable state achieved by the users [NH09].

The association of users to APs is managed by an EGT-based algorithm, which includes the following key components:

- *Players:* The users participating in the network.
- *Population:* The population $\mathcal{U}_a$ represents the set of users associated to an AP a, where the number of users assigned to AP a is denoted as N_a.
- *Strategy:* Players can choose one among all the LiFi and WiFi APs available to serve them.

- *Payoff:* Players adjust their strategy in each iteration to maximize their payoff. The payoff for a user u is based on the link rate $R_{u,a}$ offered by the AP a that it is associated to. The payoff is given by:

$$F_{u,a} = \frac{y_{u,a}\tilde{R}_{u,a}}{w_u} \tag{4.31}$$

  where $\tilde{R}_{u,a} = R_{u,a}\eta_{a',a}$ when the handover overhead is considered in the proposed optimization. If the handover overhead is not considered, $\tilde{R}_{u,a} = R_{u,a}$. We use this as a comparative formulation to establish the benefits of considering the handover losses.

The AP association that maximizes the payoff is selected for each user using the game process. During the game, each player selected the AP to which it associates based on its own payoff and the global average payoff of all players. The global average payoff of users in the i^{th} iteration is calculated as:

$$\bar{F}^i = \frac{1}{M^U}\sum_u F_{u,a}. \tag{4.32}$$

where M^U is the total number of users in the system. Users with lower payoff values compared to the global average have a higher probability of switching their association to another AP (mutation). This switching probability is determined by:

$$p_u^i = \begin{cases} 1 - \frac{F_{u,a}{}^{(i-1)}}{\bar{F}^{(i-1)}} & F_{u,a}^{(i-1)} < \bar{F}^{(i-1)} \\ 0 & \text{otherwise.} \end{cases} \tag{4.33}$$

If the user is selected to mutate, the new AP for association is the one that is estimated to offer the maximum payoff i.e., $a = \arg\max_{a'} F_{u,a'}$.

After the AP association is complete with the iterative game, the AP allocates wireless resources to the users based on the alpha-fairness rule. For a given AP assignment, the optimal resource proportion $y_{u,a}$ for a user u associated with AP a is determined by:

$$y_{u,a} = \begin{cases} \begin{cases} 1 & \text{if } u = \arg\max_{u'} w_{u'}\tilde{R}_{u',a} \\ 0 & \text{elsewhere} \end{cases} & \text{if } \alpha = 0 \\ \frac{w_u^{\frac{1}{\alpha}}\tilde{R}_{u,a}^{\frac{1}{\alpha}-1}}{\sum_{u'\in\mathcal{U}_a} w_{u'}^{\frac{1}{\alpha}}\tilde{R}_{u',a}^{\frac{1}{\alpha}-1}} & \text{if } \alpha > 0 \end{cases} \tag{4.34}$$

For $\alpha > 0$, it can be shown that selecting the resource proportion $y_{u,a}$ as in (4.34) is optimal for a given AP assignment. With the AP association $x_{u,a}$ fixed by the result of the game, the optimization problem is formulated as:

$$\max_{y_{u,a}} \quad \sum_a \sum_{u\in\mathcal{U}_a} f_\alpha(y_{u,a}) \tag{4.35}$$

$$\text{subject to} \quad \sum_{u\in\mathcal{U}_a} y_{u,a} \le 1 \qquad \forall a \tag{4.36}$$

$$x_{u,a} \in \{0,1\},\ y_{u,a} \in (0,1] \qquad \forall u, a \tag{4.37}$$

Here, $\mathcal{U}_a$ represents the set of all users associated with AP a, and the alpha-fairness function $f()$ is defined as:

$$f_\alpha(y_{u,a}) = \begin{cases} w_u \log(y_{u,a}\tilde{R}_{u,a}) & \alpha = 1 \\ \frac{w_u(y_{u,a}\tilde{R}_{u,a})^{1-\alpha}}{1-\alpha} & \alpha \ge 0, \alpha \ne 1 \end{cases} \tag{4.38}$$

After fixing the AP assignment, the resource allocation can be independently performed at each AP for its associated users, transforming the problem for a single AP into:

$$\max_{y_{u,a}} \quad \sum_{u \in \mathcal{U}_a} f_\alpha(y_{u,a}) \tag{4.39}$$

$$\text{subject to} \quad \sum_{u \in \mathcal{U}_a} y_{u,a} \leq 1 \tag{4.40}$$

The objective function (4.39) is concave and can be solved using the Lagrangian multiplier method. The Lagrangian is given by

$$L = \sum_{u \in \mathcal{U}_a} f_\alpha(y_{u,a}) - \lambda \left(\sum_{u \in \mathcal{U}_a} y_{u,a} - 1 \right) \tag{4.41}$$

where λ is the Lagrangian multiplier. The optimal $y_{u,a}$ can be calculated by setting the gradient of the Lagrangian to 0 and solving for $y_{u,a}$:

$$\frac{\partial L}{\partial y_{u,a}} = \frac{w_u \tilde{R}_{u,a}}{(y_{u,a} \tilde{R}_{u,a})^\alpha} - \lambda = 0 \tag{4.42}$$

$$y_{u,a} = \frac{w_u^{\frac{1}{\alpha}} \tilde{R}_{u,a}^{\frac{1}{\alpha}-1}}{\lambda^{\frac{1}{\alpha}}} \tag{4.43}$$

Since $\sum_{u \in \mathcal{U}_a} y_{u,a} \to 1$ if the objective is to be maximized,

$$\lambda^{\frac{1}{\alpha}} = \sum_{u \in \mathcal{U}_a} w_u^{\frac{1}{\alpha}} \tilde{R}_{u,a}^{\frac{1}{\alpha}-1} \tag{4.44}$$

This results in the optimal resource proportion being

$$y^*_{u,a} = \frac{w_u^{\frac{1}{\alpha}} \tilde{R}_{u,a}^{\frac{1}{\alpha}-1}}{\sum_{u' \in \mathcal{U}_a} w_{u'}^{\frac{1}{\alpha}} \tilde{R}_{u',a}^{\frac{1}{\alpha}-1}} \qquad \text{for } \alpha > 0. \tag{4.45}$$

In the case of $\alpha = 1$, which corresponds to proportional fairness, $y_{u,a} = \frac{1}{N_a}$, where N_a is the number of users connected to the same AP.

To implement this algorithm, the central controller collects wireless channel statistics from users, assuming these statistics remain constant within an optimization state. Initially, users are allocated to AP based on the highest available link rate. During each iteration, users select a new AP with a probability defined by (4.33). Once an AP is chosen, resources are allocated according to (4.34). This iterative process continues until the algorithm converges. The entire procedure is centrally managed by the controller, which then relays the final decision back to the network upon convergence. An overview of this algorithm can be found in Algorithm 3.

Algorithm 3 EGT-based algorithm for alpha-fair wireless resource allocation

Users send their channel statistics and positions to the controller, which then predicts the future channels.
Initialization: $i = 0$. Assign each user to the AP offering the best link rate. Allocate resources to the connected users using (4.34). Calculate the initial global average payoff $\bar{F}^0$ as in (4.32).
repeat
  for each user u **do**
    The controller calculates the probability p_u^i for each user to switch APs as in (4.33).
    The user u is assigned to the AP that offers the maximum payoff $F_{u,a}$ with this probability. The controller updates the number of users connected to each AP as N_a.
  end for
  for each AP a **do**
    The controller allocates the wireless resources $y_{u,a}$ for each user connected to the AP as in (4.34)
  end for
  $i \leftarrow i + 1$
until no user changes their AP assignment
The controller communicates the final allocation decisions to the APs, which then re-associate the users accordingly.

The complexity of the algorithm scales with the total number of users and APs, described by $O(M^U M^A I)$, where I is the number of iterations needed for convergence. Due to the independence of many operations, parallel processing can be utilized to further reduce complexity. Typically, the algorithm converges within a few tens of iterations in a 10 x 10 m scenario.

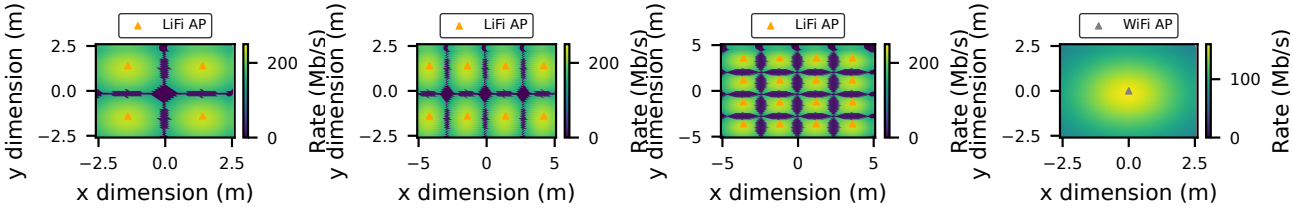
4.2.8 Evaluation Methods and Quality Metrics

Baseline Method for Benchmarking:
We develop a baseline model to benchmark the performance of our optimization framework for wireless resource allocation in LiFi-WiFi networks. User association with APs is determined by selecting the AP that provides the highest link rate, as given by $a = \arg\max_{a'} \tilde{R}_{u,a'}$. This process is executed locally at the user level, eliminating the need for a central controller. After association, the resource proportion $y_{u,a}$ is allocated according to (4.34) and can be managed locally at each AP with knowledge of its associated users. The proposed resource allocation is evaluated using a simulation framework implemented in Python 3.10.6 with Gurobi v11.0.0.

To evaluate the effectiveness of the proposed resource allocation approach, we examine their performance in various indoor environments classified by size: Small, Medium, and Large. Each topology includes a centrally mounted WiFi AP located at coordinates (0,0,3) m, with all LiFi APs also positioned at a height of 3 m. The specific parameters for these scenarios are provided in Table 4.4. The Small topology is the focus of the stability analysis of the evaluations. The effectiveness of these architectures is visualized through data rate coverage maps, as depicted in Figure 4.7, showcasing the reach and efficiency of both LiFi and WiFi across the various scenarios. The WiFi coverage area is shown only for the Large topology since the AP is always positioned at the center of the room across all scenarios.

Table 4.4 Network Topologies under evaluation for the weighted alpha-fair utility maximization

Topology	Dimension	Users	LiFi APs	Position of LiFi APs
Small	5 x 5 x 3 m	10	4	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4)
Medium	10 x 5 x 3 m	15	8	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4), (-1.4, 4.2), (1.4,-4.2), (-1.4,-4.2), (1.4, 4.2)
Large	10 x 10 x 3 m	20	16	(-3.6,-3.6), (-3.6,-1.2), (-3.6, 1.2), (-3.6, 3.6), (-1.2,-3.6), (-1.2,-1.2), (-1.2, 1.2), (-1.2, 3.6), (1.2,-3.6), (1.2,-1.2), (1.2, 1.2), (1.2, 3.6), (3.6,-3.6), (3.6,-1.2), (3.6, 1.2), (3.6, 3.6)



(a) LiFi Coverage for Small architecture **(b)** LiFi Coverage for Medium architecture **(c)** LiFi Coverage for Large architecture **(d)** WiFi Coverage for Large architecture

Figure 4.7 Data rate coverage achieved for various network architectures under evaluation

Performance testing was conducted using an 11th Generation Intel® Core™ i7-11700 16-Core Processor. Each set of results was derived from 40 independent simulation runs, with each run consisting of 240 time steps. To validate our findings, we employed the Mann-Whitney U test [MW47], operating under the null hypothesis that there is no difference in the distributions of the two parameters being compared. The outcomes of these tests are annotated in the corresponding figures with symbols based on the p-values. Furthermore, due to multiple hypothesis testing on the same dataset, we applied the Benjamini-Hochberg procedure [BH95] to adjust for the false discovery rate, ensuring the reliability of our statistical inferences. Simulation results are evaluated using various quality metrics to analyze the performance of the resource allocation framework.

Utility:
This represents the alpha-fair utility objective that the optimization problem aims to maximize

$$\text{Utility} = \begin{cases} \frac{1}{M^U} \sum_u \sum_a \log(R_u) & \text{if } \alpha = 1 \\ \frac{1}{M^U} \sum_u \sum_a \frac{R_u^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1 \end{cases} \tag{4.46}$$

where R_u is the rate achieved by the user u after optimization.

We present "All-time" quality metrics that offer a long-term view of the network performance by considering the quality metric during the entire duration of a user's trajectory.

All-time Average Rate (Mb/s):
This metric is calculated as the average of the user rate across all time steps in the simulation duration.

All-time Minimum User Rate (Mb/s):
This metric is calculated as the worst of the user rates across all time steps in the simulation duration.

All-time Fair-rate:
This metric is calculated as the product of the Jain Fairness index [JCH84] and average rate metrics.

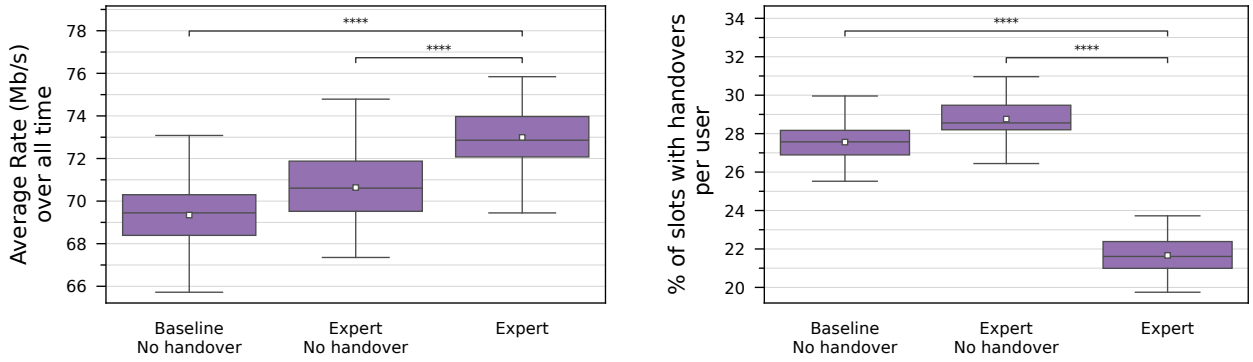
% of Handovers per User:
This metric measures the frequency of handovers, reflecting the percentage of time slots in which a handover occurs for each user.

Time to solve (s):
The efficiency of the optimization algorithms, specifically the Expert, Lagrangian, and EGT solvers, is gauged by the time required to solve the optimization problem.

The evaluations first begin with the analysis of the resource allocation with and without handover considerations for the goal of providing a stable LiFi-WiFi heterogeneous network. The stability analysis is followed by detailed evaluations of weighted alpha-fair resource allocation with a focus on $\alpha = 1$ proportional fairness, for the special case of equal weights assigned to all users. Finally, we end the evaluation with resource allocation where the users are assigned varying weights based on their resource requirements.

4.2.9 Stability Analysis

In all stability analysis simulations, we focus on the Small topology with 10 users, α=1-fairness with equal weights assigned to all users. We evaluate the proposed Expert solver with and without handover considerations and compare it to the other proposed solvers - EGT and Lagrangian. We benchmark evaluations with the baseline algorithm without handover considerations.



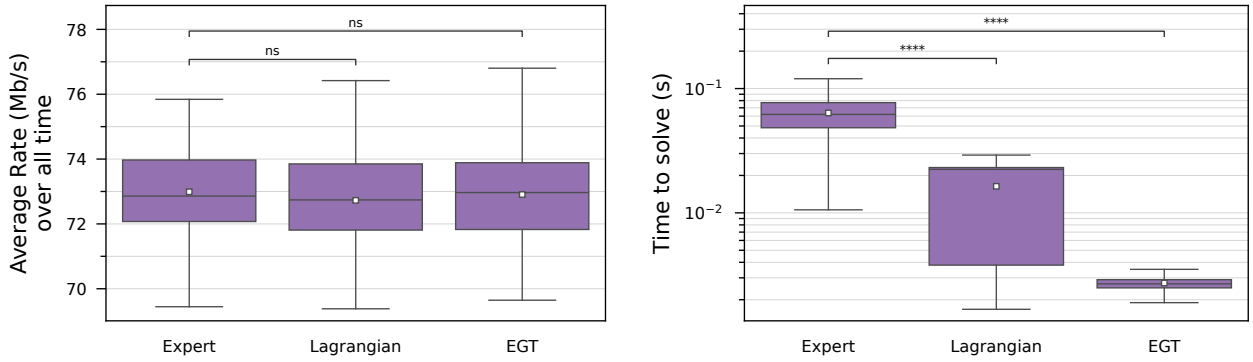
(a) Average rate for user

(b) % of handovers per user

Figure 4.8 Quality metrics for baseline and optimized resource allocation solved with the Expert with and without handover considerations showing the better performance of the Expert in general and even further improvements by considering the loss due to handovers. The significance of the differences are annotated into the figures with the results of the statistical tests.

Figure 4.8 compares the baseline algorithm with the Expert algorithm with and without considering the losses in data rate due to handovers. The Expert solver differs from the Baseline solver in the way that it assigns users to APs and this brings a significant improvement to the average rates of the users. Further gains can be achieved by considering the handover losses which leads to a more accurate prediction of the rates that would be experienced by a user and, hence, a better allocation of resources. Handovers both vertical and horizontal lead to losses in rate and overhead due to additional signalling, with vertical handovers incurring larger losses. Therefore, we look into this

metric for the users in Figure 4.8b under the condition of instantaneous blockages. We observe that the Expert with no handover considerations causes increased handovers because it greedily aims to increase the rate for the users while not being aware of the losses that these handovers would introduce. When the handover efficiency is considered, the Expert significantly reduces the number of handovers by 24.6%, thus leading to a more stable network. In this way, we can ensure stability in link connectivity for users in a heterogeneous network even without the support of multi-homing (or multi-path) devices.

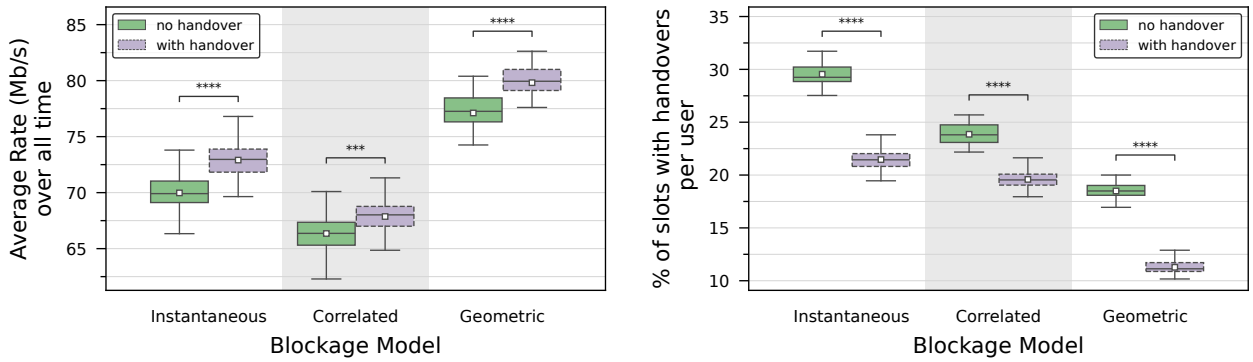


(a) Average rate for user

(b) Time to solve the optimization problem

Figure 4.9 Quality metrics for various proposed solvers showing the comparable network performance of all solvers and the real-time implementation potential of EGT

While the Expert solver has proven its capability in maintaining a stable network, it may not be practical for a real-time implementation due to the time it takes to solve the optimization problem. Therefore, we explore alternative approaches such as the EGT and Lagrangian approaches and compare them to the Expert solver in Figure 4.9. The results show remarkable similarity in the average rate performance of all the solvers with non-significant differences. The EGT solver, additionally, displays the least average time to solve the optimization problem with the lowest variance as well. This highlights the suitability of this solver for a real-world implementation. Therefore, we continue the rest of the evaluations with this solver.



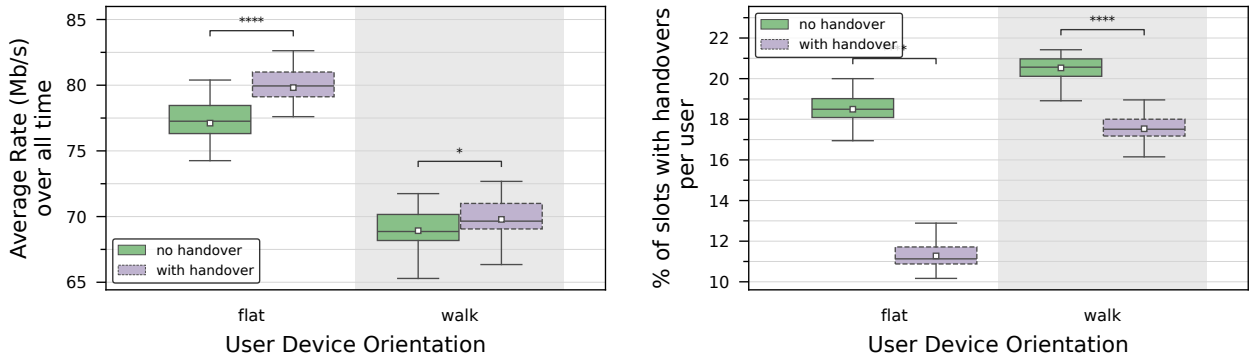
(a) Average rate for user

(b) % of handovers per user

Figure 4.10 Quality metrics for varying blockage models with the EGT solver with and without handover considerations showing the better performance when considering the loss due to handovers across all types of blockages

The previously discussed results were all considering the instantaneous blockage models. To understand the impact of the blockage models on the network performance we compare the instantaneous, correlated and geometric blockage models, as explained in Section 2.4. We observe improvements

in average rates across all blockage models when considering the effects of handover overhead. The instantaneous blockage model shows the highest number of handovers while the blockage model shows the least since it only considers other users as blockages. The correlated model shows the lowest potential for improvement by considering the overhead of handovers since it is the most difficult model for preventing handover. When the blockages last for a continuous period of time as they do in the correlated model, handovers become unavoidable.

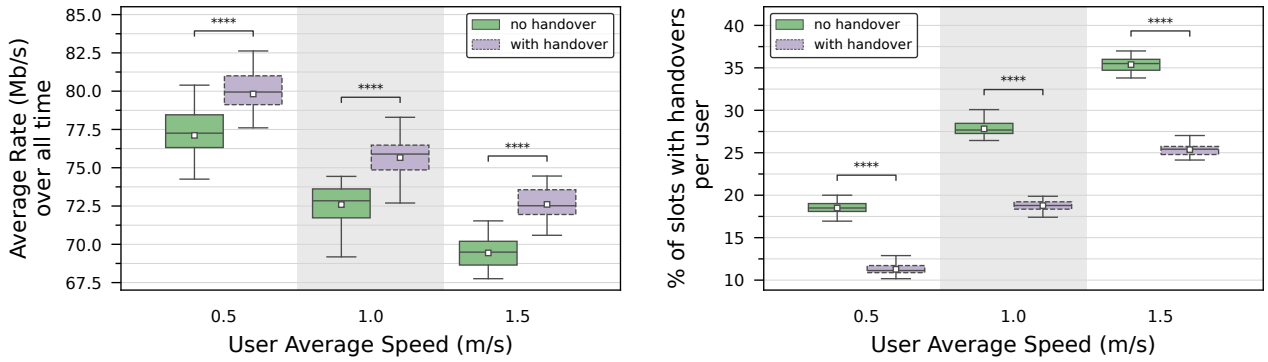


(a) Average rate for user

(b) % of handovers per user

Figure 4.11 Quality metrics for varying user device orientations with the EGT solver with and without handover considerations

Apart from blockages, user device orientation also causes handovers, so we evaluate the stability of the network when no blockages are present and the instability is only caused due to the changing orientation of the user devices. We compare the flat device orientation which is when the user device is parallel to the ceiling facing up, and the walk orientation which is when the user device is tilted at an angle of 28°, a typical orientation for handheld devices. The evaluation in Figure 4.11 shows that although improvements can be achieved, it is more difficult to do so for the walk orientation of the user device since it is more difficult to avoid the handovers that are due to the APs just being out of the Field of View of the user device.



(a) Average rate for user

(b) % of handovers per user

Figure 4.12 Quality metrics for varying user speeds in the RWP mobility model with the EGT solver with and without handover considerations

Another factor that causes instability in the network, is the user movement. So we evaluate the stability of the network when no blockages are present and the instability is only caused due to the changing positions of the users. We compare different speeds of the users following the RWP mobility model in Figure 4.12. With increasing speed, more handovers occur and the potential for reducing the handovers by considering the losses decreases. This decreasing potential is evident by the increasing

gap between the % of handovers with and without handover considerations, as the speed increases. In situations like these, when the handovers cannot be avoided, it becomes even more important to reduce the overhead caused by the handover procedure.

In summary, this analysis shows the necessity of considering the losses associated with handovers to improve network performance in terms of the average rate achieved by users and number of user handovers. We can also conclude that it is not always possible to avoid handovers like scenarios with fast moving users or blockages that extend over time. This brings forth the need to reduce the overhead due to handovers, especially vertical handovers, and we address in the next section.

4.2.10 System Approach to reducing handover overhead

This section presents the proposed system approach to mitigating handover overhead when handovers are unavoidable. Previously, we established that high-speed user movement results in numerous unavoidable handovers. Additionally, the overhead from vertical handovers is significantly higher due to the interface switch [ABG14]. This motivates the need for an efficient interface switching protocol to minimize handover overhead and ensure seamless connectivity, thus maintaining a stable network. Designing such an approach is challenging, particularly in networks where APs are not open for control, which can occur when different technologies have separate operators.

Our approach assigns the same IP address to both LiFi and WiFi interfaces on user devices, altering the routes between the source and destination at the user device and in the downlink to direct traffic through the desired interface. By changing routes at the network layer, the TCP continues to send data to the destination without resetting the connection, ensuring seamless connectivity. Although MPTCP [BPB11] offers a method for seamless connectivity by splitting data flow between two wireless paths and serving the user via both technologies simultaneously, this paper focuses on scenarios where user devices do not support MPTCP. This includes many devices, such as those on-board aircraft, which currently lack MPTCP support. Our goal is to provide an approach that can be readily integrated with all existing user devices.

To manage vertical handovers, a central controller, which can operate on the data server, communicates with agents on the user device and downlink. The agents establish a connection with the controller by exchanging Hello messages at startup. Users are assigned the same IP address on both LiFi and WiFi interfaces. The agents then await the handover command. Upon receiving this command, the agent on the user device selects a new route through a different interface while maintaining the same IP destination. Simultaneously, the downlink switches to the new interface. Because the user's IP address remains unchanged after the switch, downlink data traffic is redirected without requiring a TCP re-connection. The vertical handover protocol is detailed in Algorithm 4.

Algorithm 4 Vertical Handover Protocol

```
Initial configuration: Assign a fixed IP address to one initial interface on the user device.
The controller communicates control information to the agents via an out-of-channel connection.
The user agent and downlink agent establish a connection with the controller.
repeat
    if the user or downlink agent receives a handover command from the controller then
        The agent assigns the fixed IP address to the new interface (LiFi or WiFi) and remove it from the old interface
        The agent updates the routing table, replacing the current route with the desired interface
    else
        The agent continues to wait for a command from the controller
    end if
until termination
```

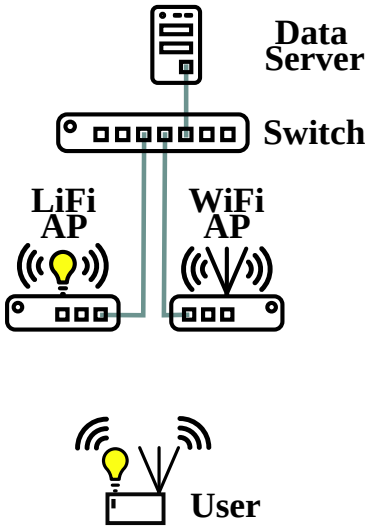


Figure 4.13 Measurement setup with a data server/controller, LiFi and WiFi APs and a user with two network interfaces

The measurement setup for evaluating the proposed protocol is shown in Figure 4.13 where the vertical distance between the APs and the user is 1 m, and the elevation or pitch angle of the user device is 0°.

A TCP iPerf3 server is configured on the user device to continuously receive data. The iPerf3 is set to maintain a constant packet rate with Nagle's algorithm disabled. To gather handover results, the controller enforces a handover every 5 seconds. Wireshark [Wir] is used to analyze the incoming and outgoing packets on each interface. The overhead is measured as the time interval between the last packet on the old interface and the first packet on the new interface following the switch. We name our approach the "Same IP" approach in the results and compare it to a baseline approach where the new interface comes up at the user device when the old interface goes down and the data switches to the new IP address after the Address Resolution Protocols (ARPs) messages are re-sent on the new interface. This baseline approach is named the "Iface Down" approach and represents the default behavior of user devices. This approach, however, leads to TCP disconnections and a failed handover as we have observed in our testbed. Therefore, we also include another approach to vertical handovers that changes the routing table at the user device like the "Change IP" approach but does not transfer the IP address to the new interface. We name this the "Change Route" approach.

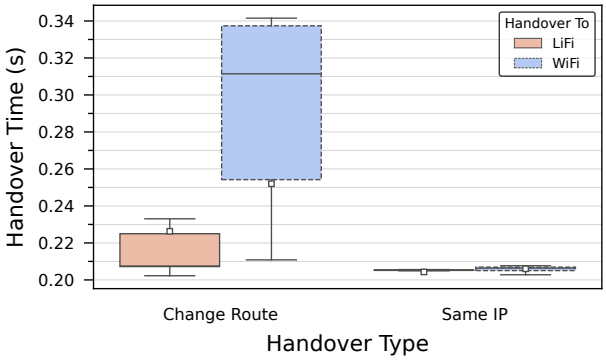


Figure 4.14 Vertical handover overhead comparing our proposed "Same IP" approach and the baseline approach "Change Route" when switching from WiFi to LiFi or vice versa during TCP data transfers

Figure 4.14 presents the vertical handover overhead measured for our proposed "Same IP" approach and the baseline "Change Route" approach for TCP data streams. A total of 1000 measurements were

obtained for both WiFi to LiFi and LiFi to WiFi switches. To ensure the reliability of the data, the same measurements were repeated on multiple days to mitigate any environmental effects.

The results indicate that the overhead is considerably reduced, as TCP re-connections were not observed. When an upper-layer application sends data, the packets are first copied to a kernel write queue. The kernel then transfers the packets from the write queue to the network interface controller, which finally sends them. TCP queues packets to the driver's queue (write buffer), transmitting them sequentially. Each interface, in this case, LiFi and WiFi, has its own queue. During a handover, the interfaces switch, and the new interface's queue is filled. The average of 300 ms of overhead as seen for the "Change Route" approach is the value that we consider in our simulations of handover losses since this approach to vertical handovers can be implemented on user devices without any special requirements from the IP address.

We perform the same analysis for User Datagram Protocol (UDP) data streams which would incur a much lower handover overhead and in this case, the "Iface Down" approach manages to complete the handover at a very high cost. We visualize the results for the vertical handover overhead in Figure 4.15. The overhead time is presented in the log scale due to the large range of values.

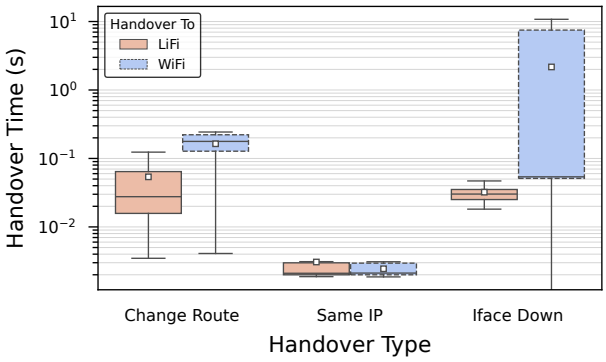


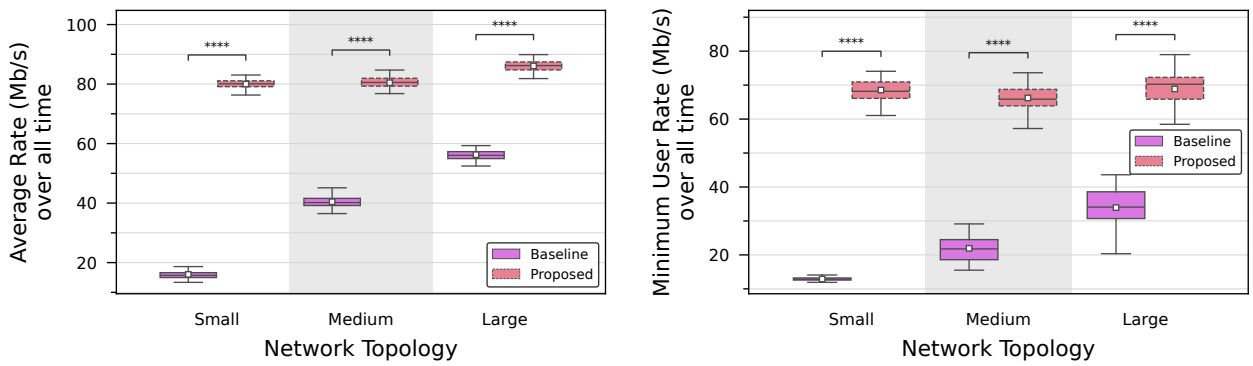
Figure 4.15 Vertical handover overhead comparing our proposed "Same IP" approach and the baseline approach "Change Route" when switching from WiFi to LiFi or vice versa during TCP data transfers

Once again, we see that the proposed "Same IP" approach results in the least overhead while the default "Iface Down" approach takes up to 10 s to complete the handover. In summary, our proposed vertical handover technique does not necessitate TCP re-connections, as opposed to the default "Iface Down" approach, ensuring the stability of a LiFi-WiFi network during unavoidable handovers.

Now that we have established stability in a heterogeneous network, we perform detailed evaluations of the effects of the alpha-fair utility maximization on the network performance. In all the forthcoming evaluations, we include the consideration of handover losses and the blockage model under test is the geometric model where all other interfering users are modeled as cylindrical blockages.

4.2.11 Equally weighted alpha-fair resource allocation

In Section 4.2.9, we have already shown the benefits of an optimized resource allocation over the baseline approach considering an instantaneous blockage model for the Small network topology, here we show the same result for all network topologies with the geometric blockage model in Figure 4.16.

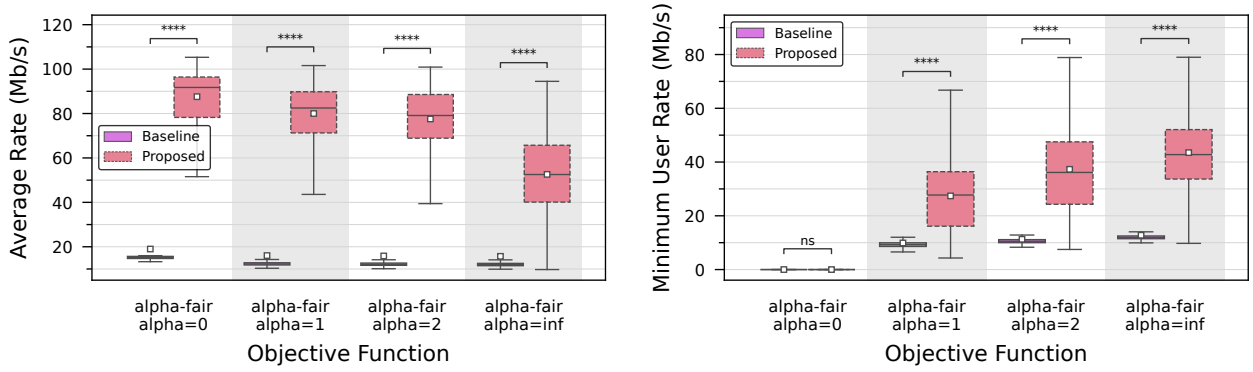


(a) Average rate for user

(b) Minimum Rate for user

Figure 4.16 Quality metrics for baseline and optimized resource allocation solved with Expert for various network topology showing the need for optimization

In LiFi-WiFi networks, the availability of diverse AP association options creates potential for better load balancing that can be achieved by our proposed alpha-fair resource allocation approach. Our proposed model significantly outperforms the baseline, with average user rate increases of 398%, 98.9%, and 52.9% for Small, Medium, and Large topologies, respectively, while the worst user rates improve by 431%, 201%, and 102% for Small, Medium, and Large topologies, respectively.



(a) Average rate for user

(b) Minimum Rate for user

Figure 4.17 Quality metrics for baseline and optimized resource allocation solved with Expert for various alpha-fair objectives

Our analysis turns to the performance implications of various alpha-fair objectives, focusing on average rate and minimum rate for the Small topology as shown in Figure 4.17. The metrics here are per time slot of optimization, since that gives the best picture of how the α parameter can tune the objective of the optimization towards network goals. The results show that the worst user rate increases at the cost of the average user rate as the α value increases. This is in accordance with the goal of the α parameter. In all cases, we see that the proposed approach performs better than the baseline approach. The exception is the minimum user rate when $\alpha = 0$ which is due to the fact that this objective only aims to improve the network sum rate which means that it would allocate all the wireless resources to the user with the best channel quality to utilize the full capacity at the AP.

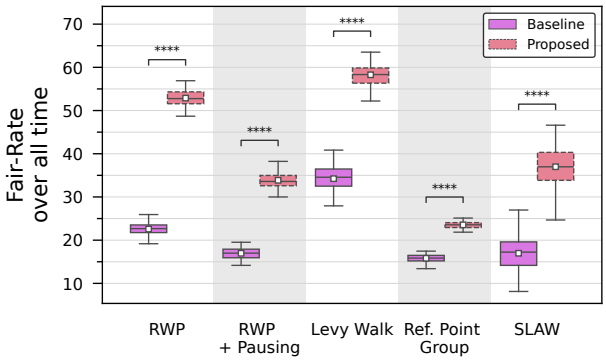
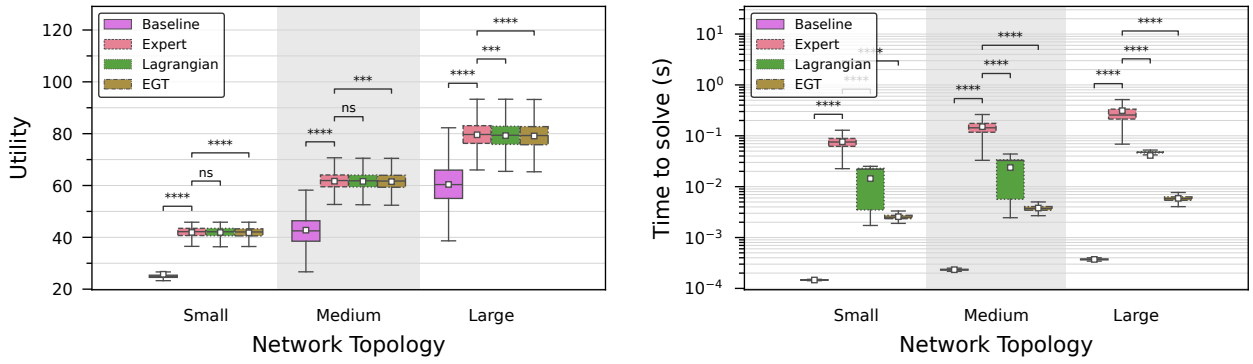


Figure 4.18 Fair-rate metric for baseline and optimized resource allocation solved with Expert for various indoor mobility models under consideration

Since the pattern of user movements highly affects the association of users to APs, we evaluate the performance of our proposed resource allocation approach for different mobility models in the Large topology. For this we compare the fair-rate metric, which combines the user average rate and user rate fairness, in Figure 4.18. Mobility models incorporating pausing periods (e.g., RWP with Pausing) or group movement tendencies (e.g., RPG or SLAW) typically result in lower rates. Pausing intervals cause users to linger in sub-optimal channel conditions for a longer duration. Group movements concentrate resource demands at specific APs, reducing overall network performance. Despite these challenges, the proposed approach consistently outperforms the baseline method, optimizing network performance even under unfavorable conditions.

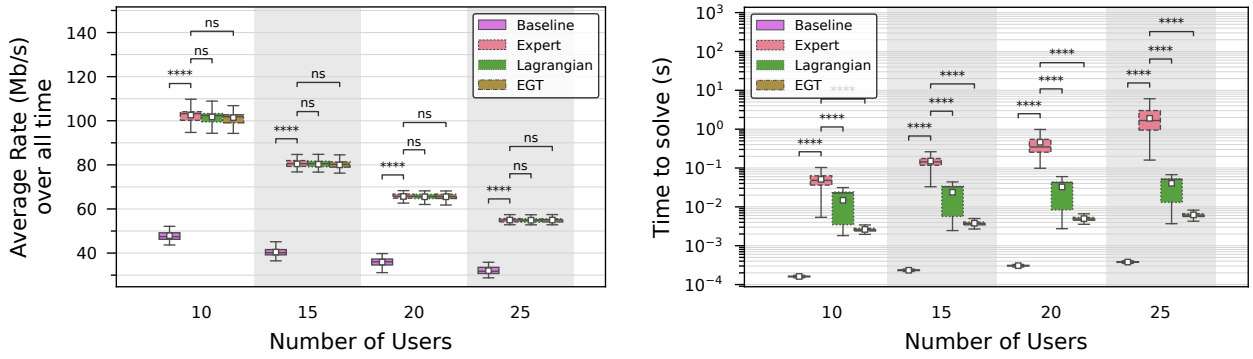


(a) Alpha-fair utility objective

(b) Time to solve the optimization problem

Figure 4.19 Quality metrics for baseline and optimized resource allocation solved with Expert for various network topology showing the need for optimization

Similar to the stability analysis, we still need to find the solver that can reduce the solve time further than the Expert solver while still showing comparable performance. For the Small topology we showed that this is the EGT solver. To confirm that finding for all the network sizes we compare them in Figure 4.19. Instead of the average rate metric, we directly look at the alpha-fair utility objective function that is being optimized as this gives the best comparison between the solvers. While the utility metric shows a statistically significant difference between the Expert and EGT solvers, practically this difference is very small and still shows a considerable improvement over the baseline approach. The solve time of the EGT in the order of a few milliseconds furthers cements it as the best choice for a real-world implementation.



(a) Average rate for user

(b) Time to solve the optimization problem

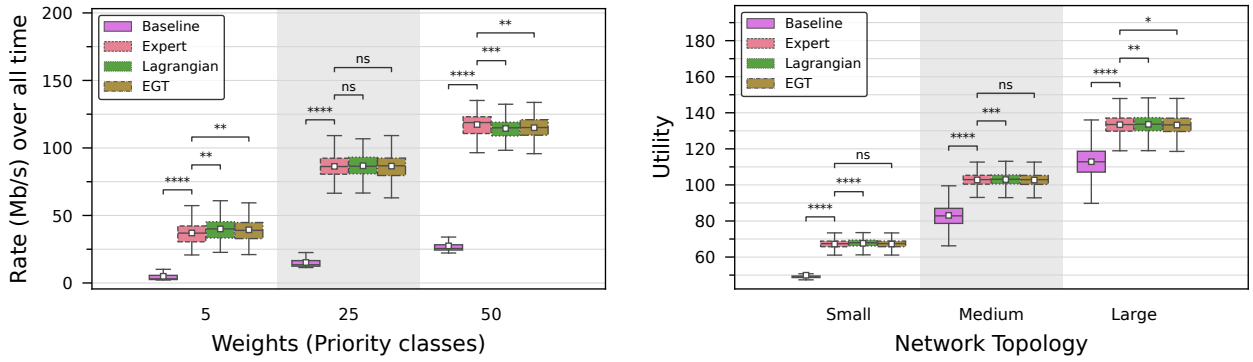
Figure 4.20 Quality metrics for baseline and optimized resource allocation solved with Expert for varying number of users showing the scalability of the proposed solution approaches

To analyze the impact of the network performance of the various proposed solvers when the number of users in the network scales higher, we increase the number of users in a Medium network size from 10-25 and visualize the results in Figure 4.20. The average rate of the users shows no significant different amongst the proposed solvers while showing a significant improvement compared to the baseline approach across all user counts. This confirms that our proposed alpha-fair wireless resource allocation approach can be easily scaled to larger networks while maintaining network optimization benefits. The scalability is also evident in the solve times needed by each approach. The EGT solver, again, displays the least needed time to solve the optimization problem while the Lagrangian and the EGT show less steep slopes in solve time compares to the Expert solver.

Now that we have clearly shown the need for using our proposed wireless resource allocation approach when all users are equally weighted, we turn to analyzing the behavior of the network under varying user weights.

4.2.12 Weighted alpha-fair resource allocation

To accommodate the varying transmission/reception data rate needs of users, the weighted-alpha fair utility integrates different weights or priorities for each user based on their resource demands.



(a) Rate per user priority group

(b) Alpha-fair utility objective

Figure 4.21 Quality metrics for weighted alpha-fair resource allocation with users assigned weights of 5, 25, and 50 with higher weight indicating the demand for more resources

Figure 4.21 shows the results when users are assigned weights of 5, 25, and 50, representing their varying resource requirements with the higher weight indicating a higher demand for data rate.

Figure 4.21a shows an increase in data rate corresponding to user weight, effectively meeting individual user needs. Figure 4.21b displays the network's utility function, indicating that our proposed strategy consistently outperforms the baseline, even under weighted conditions. This finding applies for all proposed solvers. Adjusting user weights allows prioritization based on resource demands and priorities, enabling targeted optimizations that align with various user objectives.

In summary, a stable LiFi-WiFi network is established by considering the losses due to handovers and minimizing the number of handovers that occur in the network. After this, the proposed optimization problem can be solved to optimize the network according to various user needs and network requirements resulting in quality metrics that are better than for the baseline approach. The EGT-based solution approach emerges as best compromise between network performance and solve times suitable for real-world implementations.

4.3 Summary and Conclusions

This chapter has addressed the complex issue of optimizing wireless resource allocation in LiFi-WiFi heterogeneous networks. Our research focused on three primary objectives: minimizing average network packet delay, ensuring network stability, and maximizing alpha-fair utility.

We tackled the problem of wireless resource allocation and user association to minimize average network packet delay, which is crucial for delay-sensitive applications. We explored this optimization in both singlepath and multipath LiFi-WiFi heterogeneous networks. Our key finding from this work is that our proposed methods not only significantly reduce delay compared to existing max-SNR-based techniques but also demonstrate that LiFi-WiFi networks are highly effective for low-latency applications. A key finding is that these networks can achieve sub-millisecond latency, particularly when multi-homing user devices are utilized.

Next, we addressed the stability of heterogeneous networks by incorporating handover losses due to mobility, changes in receiver orientation, and signal blockages into a throughput maximization resource allocation problem. Our proposed approach successfully reduces the number of handovers by 24.6%. To minimize the overhead caused by unavoidable handovers, we proposed a system protocol for managing vertical handovers or interface switches, which typically incur significant overhead. This protocol was implemented on a hardware setup and evaluated through measurements, demonstrating that our approach effectively avoids TCP re-connections.

After establishing network stability, we focused on further improving network performance using a weighted alpha-fair utility maximization model. Our proposed model significantly outperforms the baseline, with average user rate increases of 398%, 98.9%, and 52.9% for Small, Medium, and Large topologies, respectively. Additionally, the worst user rates improved by 431%, 201%, and 102% for Small, Medium, and Large topologies, respectively.

In summary, we have demonstrated the challenges in wireless resource allocation in LiFi-WiFi networks and proposed various resource management approaches that address these challenges and enhance network performance. These resource allocation strategies have proven useful for managing immediate network conditions.

In the next chapter, we will build on these findings to explore the potential for improving long-term network performance. We aim to develop proactive resource allocation strategies that can adapt to changing user behaviors and network environments. This approach will involve the integration of predicted future positions and channels of users, enabling more efficient and dynamic management of network resources over time.

Chapter 5

Mobility-aware Proactive Wireless Resource Allocation in LiFi-WiFi Networks

Resource management in a Light-Fidelity (LiFi)-Wireless-Fidelity (WiFi) heterogeneous network offers unprecedented data transmission capabilities but also faces challenges, as outlined in the previous chapter. Challenges stem from multiple factors including the possibility of assigning from among multiple Access Points (APs) to a user [Wan+17], an increased load on WiFi APs, frequent handovers between densely located LiFi cells [WSH17], and the risk of ping-pong handovers posing a threat to network stability and efficiency. In response to these challenges, strategies for reactive resource allocation were discussed in the previous chapter. These strategies adapt AP assignments and bandwidth share based on current network and user demands [Vij+21; VK21; Ahm+22].

Reactive allocation strategies, while useful for managing immediate network conditions, fail to leverage predictive technologies that could improve efficiency. With advancements in LiFi-based indoor positioning [Arf+21], it is possible to foresee user movements and upcoming network conditions [Bui+17] rather than react upon their occurrence. Proactive resource allocation, thus emerges as important especially in scenarios with LiFi blockages or reduced channel quality due to user mobility. By predicting the trajectory of user movements, it becomes possible to optimize resource allocation, aligning the bulk of data transfers with better channel conditions, and reducing service interruptions, thereby maintaining consistent service quality.

This chapter introduces a proactive resource allocation framework for LiFi-WiFi networks that considers rate fairness, user mobility, Line-of-Sight (LoS) blockages, and different models for potential errors in the prediction of user positions. The content of this chapter is based on the journal contribution [VK24].

Proactive allocation enhances user satisfaction and presents as a vital area of study in heterogeneous network environments. Nevertheless, it is important to recognize the limitations due to inaccuracies in predicting user location, which could diminish the benefits of proactive strategies. This recognition drives the need to explore various models of prediction errors and assess the effectiveness of proactive resource allocation amidst these uncertainties.

Through these discussions, the chapter sets the stage for a deep dive into the architecture and methodology of MobiFi, the proposed proactive wireless resource management framework that represents a significant step towards fully autonomous LiFi-WiFi heterogeneous networks that can adapt to user behaviors and environmental changes.

5.1 State-of-the-art Analysis

The body of research on wireless resource allocation in LiFi and Radio Frequency (RF) networks has made substantial strides, particularly in anticipatory mobile networking and proactive resource allocation. Nevertheless, challenges specific to LiFi-WiFi networks remain inadequately addressed.

A comprehensive survey by [Bui+17] examines prediction and optimization within anticipatory mobile networking but fails to address the unique issues of LiFi-WiFi networks, such as blockages and AP density. This gap shows the need for targeted research in this area.

In the field of indoor LiFi systems, [Arf+24] focuses on proactive optimization using deep learning, specifically considering channel aging. The work introduces a Long Short-Term Memory (LSTM) network to predict the positions, orientations, and channel coefficients of mobile users, allowing for the optimal precoding matrix for sum rate maximization. However, this research is limited to beamforming, while our work also includes AP assignment and wireless resource allocation. Their primary focus was on prediction accuracy and time which leaves a gap in understanding its overall impact on resource allocation.

In the Visible Light Communication (VLC) domain, [Zha+18] explores anticipatory association for user-to-AP connections, utilizing predicted user locations and traffic dynamics to prepare APs in advance. However, it does not address proactive wireless resource allocation for mobility awareness. Our research emphasizes both wireless resource allocation and user-to-AP association, essential for load balancing in heterogeneous networks.

Reference [Das+18] introduces mobility-aware resource optimization using a look-ahead policy in VLC networks, proposing a resource allocation algorithm based on anticipated Channel State Information (CSI) from predicted user locations. However, it does not consider the impact of prediction errors on resource allocation. Our work addresses this by modeling prediction errors and analyzing their effects on proactive resource allocation.

Several works have focused on proactive optimization in RF-only heterogeneous networks. [CDS18] develops a deep learning method for resource allocation in Long Term Evolution (LTE)-License Assisted Access (LAA) small base stations, ensuring fair coexistence with WiFi. This method relies on traffic history to predict resource allocation actions but does not account for predicted channel conditions, crucial for optimizing short-range LiFi networks susceptible to blockages. Our approach includes predicted channel conditions to enhance long-term network performance.

Another important contribution is from [Li+20], which introduces a dynamic, proactive resource allocation approach using deep reinforcement learning in heterogeneous networks. While the work emphasizes robustness, it focuses on traditional RF heterogeneous networks. Our research complements this by proposing resource allocation strategies in LiFi-WiFi networks, addressing prediction uncertainties and improving network performance across different technologies. Unlike [Li+20], which optimizes wireless resource allocation with fixed AP association, our approach optimizes AP association for maximum benefits.

In the context of LiFi-WiFi network management, both [Fró+24] and our work incorporate user trajectory predictions for effective handover strategies. They optimize handovers by selecting the best LiFi or WiFi AP or maintaining the current connection based on future channel conditions predicted using reinforcement learning, aiming to maximize user rate satisfaction. [TQS23] also addresses proactive handovers based on blockage prediction using beam tracking. However, our research extends beyond single AP choices, considering all potential allocations, and integrates wireless resource allocation with AP selection. Additionally, our optimization extends to future time blocks rather than just current time slot allocations.

In conclusion, while existing research has significantly advanced proactive handovers and mobility-aware networking, our proposal distinguishes itself by presenting a unified resource allocation framework tailored to LiFi-WiFi heterogeneous networks. By addressing anticipatory allocation, incorporating predictive modeling, and utilizing advanced optimization techniques, our work makes a substantial contribution to the evolving field of heterogeneous network optimization.

5.2 Key Contributions

We introduce the architecture and methodology of MobiFi, a proactive wireless resource management framework for LiFi-WiFi heterogeneous networks. Specifically this chapter makes the following significant contributions:

1. We develop a proactive resource allocation framework that specifically considers mobility of users, LoS blockages, and variations in channel quality. This approach is designed to proactively manage wireless resources in both LiFi-only and LiFi-WiFi networks.

2. We propose tackling the challenge of proactive resource allocation by setting up an optimization problem that assigns AP and allocates wireless resources using an alpha-fairness objective over time. This approach not only optimizes network efficiency but also consistently outperforms traditional reactive models in terms of average data rates.

3. To solve this complex optimization problems, we utilize advanced solution techniques, including a Branch and Bound-based Mixed Integer Nonlinear Programming (MINLP) solver and iterative Evolutionary Game Theory (EGT)-based algorithms.

4. We incorporate models that simulate errors in predicting user positions to test the robustness of our proactive allocation strategy. This simulation is necessary to evaluate how well the proactive strategy performs under the uncertainty of user position predictions, ensuring the reliability of the resource management processes.

5.3 System Model

This work explores resource management in a LiFi-WiFi heterogeneous network comprising M^L LiFi APs and M^W WiFi APs, with a combined total of M^A APs. The LiFi APs, which are Light Emitting Diodes (LEDs) mounted at a height of 3 m, operate at the same frequency, leading to co-channel interference in overlapping areas. Conversely, WiFi APs operate on distinct, non-overlapping frequencies.

The network serves a total of M^U users, each equipped with both LiFi photodiode and WiFi receivers, although connectivity to only one AP is possible at any time. All APs are connected to a central network controller that facilitates instantaneous, error-free communication and monitors the load on each AP by collecting and analyzing the wireless CSI from all users.

Resource allocation decisions are made by a centralized algorithm that operates at intervals denoted by τ. This algorithm uses global network data to allocate resources and communicate these decisions back to the users. Additionally, a positioning system captures user locations, which the controller uses to predict future movements and channel conditions over a defined period. Users' devices typically come equipped with GPS receivers that can share location data with the centralized controller. Alternatively, an AP-based positioning system can also be used, requiring the user to be within the coverage area of multiple APs simultaneously. While our network design approach in Chapter 3 does not incorporate this consideration, the resource allocation method discussed in this chapter remains independent of the source of user position data. We assume a proactive resource allocation window spanning five service time slots ($T^{\text{win}} = 5$). This setup means the resource allocation algorithm is executed once at the beginning of this window, managing network resources across the specified slots.

The channel model and geometric blockage model for LiFi are elaborated in Section 2.3 and Section 2.4.3, respectively, while the channel model for WiFi is described in Section 2.6. To evaluate the performance of our proactive resource allocation strategy with mobile users, we include four mobility models: Random Waypoint (RWP), Truncated Levy Walk, Self-Similar Least Walk (SLAW) [Lee+09], and Reference Point Group (RPG) [Hon+99] which also capture the changing orientation of

a user device. Further details are provided in Section 2.7. The user experiences losses due to vertical (inter-technology) and horizontal (intra-technology) handovers that occur when the user changes its AP association from one time slot τ=500 ms to the next. Further details are provided in Section 4.2.4.

Table 5.1 provides a summary of the notations used in this chapter, where bold lowercase letters represent vectors and cursive capital letters denote sets.

Table 5.1 List of Notations used in Proactive Resource Allocation

Notation	Description
$l, M^L, \mathcal{L}$	Index, number, and set of LiFi APs
$w, M^W, \mathcal{W}$	Index, number, and set of WiFi APs
$a, M^A, \mathcal{A}$	Index, number, and set of total APs
$u, M^U, \mathcal{U}$	Index, number, and set of users
$N_a, \mathcal{U}_a$	Number and set of users associated to AP a
t, T^{win}	Index and length of Optimization/Future window
$\boldsymbol{p}(t), \hat{\boldsymbol{p}}(t)$	Real and predicted (x, y, z) user position
$\boldsymbol{\Theta}(t), \hat{\boldsymbol{\Theta}}(t)$	Real and predicted (yaw, pitch, roll) user orientation
σ	Standard deviation of error in user prediction
τ	Length of one time slot
$\eta_{a',a}$	Handover efficiency when switching from AP a' to a
$R^{(t)}_{u,a}$	Achievable rate of user u connected to AP a at time t
$\tilde{R}^{(t)}_{u,a}$	Handover-reduced rate for user u to AP a at time t
$R^{(t)}_u, R^{\text{win}}_u$	Achieved rate of user u at time slot t and in a window
$f_\alpha()$	Alpha-fairness function
α	Parameter to control the alpha-fairness
$x^{(t)}_{u,a}$	Optimization variable indicating association
$y^{(t)}_{u,a}$	Optimization variable indicating resource proportion
$U_\alpha(x^{(t)}_{u,a}, y^{(t)}_{u,a})$	Utility function to be maximized
i	Index of iteration of the EGT algorithm
$F_{u,\boldsymbol{a}}$	Payoff for user u to APs $\boldsymbol{a}$ over time
$\bar{F}^i$	Global average Payoff in i^{th} iteration
p^i_u	Mutation probability of user u
L, λ	Lagrangian function and multiplier
w_u	Weights assigned to user for a weighted alpha-fairness

5.3.1 Models for Errors in User Position Prediction

The accuracy of user position predictions is essential for effective proactive resource allocation. To assess our strategy's resilience to prediction inaccuracies, we use several models to simulate prediction errors.

Known Position Model

The *known* position model assumes perfect knowledge of users' future locations, representing an ideal scenario with zero prediction error. This model serves as a benchmark for evaluating the performance of more realistic error models.

Assumed Position Model

The *assumed* position model adopts a basic approach wherein only the initial location of the user is known, with the presumption that the user remains stationary throughout the prediction period. If the user's initial position at time t_0 is $\boldsymbol{p}(t_0)$, then the predicted position $\hat{\boldsymbol{p}}(t)$ for any subsequent time t within the prediction window is expressed as:

$$\hat{\boldsymbol{p}}(t) = \boldsymbol{p}(t_0), \quad \forall t > t_0$$

This assumption also applies to the orientation of the user's device, represented as:

$$\hat{\boldsymbol{\Theta}}(t) = \boldsymbol{\Theta}(t_0), \quad \forall t > t_0$$

Error Model

The *error* model incorporates prediction errors based on the statistical characteristics of typical errors. **Drift Error Model**: The *drift* error model introduces prediction errors that accumulate and drift over time, based on previous predictions. For a 2D position vector $\boldsymbol{p}(x, y, t)$ at time t, the error is modeled as a normal distribution where the mean is the error at time $t-1$ and the standard deviation remains constant. The mathematical representation is as follows:

$$\Delta \boldsymbol{p}_t = \Delta \boldsymbol{p}_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(\Delta \boldsymbol{p}_{t-1}, \sigma)$$

$$\hat{\boldsymbol{p}}(x, y, t) = \boldsymbol{p}(x, y, t) + \Delta \boldsymbol{p}_t$$

where $\Delta \boldsymbol{p}_t$ is the error at time t with $\Delta \boldsymbol{p}_0 = 0$, ϵ_t is a random error drawn from a normal distribution with a standard deviation σ, and $\hat{\boldsymbol{p}}(x, y, t)$ is the predicted (x,y) position at time t and $\boldsymbol{p}(x, y, t)$ is the actual (x,y) position at time t.

Standard Deviation Error Model: The *std* error model generates prediction errors where the standard deviation of the error increasing over time. Each time step's error is independent and normally distributed, with a mean of zero and a standard deviation that increases with the time step. The error model is given by:

$$\hat{\boldsymbol{p}}(x, y, t) = \boldsymbol{p}(x, y, t) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma \frac{t}{T^{\text{win}}})$$

where T^{win} is the total duration of the prediction window, and σ is the scaling factor for the standard deviation.

In both *error* models, the predicted (x,y) positions are utilized to determine the direction of movement, which provides the predicted yaw angle of orientation. The z position and the pitch and roll orientations of the user device are assumed to remain constant throughout the prediction window, using the values from the initial time step within the window.

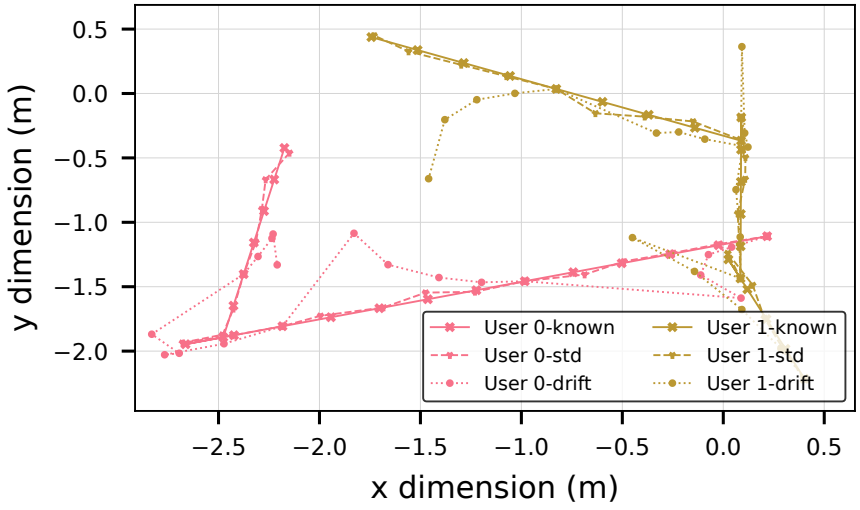


Figure 5.1 Real (or known) and predicted trajectories for 2 exemplary users following the RWP mobility model for the *drift* and *std* error models

Figure 5.1 shows the impact of prediction errors on user trajectories. The ground truth for user movements is represented by the known trajectory, which is derived from the RWP model. The figure also depicts the trajectories generated by two prediction error models: the *drift* error model and the *std* error model.

5.4 Optimization Problem Formulation for the Proactive Resource Allocation

The proactive resource allocation strategy for LiFi-WiFi heterogeneous networks incorporates a future time dimension to predict user positions and network conditions. This approach seeks to optimize the allocation of wireless resources over a specified time window, aiming to maximize the alpha-fairness utility function, as detailed in Section 4.2.5 and based on the principles outlined in [MW00].

The optimization problem is formulated as follows:

$$\max_{x_{u,a}^{(t)}, y_{u,a}^{(t)}} \quad U_\alpha(x_{u,a}^{(t)}, y_{u,a}^{(t)}) \tag{5.1}$$

$$\text{subject to} \quad \sum_a x_{u,a}^{(t)} = 1 \qquad \forall u, t \tag{5.2}$$

$$\sum_u x_{u,a}^{(t)} y_{u,a}^{(t)} \leq 1 \qquad \forall a, t \tag{5.3}$$

$$x_{u,a}^{(t)} \in \{0, 1\} \qquad \forall u, a, t \tag{5.4}$$

$$0 < y_{u,a}^{(t)} \leq 1 \qquad \forall u, a, t \tag{5.5}$$

where $x_{u,a}^{(t)}$ is the binary association variable that is 1 if the user u is associated to AP a at time slot t and $y_{u,a}^{(t)}$ is the continuous resource proportion variable that denotes the fraction of wireless resources allocated to the user. Constraint (5.2) requires that each user is associated to a single AP per time slot. Constraint (5.3) prevents the resource load at any AP from exceeding its capacity.

The utility function of alpha-fairness U_α considering different values of α is:

$$U_\alpha(x_{u,a}^{(t)}, y_{u,a}^{(t)}) = \begin{cases} \sum_u \sum_t \sum_a x_{u,a}^{(t)} \log\left(y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}\right) & \text{if } \alpha = 1 \\ \sum_u \sum_t \sum_a x_{u,a}^{(t)} \frac{\left(y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}\right)^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1 \end{cases} \tag{5.6}$$

Here, $\tilde{R}_{u,a}^{(t)}$ represents the predicted link rate from AP a to user u at time t. This rate accounts for handover efficiency, denoted as $\eta_{a',a}^{(t)}$, which affects the handover-reduced rate. The value of $\eta_{a',a}^{(t)}$ is known at $t = 0$ within the window, and for subsequent slots, it depends on the optimization variable of the user-AP association from the previous time slot, $t - 1$.

$$\tilde{R}_{u,a}^{(t)} = R_{u,a}^{(t)} \sum_{a'} x_{u,a'}^{(t-1)} \eta_{a',a} \tag{5.7}$$

Therefore, the rate $\tilde{R}_{u,a}^{(t)}$ for $t \geq 2$ is treated as an auxiliary optimization variable. The alpha value of 1 represents a proportional fairness objective among the users and has a special form. Each of the other alpha values represent varying degrees of fairness among the user achieved rates.

This formulation is applicable to both LiFi-only and LiFi-WiFi networks. The difference lies in the source set for user association: $\mathcal{L}$ for LiFi-only networks and $\mathcal{A}$ for LiFi-WiFi networks. Additionally, handover efficiency considerations vary, with potential vertical handovers in LiFi-WiFi networks. Both optimization formulations result in MINLP problems due to the presence of binary and continuous variables, rendering them NP-hard and mathematically intractable.

5.4.1 Convex relaxation

To convert the problem into a convex form, the binary variable $x_{u,a}^{(t)}$ can be relaxed to take fractional values in the range $[0, 1]$. This relaxation transforms the problem formulation as follows:

$$\max_{x_{u,a}^{(t)}, y_{u,a}^{(t)}} \quad \sum_u \sum_t \sum_a x_{u,a}^{(t)} \frac{\left(y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}\right)^{1-\alpha}}{1-\alpha} \tag{5.8}$$

$$\text{s.t.} \quad \tilde{R}_{u,a}^{(t)} = R_{u,a}^{(t)} \sum_{a'} x_{u,a'}^{(t-1)} \eta_{a',a} \qquad \forall u, a, t \geq 2 \tag{5.9}$$

$$\sum_a x_{u,a}^{(t)} = 1 \qquad \forall u, t \tag{5.10}$$

$$\sum_u x_{u,a}^{(t)} y_{u,a}^{(t)} \leq 1 \qquad \forall a, t \tag{5.11}$$

$$x_{u,a}^{(t)} \in [0, 1],\ y_{u,a}^{(t)} \in (0, 1] \qquad \forall u, a, t \tag{5.12}$$

This relaxed problem can be proven to be a concave optimization problem for $\alpha > 0$ since the objective is concave and the constraints are affine. Let the objective function be denoted as $F\left(x_{u,a}^{(t)}, y_{u,a}^{(t)}, R_{u,a}^{(t)}\right)$ which is a function of three variables. The Hessian matrix is written as

$$\mathcal{H}(F) = \begin{bmatrix} \frac{\partial^2 F}{\partial (x_{u,a}^{(t)})^2} & \frac{\partial^2 F}{\partial x_{u,a}^{(t)} \partial y_{u,a}^{(t)}} & \frac{\partial^2 F}{\partial x_{u,a}^{(t)} \partial R_{u,a}^{(t)}} \\ \frac{\partial^2 F}{\partial y_{u,a}^{(t)} \partial x_{u,a}^{(t)}} & \frac{\partial^2 F}{\partial (y_{u,a}^{(t)})^2} & \frac{\partial^2 F}{\partial y_{u,a}^{(t)} \partial R_{u,a}^{(t)}} \\ \frac{\partial^2 F}{\partial R_{u,a}^{(t)} \partial x_{u,a}^{(t)}} & \frac{\partial^2 F}{\partial R_{u,a}^{(t)} \partial y_{u,a}^{(t)}} & \frac{\partial^2 F}{\partial (R_{u,a}^{(t)})^2} \end{bmatrix} \tag{5.13}$$

where

$$\frac{\partial^2 F}{\partial x_{u,a}^{(t)^2}} = 0 \tag{5.14}$$

$$\frac{\partial^2 F}{\partial y_{u,a}^{(t)^2}} = -\alpha \frac{x_{u,a}^{(t)} R_{u,a}^{(t)^2}}{\left(y_{u,a}^{(t)} R_{u,a}^{(t)}\right)^{\alpha+1}} \leq 0 \tag{5.15}$$

$$\frac{\partial^2 F}{\partial R_{u,a}^{(t)^2}} = 0, \qquad t = 1 \tag{5.16}$$

$$\frac{\partial^2 F}{\partial R_{u,a}^{(t)^2}} = -\alpha \frac{x_{u,a}^{(t)} y_{u,a}^{(t)^2}}{\left(y_{u,a}^{(t)} R_{u,a}^{(t)}\right)^{\alpha+1}} \leq 0, \qquad \forall t > 1 \tag{5.17}$$

Since the leading principal minors are non-positive, the Hessian matrix is negative semi-definite, rendering the objective function concave with respect to the variables $x_{u,a}^{(t)}$, $y_{u,a}^{(t)}$, and $R_{u,a}^{(t)}$.

Consequently, this relaxed problem can be solved using any convex optimization technique, such as the Lagrangian approach. The binary variables can then be recovered by associating the user with the AP that has the highest value for the binary variable at each time step. However, in this chapter, we employ the off-the-shelf solver Gurobi [Gur23], which utilizes the Branch and Bound algorithm [LD10] and approximates non-linear functions with piecewise linear functions to solve the problem. The Gurobi solver implementation is referred to as the “Expert" model from here on. Additionally, to reduce computational complexity and ensure solutions within the specified optimization interval τ, we implement an algorithm using EGT.

5.5 Proposed Resource Allocation with Evolutionary Game Theory

EGT models user interactions as a competitive game for the shared wireless resources. Evolutionary games are advantageous for distributed implementations but, in this work, the algorithm is executed at the central controller. This approach ensures that the controller maintains a global understanding of the entire network and can communicate decisions to users after the algorithm's convergence. The algorithm converges to an evolutionary equilibrium, representing a stable state achieved by the users.

5.5.1 Algorithm Setup

The association of users to APs is managed by an EGT-based algorithm, which includes the following key components:

- *Players:* The users participating in the network.
- *Population:* The population $\mathcal{U}_a$ represents the set of users associated to an AP a at each time slot, where the number of users assigned to AP a is denoted as N_a.
- *Strategy:* Players can choose one among all the LiFi and WiFi APs available to serve them in each time slot within the window of optimization.
- *Payoff:* Players adjust their strategy in each iteration to maximize their payoff. The payoff for a user u is an aggregation of achievable rates $\tilde{R}_{u,a}^{(t)}$ over the window. Given an association vector, $\boldsymbol{a}$ containing the user u's associated AP for all time steps t within the window, the payoff is calculated as

$$F_{u,\boldsymbol{a}} = \sum_{t} y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}. \tag{5.18}$$

5.5.2 Access Point Association

Based on the previously given game setup, the algorithm identifies the AP association that maximizes the payoff for all users. Each user adjusts their AP association strategy by considering the global average payoff of all players. The global average payoff of users connected to APs $\boldsymbol{a}$ in the i^{th} iteration is calculated as:

$$\bar{F}^i = \frac{1}{M^U} \sum_{u} F_{u,\boldsymbol{a}}. \tag{5.19}$$

where M^U is the total number of users in the system. Users with lower payoff values compared to the global average have a higher probability of switching their association to another AP (mutation). This switching probability is determined by:

$$p_u^i = \begin{cases} 1 - \frac{F_{u,\boldsymbol{a}}{}^{(i-1)}}{\bar{F}^{(i-1)}} & F_{u,\boldsymbol{a}}^{(i-1)} < \bar{F}^{(i-1)} \\ 0 & \text{otherwise.} \end{cases} \tag{5.20}$$

When a user mutates, the new AP to be associated to is chosen by maximizing the estimated payoff for the current iteration, i.e., $\boldsymbol{a} = \arg\max_{\boldsymbol{a}'} F_{u,\boldsymbol{a}'}$. This involves evaluating each AP for every time slot t within the window as a potential choice for the user, and then choosing the optimal allocation $\boldsymbol{a}$.

5.5.3 Resource Allocation

After the AP association is completed, the AP allocates wireless resources to the users based on the alpha-fairness rule. For a given AP assignment, the optimal resource proportion $y_{u,a}^{(t)}$ for a user u associated with AP a at time slot t is determined by:

$$y_{u,a}^{(t)} = \begin{cases} \begin{cases} 1 & \text{if } u = \arg\max_{u'} \tilde{R}_{u',a}^{(t)} \\ 0 & \text{elsewhere} \end{cases} & \text{if } \alpha = 0 \\ \dfrac{\left(\tilde{R}_{u,a}^{(t)}\right)^{\frac{1}{\alpha}-1}}{\sum_{u' \in \mathcal{U}_a} \left(\tilde{R}_{u',a}^{(t)}\right)^{\frac{1}{\alpha}-1}} & \text{if } \alpha > 0 \end{cases} \tag{5.21}$$

The proactive strategy allocates resources according to (5.21), independently considering each time slot in the window. Although resource allocation is carried out independently for each time slot, the AP association process takes the entire window into account when calculating payoffs and reassigning users.

5.5.4 Optimality of alpha-fair resource allocation strategy

For $\alpha > 0$, it can be shown that selecting the resource proportion $y_{u,a}^{(t)}$ as in (5.21) is optimal for a given AP assignment. When the AP assignment is fixed, the resource proportion allocation does not depend on the time slot and can be separated into T^{win} different sub-problems, allowing the time index (t) to be omitted. Consequently, resources are allocated independently in each slot. Therefore, the proof follows exactly as previously explained in Section 4.2.7 for the reactive strategy by setting all user weights w_u to 1.

5.5.5 Algorithm Implementation

The controller collects wireless channel statistics from users, assuming these statistics remain constant within an optimization state. Initially, users are allocated to APs based on the highest available link rate. During each iteration, users select a new AP with a probability defined by (5.20). Once an AP is chosen, resources are allocated according to (5.21). This iterative process continues until the algorithm converges. The entire procedure is centrally managed by the controller, which then relays the final decision back to the network upon convergence. An overview of this algorithm can be found in Algorithm 5.

Algorithm 5 EGT-based Algorithm for Proactive Wireless Resource Allocation

1: Users send their channel statistics and positions to the controller, which then predicts the future channels.
2: Initialization: $i = 0$. Assign each user to the AP vector $\boldsymbol{a}$ offering the best link rate. Allocate resources to the connected users using (5.21). Calculate the initial global average payoff $\bar{F}^0$ as in (5.19).
3: **repeat**
4: **for each** user u **do**
5: The controller calculates the probability p_u^i for each user to switch APs as in (5.20).
6: The user u is assigned to the set of APs that offers the maximum payoff $F_{u,\boldsymbol{a}}$ with this probability. The controller updates the number of users connected to each AP as N_a.
7: **end for**
8: **for each** AP vector $\boldsymbol{a}$ **do**
9: The controller allocates the wireless resources $y_{u,a}^{(t)}$ for each user connected to the AP as in (5.21).
10: **end for**
11: $i \leftarrow i + 1$
12: **until** no user changes their AP assignment
13: The controller communicates the final allocation decisions to the APs, which then re-associate the users accordingly.

The complexity of the algorithm scales with the total number of users, APs, and window length, described by $O(M^U M^A T^{\text{win}} I)$, where I is the number of iterations needed for convergence. Due to the independence of many operations, parallel processing can be utilized to further reduce complexity. Typically, the algorithm converges within a few tens of iterations in a 10 x 10 m scenario.

5.6 Evaluation Methods and Quality Metrics

Our proposed MobiFi proactive resource allocation is evaluated through a simulation framework implemented in Python 3.10.6 with Gurobi v11.0.0. To benchmark the performance of our proactive resource allocation framework, we compare all the results to the reactive approach introduced in the previous chapter Section 4.2. To recollect, the reactive approach manages the resources of a LiFi-WiFi network through an alpha-fair utility optimization problem that considers the network slot-by-slot with no knowledge of the users' positions or channel conditions in the future.

Table 5.2 Network Topologies under evaluation

Topology	Dimension	Users	LiFi APs	Position of LiFi APs
Small	5 x 5 x 3 m	10	4	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4)
Medium	10 x 5 x 3 m	15	8	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4), (-1.4, 4.2), (1.4,-4.2), (-1.4,-4.2), (1.4, 4.2)
Large	10 x 10 x 3 m	20	16	(-3.6,-3.6), (-3.6,-1.2), (-3.6, 1.2), (-3.6, 3.6), (-1.2,-3.6), (-1.2,-1.2), (-1.2, 1.2), (-1.2, 3.6), (1.2,-3.6), (1.2,-1.2), (1.2, 1.2), (1.2, 3.6), (3.6,-3.6), (3.6,-1.2), (3.6, 1.2), (3.6, 3.6)

To assess the efficacy of the algorithms we introduced, we analyzed their performance across a range of indoor environments categorized by size — Small, Medium, and Large — which correspond to different network topologies as already introduced in the previous chapter. To recall, each topology features one WiFi AP mounted centrally on the ceiling at coordinates (0,0,3) m, with all LiFi APs also positioned at the same height of 3 m. Additional parameters relevant to these scenarios are detailed in Table 5.2. Performance testing was conducted using an 11th Generation Intel® Core™ i7-11700 16-Core Processor. Each set of results was derived from 40 independent simulation runs, with each run consisting of 240 time steps. To validate our findings, we employed the Mann-Whitney U test [MW47], operating under the null hypothesis that there is no difference in the distributions of the two parameters being compared. As in the previous chapters, the outcomes of these tests are annotated in the corresponding figures with symbols based on the p-values [Cha+22]. To recollect, the star notation is given as:

$$\begin{aligned} \text{ns} &: p > .05 \\ * &: .01 < p <= .05 \\ ** &: .001 < p <= .01 \\ *** &: .0001 < p <= .001 \\ **** &: p <= .0001 \end{aligned}$$

Furthermore, due to multiple hypothesis testing on the same dataset, we applied the Benjamini-Hochberg procedure [BH95] to adjust for the false discovery rate, ensuring the reliability of our statistical inferences.

Simulation results are evaluated using various quality metrics to analyze the performance of the resource allocation framework. These metrics are grouped into two categories: windowed evaluations and all-time evaluations.

Each windowed metric involves calculating the sum of rates for each user over the time window T^{win}, termed as the windowed rate per user $R_u^{\text{win}} = \sum_{t=1}^{T^{\text{win}}} R_u^{(t)}$. Each window yields a single metric value.

Windowed Fairness:
The Jain fairness index [JCH84], a measure of resource allocation fairness among users, is calculated for each window as follows:

$$\text{Jain Fairness} = \frac{\left(\frac{1}{M^U}\sum_{i=u}^{M^U} R_u^{\text{win}}\right)^2}{\frac{1}{M^U}\sum_{u=1}^{M^U} {R_u^{\text{win}}}^2} \tag{5.22}$$

where M^U is the total number of users.

Average Windowed Rate (Mb/s):

$$\text{Average Rate} = \frac{\sum_{u=1}^{M^U} R_u^{\text{win}}}{M^U T^{\text{win}}} \tag{5.23}$$

Windowed Fair-rate:
This metric integrates the fairness and average rate by multiplying them for each window.

Minimum Windowed Rate (Mb/s):

$$\text{Minimum Rate} = \min_u \frac{R_u^{\text{win}}}{T^{\text{win}}} \tag{5.24}$$

Next, we look into "All-time" metrics that extend the windowed evaluations to the entire duration of the simulation, offering a long-term view of the network performance.

All-time Fairness:
This metric is calculated as the average of the Windowed Jain Fairness across all windows in the simulation duration.

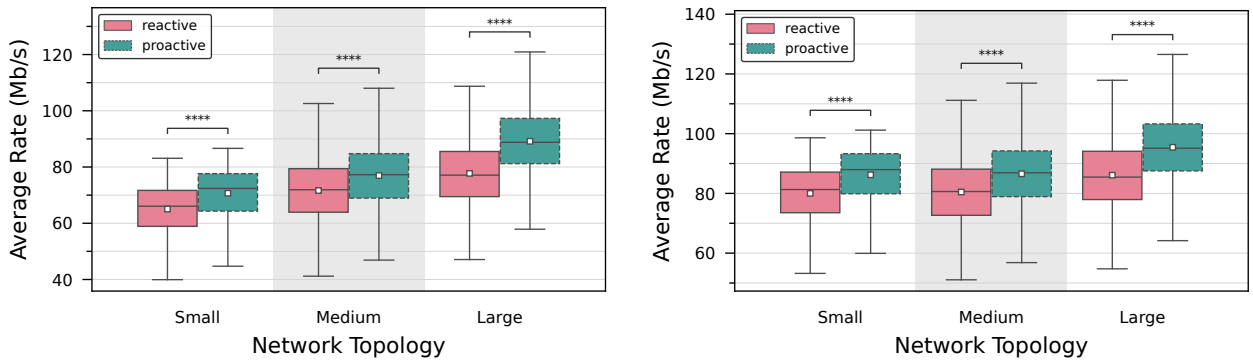
All-time Average Rate (Mb/s):
This metric is calculated as the average of the Windowed Average Rate across all windows in the simulation duration.

All-time Fair-rate:
This metric is calculated as the product of the all-time fairness and average rate metrics.

% of Handovers per User:
This metric measures the frequency of handovers, reflecting the percentage of time slots in which a handover occurs for each user.

Time to solve (s):
The efficiency of the optimization algorithms, specifically the Expert and EGT solvers, is gauged by the time required to solve the optimization problem.

5.7 Results and Comparative Analysis



(a) LiFi-only Network

(b) LiFi-WiFi Network

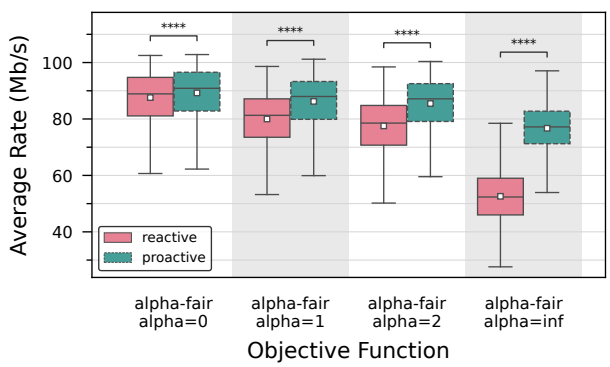
Figure 5.2 Windowed Average rate for reactive and proactive optimized resource allocation solved with Expert for various network topology showing the need for the proactive resource optimization

In Figure 5.2, we conduct a comparative analysis between the reactive resource allocation strategy and our proposed proactive strategy, looking into the proportional fair objective with $\alpha = 1$. The proactive model consistently outperforms the reactive approach in terms of average user rate across both network types. This enhancement can be attributed to the proactive model's ability to utilize future channel predictions, thus providing improved Quality of Service (QoS) over a specified time period without the shortsightedness of the reactive strategy. Specifically, in the LiFi-only network, we can observe improvements of 8.6%, 7.4%, and 14.6% in the average rate per user for the Small, Medium, and Large topologies, respectively. Similarly, in the LiFi-WiFi network, gains of 7.7%, 7.5%, and 10.8% are observed for the Small, Medium, and Large topologies, respectively.

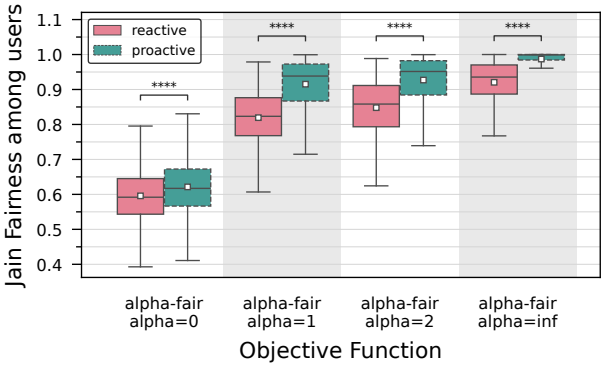
Within the LiFi-WiFi networks, our investigation moves to the performance implications of various alpha-fair objectives, focusing on average rate, minimum rate, and Jain fairness for the Small topology as illustrated in Figure 5.3. Setting $\alpha = 2$ optimizes fairness but adversely impacts other performance metrics. In comparison, other objectives manage to maintain high fairness without degrading overall network performance.

Our analysis consistently demonstrates the superiority of the proactive approach over the reactive method across different objectives, showing its robustness even when the quality metric evaluated is not directly the optimization objective. When $\alpha = 0$, which prioritizes maximizing the sum user rate, the minimum user rate shows negligible difference, as this objective disregards individual user rates.

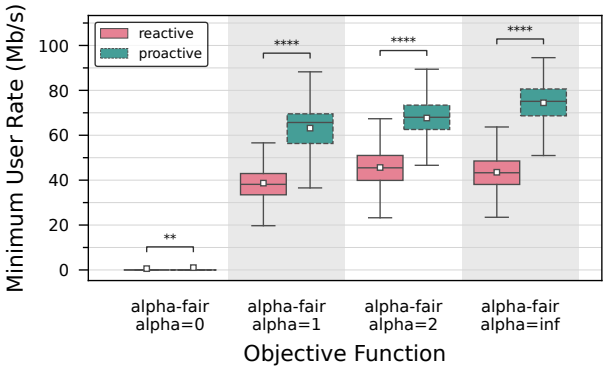
In the reactive strategy, an interesting finding is the sub-optimal minimum rate when $\alpha = \infty$. While intuitively this should yield the best results for the worst-case user rate within each time slot, it fails to translate these short-term gains into long-term benefits. In contrast, the proactive approach, with its foresight into future conditions, achieves more favorable outcomes in enhancing long-term network performance.



(a) Average Rate

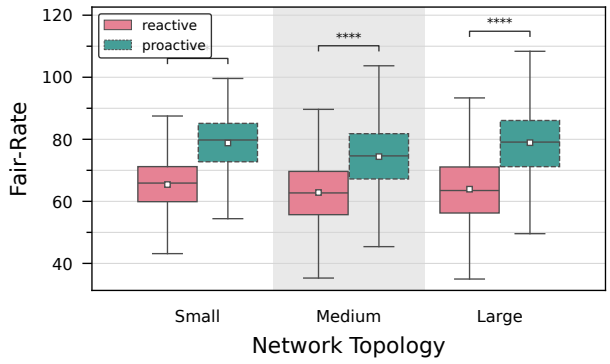


(b) Jain Fairness

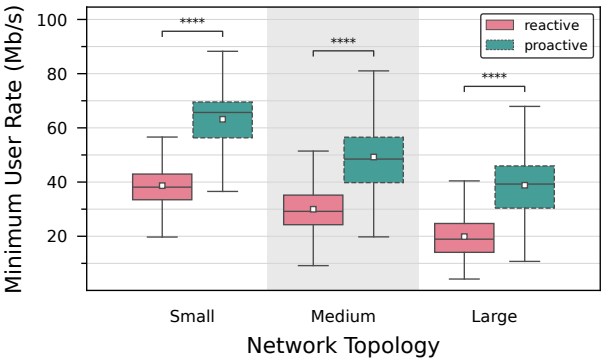


(c) Minimum Rate

Figure 5.3 Quality metrics for various alpha-fair objectives using the reactive and proactive resource allocation strategies solved with Expert for the Small scenario in a LiFi-WiFi Network



(a) Fair-rate

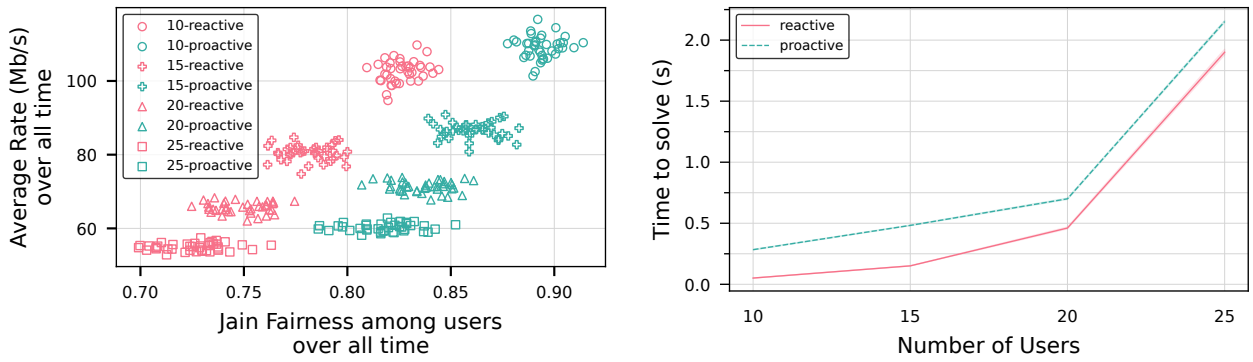


(b) Minimum Rate

Figure 5.4 Quality metrics for various network topology using the reactive and proactive resource allocation strategies solved with Expert for a LiFi-WiFi Network

We now move to a deeper analysis of the different network topologies in Figure 5.4 for which the average rate metric was already discussed in Figure 5.2b. The box plots demonstrate the superior performance of the proactive strategy over the reactive approach, regardless of network size. This is evident as the proactive strategy consistently achieves a higher median fair-rate. Moreover, the interquartile range for the proactive strategy is narrower, indicating less variance and more consistent performance across measurements. The proactive approach also improves upon the reactive strategy in terms of the minimum rate metric across all network configurations. Specifically, the improvement

in the worst user rate when going from the reactive to the proactive strategy is significant, as indicated by the annotation of the results of the tests for statistical significance. In the LiFi-WiFi network, improvements are 63.3%, 64.3%, and 95.3% for the Small, Medium, and Large network topologies, respectively. In the LiFi-only network (which is not shown in the figure), the respective gains are 69.2%, 82.9%, and 68.1% for the same topologies.



(a) Fairness vs Average Rate

(b) Time to Solve

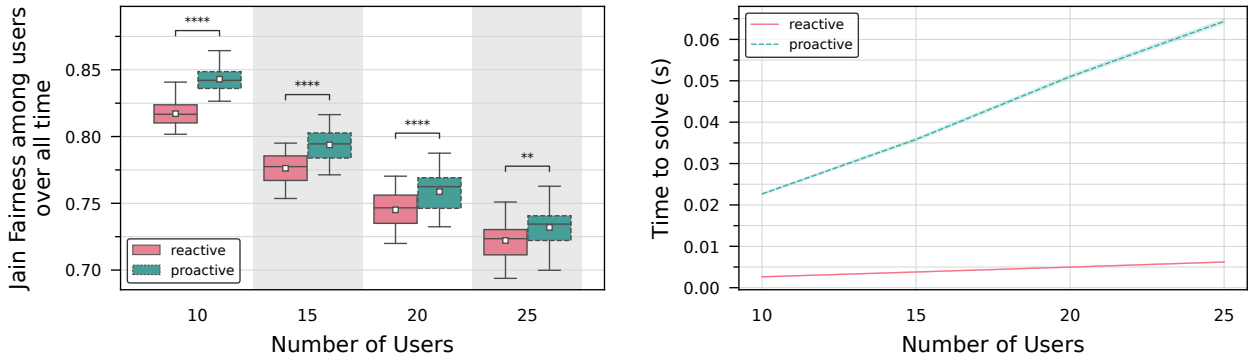
Figure 5.5 Quality metrics for varying numbers of users using the reactive and proactive resource allocation strategies solved with Expert for a 10 x 5 m scenario in a LiFi-WiFi Network

Our investigation now turns to the impact of varying each simulation parameter, starting with the number of users in a fixed topology. We focus on a Medium scenario configured with a 10 x 5 m room and 8 APs, where we analyze the effects of increasing the user count from 10 to 25. Figure 5.5a plots the relationship between user fairness and the average rate achieved during each simulation run. With the rise in user numbers, a decrease in average rates is observed due to bandwidth being shared among more users. Despite this, the proactive allocation strategy achieves better results in both average rate and fairness compared to the reactive approach, as evidenced by the positioning of data points in the scatter plot which appear higher and further to the right for the proactive strategy. This visualization highlights the proactive approach's capability to effectively optimize network performance, even under conditions of increased user density.

We evaluate the run time performance when the number of users are increased in Figure 5.5b. The displayed trend lines indicate a monotonically increasing function, with the time required to solve the problem increasing as the network accommodates more users. Although the absolute run time values are manageable, they may present challenges in real-world implementation. These run times exceed the optimization interval of 500 ms, suggesting the need for exploring alternative optimization strategies that can satisfy the stringent real-time demands of practical deployments.

Integrating EGT into our optimization framework shows advantages, particularly in terms of run time efficiency. As demonstrated in Figure 5.6b, the execution times with EGT remain within the optimization interval, addressing the challenge associated with real-time deployment.

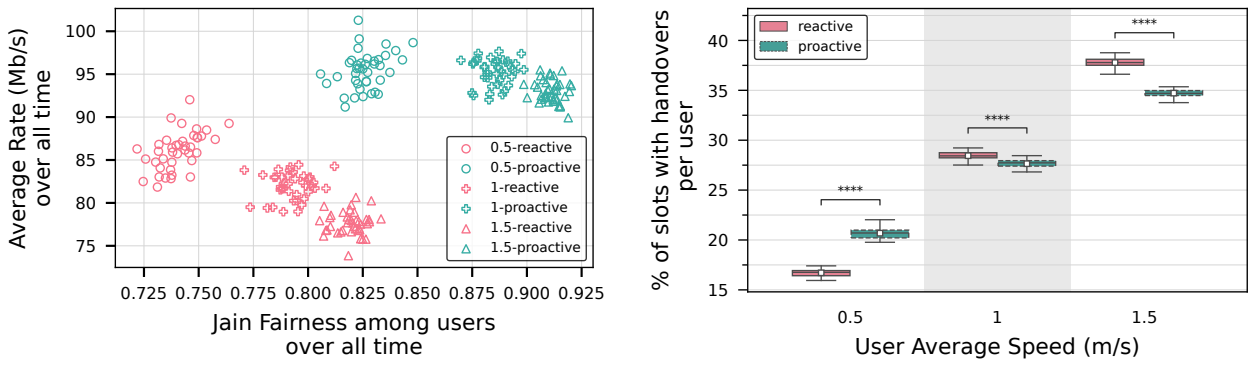
When evaluating the performance metrics, noticeable improvements in fairness are evident with the proactive strategy over the reactive approach, as shown in Figure 5.6a. However, we do not observe similar increases in the average rate metric. Although these improvements are less marked than those achieved using the Expert solver, it is important to understand that this does not undermine the value of the EGT solver. Rather, it emphasizes the practical trade-offs that must be considered when implementing optimization solutions in real-world environments. In situations where the need for a fast response is as important as the quality of the optimization, the advantages of employing EGT become clear.



(a) Jain Fairness

(b) Time to Solve

Figure 5.6 Quality metrics for varying numbers of users using the reactive and proactive resource allocation strategies solved with Evolutionary Game Theory (EGT) for a 10 x 5 m scenario in a LiFi-WiFi Network



(a) Fairness vs Average Rate

(b) % of the simulation time slots with a handover for a user

Figure 5.7 Quality metrics for varying user velocities in the RWP model using the reactive and proactive resource allocation strategies solved with Expert for a Large scenario in a LiFi-WiFi Network

Having demonstrated the advantages of a proactive approach that anticipates future channel conditions, it becomes necessary to explore how these benefits are influenced by the mobility model and user speed, given that these factors determine future user positions. We first investigate the impact of different user speeds on the effectiveness of our proactive strategy using the RWP model in the Large network topology.

In Figure 5.7a, where we assess the trade-offs between Jain fairness and average rate, we note the most significant gains at higher user speeds. This suggests a positive correlation between user speed and the effectiveness of the proactive strategy. The rationale is that as user positions change more rapidly over the evaluation window, the benefit of anticipating these changes increases. Specifically, the average rate improvement ranges from 10.8% at 0.5 m/s to 19.6% at 1.5 m/s.

An interesting aspect to note is the apparent upper limit on proactive average rates, attributed to the capacity limits at the APs. This indicates that the proactive strategy maximizes the potential benefits from the capacity of APs. Moreover, as user speed increases, average rates begin to fall due to the frequency of handovers and associated losses. Nonetheless, rate fairness among users improves under these conditions, suggesting that the higher mobility levels out the channel conditions across users on average, leading to more uniform rates.

Further, in our analysis of handover frequencies in Figure 5.7b, we observe that at the slowest speed of 0.5 m/s, handovers per user increase from the reactive to the proactive approach. However, this increase does not adversely affect the average rate. At higher speeds, the proactive model

manages to reduce the number of handovers, particularly vertical handovers, which are known for their substantial handover losses.

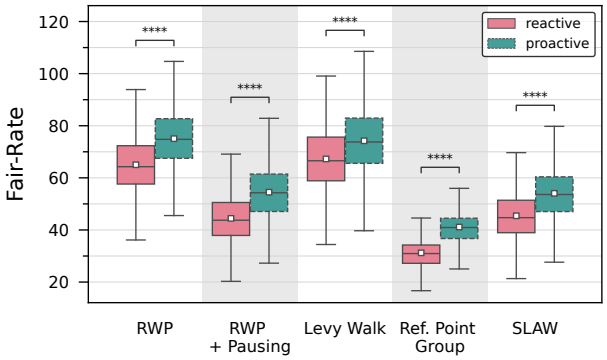


Figure 5.8 Fair-rate for varying mobility models using the reactive and proactive resource allocation strategies solved with Expert for a Large scenario in a LiFi-WiFi Network

Figure 5.8 provides a comparative analysis across different mobility models for the fair-rate metric. The proactive strategy demonstrates better performance relative to the reactive approach across all the mobility models examined. This highlights the adaptability of proactive resource allocation in managing the varied movement patterns of users within indoor environments.

Mobility models that incorporate pausing periods (such as RWP with Pausing) or exhibit tendencies for users to move in groups (such as RPG or SLAW) show a tendency towards lower rates. The introduction of pausing intervals means that users may remain in sub-optimal channel conditions for extended periods. Similarly, when users move as a group, it leads to concentrated resource demand at specific AP, which can reduce the overall network performance. Despite these challenges, the proactive strategy continues to outshine the reactive approach, maintaining its edge in optimizing network performance even under less favorable conditions.

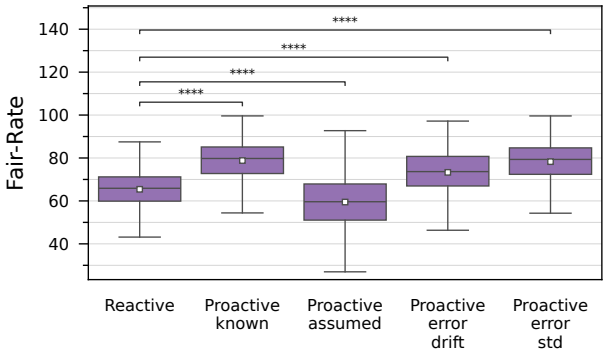
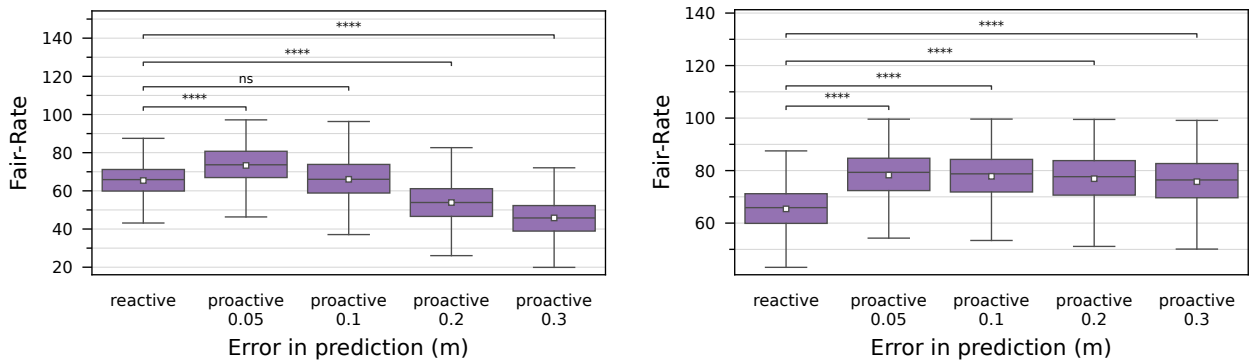


Figure 5.9 Fair-rate for varying future position prediction models solved with Expert for a small scenario in a LiFi-WiFi Network

The improvements observed with the proactive approach in all previous evaluations show the importance of forecasting future positions and channel conditions to boost network performance. A critical question arises: How precise must our predictions be to achieve benefits over the reactive strategy? To address this, we evaluate various prediction models in Figure 5.9.

The *assumed* prediction model, which assumes same positions throughout the prediction window, yields the lowest fair-rate. This highlights the need for more sophisticated prediction capabilities, even in scenarios with low user mobility, to surpass the performance of the reactive model. Consequently, our analysis shifts toward the *error* models that incorporate errors to simulate real-world prediction

errors. The *std* model, which increases the standard deviation of the prediction error throughout the window, performs better than the *drift* model, where errors cumulatively increase as the predicted position gradually deviates over time. Despite the introduced errors in user positions, the proactive strategies under both error models still outperform the reactive approach.



(a) *Drift* Error Model

(b) *std* Error Model

Figure 5.10 Fair-rate for varying future position prediction errors solved with Expert for a small scenario in a LiFi-WiFi Network

It is essential to recognize that the σ for both *error* models in Figure 5.9, as described in Section 5.3.1 is set at 0.05 m. This parameter setting leads us to explore the resilience of the benefits provided by the proactive approach under worse predictions with varying σ. As shown in Figure 5.10, the required precision for prediction varies between models. In the *std* model, predictions with an error margin of up to 0.3 m still yield advantages over the reactive strategy. Conversely, the *drift* model indicates that an error margin beyond 0.1 m starts to negate the benefits.

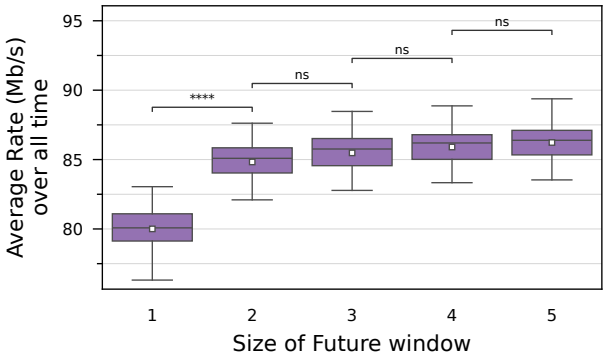


Figure 5.11 Average Rate for varying prediction windows solved with Expert for a small scenario in a LiFi-WiFi Network

To evaluate whether a reduction in prediction window length could compensate for decreased accuracy in user position prediction, we investigate the effects of shorter prediction windows on performance benefits. In this analysis, various window sizes are compared, with a window size of 1 representing the reactive strategy, as shown in Figure 5.11.

The results reveal that, for this particular scenario, a window size of 2 already delivers the anticipated benefits for the average rate metric, with larger window sizes failing to produce large improvements. This indicates that in specific situations, shorter prediction windows could suffice to preserve the advantages offered by a proactive resource allocation strategy.

The findings across various scenarios demonstrate that the proactive approach to wireless resource allocation significantly outperforms the reactive strategy in enhancing network performance metrics in both LiFi-only and LiFi-WiFi networks.

5.8 Discussions on Feasibility of Real-World Implementation

Both the reactive and proactive resource allocation strategies would definitely benefit from a field test on real hardware. In this section, we discuss how our simulations incorporate real-world parameters while acknowledging the factors that challenge the direct implementation of a proof of concept. A known limitation in the field is the scarcity of open-source LiFi hardware that permits access to the Medium Access Control (MAC) layer, necessary for deploying and evaluating our scheduling algorithms, as noted in existing literature [Wu+21b]. While our LiFi-WiFi testbed, detailed in Section 2.1 and constructed using commercially available technology, confirms aspects of our simulation models, it does not support the direct application of our proposed resource allocation strategies. Nonetheless, we remain hopeful that the recent introduction of the IEEE 802.11bb standard [IEE23] will stimulate the creation of new open-source hardware, facilitating future implementations.

Our simulation of the LiFi channel incorporates realistic parameters, including a maximum data rate of 250 Mb/s and limits the coverage based on the receiver's field of view, with signal coverage dropping to zero outside this range. These parameters stem from empirical data obtained through our in-house hardware testbed, which corroborates that our simulations replicate real-world conditions.

The efficiency of vertical handovers, η, is established at 0.4, reflecting a handover time or associated loss of 300 ms, a number also corroborated by results from our testbed as depicted in Figure 5.12.

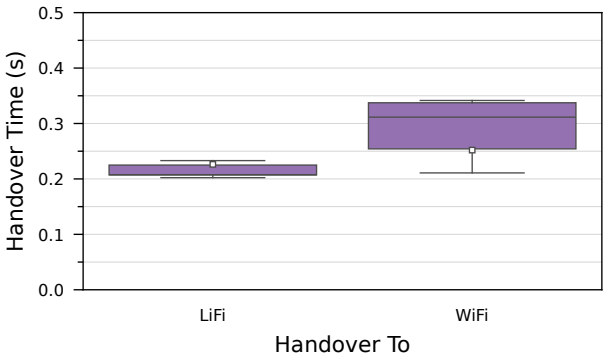


Figure 5.12 Vertical Handover Overhead observed when switching from WiFi to LiFi and vice versa

This specific result was further analyzed in the previous chapter (see Section 4.2.10). Additionally, our simulations integrate realistic LoS blockages, modeling the loss of signal that significantly impacts the LiFi channel. This ensures that our resource allocation strategies consider the realistic environmental factors affecting signal quality.

In our simulations, we employ the SLAW mobility model, which draws on real-world human mobility patterns derived from GPS traces of pedestrian movements [Lee+09]. Furthermore, acknowledging the inherent uncertainties in predicting future user positions, our simulations incorporate prediction error models. These models replicate the inaccuracies typically encountered [PVK21] in real-world prediction algorithms. By integrating these error models, we can assess the resilience of our proactive optimization strategy in the face of such prediction errors.

Our proactive resource allocation strategy is particularly effective for real-world scenarios like video streaming, where a full buffer model of transmission applies. In such use cases, data transmission is ongoing and predictable, which supports the feasibility of proactive resource management. This approach allows for the pre-emptive allocation of resources. Thus, when anticipated network condi-

tions deteriorate, the strategy can adjust resource distribution to sustain a consistent quality of service by pre-filling the receive buffer.

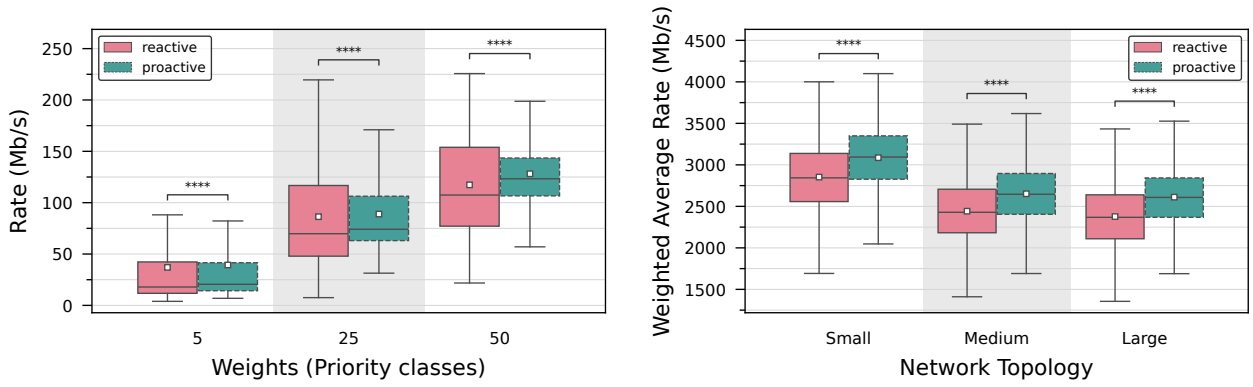
Although the full buffer model applies for the specific use case of continuous video data streaming, our framework is designed with the flexibility to adapt to a variety of real-world conditions. It accommodates the diverse transmission needs of users by integrating different priorities or weights based on their resource demands. This adaptability is facilitated through the use of the weighted alpha-fair resource allocation method [MW00], which enables differentiated treatment of users according to their specific requirements. Consequently, our framework ensures efficient and conservative resource allocation, minimizing the potential for wasting radio resources.

Incorporating user-specific weights w_u modifies the utility function in our proactive optimization framework as outlined in (5.6) into

$$U_\alpha(x_{u,a}^{(t)}, y_{u,a}^{(t)}) = \begin{cases} \sum_u \sum_t \sum_a w_u x_{u,a}^{(t)} \log\left(y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}\right) & \text{if } \alpha = 1 \\ \sum_u \sum_t \sum_a w_u x_{u,a}^{(t)} \frac{\left(y_{u,a}^{(t)} \tilde{R}_{u,a}^{(t)}\right)^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1 \end{cases} \tag{5.25}$$

With this redefinition, the resource allocation strategy executed by the EGT algorithm, as originally described in (5.21) is adapted to accommodate these weighted user considerations.

$$y_{u,a}^{(t)} = \begin{cases} \begin{cases} 1 & \text{if } u = \arg\max_{u'} w_{u'} \tilde{R}_{u',a}^{(t)} \\ 0 & \text{elsewhere} \end{cases} & \text{if } \alpha = 0 \\ \frac{w_u^{\frac{1}{\alpha}} \left(\tilde{R}_{u,a}^{(t)}\right)^{\frac{1}{\alpha}-1}}{\sum_{u' \in \mathcal{U}_a} w_{u'}^{\frac{1}{\alpha}} \left(\tilde{R}_{u',a}^{(t)}\right)^{\frac{1}{\alpha}-1}} & \text{if } \alpha > 0 \end{cases} \tag{5.26}$$



(a) Rate per user priority group

(b) Weighted Average Rate representing the utility function

Figure 5.13 Quality metrics for weighted alpha-fair resource allocation in a LiFi-WiFi Network with users assigned weights of 5, 25, and 50 with higher weight indicating the need for more resources

Figure 5.13 shows a exemplary outcome when users are assigned random weights of 5, 25, and 50, reflecting their different resource requirements. Figure 5.13a displays that the data rate increases in accordance with the user weight, effectively responding to individual user needs. Furthermore, Figure 5.13b presents the utility function for the entire network, demonstrating that our proactive strategy surpasses the reactive even in this weighted scenario. These weights can be adjusted to give precedence to users according to their resource demands and priorities, facilitating targeted optimizations that align with different network objectives.

All the aspects presented here highlight the practical applicability of our resource allocation framework, offering a realistic evaluation of their deployment potential in real-world environments.

5.9 Summary and Conclusions

In conclusion, this chapter enhances the resource management of mobile LiFi-WiFi heterogeneous networks through the adoption of proactive resource allocation strategies. Our approach not only achieves an 7.7% increase in average rates but also a remarkable 63.3% improvement in the worst user rates compared to the reactive approach, demonstrating its ability to deliver sustained QoS.

Challenges in scalability associated with the Branch and Bound solver run times are addressed by incorporating EGT, which shows promise in enhancing run time efficiency and meeting the demands of real-world deployments. Additionally, the proactive approach's consistent superiority across varying user densities, speeds, and mobility models attests to its adaptability in diverse indoor network environments.

A pivotal element in proactive allocation is the accuracy of user position predictions. Our findings indicate that inaccurate predictions can, in some scenarios, render the proactive strategy less effective than reactive approaches. However, our evaluations of simulated prediction models demonstrate the resilience of the proactive model to a degree, suggesting its practicality with realistic prediction methods. A shorter prediction window, which typically yields better prediction accuracy, was also investigated, confirming that our proactive resource allocation retains its benefits even with reduced prediction windows.

In summary, this research strongly supports the implementation of proactive resource allocation in LiFi-WiFi networks, advocating for its capacity to enhance long-term QoS despite challenges such as blockages, sub-optimal channel conditions, and user mobility.

So far we have shown the challenges in wireless resource allocation in LiFi-WiFi networks and proposed various resource management approaches that address these challenges and improve the network performance for the various goals of minimizing network latency, achieving stability, and maximizing the alpha-fair throughput of users in the network for both short term and long term network performance. In future networks, which are increasingly becoming dependent on computational resources as part of Multi-access Edge Computing (MEC) systems, there is a need to also manage the allocation of computational resources. The next chapter deals with the challenge of managing wireless and computation resources towards task offloading in LiFi-WiFi networks.

Chapter 6

Task Offloading in Multipath Multihop LiFi-WiFi Networks

The previous chapter discussed proactive wireless resource allocation strategies in mobile Light-Fidelity (LiFi)-Wireless-Fidelity (WiFi) heterogeneous networks, demonstrating significant improvements in average and worst-user rates and mitigating the adverse effects of mobility and blockages. This strategy demonstrated the potential of proactive wireless resource allocation approaches in sustaining Quality of Service (QoS) over time and adapting to diverse network environments.

While addressing wireless resource allocation is crucial, the growing demand for computational resources to process latency-critical tasks like Ultra Reliable Low Latency Communications (URLLC), factory automation, and Artificial Intelligence (AI) image classification has introduced new challenges. In response to these challenges, we propose a task offloading framework called ComputiFi, which aims to minimize the task completion latency, network energy, and user device energy in multipath, multihop LiFi-WiFi networks. This chapter, in part, is based on the contribution published in [VMK24] which is a journal publication that aimed to minimize the task completion latency.

6.1 Introduction

The rapid evolution of wireless communications technologies is driven by the growing needs of latency-sensitive applications such as URLLC, factory automation, and AI image classification [3GP22a; 3GP22b]. These applications have stringent requirements for low latency and uninterrupted data processing, which are challenging for networks relying solely on traditional technologies like WiFi due to their susceptibility to interference and bandwidth limitations [Ayy+16]. Apart from requiring minimal latency for effective real-time processing, these applications also demand high energy efficiency to reduce operational costs and environmental impacts. Addressing all these objectives is essential as networks expand to accommodate a wide range of devices and data-intensive activities. Consequently, there is a need to develop frameworks that can effectively integrate and manage multiple network technologies.

LiFi, used alongside WiFi, is beneficial in environments such as manufacturing facilities, where high data rates, low latency, and energy efficiency are crucial. Operating on the visible light and infrared spectrum, LiFi provides an interference-free communication channel that offers advantages in security over WiFi, which often suffers from bandwidth saturation in densely populated settings [Haa+16].

The integration of LiFi with WiFi into a multipath network architecture is particularly advantageous [Wu+21a]. Current offloading methods that rely on a singlepath approach [Fan+22; XS23] do not fully utilize available network resources, often leading to increased latency and higher energy consumption. This shows the importance of adopting multipath transmission techniques, which utilize several simultaneous channels to improve the throughput and reliability of data transmission [Yun+22]. Such integration of technologies allows for the strategic routing of tasks through the most efficient paths optimizing for latency. Network resources can also be dynamically allocated to minimize the

transmission energy by selecting shorter or less congested paths and to optimize computation energy by offloading tasks to the most efficient compute nodes available.

Traditional approaches to task offloading often overlook the energy dimensions involved in network and device operations. Typically, these solutions focus on reducing latency by selecting single fixed offloading destinations, which inadvertently increases network congestion and energy consumption [Zha+20]. To overcome these limitations, a multi-tiered approach to Multi-access Edge Computing (MEC) destinations is suggested. This includes processing tasks locally on user devices, offloading to proximal LiFi and WiFi Access Point (AP)s, utilizing routers that link these access points, and leveraging cloud servers through the network backbone. By doing so, it fosters a multihop network infrastructure that not only reduces transmission distances [SKO23; Fan+22], but also conserves energy.

The design of this multihop infrastructure necessitates continuous adjustments to routing strategies and the dynamic allocation of resources, considering real-time network conditions, traffic volumes, and computational demands at each node [Man+23]. Enabling local processing at user devices reduces the energy required for data transmission to distant centralized servers and decreases the latency associated with such transfers. However, this is counter-intuitive for the energy consumption of the battery operated mobile and Internet of Things (IoT) devices. Therefore, an optimization of both energy consumption and latency is critical for sustainable wireless network operations. Although many existing frameworks employ multihop architectures to improve network coverage and connectivity, they often simplify latency assessments by merely summing transmission latencies across hops without accounting for the complexities involved in handling multiple data packets and managing data flows through intermediate nodes [Fan+22; Fen+21; Den+24]. Such methods fail to capture the properties of the task's traffic when it is composed of multiple data packets.

In response, we introduce ComputiFi, a task offloading framework designed to optimize latency and improve energy efficiency in multipath, multihop LiFi-WiFi networks. ComputiFi intelligently determines the most suitable offloading destinations, dynamically distributes data across available technologies, and efficiently manages computational resources among all potential computing destinations towards optimizing latency, network energy and user device energy consumption.

To achieve these objectives, ComputiFi incorporates a range of optimization tools including Mixed Integer Nonlinear Programming (MINLP) solvers, meta-heuristics, Deep Reinforcement Learning (DRL), and black-box optimization techniques. This array of solutions equips the framework to not only fulfill the stringent performance requirements of latency-sensitive applications but also aligns with the sustainability objectives crucial for future network operations.

6.2 State-of-the-art Analysis

Here, we provide an overview of relevant literature related to task offloading, highlighting gaps that motivate our research.

In [Fen+22], a detailed survey on computation offloading in edge computing networks is provided. This work reviews various objectives and methods for offloading, aiming to improve computational efficiency, and identifies key trends, challenges, and open research questions. However, it does not address the specific issues associated with LiFi-WiFi heterogeneous networks, such as blockages and the density of placement of LiFi APs. Although it mentions some works on multipath task routing, it does not evaluate this concept in depth.

Similarly, [Sal+21] proposes a mobility-aware framework for coordinating task scheduling and resource allocation within a cooperative device-to-device computing network. While it does not consider a multipath network, it does account for user mobility, which is pertinent to our research where user mobility is also a factor for the dynamic nature of the network.

In [Liu+24a], a method leveraging deep learning for online task offloading is introduced, aiming to minimize latency in heterogeneous mobile edge environments. This adaptive technique responds to changing network conditions and user demands, focusing on latency reduction. Although this work

does not explore multipath networks, its approach to using DRL for real-time implementation and comparison with the Gurobi solver for optimality is aligned with our research methodology. While the authors only focus on latency minimization, we also include the objective of energy minimization for network sustainability.

Shifting the focus to heterogeneous networks, [Fan+22] presents a framework for task offloading and service caching in WiFi-cellular heterogeneous networks. Their approach optimizes resource and channel allocation while reducing latency and energy consumption. Despite considering diverse technologies, it does not evaluate multipath transmissions in this context.

In [Yun+22], an approach using Genetic Algorithm (GA) and DRL is proposed for managing task offloading and resource allocation across multiple Radio Access Technologies (RATs) for URLLC and enhanced Mobile Broadband (eMBB). The model dynamically adapts to varying network conditions, optimizing offloading success rates and maximizing spectral efficiency. While the authors investigate multipath transmission for eMBB tasks, further research is needed to address the use of two technologies like LiFi and WiFi which exhibit stark differences in properties. Additionally, the work does not consider multiple offloading destinations, which can further reduce latency by distributing tasks across the network.

Taking a step towards multiple offloading destinations, [CL21] explores the collaboration between adjacent MEC servers for task processing. Although this setup involves forwarding tasks to adjacent servers due to resource constraints, it does not represent a true multihop architecture and lacks support for multipath transmissions.

In [SKO23], the authors extend the concept of computing destinations to multiple edge and cloud servers, using multi-agent DRL for cooperative computing offloading and route optimization. While they explore multiple transmission paths, the work does not address scenarios involving mobile users who frequently change their wireless associations.

Combining both multipath and multihop aspects, [Fen+21] investigates task offloading strategies that balance communication latency and energy efficiency in mobile ad-hoc networks, utilizing an adaptive path selection algorithm that accounts for drone mobility. However, their communication model is not fully suitable for multi-packet tasks relayed across multiple hops.

In conclusion, unlike existing literature that typically focus on single-hop or single-path offloading strategies, ComputiFi uses a set of optimization tools — including MINLP solvers, meta-heuristics, DRL, and black-box techniques — to dynamically adjust offloading destinations and resource allocations across various offloading destinations and wireless technologies. These tools are assessed for their effectiveness in task offloading problems. While numerous state-of-the-art papers explore reinforcement learning in edge computing, they often do not tackle the specific latency, energy and resource allocation challenges presented by LiFi-WiFi systems.

6.3 Key Contributions

We introduce the architecture and methodology of ComputiFi, a task offloading framework, showcasing its capabilities to reduce both the latency and energy consumption of tasks in multipath, multihop LiFi-WiFi networks. Specifically this chapter makes the following significant contributions:

1. Comprehensive Task Offloading Framework: We present ComputiFi, a framework designed for offloading latency-sensitive tasks across multipath, multihop LiFi-WiFi networks. The framework not only determines the most efficient offloading destinations but also manages data distribution between LiFi and WiFi. It optimizes both network energy usage and user device energy consumption, employing a multihop latency formulation that is particularly suited for handling tasks with multiple data packets.

2. Advanced Optimization Solvers: ComputiFi integrates a variety of sophisticated optimization tools, including MINLP solvers, meta-heuristics, DRL, and black-box optimization techniques.

These solvers are tailored to adapt dynamically to the fluctuating network conditions and varying traffic demands, ensuring effective and efficient task offloading.

3. Resource Allocation Strategy: The framework implements a dynamic strategy for allocating resources across a diverse range of network devices — local devices, LiFi and WiFi access points, routers, and cloud servers. This strategy optimizes task routing to minimize energy consumption at both the network and user device levels while providing guaranteed latency, as well as minimizing task completion latency.

4. Benchmarking and Performance Validation: ComputiFi is rigorously benchmarked against various baseline solvers to validate its enhanced performance in reducing latency and energy consumption. The results confirm the effectiveness of ComputiFi in managing latency-sensitive and energy-intensive tasks, establishing its suitability for deployment in heterogeneous network environments.

6.4 System Model

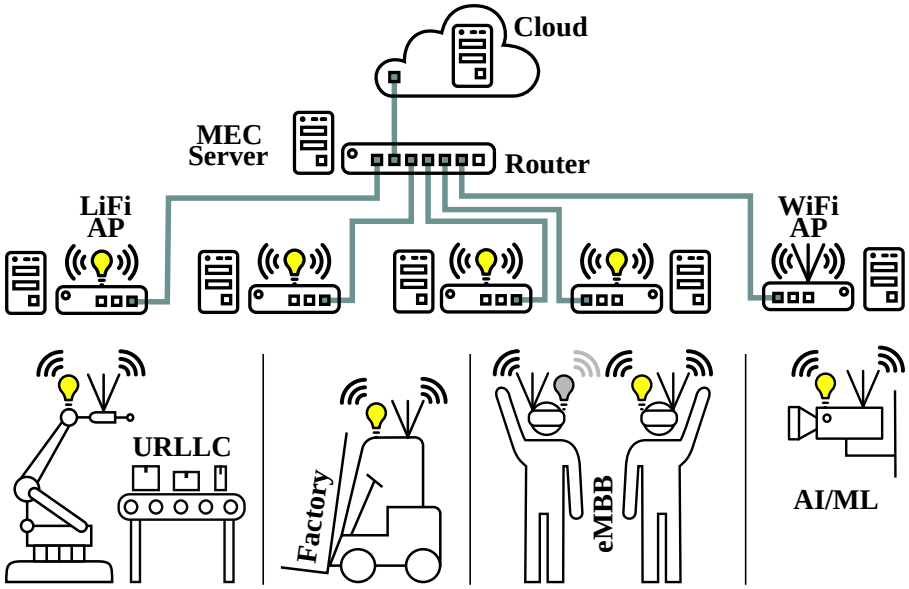


Figure 6.1 Network architecture for a multipath multihop Task offloading system

The architecture of the network designed for task offloading in multipath, multihop LiFi-WiFi networks includes M^L LiFi APs and M^W WiFi APs. Within the network, LiFi APs experience co-channel interference when their coverage areas overlap, while WiFi APs operate interference-free due to the utilization of distinct frequency channels.

The network serves M^U users, each equipped with LiFi and WiFi capabilities for both sending and receiving data, with users denoted by $u \in \{1, 2, \ldots, M^U\}$. Every user is connected to one LiFi and one WiFi AP simultaneously, based on the highest Signal to Interference and Noise Ratio (SINR) values available, making a total of $M^N = 2$ wireless connections per user.

Routers facilitate connectivity between APs, with one router serving every five APs. These routers are also directly linked to cloud servers through wired connections, ensuring efficient backhaul operations. The layout of this network is depicted in Figure 6.1, and all relevant terminology and symbols used in this chapter are detailed in Table 6.1.

Table 6.1 List of Notations used in Task Offloading in Multipath Multihop LiFi-WiFi Networks

Notation	Description
l, M^L	Index and number of LiFi APs
w, M^W	Index and number of WiFi APs
u, M^U	Index and number of users
l_u	Index of LiFi AP associated to user
w_u	Index of WiFi AP associated to user
n, M^N	Index and number of technologies
m, $\mathcal{M}$	Index, and set of computing destinations
$R_{u,l}$, $R_{u,w}$	Achievable rate of user for LiFi and WiFi
ij	Link between two nodes i and j in the network
BW_{ij}	Capacity of link ij in bits per second
F_m	Computing server's capacity in instructions per second
E_m	Computing server's energy consumption in Joules per instruction
F_u	User's processing capacity in instructions per second
E_u	User's energy consumption in Joules per instruction
t, N_u^T	Type and number of tasks at a user
τ_u^{req}	User's task delay requirement in seconds
C_u	Number of instructions to process user's task
S_u	Data size of user's task in bits
S_u^{pkt}	Size of one packet of the user's task in bits
$P_{u,n}$	Transmission Power of user device u with technology n
$\lambda_{ij}^{u,n,m}$	Binary indicator of utilization of links
$x_{u,m}$	Optimization variable for the destination selected for user
k_u^n	Optimization variable for data split proportion
$c_{u,m}$	Optimization variable for compute resource proportion
p_u	Optimization variable to process the task locally or not
τ_u^L	Local computation latency
$\tau_u^{\mathrm{compute}}$	Edge computation latency
$\tau_u^{\mathrm{Tx},n}$	Transmission latency over one technology
$\tau_{u,n,m}^{\mathrm{flow}}$	Auxiliary variable for the latency of the bottleneck link
τ_u^{Tx}	Total transmission latency
$\tau_{ij}^{\mathrm{prop}}$	Propagation latency for a link ij
τ_u^E	Total latency when the task is offloaded
τ_u	Task completion latency
e_u^{Tx}	Total transmission energy
e_u^L	Local computation energy
e_u^{compute}	Edge computation energy
e_u	Network energy consumption for task completion
$e_u^{\mathrm{Tx},D}$	Transmission energy consumed by user
e_u^D	User device energy consumption for task completion
G	Reward for DRL-based algorithm
V_{τ_u}	Latency Violation
N_u^V	Number of users violating the latency bound

In this network configuration, the bandwidth for each link ij is specified as BW_{ij} and is assumed to be adequate, ensuring that the capacity of the backhaul connections either matches or exceeds that of the fronthaul. Additionally, the capacity of each router is engineered to be at least equivalent to the aggregate bandwidth of its associated links. Details of these capacities are presented in Table 6.2, with the wireless capacities derived from channel models. Although the propagation delay is typically negligible for most connections due to their proximity in indoor environments, the latency to the cloud server is set at 7.5 ms, as documented in previous research [Bra+20].

Table 6.2 Properties of Links between Network Nodes

From/To	To/From	Capacity (Mb/s)	Propagation Delay (ms)	Transmission Power W
Router	Router	500	–	–
Router	Cloud	2500	7.5	–
Router	WiFi AP	500	–	–
Router	LiFi AP	500	–	–
User	LiFi AP	max. 250	–	5
User	WiFi AP	max. 160	–	0.1

A central network controller is integral to the network's architecture, facilitating the orchestration of communication resources and data flows throughout the system. This controller is needed for acquiring Channel State Information (CSI) from users and monitoring server load, thereby enabling the implementation of a centralized resource allocation algorithm. This algorithm capitalizes on comprehensive network data to optimize decisions regarding MEC server destinations, packet split ratios, and the allocation of computational resources.

The channel model and geometric blockage model for LiFi are elaborated in Section 2.3 and Section 2.4.3, respectively, while the channel model for WiFi is described in Section 2.6. These channel models determine both the wireless link capacity as well as the transmission power P_u of the user device in the uplink. To assess the efficacy of our task offloading strategy in environments characterized by user mobility, the Random Waypoint (RWP) mobility model is employed. This model also considers changes in the orientation of user devices and is applicable to a variety of mobile entities, including robots, automated systems, and manually operated machines, with further details provided in Section 2.7.

6.4.1 Computing Server Model

In the ComputiFi framework, the MEC architecture is instrumental in minimizing latency by situating computational resources proximate to the users. The processing server locations, represented as $m \in \mathcal{M}$, encompass cloud servers, routers, and both LiFi and WiFi APs. While user devices also participate in processing, they are not server nodes. The computational capabilities of each server component, including processor speed in instructions per second and core count, are important for enhancing network performance. The aggregate processing power of all cores within a component is denoted as F_m, whereas the processing capacity of a local user device is indicated by F_u.

Table 6.3 displays the computational specifications for each server category within the network. The processing capacity for user devices is taken from the Qualcomm Snapdragon 720G processor. APs for both LiFi and WiFi incorporate processing capabilities as reported in literature [Zha+17]. Routers are equipped with Ampere Altra Max M128-30 processors, which are capable of handling intensive computational demands. The model for a cloud server includes five such processors, reflecting a high-capacity computational environment.

Table 6.3 Computational Properties of Processing Servers

Component	Processing Capacity (Instructions per core per second)	Cores	Energy Consumption (Joules per instruction)
User	1.9×10^9	1	3.29×10^{-10}
WiFi AP	5×10^9	8	1×10^{-9}
LiFi AP	5×10^9	8	1×10^{-9}
Router	3×10^9	128	6.51×10^{-10}
Cloud	3×10^9	128 x 5	6.51×10^{-10}

6.4.2 Task Model

This section outlines the framework used for tasks that need to be offloaded and subsequently processed within the system. Users encountering tasks are required to process these tasks locally or offload them to a server for processing. Each user has interdependent tasks of a single application that are collectively processed at a single designated server. Furthermore, each task is characterized by a delay requirement that must be satisfied within the given optimization interval. Depending on the network's current conditions and the urgency of the task, users may utilize either LiFi or WiFi technologies for transmitting these tasks to the server.

The task parameters, such as arrival rates, size, and delay requirements, differ among various applications including eMBB like Video Streaming, URLLC, Factory Automation, and AI like Image Classification. The task arrival rate is modeled using a Poisson distribution, with the mean rates detailed in Table 6.4. Additional task characteristics such as size, latency requirements, and processing instructions for each application are also provided in Table 6.4.

The task attributes for eMBB and URLLC are adapted from [Yun+22], while the specifications for Factory Automation are sourced from [Jan+23], and those for AI/Image Classification are based on [Liu+24b]. Moreover, the specific delay requirement for any given task is denoted as τ_u^{req}, the processing demand in terms of instructions is indicated by C_u, the overall size of the task in bits is denoted by S_u, and the size of an individual packet of the task in bits is denoted as S_u^{pkt}.

Table 6.4 Task Properties for Various Applications

App	Arrival Rate (tasks/s)	Size (pkts)	Pkt. Length (bytes)	Processing (Instructions)	Latency (ms)
eMBB	25	33	1500	16.5×10^6	500
URLLC	150	1	32	10 560	1
Factory	100	20	1500	1.25×10^6	100
AI	2	417	1500	569×10^6	500

6.5 Task Offloading Problem Formulation

This section develops the problem formulation for task offloading in LiFi-WiFi multipath, multihop networks aiming to minimize latency and energy consumption.

6.5.1 Communication Model

Following the optimization process, when a destination m is assigned to user u's tasks, the system supports n potential pathways to this destination. The value n specifies the wireless technology determining the initial hop of the transmission. Once this initial hop is established, no additional routing decisions are required due to the network's tree topology configuration, which dictates a predetermined route to the destination.

In the network, the usage of transmission links, represented by the binary indicators $\lambda_{ij}^{u,n,m}$, specifies whether link ij is utilized for offloading tasks from user u via technology n to destination m. The bandwidth $\text{BW}_{ij}^{u,n}$ allocated to user u on this link is influenced by the task's data volume and the number of users sharing the same link. The allocation of link capacity among users is expressed as follows:

$$\text{BW}_{ij}^{u,n} = \frac{\text{BW}_{ij} k_u^n S_u}{\sum_{u'} \sum_{m'} x_{u',m'} \sum_{n'} \lambda_{ij}^{n',u',m'} k_{u'}^{n'} S_{u'}} \tag{6.1}$$

In this equation, BW_{ij} represents the total capacity of the link, k_u^n denotes the optimization decision variable reflecting the fraction of user u's data transmitted using technology n, and S_u is the total size

of the user's data. The denominator aggregates the total data volumes across all tasks transmitted over that link, where $x_{u',m'}$ is a binary decision variable indicating whether user u' offloads their task to destination m'.

This model shows that while the wireless resources themselves are not direct variables in the optimization, the proportion of data routed through a specific technology is a decision variable which determines how wireless resources are allocated to each user..

Multihop Transmission Latency for a Single Technology:
The latency involved in transmitting a user's task $\tau_u^{\mathrm{Tx},n}$ through a single technology n from user u to the chosen offloading destination m is expressed by the following equation:

$$\tau_u^{\mathrm{Tx},n} = \sum_m x_{u,m} \left[\sum_{ij} \lambda_{ij}^{n,u,m} \left(\frac{S_u^{\mathrm{pkt}}}{\mathrm{BW}_{ij}^{u,n}} \right) + \max_{ij} \left(\frac{\lambda_{ij}^{n,u,m} \left(k_u^n S_u - S_u^{\mathrm{pkt}} \right)}{\mathrm{BW}_{ij}^{u,n}} \right) + \sum_{ij} \lambda_{ij}^{n,u,m} \tau_{ij}^{\mathrm{prop}} \right] \tag{6.2}$$

This equation decomposes the total latency $\tau_u^{\mathrm{Tx},n}$ into three components:

1. Packet Transmission Latency: The first summation, $\sum_{ij} \lambda_{ij}^{n,u,m} \left(\frac{S_u^{\mathrm{pkt}}}{\mathrm{BW}_{ij}^{u,n}} \right)$, calculates the latency for the transmission of a single packet across all utilized hops from the user to the destination. The latency at each hop is influenced by the packet size S_u^{pkt} and the bandwidth $\mathrm{BW}_{ij}^{u,n}$ available on that link.

2. Bottleneck Link Latency: The second term, $\max_{ij} \left(\frac{\lambda_{ij}^{n,u,m} \left(k_u^n S_u - S_u^{\mathrm{pkt}} \right)}{\mathrm{BW}_{ij}^{u,n}} \right)$, quantifies the latency induced by the bottleneck link, which is the link with the minimum bandwidth within the chosen transmission route. This specific link establishes the upper limit on the time required to transmit the remaining bulk of the task's data, as it dictates the throughput of the entire path. The expression $k_u^n S_u - S_u^{\mathrm{pkt}}$ quantifies the volume of data left to be sent following the initial packet. When this remainder is divided by the bandwidth $\mathrm{BW}_{ij}^{u,n}$ of the bottleneck link, it yields the duration needed to transmit this residual data. As this link operates at the slowest speed among all, the total transmission duration is dependent on the time it takes for this link to transmit all its data and clear its data queue.

3. Propagation Latency: The final term, $\sum_{ij} \lambda_{ij}^{n,u,m} \tau_{ij}^{\mathrm{prop}}$, includes the propagation delays over all the links utilized in the task's transmission path.

In our proposed framework, we proactively allocate both bandwidth and computational resources to users, which helps in avoiding any queuing latency at the task level. However, at the packet level, queuing might still occur, particularly at the destination computing server. Here, packets are held until the entire task's packets arrive, at which point processing begins. Additionally, queuing latency is accounted for at the relay nodes within the network, as described in the Bottleneck Link Latency section. These relay nodes serve as interim stations where packets are buffered before being forwarded to their next hop. We ensure that these relay points have adequate buffering capacity to efficiently manage any incoming packet traffic, thus mitigating potential queuing delays. Our approach simplifies queuing management by assigning each user to a specific application, which allows us to implement a First Come First Serve (FCFS) queuing mechanism at the user level rather than the application level. These factors are incorporated into our latency model, which combines transmission and computing latencies to reflect that processing starts only following the completion of all packet transmissions. Furthermore, the queuing at relay nodes is managed within the framework of Bottleneck Link Latency.

Total Multipath Transmission Latency:
After computing the latency for each path using different technologies, the overall transmission latency, τ_u^{Tx}, is defined by the latency of the slower path:

$$\tau_u^{\text{Tx}} = \max(\tau_u^{\text{Tx},n}, n \in \{\text{LiFi}, \text{WiFi}\}). \tag{6.3}$$

This work focuses on uplink traffic resource management, based on the assumption that the data transmitted upstream is substantially larger than the data received downstream after task processing.

6.5.2 Task Processing Model

For tasks processed on the user's device, the local processing latency, denoted as τ_u^L, is calculated using the formula:

$$\tau_u^L = \frac{C_u}{F_u} \tag{6.4}$$

Here, C_u represents the number of instructions required to process the task, and F_u indicates the processing power of the user's device, measured in instructions per second.

For tasks processed at a server, the processing latency, τ_u^{compute}, is expressed as:

$$\tau_u^{\text{compute}} = \frac{C_u}{\sum_m x_{u,m} c_{u,m} F_m} \tag{6.5}$$

In this equation, C_u again is the number of instructions needed for the task, F_m is the aggregate processing capacity of the server in instructions per second, and $c_{u,m}$ denotes the fraction of the server's computing resources allocated to user u at server m.

6.5.3 Task Completion Latency

The total task completion latency is the sum of the transmission and processing latencies. When the task is processed on the user's device, the completion latency is straightforwardly defined as τ_u^L. Conversely, if the task is offloaded for processing at the edge, the edge processing latency, τ_u^E, is computed as follows:

$$\tau_u^E = \tau_u^{\text{Tx}} + \tau_u^{\text{compute}} \tag{6.6}$$

The overall task completion latency, denoted τ_u, is then determined by:

$$\tau_u = p_u \tau_u^L + (1 - p_u)\tau_u^E \tag{6.7}$$

In this formula, τ_u^L is applied if the binary decision variable $p_u = 1$, indicating local processing, and τ_u^E is applied if $p_u = 0$, indicating processing at the edge. In the ComputiFi framework, this task completion latency is used as the objective to be minimized besides the energy consumption objective. The task completion latency is also constrained to be less than or equal to the latency requirement of an application, in all formulations of the optimization problem.

6.5.4 Network Energy Consumption Model

To address the sustainability of future task processing networks, we introduce the model for network energy consumption with the goal of minimizing it.

The energy consumed by the network in transmitting a task from the source user to the destination server, if the task is offloaded, is given by

$$e_u^{\text{Tx}} = \sum_n \sum_m x_{u,m} \left[\sum_{ij} P_{i,n} \left(\lambda_{ij}^{n,u,m} \left(\frac{k_u^n S_u}{\text{BW}_{ij}^{u,n}} \right) + \tau_{ij}^{\text{prop}} \right) \right] \tag{6.8}$$

where the energy consumed per link ij is power $P_{i,n}$ used by node i to transmit the task over the hop multiplied by the time that the node is transmitting, which translates to the latency experienced by the entire task on that link. The entire transmission energy is this per link energy summed over all links in the path and all paths from the source to destination. However, since we only consider the wireless transmission energy, only the first hop of the path is included in the energy calculations and this link does not include a propagation latency. Therefore, the transmission energy consumed by the network transforms to the wireless transmission energy consumed by the user device

$$e_u^{\mathrm{Tx},D} = \sum_n \sum_m x_{u,m} \left[\sum_{uj} P_{u,n} \left(\lambda_{uj}^{n,u,m} \left(\frac{k_u^n S_u}{\mathrm{BW}_{uj}^{u,n}} \right) \right) \right] \tag{6.9}$$

Besides the transmission energy, a task also consumes computation energy while being processed and is given by

$$e_u^{\mathrm{compute}} = C_u E_m. \tag{6.10}$$

where E_m is the energy consumed by a MEC server m to process one instruction. If the task is processed entirely locally, then the computation energy consumption is given by

$$e_u^L = C_u E_u \tag{6.11}$$

where E_u is the energy consumed by a user u to process one instruction. The total network energy consumption e_u can then be given by the local computation energy if the task is processed locally or the transmission and edge computation energy if the task is offloaded,

$$e_u = p_u e_u^L + (1 - p_u)(e_u^{\mathrm{Tx}} + e_u^{\mathrm{compute}}). \tag{6.12}$$

6.5.5 User Device Energy Consumption Model

While aiming to reduce network energy consumption supports the overall sustainability of the network, it may inadvertently lead to favoring local processing. This is because edge computing servers, being more powerful, consume more energy. However, this approach is not suitable for user devices that rely on battery power or are integral to emerging IoT applications. Therefore, we also consider the goal to minimize energy consumption on these user devices while ensuring that task completion meets required latency bounds. The energy to transmit the task as consumed by the user is the same as in (6.8) and the energy process a task locally at the user device is as in (6.11). Therefore, the total energy consumed by a user device e_u^D is either the local processing energy or the transmission energy, if the task is offloaded, and can be given by

$$e_u^D = p_u e_u^L + (1 - p_u) e_u^{\mathrm{Tx},D}. \tag{6.13}$$

6.5.6 Optimization Problem

This section presents the formulation of the optimization problem which seeks to minimize either latency, network energy consumption or user device energy consumption within a multipath, multihop LiFi-WiFi network, while guaranteeing the maximum latency of the tasks.

The optimization problem formulated here aims to minimize the total task completion latency across all users. The mathematical representation of this objective is:

$$\min_{p_u, x_{u,m}, k_u^n, c_{u,m}} \sum_u \tau_u \tag{6.14}$$

$$\text{subject to} \quad \tau_u <= \tau_u^{\text{req}} \qquad \forall u \tag{6.15}$$

$$\sum_m x_{u,m} = 1 \qquad \forall u \tag{6.16}$$

$$\sum_n k_u^n = 1 \qquad \forall u \tag{6.17}$$

$$\sum_u (1 - p_u) x_{u,m} c_{u,m} <= 1 \qquad \forall m \tag{6.18}$$

$$p_u \in \{0, 1\} \qquad \forall u \tag{6.19}$$

$$x_{u,m} \in \{0, 1\} \qquad \forall u, m \tag{6.20}$$

$$0 < k_u^n \leq 1 \qquad \forall u, n \tag{6.21}$$

$$0 < c_{u,m} \leq 1 \qquad \forall u, m \tag{6.22}$$

The decision variables within this formulation include p_u, a binary indicator where 1 represents local computing and 0 indicates edge computing; $x_{u,m}$, a binary variable where 1 indicates that the task of user u is offloaded to server m; $k_u^n \in [0, 1]$, the proportion of data split across the available wireless technologies; and $c_{u,m} \in [0, 1]$, the fraction of computational resources allocated to user u at server m.

The constraint (6.15) sets an upper limit on the latency for each application's task. Constraint (6.16) requires that each user's task is processed at a single destination, ensuring that all packets of the task reach the same destination. Constraint (6.17) ensures that when a task is offloaded, the sum of the data split across all available technologies equals one, confirming that the entire task is transmitted. Finally, constraint (6.18) prevents the computational load at any server from exceeding its capacity.

To support network sustainability, we include the objective of minimizing network energy

$$\min_{p_u, x_{u,m}, k_u^n, c_{u,m}} \sum_u e_u \tag{6.23}$$

and user device energy

$$\min_{p_u, x_{u,m}, k_u^n, c_{u,m}} \sum_u e_u^D \tag{6.24}$$

For these objectives, we continue to uphold the latency bound constraint (6.15), as we are addressing latency-critical tasks. Our goal is to maintain latency guarantees while also focusing on minimizing the energy consumption. Besides the latency constraint, all constraints from (6.15) - (6.18) are still part of the optimization problem. Considering the complex nature of MINLP problems, which arise from the non-linear aspects of both the objective function and several constraints, and the extensive scale of the network, it is necessary to employ advanced optimization techniques.

6.6 Methods to solve the Task Offloading Problem

6.6.1 Baselines

To evaluate the effectiveness of the ComputiFi framework, we compare it against various baseline strategies designed to address the task offloading problem in heterogeneous LiFi-WiFi networks. These baselines represent different un-optimized approaches to task distribution and provide a benchmark to assess the enhancements brought by specifically optimizing for the objectives of ComputiFi.

1. **Local-Only:** This baseline processes all computational tasks entirely on the local device. While it minimizes data transfer latency, this approach can lead to excessive processing delays due to limited local computational resources as well as increased energy consumption on user devices.

2. **AP-Only (Singlepath):** In this strategy, all computational tasks are directed to a single AP (either LiFi or WiFi) for processing. If a LiFi AP is available, it is prioritized over WiFi due to lower user density associated to one LiFi AP which would lead to higher data rates and lower latency. We also consider a variant that prefers WiFi over LiFi but this is only useful for energy optimization since the transmission power over WiFi is lower.

3. **AP-Only (Multipath):** This variation of the AP-Only strategy takes advantage of both LiFi and WiFi wireless interfaces by splitting data packets across both technologies. Data transmission is divided equally among the available technologies to reduce processing delays and congestion while maximizing the network's multipath capabilities. However, the entire task has to reach the same destination to be processed.

4. **Local-First:** The Local-First baseline attempts to process tasks on the local device initially. However, if processing the task locally would violate latency requirements, it offloads the task to an external AP following the AP-Only (Singlepath) strategy. This approach balances local computing resources with external offloading to optimize latency while maintaining reliability.

5. **URLLC-Local (Singlepath):** The URLLC-Local baseline prioritizes URLLC tasks by processing them locally on the user's device, ensuring that critical latency requirements are met. Other types of tasks are routed to a single AP following the AP-Only (Singlepath) strategy.

6. **URLLC-Local (Multipath):** Similar to the singlepath variant, this strategy ensures URLLC tasks are processed locally but distributes the traffic of other tasks across multiple technologies via a multipath approach. Data is split across LiFi and WiFi technologies to optimize resource utilization while meeting URLLC latency requirements.

6.6.2 MINLP Optimizer

A key method for solving the task offloading problem is the use of MINLP solvers, such as Gurobi [Gur23]. These solvers employ advanced branch-and-bound techniques to handle optimization problems involving both discrete and continuous variables. Gurobi systematically explores the solution space by branching into subproblems based on variable constraints and bounding suboptimal solutions to refine its search. This approach provides an optimal solution for the given problem.

However, due to the intrinsic complexity of MINLP formulations, solving the task offloading problem with such methods becomes computationally expensive for larger network scenarios. The branching process grows exponentially with the problem size, which makes MINLP solvers less practical in large, real-world networks. Instead, Gurobi is used primarily to establish an optimal solution for smaller scenarios, providing a benchmark for validating the optimality of our proposed learning-based and meta-heuristic approaches in larger and more complex cases. By comparing the results obtained from other approaches to the MINLP solution, we demonstrate the efficacy of our framework while ensuring that the proposed solutions remain within an acceptable margin of optimality.

The complexity of the full optimization formulation makes it impractical to solve directly; therefore, we have modified the problem to render it more tractable by transforming functions into linear or convex forms. Particularly, in the equations (6.2) and (6.3), we need to transform the max operations. To handle this within a linear framework, we introduce an auxiliary variable, $\tau^{\text{flow}}_{u,n,m}$, which represents the maximum latency of data flow across all hops for a specific set of user u, technology n, and

destination m. We also implement a constraint ensuring that this variable is at least as great as the latency of each individual hop, expressed as:

$$\tau^{\text{flow}}_{u,n,m} \geq \left(\frac{\lambda^{n,u,m}_{ij} \left(k^n_u S_u - S^{\text{pkt}}_u \right)}{\text{BW}^{u,n}_{ij}} \right) \qquad \forall ij \tag{6.25}$$

This adjustment is beneficial because the objective functions seek to minimize $\tau^{\text{flow}}_{u,n,m}$, which naturally drives it to the lowest possible value while this constraint (6.25) ensures that it is greater than or equal to individual hop latency.

Similarly, we apply transformations to (6.3) to linearize the max operation. These transformations ensure that the problem can be simplified without sacrificing the optimality of the model. Following these modifications, the revised MINLP solver implementation is referred to as the "Expert" model.

6.6.3 Meta-heuristics

In addition to benchmarking with MINLP solvers, we implement multiple meta-heuristic algorithms to solve the task offloading problem efficiently in larger and more complex network scenarios. These solvers strike a balance between solution quality and computational efficiency, offering near-optimal solutions without the prohibitive time costs of exact optimization methods. The following meta-heuristic techniques are employed:

1. **Genetic Algorithm (GA):** GA [Gol89] is motivated by the principles of natural selection. It evolves a population of candidate solutions over multiple generations through selection, crossover, and mutation operations. This approach enables GA to explore a wide search space, converging towards optimal solutions by favoring well-performing individuals.

2. **Differential Evolution (DE):** DE [PSL05] is an evolutionary algorithm that optimizes complex functions by iteratively improving a set of candidate solutions. It uses vector differences to generate new candidate solutions and maintains diversity within the population.

3. **Pattern Search:** This derivative-free optimization technique [HJ61] iteratively explores the solution space by evaluating neighboring points based on a predefined pattern. It adjusts the pattern size dynamically based on progress, allowing the search to adapt its exploration-exploitation balance for faster convergence.

4. **Particle Swarm Optimization (PSO):** PSO [KE95] is a swarm-based algorithm where particles represent candidate solutions moving through the search space based on their own experiences and those of their neighbors. The particles adjust their velocity towards the best positions found, leading to efficient convergence.

5. **Stochastic Ranking Evolutionary Strategy (SRES):** SRES [RY00] combines evolutionary principles with a ranking mechanism to handle constrained optimization problems. It ranks candidate solutions based on objective values and constraint violations, allowing it to balance feasibility and optimality.

All the meta-heuristic algorithms used in this work start with an initial population of 100 and are designed to run through a maximum of 100 generations. However, these algorithms may end earlier if convergence is reached before the maximum number of generations.

6.6.4 Black-box optimizers

Alongside meta-heuristic algorithms, we also utilize black-box optimization techniques, which are advantageous for addressing problems characterized by complex or unknown structure. These

techniques consider the optimization function as a black box, using statistical learning to map input-output relationships effectively. Our framework integrates the following black-box optimization methods:

1. **Random Forest:** Originating from ensemble learning theory, Random Forest, [Bre01] constructs multiple decision trees and aggregates their predictions. Each tree is trained on a random subset of the dataset, providing a diverse collection of decision rules that improve generalization. This diversity allows Random Forest to effectively handle complex, nonlinear task offloading problems.

2. **Extra Trees:** Extra Trees [GEW06], also known as Extremely Randomized Trees. Unlike Random Forest, where each tree is trained on a random data subset, Extra Trees uses the entire training set and applies a randomized selection of split points. This method often leads to greater diversity in tree structures and improved predictive accuracy.

Additionally, our framework proposes a Greedy method that uniformly samples within variable bounds and iteratively repeats this selection process to obtain the solution.

6.6.5 Deep Reinforcement Learning

The task offloading challenge involves multiple decision variables and constraints, necessitating a solution that adapts effectively to the changing dynamics of network conditions and diverse task demands. DRL is employed to train an agent through repeated interactions with the network, enabling the development of policies that optimize network performance. This approach achieves a balance between optimal solutions and practical real-time implementations.

Among various DRL techniques, Proximal Policy Optimization (PPO) [Sch+17], is recognized for its reliability and stability. PPO maintains policy updates within a small range to prevent significant changes in performance after updates. Its capability to handle complex decision-making in environments with high-dimensional action spaces makes it well-suited for managing multipath, multihop network configurations.

This section outlines the problem formulation used in the DRL model, including the state space, action space, and reward function.

State space $\mathcal{S}$ comprises a set of features that capture the current status of the network and the attributes of the tasks involved:

$$\mathcal{S} = \{l_u, w_u, R_{u,l}, R_{u,w}, t_u, N_u^T\} \tag{6.26}$$

In this formulation: w_u and l_u indicate the WiFi and LiFi APs, respectively, to which each user is connected. $R_{u,w}$ and $R_{u,l}$ represent the data rates of WiFi and LiFi connections for each user. t_u identifies the type of task being handled (e.g., eMBB, URLLC, factory automation, AI/image classification). N_u^T specifies the total number of tasks generated by each user.

Action space $\mathcal{A}$ defines the possible actions taken by the learning agent:

$$\mathcal{A} = \{x_u, k_u^W, c_u\} \tag{6.27}$$

Here, x_u determines the computational destination for each user's task. Rather than using a binary variable or action for each possible user-destination pair ($x_{u,m}$), we simplify this by directly specifying the destination index for each user (x_u). This approach reduces the complexity of the variable space and facilitates learning. Consequently, the need for the single destination per user constraint (6.16) is made unnecessary.

k_u^W represents the fraction of data transmitted via the WiFi connection, with the remainder $(1 - k_u^W)$ either transmitted through LiFi or solely via WiFi if LiFi connectivity is not available. This setup removes the requirement for the sum over technologies constraint (6.17).

c_u indicates the proportion of computing resources allocated to each user at the server. After an action is chosen, the proportion of computing resources allocated is normalized manually to ensure that the total does not exceed the server's capacity, effectively making constraint (6.18) redundant.

In the discrete model, these proportions are quantized into 20 distinct values, while in the continuous model, they vary continuously from 0 to 1.

Reward function G is designed to minimize either latency or energy consumption and impose penalties for exceeding latency limits. Latency violations are quantified by

$$V_{\tau_u} = \max(0, \tau_u - \tau_u^{\text{req}}) \tag{6.28}$$

Here, τ_u represents the actual task completion latency for a user, and τ_u^{req} denotes the latency threshold. The reward for each user, G_u, inversely correlates with their task completion latency, network energy consumption, or user device energy usage, defined as follows:

$$G_u = \frac{1}{\tau_u} \tag{6.29}$$

$$= \frac{1}{e_u} \tag{6.30}$$

$$= \frac{1}{e_u^D} \tag{6.31}$$

These rewards are normalized to a scale of $[0, 10]$ to facilitate a stable and efficient learning process:

$$G_u^s = \frac{G_u}{\max(G_u)} \times 10 \tag{6.32}$$

Normalizing the reward helps balance the agent's exploration and exploitation, enhancing overall performance and accelerating convergence.

In scenarios with latency violations, the overall reward, G, is adjusted by penalizing based on the number of users who exceed their latency limits:

$$G = -N_u^V \times 10 \tag{6.33}$$

where N_u^V indicates the number of users with latency violations. Conversely, if there are no violations, the total reward is the sum of all individual scaled rewards:

$$G = \sum_u G_u^s \tag{6.34}$$

To summarize, the reward is:

$$\text{Reward} = \begin{cases} -N_u^V \times 10 & \text{if latency violation} \\ \sum_u G_u^s & \text{else} \end{cases} \tag{6.35}$$

This reward framework is designed to motivate the agent to minimize the objective effectively through optimal task distribution and resource allocation while complying with latency constraints of each task.

The algorithm operates by first gathering experiences from the environment under the existing policy, and calculating the advantages of these experiences using Generalized Advantage Estimation (GAE). PPO then updates the policy network by aiming to maximize these computed advantages while keeping the adjustments within a small range. This control is achieved through a clipping parameter set at 0.2. Furthermore, PPO updates a value network that estimates the expected returns, which are used in determining the advantages. Optimization of both the policy and value networks is conducted through gradient descent using the Adam optimizer. For this implementation, the networks for the actor (policy) and the critic (value function) are structured as a MultiInputPolicy with

a three-layer neural network configuration, consisting of [64, 128, 64] neurons in each layer respectively. The PPO algorithm was executed using the stable-baselines3 package [Raf+21], version 2.3.0, and the environment was configured using the gymnasium library version 0.29.1. All hyperparameters involved in training the model were fine-tuned with the aid of the Optuna framework [Aki+19]. The training process involved a total of 2.34×10^5 samples, employing a learning rate of 0.0001. A detailed explanation of the PPO algorithm's operation is provided in Algorithm 6.

Algorithm 6 Optimization using the PPO Algorithm

Initialize: Policy network π_θ, value network V_ϕ, replay buffer $\mathcal{B}$, learning rate $\alpha = 0.0001$, discount factor $\gamma = 0.9$, PPO clipping parameter $\epsilon = 0.2$, Factor for trade-off of bias vs variance for GAE $\lambda = 0.985$
Input: State space $\mathcal{S}$, action space $\mathcal{A}$, reward G
while not converged **do**
    Reset environment, get initial state s_0
    for each episode **do**
        for each time step t **do**
            Select action $a_t \sim \pi_\theta(s_t)$
            Execute action a_t, observe reward G_t and next state s_{t+1}
            Store (s_t, a_t, G_t, s_{t+1}) in $\mathcal{B}$
        end for
    end for
    Compute advantages $\hat{A}_t$ using GAE:

$$\hat{A}_t = \sum_{l=0}^{T} (\gamma\lambda)^l \delta_{t+l}$$

$$\delta_t = G_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t)$$

    Update policy network θ by maximizing:

$$L^{\text{CLIP}}(\theta) = \hat{\mathbb{E}}_t \left[\min \left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \hat{A}_t, \right. \right.$$

$$\left. \left. \text{clip} \left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, 1-\epsilon, 1+\epsilon \right) \hat{A}_t \right) \right]$$

    Update value network ϕ by minimizing:

$$L^{\text{value}}(\phi) = \hat{\mathbb{E}}_t \left[(G_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t))^2 \right]$$

    Perform gradient descent on θ and ϕ using Adam optimizer
    Clear replay buffer $\mathcal{B}$
end while
Return: Optimized policy network π_θ and value network V_ϕ

6.7 Evaluation Methods and Quality Metrics

The ComputiFi framework is evaluated through detailed simulations conducted using Python 3.10.12. The framework is tested across three distinct network architectures — Small, Medium, and Large — each differing in size and configuration to reflect various potential deployment environments. The user and AP configurations of these architectures are as already presented in the wireless resource allocation chapters. The network setups are detailed in Table 6.5, where the task offloading capabilities

are explored in terms of network layout, the number of users, and AP distribution. Each setup includes LiFi APs installed at a ceiling height of 3 m, and a single WiFi AP centrally located on the ceiling at coordinates (0,0,3) m. Additionally, each architecture integrates a cloud server with all network devices equipped with processing capabilities.

Table 6.5 Task offloading architectures under evaluation

Scenario	Size	Users	LiFi APs	(x,y) Coordinates LiFi AP	Routers
Small	5 x 5 x 3 m	10	4	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4)	1
Medium	10 x 5 x 3 m	15	8	(-1.4, 1.4), (1.4,-1.4), (-1.4,-1.4), (1.4, 1.4), (-1.4, 4.2), (1.4,-4.2), (-1.4,-4.2), (1.4, 4.2)	2
Large	10 x 10 x 3 m	20	16	(-3.6,-3.6), (-3.6,-1.2), (-3.6, 1.2), (-3.6, 3.6), (-1.2,-3.6), (-1.2,-1.2), (-1.2, 1.2), (-1.2, 3.6), (1.2,-3.6), (1.2,-1.2), (1.2, 1.2), (1.2, 3.6), (3.6,-3.6), (3.6,-1.2), (3.6, 1.2), (3.6, 3.6)	4

The evaluations were performed using an 11th Generation Intel® Core™ i7-11700 16-Core Processor and an NVIDIA GeForce RTX 3070 GPU. Each set of results was derived from 20 independent simulation runs, with each run consisting of 120 time steps. To validate our findings, we employed the Mann-Whitney U test [MW47], operating under the null hypothesis that there is no difference in the distributions of the two parameters being compared. As in the previous chapters, the outcomes of these tests are annotated in the corresponding figures with symbols based on the p-values [Cha+22], using the following star notation:

$$\begin{aligned} \text{ns} &: p > .05 \\ * &: .01 < p <= .05 \\ ** &: .001 < p <= .01 \\ *** &: .0001 < p <= .001 \\ **** &: p <= .0001 \end{aligned}$$

Furthermore, due to multiple hypothesis testing on the same dataset, we applied the Benjamini-Hochberg procedure [BH95] to adjust for the false discovery rate, ensuring the reliability of our statistical inferences.

Simulation results are assessed using a variety of quality metrics to analyze the performance of the offloading framework comprehensively. For latency-related metrics, we calculate the task completion latency for each user's task, denoted as τ_u.

Task Latency (ms):

$$\text{Task Latency} = \frac{1}{M^U} \sum_{u=1}^{M^U} \tau_u \tag{6.36}$$

Latency QoS: This metric is derived by comparing the required latency against the actual latency for

each user, averaging these ratios across all users. This metric serves as an indicator of how well the system meets the latency expectations of its users.

$$\text{Latency QoS} = \frac{1}{M^U} \sum_{u=1}^{M^U} \frac{\tau_u^{\text{req}}}{\tau_u} \tag{6.37}$$

A higher Latency QoS score indicates superior system performance, showing that the latencies experienced are close to or lower than the expected latencies, thereby fulfilling user expectations. On the other hand, a lower Latency QoS score points to deficiencies in meeting latency demands, potentially affecting the user experience.

Latency QoS per application: We also calculate the Latency QoS for different types of applications by categorizing users based on their specific applications. For instance, for URLLC tasks,

$$\text{Latency QoS}^{\text{URLLC}} = \frac{\tau_{u'}^{\text{req}}}{\tau_{u'}} \qquad \forall u' \in \mathcal{U}^{\text{URLLC}} \tag{6.38}$$

For the energy-related metrics we first calculate the energy consumed by the network e_u in completing the task and also the user device energy consumed e_u^D.

Task Energy (J):

$$\text{Task Energy} = \frac{1}{M^U} \sum_{u=1}^{M^U} e_u \tag{6.39}$$

User Device Energy (J):

$$\text{User Device Energy} = \frac{1}{M^U} \sum_{u=1}^{M^U} e_u^D \tag{6.40}$$

Additional metrics are provided to assess the functionality of the optimization algorithms used:

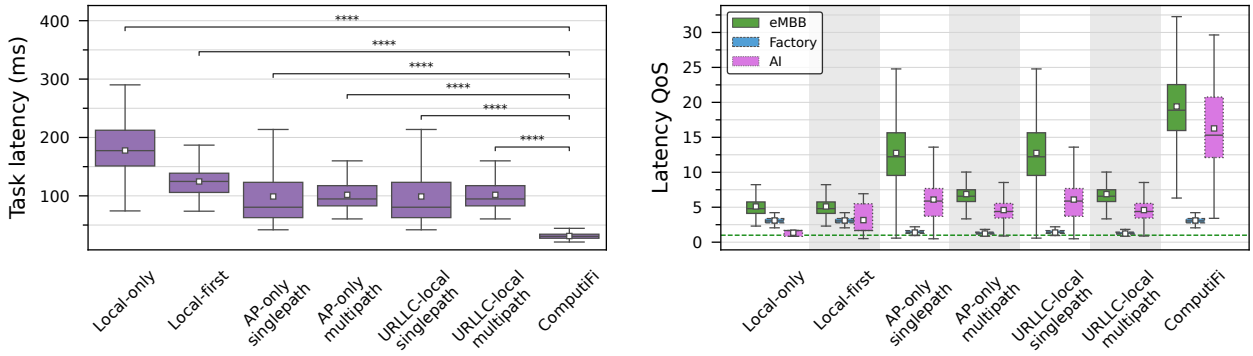
MEC destination: We track the MEC server destination to which each user's task is offloaded to understand the distribution and usage of network resources.

Proportion of data flow: To gauge the effectiveness of utilizing both LiFi and WiFi in a multipath setting, we examine the proportion of each user's data that is offloaded via each technology, with values ranging from 0 to 1. This analysis excludes local processing data which does not contribute to the metric.

Time to solve (s): Lastly, we measure the time required to solve the optimization problem with the implemented algorithms.

6.8 Results and Comparative Analysis

The first part of this evaluation deals with the results obtained from the latency minimization objective.



(a) Task completion latency

(b) Latency QoS per application where a QoS of 1 (green dotted line) indicates user's latency requirement is exactly satisfied

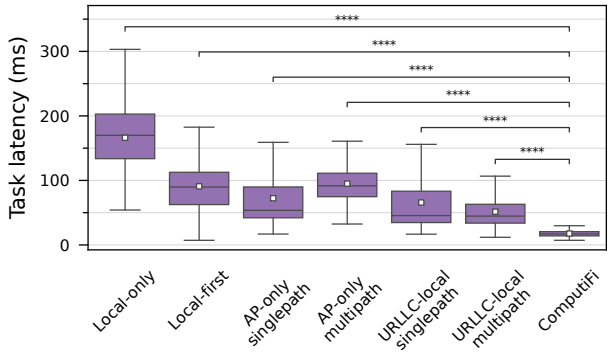
Figure 6.2 Network quality metrics for the baseline algorithms and ComputiFi's optimized task offloading solved with the Expert for the Small scenario with eMBB, Factory, and AI applications showing the need for our proposed optimization

The analysis begins with Figure 6.2, comparing various baseline methods and the proposed optimization employing the MINLP solver. This small network configuration assigns ten users randomly across four applications: eMBB, URLLC, Factory, and AI. In this scenario, where random allocation and limited user count are factors, only the eMBB, Factory, and AI applications are assigned.

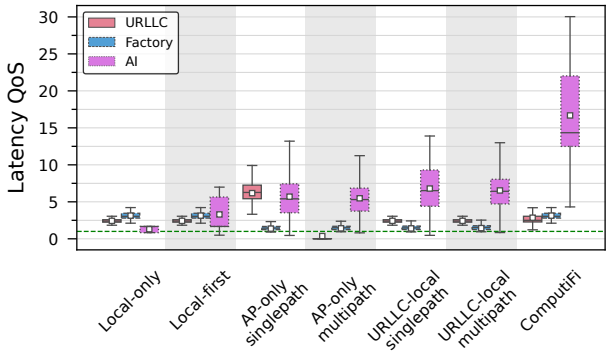
In Figure 6.2a, the Expert optimizer provides the lowest latency for user tasks, highlighting the effectiveness of our optimization strategy. It shows a 69.3% reduction in average latency compared to the best baseline method. The inclusion of statistical annotations supports the significant performance differences between ComputiFi and baseline approaches, affirming the optimization's importance in computational task offloading.

Further examination indicates that the Local-Only strategy experiences significant latency with substantial variance, illustrating the drawbacks of processing all tasks locally. The Local-First strategy marginally decreases this latency by offloading tasks only when local processing fails to meet latency requirements. Examining singlepath and multipath strategies for both AP-only and URLLC-local, a small decrease in average latency is observed with the singlepath approach, although the upper limit remains higher than that of the multipath approach. This suggests that multipath transmission reduces latency, although this is not consistently the case as part of the data flow must navigate through the backhaul to the destination AP, adding extra delay. Additionally, in this instance, both the URLLC-local and the AP-only strategies are identical due to the absence of URLLC application users.

Upon evaluating the overall performance of different baseline strategies and the proposed Expert solution in reducing task latency, we perform a per-application analysis by examining the QoS metric. Figure 6.2b offers detailed insights into the latency performance for three specific application types: eMBB, Factory, and AI. A QoS value of 1 or above indicates that latency requirements are met, with the green line across the plot representing this threshold. For all applications, the Expert optimizer consistently achieves a QoS value above 1, ensuring it meets the latency requirements. Typically, the latency requirements for eMBB application are always satisfied even with the baseline approaches due to its relatively lenient limit of 500 ms. For the Factory application, most baselines struggle to reach the satisfaction level of 1. For the AI application, given its large task size, baseline approaches often fail to consistently meet latency requirements. The Expert optimizer reliably maintains a QoS value above 1 for all application types.



(a) Task completion latency

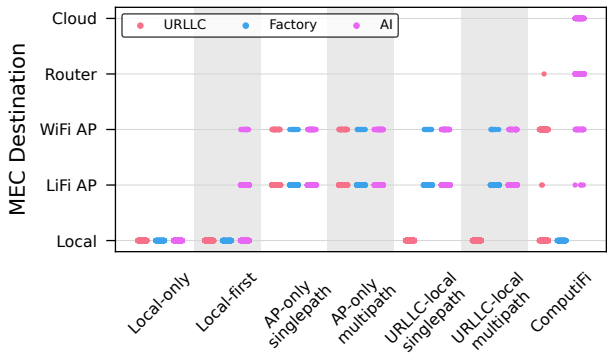


(b) Latency QoS per application where a QoS of 1 (green dotted line) indicates user's latency requirement is exactly satisfied

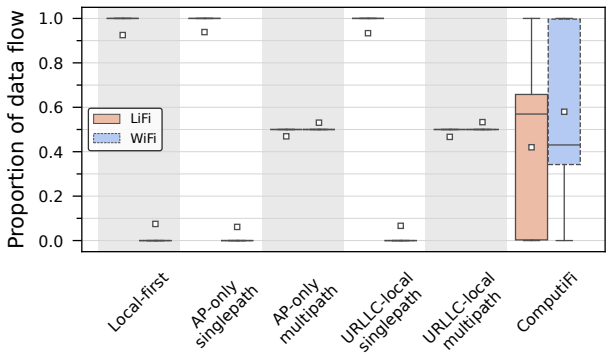
Figure 6.3 Network quality metrics for the baseline algorithms and ComputiFi's optimized task offloading solved with the Expert for the Small scenario with URLLC, Factory, and AI applications showing the efficacy of our proposed optimization for various application scenarios

To corroborate these findings in a network incorporating the more stringent URLLC application, tasks for users are exclusively generated from URLLC, Factory, and AI, with results depicted in Figure 6.3. Here again, the optimized approach achieves the lowest latencies as shown in Figure 6.3a, with a 65.8% reduction in average latency compared to the top-performing baseline (URLLC-local multipath).

Interestingly, multipath transmission enhances performance for the URLLC-local method compared to the AP-only strategy. This advantage is better explained when examining the latency QoS per application. Local processing of the strictly bounded URLLC application mitigates congestion on wireless links, significantly benefiting the URLLC-local method via multipath transmission. This also results in complete user satisfaction across almost all instances.



(a) MEC destination to which each task is offloaded



(b) Proportion of the traffic of each task transmitted with each technology

Figure 6.4 A detailed look into the working of the baseline algorithms and ComputiFi's optimized task offloading solved with the Expert for the Small scenario with URLLC, Factory, and AI applications

An in-depth analysis of algorithm performance under the URLLC-focused scenario is presented in Figure 6.4. Here, the chosen destinations for task processing by the algorithms are illustrated in Figure 6.4a. The AI application, representing the largest task, is consistently offloaded except in scenarios employing the Local-only strategy. The optimized solution efficiently leverages all categories of MEC servers to process AI tasks, significantly reducing latency. Interestingly, the optimized solution offloads URLLC tasks further, indicating that multipath LiFi-WiFi networks can reliably

deliver ultra-low latency. The Factory task is entirely processed locally by the Expert algorithm, which may appear counter-intuitive since both URLLC and Factory tasks could potentially be offloaded while still adhering to latency constraints. Nonetheless, the primary goal of the optimization problem extends beyond merely satisfying the latency requirement to minimizing it as much as possible. The URLLC task, with a stringent 1 ms requirement, already operates at a minimal latency, thus no further reduction is possible. In contrast, the Factory task's 100 ms threshold allows for a reduction to 20-30 ms, considerably lowering overall network delay. Therefore, our framework opts to process Factory tasks locally to significantly conserve time, while strategically offloading some URLLC tasks to maintain them just below the 1 ms threshold. Another factor influencing this decision is the data stream size. The URLLC tasks comprise a single 32-byte packet, facilitating fast offloading without overwhelming the link. Conversely, Factory tasks involve 20 packets of 1500 bytes each, making them slower to offload. Coupled with a higher packet arrival rate, local processing of Factory tasks is essential for optimal performance.

The usage of multiple paths for offloading is evidenced by the proportion of data flow shown in Figure 6.4b. The baseline multipath algorithms always split the data equally between available technologies. Due to the possibility of blockages in LiFi, the proportion of data flow through it is not always exactly 0.5. The optimized solution balances the use of both LiFi and WiFi networks. The median transmission is higher for LiFi due to its higher data rate, but the average is higher for WiFi due to some samples with LiFi blockages and when the offered rate through LiFi is low. This approach ensures stability and consistently low latency by using multipath offloading.

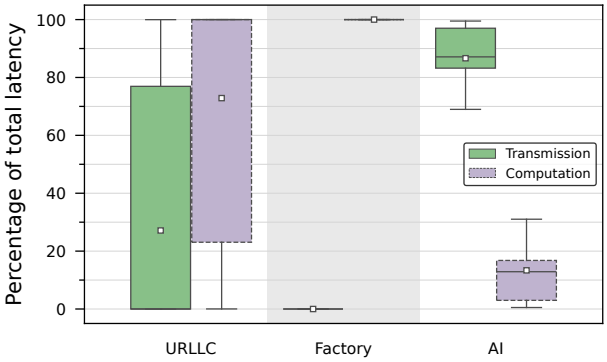


Figure 6.5 Percentage of the individual components of latency per application using ComputiFi's optimized task offloading solved with the Expert for the Small scenario with URLLC, Factory, and AI applications

Figure 6.5 shows how latency is split between transmission and computation for URLLC, Factory, and AI applications. For the URLLC, transmission latency varies from minimal percentages reflecting local processing, with a median at zero percent, to a high of 99.7%, indicative of efficient task offloading to remote servers with better computational power. For Factory applications, processed solely locally, computation latency comprises the total delay. Conversely, the AI application, predominantly offloaded, shows a significantly reduced computation latency. This analysis shows the need for minimizing both transmission and computation delays to achieve the best performance balance in real-time applications, a target ComputiFi successfully meets.

Despite confirming the advantages of our optimized task offloading approach, its practicality for real-time applications is limited due to long solution times, typically spanning several minutes. Therefore, we explore the performance of meta-heuristics as alternatives and evaluate their outcomes against the optimal solver in Figure 6.6. The discrete versions of both the GA and DE algorithms perform remarkably well, closely approaching the optimal solution. Upon detailed analysis, a slight yet statistically significant difference is observed between the GA and DE in terms of the QoS metric. Consequently, the GA is selected as the most effective algorithm, showing non-significant difference in performance when compared to the Expert solver-based approach.

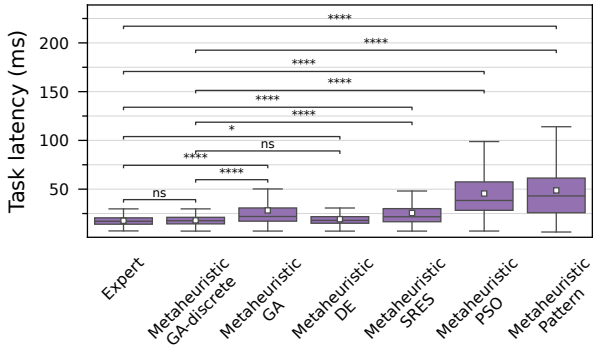
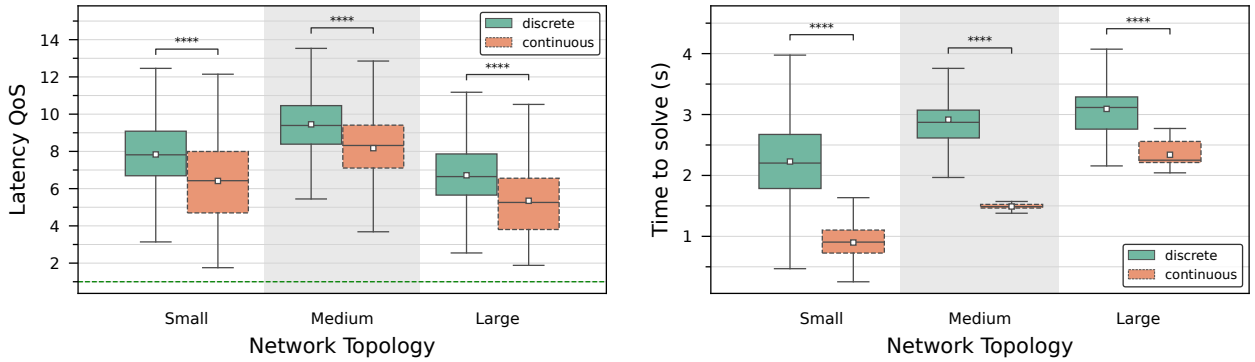


Figure 6.6 Task completion latency achieved by all the meta-heuristic algorithms under test in comparison with the optimal Expert solution for the Small scenario with URLLC, Factory, and AI applications

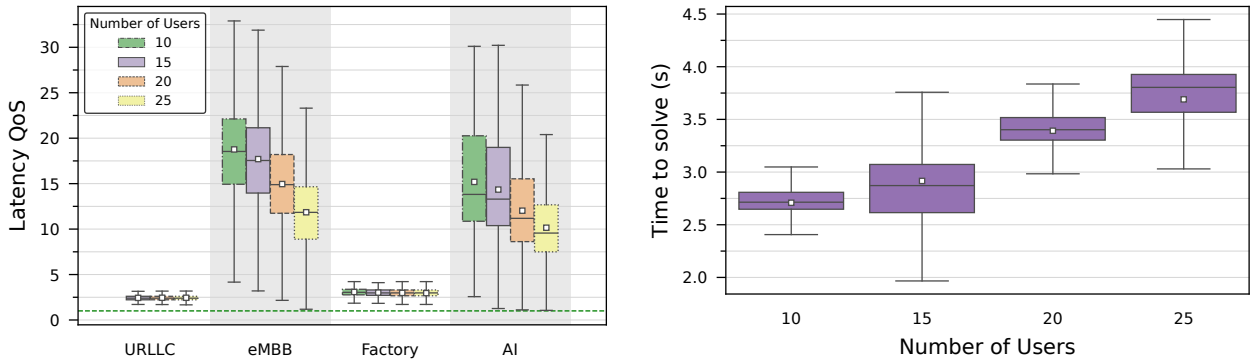


(a) Latency QoS

(b) Time to solve

Figure 6.7 Quality metrics for various network architectures using the discrete and continuous GA showing the superiority of the model with discrete variables

Motivated by the performance of the GA algorithm, further investigations are conducted to assess its efficiency across various network architectures as shown in Figure 6.7a. These evaluations confirm that the discrete model of the GA surpasses the continuous one in performance. Nevertheless, both models reliably meet user demands in every network configuration. Despite the better performance of the discrete model, the continuous model may still be preferable in situations where faster solution times are crucial, as it demands less computational effort.



(a) Latency QoS per application

(b) Time to solve

Figure 6.8 Quality metrics for increasing number of users using the discrete GA in a Medium architecture

To assess the scalability of the proposed GA algorithm as the number of users increases, we maintained the number of offloading destinations constant within a Medium network architecture and visualize the results in Figure 6.8. With only 10 users, the URLLC application is not included due to random application assignments. For URLLC and Factory applications, QoS remains stable even as the number of users grows, as illustrated in Figure 6.8a. This stability indicates that the algorithm manages additional workload efficiently without degrading these applications' performance. In contrast, eMBB and AI applications show a declining QoS trend as more users are added. This decrease reflects the increasing demand on network resources. Despite this, the system continues to meet all user requirements even with up to 25 users. Furthermore, analysis of the solution time in Figure 6.8b shows a linear increase. This rise is due to more variables and constraints being added as the user count grows, requiring the framework to handle more user demands with the same network setup. The linear trend indicates that the GA's performance scales effectively with user numbers. However, this also suggests that a higher number of users might lead to much longer solution times, possibly necessitating more efficient algorithms.

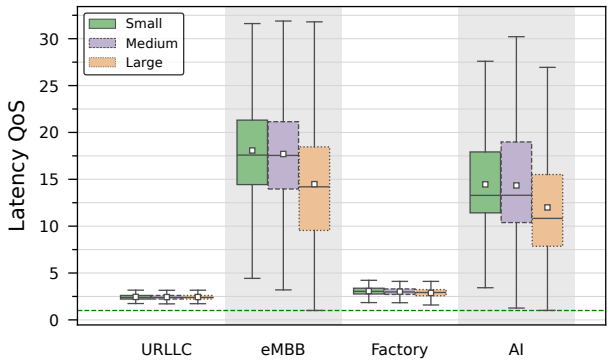


Figure 6.9 Latency QoS per application for varying the number of compute destinations while keeping the number of users fixed at 15 using the discrete GA

To further test ComputiFi's performance with an increased number of computing destinations, while keeping user numbers constant at 15, we conducted evaluations depicted in Figure 6.9. The results show a rise in latency with more LiFi APs, even though the number of users and tasks does not change. By expanding the number of compute destinations, we also had to enlarge the indoor environment to reduce interference and allow a fair comparison. Consequently, the distance between users and a WiFi AP grows, increasing transmission latency, especially when the LiFi signal is obstructed. Moreover, having more APs increases the likelihood of users being within LiFi AP interference zones.

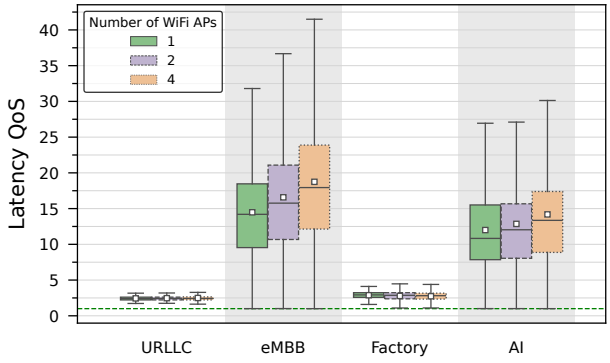


Figure 6.10 Latency QoS per application for varying the number of WiFi APs while keeping the number of users fixed at 15 using the discrete GA

Unlike LiFi APs, WiFi APs do not interfere with each other. We explore the impact of adding more WiFi APs while keeping the user count at 15 in Figure 6.10. As anticipated, QoS improves with more APs available, reducing latency due to increased offloading options. Additionally, the use of frequency reuse techniques prevents interference, ensuring that users in overlapping AP regions are not affected.

Although we have demonstrated that the meta-heuristic GA algorithm can scale effectively, its processing time of several seconds is too long for scenarios requiring optimization within 500 ms. This limitation has led us to consider a DRL-based strategy for more efficient task offloading. Initially, we explore the training process's convergence and present the training rewards in Figure 6.11, comparing the performance of discrete and continuous action models within the DRL framework. The analysis shows that the discrete model achieves a higher reward.

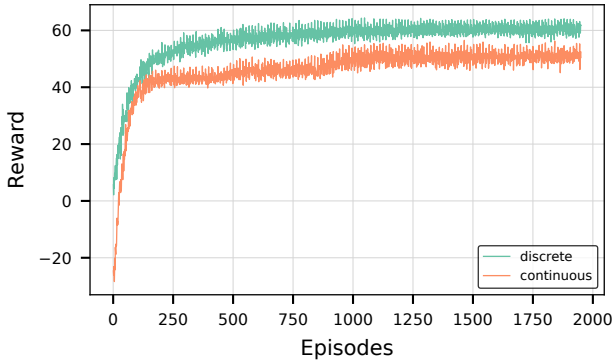


Figure 6.11 The reward obtained during the training episodes of the DRL-based PPO algorithm with discrete and continuous action space in the Small scenario with eMBB, Factory, and AI applications

Following the model's successful convergence, we apply the cross-validated model to a test set and assess the performance across various network architectures as shown in Figure 6.12. The results confirm that, similar to the training rewards, the discrete model consistently surpasses the continuous one in delivering higher QoS across all tested network layouts. This reinforces the discrete model's effectiveness in managing task offloading tasks better, while maintaining the necessary QoS for all users.

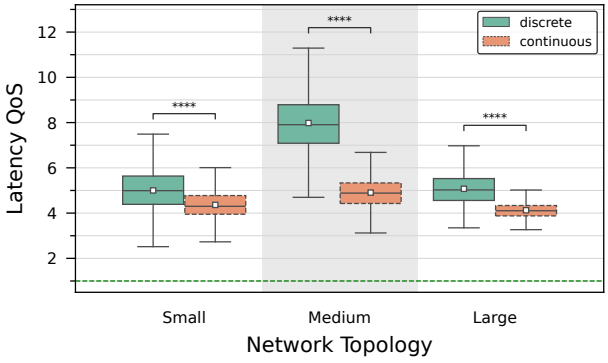


Figure 6.12 Latency QoS for various network architectures using the DRL-based algorithm with discrete and continuous actions

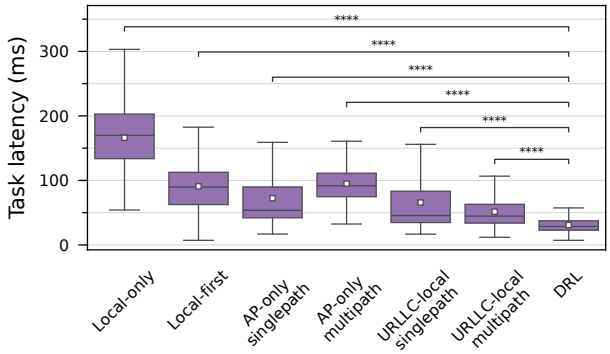
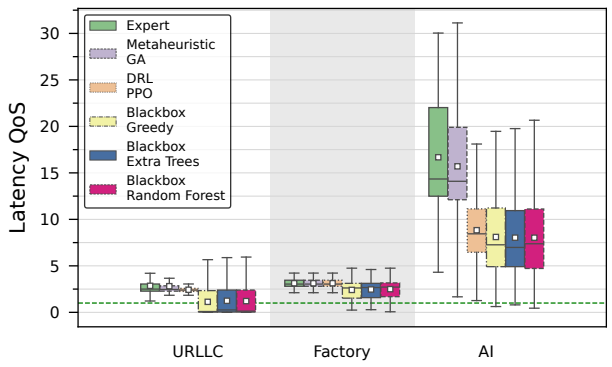


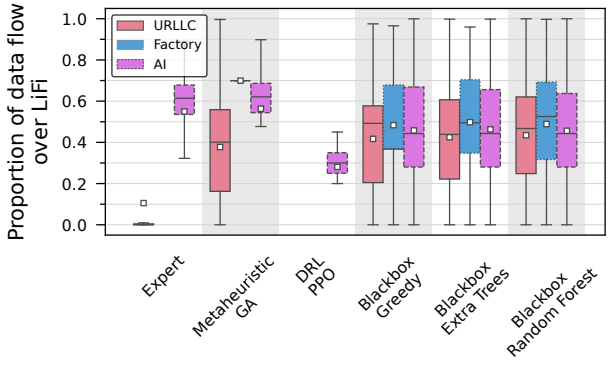
Figure 6.13 Latency for the URLLC, Factory, and AI applications comparing the DRL-based algorithm with the baseline algorithms

Furthermore, we establish that the DRL-based algorithm outperforms the baseline algorithms, as depicted in Figure 6.13. The DRL solution achieves the lowest latency, reducing the average latency by 40.23% compared to the most effective baseline method, the URLLC-local multipath.

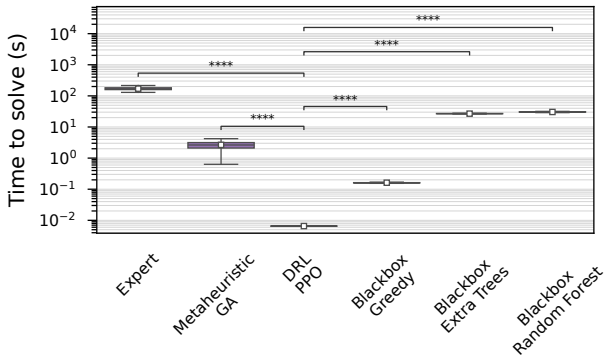
The solution time for the DRL-based algorithm, ranging from 7 ms to 12 ms across different network sizes, illustrates its practicality for real-world hardware implementation. This rapid decision-making capability of the DRL approach not only aligns with the QoS demands but also highlights its appropriateness for mobile and time-critical network environments.



(a) Latency QoS per application



(b) Proportion of the traffic of each task transmitted with LiFi



(c) Time to solve

Figure 6.14 Quality metrics for all proposed algorithms in the Small scenario with URLLC, Factory, and AI applications

In a detailed comparison of all evaluated optimization algorithms, illustrated in Figure 6.14, we include an assessment of Blackbox algorithms, which have not been previously discussed. These algorithms are crucial for determining whether a deep understanding of the problem or function being tested is necessary.

Results from Figure 6.14a show that addressing the optimization problem or objective function as a blackbox does not achieve optimal results. Additionally, apart from the basic greedy algorithm, Blackbox methods tend to take significantly longer to solve problems. This shows the advantages of employing specialized algorithms, such as the GA or DRL-based methods, which not only fulfill QoS demands more efficiently but also operate within reasonable time frames.

As seen in Figure 6.14a, the three leading algorithms in terms of network performance are the MINLP solver-based optimal solution, the meta-heuristic GA, and the DRL-based solutions. These algorithms perform similarly well for URLLC and Factory applications. However, for the AI application, the DRL solution exhibits a 47.03% reduction in QoS compared to the other two, yet it still meets all user requirements.

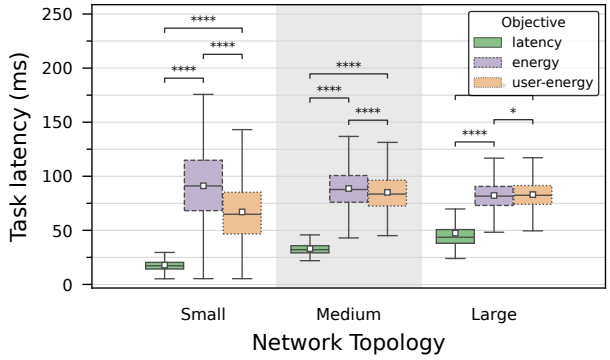
To understand the behavior of these different approaches, Figure 6.14b visualizes the data transmission proportions over LiFi, if offloaded, for the URLLC, Factory, and AI applications. The proportion of WiFi is the inverse of LiFi unless the task is processed entirely locally. Blackbox optimization techniques show similar characteristics, supporting the notion that treating the objective function as a blackbox leads to outcomes comparable to those of a simple greedy approach. The Expert strategy primarily offloads URLLC tasks through WiFi, selecting this destination unless processed locally. In contrast, AI tasks are offloaded to routers or cloud services due to their higher processing capabilities, making greater use of the LiFi connection due to its faster data rate. Factory tasks are entirely processed locally. The DRL method, meanwhile, processes URLLC and Factory tasks locally but offloads AI tasks to various locations, differing from the Expert by primarily using WiFi for offloading, as reflected in the lower latency QoS for the AI application.

While the top-performing algorithms in network performance have been identified as the MINLP solver-based optimal solution, the meta-heuristic GA, and the DRL-based algorithm, it is necessary to also evaluate them based on the time required to solve optimization problems. Figure 6.14c displays these solve times in seconds, using a logarithmic scale to accommodate the broad range of durations. The Expert algorithm, with its longest solve time, proves impractical for real-world applications. A similar conclusion applies to the Extra Trees and Random Forest algorithms. Although the meta-heuristic solutions, as shown in Figure 6.7 and Figure 3.17, resolve within a few seconds, they still fall short of efficiency for practical deployment. Conversely, the Greedy algorithm, while faster, performs inadequately in network efficiency as evidenced in Figure 6.14a. The DRL solution emerges as the most time-efficient, significantly outperforming the other methods. This efficiency is confirmed by statistical analyses comparing the DRL approach with other algorithms. This speed renders the DRL solution particularly appealing as it offers an optimal balance between network performance and processing time, making it an optimal candidate for real-time implementation on actual hardware.

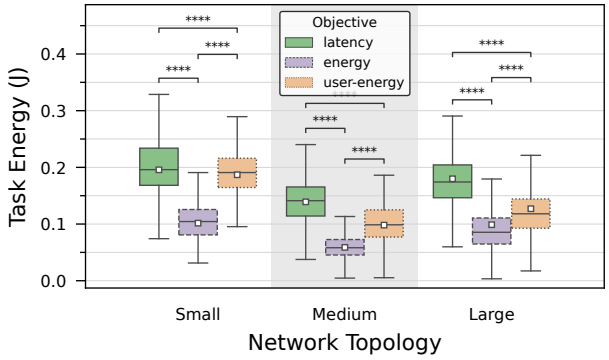
Energy consumption is another critical factor in choosing an algorithm. Excluding the MINLP solver-based and blackbox optimization methods due to their complexity and subpar network performance respectively, we compare the GA and DRL solutions. The GA consumes approximately 4.244×10^{-6} kWh per optimization run, while the DRL uses 0.0674 kWh. However, most of the DRL energy expenditure is associated with one-time training and validation. If the training and validation can be implemented on a more powerful and less energy-restricted platform, the operational phase of the DRL during runtime (on test data) consumes only about 3.989×10^{-7} kWh per run. Energy measurements were made using the Running Average Power Limit interface on Intel processors and NVIDIA-smi for GPU monitoring.

After examining the latency objective of our task offloading framework, our focus shifts to the energy objectives, which are equally crucial for the sustainability and efficiency of network operations. We compare the three objectives of minimizing task completion latency, network energy consumption and user device energy consumption in Figure 6.15. Figure 6.15a clearly shows that optimizing for the

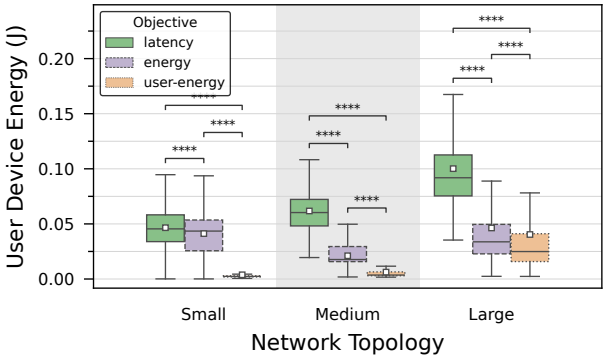
energy consumption does not automatically result in the lowest task completion latency. This finding carries over to the metrics of network and user energy as well. These results clearly show the need to tune the objective function according to the needs of the network. However, in all these cases, the latency QoS of all the applications are still met so the energy objectives still meet all user demands while reducing the energy consumption.



(a) Task completion latency

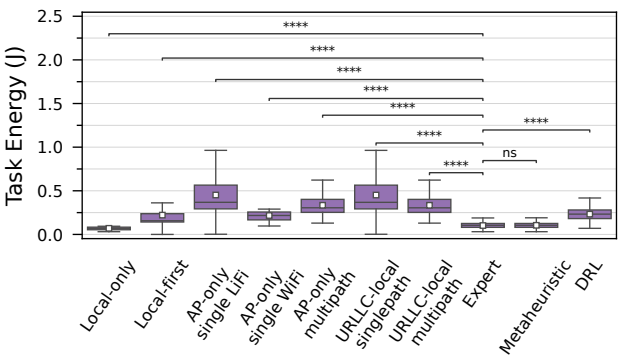


(b) Network energy consumed for task completion

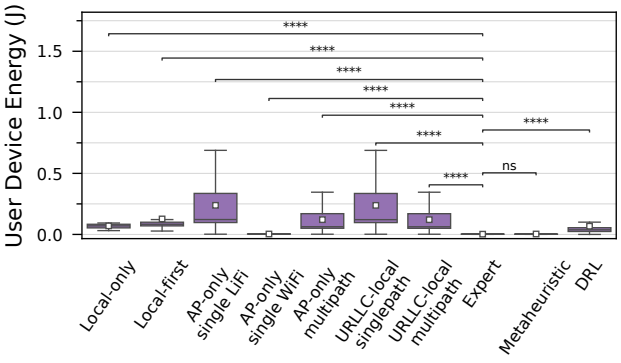


(c) Energy consumed by user device for task completion

Figure 6.15 Quality metrics comparing all proposed objective functions in the Small scenario with URLLC, Factory, and AI applications



(a) Network energy consumption when optimized for it



(b) User device energy consumption when optimized for it

Figure 6.16 Quality metrics for all proposed algorithms in the Small scenario with URLLC, Factory, and AI applications for the objective functions of minimizing network energy and user device energy

While we have shown the need to tune the objective according to network needs, we still need to compare these to the baselines to verify the need for optimization in the case of the energy objectives. This comparison is made in Figure 6.16. As expected the Local-only approach shows the best performance in terms of network energy consumption since the user devices have the low performance processors, with the optimized approaches showing comparable performance in Figure 6.16a. However, it is important to note that the Local-only approach does not satisfy the latency QoS of all users while the optimized approaches do so. Furthermore, none of the other baselines manage to always satisfy the user requirements.

Looking into the user device energy consumption in Figure 6.16b, the best performing baseline is the AP-only singlepath algorithm which prioritizes WiFi. This is because of the combination of the single technology usage, low transmission power of WiFi and the low computational energy consumption of an AP server. The Expert model still outperforms this baseline by 35.56%. However, in this user device energy minimization objective, the DRL algorithm does not satisfy the QoS requirement for the AI application in 0.6% of the cases while the average is still well above the requirement. In congruence with the results of the latency objective, the GA and Expert algorithms show non-significant differences in performance for both energy objectives.

6.9 Summary and Conclusions

This chapter has introduced ComputiFi, a framework developed to optimize task offloading within LiFi-WiFi networks, particularly focusing on reducing latency, network energy, and user device energy consumption in latency-sensitive applications. By integrating advanced optimization tools, dynamic resource allocation across diverse computational servers, and multipath transmissions, ComputiFi effectively tackles the complexities of heterogeneous network environments. Additionally, a multihop latency model appropriate for data flows involving multiple packets was discussed.

The outcomes indicate that ComputiFi not only successfully meets its designated objectives but also consistently upholds high QoS across different network configurations, highlighting its effectiveness. Specifically, our approach significantly reduced latency by 69.3% compared to the most effective baseline method in scenarios involving eMBB, Factory, and AI applications, and by 65.8% in setups with URLLC, Factory, and AI applications. Moreover, it accomplished a 35.56% reduction in average user device energy consumption against the best baseline in similar application contexts.

An extensive assessment of various optimization algorithms for task offloading within multihop, multipath LiFi-WiFi networks revealed distinct advantages and limitations of each method. Among them, the GA and DRL-based solutions stood out, with the DRL approach excelling particularly in speed, achieving prediction times in the order of milliseconds. This rapid processing capability renders it suitable for real-time applications while still managing to reduce user average latency by 40.23% compared to the best baseline approach for latency-minimized scenarios with URLLC, Factory, and AI applications. These insights highlight the importance of choosing an appropriate optimization strategy that aligns with network demands and operational conditions, ensuring optimal performance.

Chapter 7

Conclusions and Future Research

The exponential growth in wireless communication demands has necessitated the exploration of new technologies to supplement existing systems. Light-Fidelity (LiFi) technology, which uses visible light and infrared for data transmission, has emerged as a promising complement to traditional Wireless-Fidelity (WiFi) systems. The integration of LiFi and WiFi into a heterogeneous network can significantly enhance network capacity, improve coverage, and reduce latency, thereby addressing the increasing need for high-speed, reliable, and secure data transmission.

This thesis has explored the design and management of LiFi-WiFi heterogeneous networks, addressing several critical challenges like the placement of Access Points (APs) and the allocation of wireless and computational resources through novel approaches. Each research question posed at the beginning of this work has been thoroughly investigated, with solutions and methodologies developed to advance the state of knowledge in this field.

7.1 Summary

To answer **Research Question 1** on optimizing access point placement in LiFi-WiFi networks for cost efficiency, coverage, and illumination quality, we presented two significant contributions. The first contribution tackled the challenge of 3D LiFi AP placement, focusing on minimizing the number of APs while maximizing the sum rate weighted by the probability of user occurrence. By formulating this challenge as a multi-objective optimization problem with constraints on minimum data rates and required illumination levels, we solved it using Non-dominated Sorting Genetic Algorithm (NSGA-II). This analysis demonstrated that the free selection of AP heights significantly enhanced network performance compared to fixed-height placements. Moreover, incorporating expected user distribution into the sum rate evaluation was critical for optimizing network performance in densely populated user areas.

Building upon these insights, the second contribution, PlaciFi, developed a comprehensive framework for optimal 3D AP placement in LiFi-WiFi heterogeneous networks. PlaciFi aimed to maximize rate coverage while minimizing the costs of AP deployment. By utilizing the third dimension in AP placement, PlaciFi significantly improved interference management and data rate coverage over state-of-the-art 2D power optimization models. Our extensive simulations showed that PlaciFi's advanced heuristic, meta-heuristic, and black-box optimization techniques outperformed baseline solutions, particularly in scenarios with varied user distributions. This framework's adaptability was validated through its ability to tailor AP placement to different indoor application scenarios, demonstrating its effectiveness across different deployment requirements. In answering this research question, we produced the following guidelines that assist in optimizing AP placement:

1. Indoor Environment Analysis: Perform a detailed assessment of indoor spaces to pinpoint specific needs and restrictions. Analyze factors such as room dimensions and user density for both LiFi and WiFi technologies.

2. Choosing Optimization Functions: Select suitable optimization functions that align with the intended outcomes. Our work supports both Multi-Objective Optimization (MOO) and Single-Objective Optimization (SOO) techniques, utilizing the former to handle multiple competing objectives.

3. Three-Dimensional AP Deployment: Utilize our proposal's feature of three-dimensional placement of LiFi APs to optimize height variations, thereby minimizing interference and maximizing data transmission rates, ultimately enhancing network effectiveness.

4. Optimal AP Configuration: Deploy our proposal to ascertain the most strategic number and location of APs to create a balanced and effective network using different technologies.

5. Custom Deployment Strategies: Develop specific strategies for deploying technologies in varied indoor environments such as offices, corridors, and museums, ensuring that each setup meets the unique needs of the space.

6. Addressing Scalability and Complexity: Employ Dynamic Grid Explorer (DGE) for optimal results in small to medium settings. For larger environments, evaluate the system's computational intensity and storage demands, possibly incorporating meta-heuristic approaches to handle larger scales efficiently.

To address **Research Question 2** on optimizing resource management in a LiFi-WiFi network to dynamically adapt to changing user demands while ensuring minimal delay and high network stability, our work made substantial contributions across several areas of network optimization. We focused on minimizing average network packet delay for delay-sensitive applications by formulating and solving a problem to optimize wireless resource allocation in both singlepath and multipath LiFi-WiFi heterogeneous networks. The optimization, constrained by maximum allowable delays and required data rates for each user, was implemented using a Branch and Bound algorithm-based solver, alongside a meta-heuristic genetic algorithm. Extensive simulations confirmed that these methods significantly reduced network delay compared to existing max-SNR-based techniques and consistently met strict delay and data rate requirements under heavy Quality of Service (QoS) traffic.

Further, we enhanced network stability by addressing handover losses associated with mobility, changes in receiver orientation, and signal blockages. A low-complexity algorithm was developed to consider data rate loss due to handover overhead, aiming to provide stability under transient channel variations. This algorithm was extensively validated through simulations, showing a marked improvement in sum throughput and up to 24.6% reduction in number of handovers over traditional resource allocation methods. To manage unavoidable handovers, particularly in scenarios with prolonged Line-of-Sight (LoS) blockages and user mobility, we introduced a system protocol for managing vertical handovers. This protocol significantly reduced the overhead associated with these handovers, as demonstrated on a hardware setup. The results indicated that the vertical handover overhead could be minimized to avoid TCP re-connections and enhancing network stability.

Lastly, we focused on further improving the stable network's performance using a weighted alpha-fair utility maximization model, which dramatically outperformed the baseline, showing average user rate increases of 398%, 98.9%, and 52.9% for a small, medium, and large network topology, respectively.

To address **Research Question 3** on utilizing user trajectory prediction for proactive network resource management in LiFi-WiFi networks, our thesis developed the MobiFi framework. This framework significantly enhanced network performance by implementing proactive wireless resource allocation strategies that anticipate future network conditions, consistently outperforming reactive methods. The proactive strategies show a 7.7% increase in average user rate and a 63.3% improvement in the worst user rate. The MobiFi framework effectively demonstrated its adaptability and robustness across various network scenarios characterized by differing user densities and mobility patterns. A key aspect of this proactive approach is its reliance on accurate user position predictions. Our

evaluation highlighted that while inaccuracies in predictions could challenge the effectiveness of proactive strategies, the model displayed resilience to a degree, suggesting its viability with current prediction technologies. Furthermore, the work explored optimization techniques like Evolutionary Game Theory (EGT) to improve the feasibility of real-world applications by enhancing run time efficiency.

In addressing **Research Question 4** on optimizing task offloading in LiFi-WiFi networks, we developed the ComputiFi framework, aimed at minimizing latency and energy consumption while dynamically adapting to network topology changes and user demands. ComputiFi integrates advanced optimization tools, dynamic resource allocation across various computing servers, and multipath transmissions, effectively tackling the challenges within heterogeneous network environments. A specialized multi-hop latency model was particularly instrumental for capturing the behavior of data flows involving multiple packets. The framework's comprehensive approach not only met its designated objectives but also upheld high QoS across various network configurations, significantly enhancing performance in latency-sensitive applications like Ultra Reliable Low Latency Communications (URLLC), factory automation, and Artificial Intelligence (AI).

Through extensive evaluations, ComputiFi demonstrated a 65.8% improvement in latency and 35.6% reduction in user device energy consumption over the best performing baseline approach. Among the different optimization algorithms assessed for task offloading, Genetic Algorithm (GA) and Deep Reinforcement Learning (DRL)-based solutions stood out for their efficiency. The DRL strategy stood out for its quick prediction times, making it suitable for real-time applications while also achieving a significant reduction in user average latency. These results highlight the importance of selecting optimization strategies that align well with the specific demands and operational conditions of the network, ensuring optimal performance and contributing to more adaptive and energy-efficient network operations in complex LiFi-WiFi environments.

Throughout this thesis, we have systematically addressed a series of complex challenges associated with the design and management of LiFi-WiFi heterogeneous networks, paving the way for advanced 6G communications. From the strategic 3D placement of access points enhancing network coverage and efficiency in Chapter 3, to sophisticated dynamic resource management strategies that ensure minimal latency and enhance network stability in Chapter 4, each contribution has pushed the boundaries of current network capabilities. Further exploration into proactive resource management in Chapter 5 demonstrated how predictive technologies could significantly improve long-term network performance, adapting seamlessly to user behavior and environmental changes. Finally, the ComputiFi framework introduced in Chapter 6 revolutionized task offloading in multipath networks by optimizing latency and energy consumption, proving essential for real-time, energy-efficient network operations. Collectively, these contributions not only address the pressing needs of modern wireless communication systems but also set a robust foundation for future innovations in network design and management, highlighting the transformative potential of integrating LiFi and WiFi technologies in next-generation networks.

7.2 Future Work

The findings from this thesis on AP placement, resource allocation, and task offloading in LiFi-WiFi networks have opened several avenues for future research that could further refine and expand the current work in this thesis.

The optimization framework used for AP placement could be enhanced by incorporating dynamic user mobility patterns to adapt to changing environmental conditions, especially useful in outdoor scenarios. Additionally, integrating detailed energy consumption models into the optimization process could further boost energy efficiency, aligning with the goals of sustainable network design. Considering multi-operator scenarios, the framework could be expanded to accommodate the objectives of various stakeholders, allowing for a balanced optimization of network performance across different operators. There is also potential to extend the framework to include other advanced wire-

less technologies such as 5G or millimeter-wave, broadening its applicability. The incorporation of advanced machine learning techniques could significantly enhance the AP placement algorithm, enabling the system to learn from past deployments and make more intelligent decisions. This approach could lead to more effective optimization strategies in real-time, adapting to the specific needs of each deployment scenario.

The proposed resource allocation frameworks can benefit from practical implementations and real-world validations. The advent of open-source hardware compliant with emerging standards like IEEE 802.11bb will facilitate real-world testing of our resource allocation schemes. Key to this testing will be open-source LiFi modules that allow access to the Medium Access Control (MAC) layer, as well as dual-interface devices that can seamlessly switch between LiFi and WiFi. Creating a controlled lab environment with multiple LiFi and WiFi APs will enable simulation of various user scenarios, including different mobility patterns and LoS blockages. This setup will help validate the effectiveness of the resource allocation strategies through key performance metrics such as throughput and fairness. Subsequent field trials in real-world settings like office buildings or campuses will provide further insights into the strategies' practical impacts on user QoS and network performance. Collaboration with industry partners who are at the forefront of developing LiFi technology will be essential to access state-of-the-art hardware and benefit from their technical expertise. Initiatives to build such testbeds are already underway [Haa+20; Guz+23].

Looking ahead into future research in task offloading in LiFi-WiFi networks, integrating more sophisticated machine learning models that adapt based on real-time network data could dynamically adjust to changes in user behavior and network conditions. Further exploration into multi-objective formulations that minimize both energy consumption and latency could benefit battery-dependent devices and Internet of Things (IoT) applications. Enhancing bandwidth management within the task offloading framework could ensure more efficient use of network resources, addressing both throughput needs and congestion.

As we look to the future, the pathways outlined for advancing the design and optimization of LiFi-WiFi networks hold the promise of transforming wireless communication landscapes.

List of Figures

List of Tables

List of Publications

Journal Articles

[Vij+23] Hansini Vijayaraghavan, Jörg Von Mankowski, Carmen Mas-Machuca, and Wolfgang Kellerer. "PlaciFi: Orchestrating Optimal 3D Access Point Placement for LiFi-WiFi Heterogeneous Networks". In: *IEEE Access* 11 (2023), pp. 115415–115429. DOI: 10.1109/ACCESS.2023.3325097.

[VK24] Hansini Vijayaraghavan and Wolfgang Kellerer. "MobiFi: Mobility-Aware Reactive and Proactive Wireless Resource Management in LiFi-WiFi Networks". In: *IEEE Transactions on Network and Service Management* 21.6 (2024), pp. 6597–6613. DOI: 10.1109/TNSM.2024.3455105.

[VMK24] Hansini Vijayaraghavan, Jörg von Mankowski, and Wolfgang Kellerer. "ComputiFi: Latency-Optimized Task Offloading in Multipath Multihop LiFi-WiFi Networks". In: *IEEE Open Journal of the Communications Society* 5 (2024), pp. 4444–4461. DOI: 10.1109/OJCOMS.2024.3426278.

[Man+23] Jörg von Mankowski, Emre Durmaz, Arled Papa, Hansini Vijayaraghavan, and Wolfgang Kellerer. "Aerial-Aided Multiaccess Edge Computing: Dynamic and Joint Optimization of Task and Service Placement and Routing in Multilayer Networks". In: *IEEE Transactions on Aerospace and Electronic Systems* 59.3 (2023), pp. 2593–2607. DOI: 10.1109/TAES.2022.3217430.

[VVK24] Jörg Von Mankowski, Hansini Vijayaraghavan, and Wolfgang Kellerer. "RAT-mania: Advancing Multi-RAT Networks with Optimal Beamforming, Slot and Power Allocation, through Optimization, Heuristics, and Learning". In: *IEEE Access* 12 (2024), pp. 127127–127140. DOI: 10.1109/ACCESS.2024.3454968.

[Pap+24] Arled Papa, Jörg von Mankowski, Hansini Vijayaraghavan, Babak Mafakheri, Leonardo Goratti, and Wolfgang Kellerer. "Enabling 6G Applications in the Sky: Aeronautical Federation Framework". In: *IEEE Network* 38.1 (2024), pp. 254–261. DOI: 10.1109/MNET.132.2200526.

[Man+24] Jörg von Mankowski, Hansini Vijayaraghavan, Yannik Hilla, and Wolfgang Kellerer. "Characterization of the 2.4- and 5-GHz Channels in a Single-Aisle Commercial Aircraft Cabin Using Ray Tracing". In: *IEEE Transactions on Aerospace and Electronic Systems* 60.2 (2024), pp. 1386–1399. DOI: 10.1109/TAES.2023.3336854.

Conference Proceedings

[Vij+22] Hansini Vijayaraghavan, Jörg von Mankowski, Stephan Pellegrino, and Carmen Mas-Machuca. "Optimized 3D Placement of LiFi Access Points towards maximizing Wireless Network Performance". In: *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*. 2022, pp. 1278–1283. DOI: 10.1109/GLOBECOM48099.2022.10000893.

[Vij+21] Hansini Vijayaraghavan, Anna Prado, Thomas Wiese, and Wolfgang Kellerer. "Algorithmic and System Approaches for a Stable LiFi-RF HetNet Under Transient Channel Conditions". In: *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. 2021, pp. 1048–1054. DOI: 10.1109/PIMRC50174.2021.9569271.

[VK21] Hansini Vijayaraghavan and Wolfgang Kellerer. "Delay-aware Wireless Resource Allocation and User Association in LiFi-WiFi Heterogeneous Networks". In: *2021 IEEE Global Communications Conference (GLOBECOM)*. 2021, pp. 01–06. DOI: 10.1109/GLOBECOM46510.2021.9685276.

[PVK21] Anna Prado, Hansini Vijayaraghavan, and Wolfgang Kellerer. "ECHO: Enhanced Conditional Handover boosted by Trajectory Prediction". In: *2021 IEEE Global Communications Conference (GLOBECOM)*. 2021, pp. 01–06. DOI: 10.1109/GLOBECOM46510.2021.9685348.

[Man+22] Jörg von Mankowski, Hansini Vijayaraghavan, Alberto Martinez Alba, Leonardo Goratti, and Wolfgang Kellerer. "Towards the Optimal Pattern of Joint Beamforming, User Scheduling and Power Allocation in a multi-RAT Network". In: *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*. 2022, pp. 338–345. DOI: 10.1109/CCNC49033.2022.9700638.

Bibliography

[Mar17] IHS Markit. *The Internet of Things: a movement, not a market*. Accessed: 2024-07-22. 2017. URL: https://cdn.ihs.com/www/pdf/IoT_ebook.pdf.

[Nok23] Nokia. *Global Network Traffic 2030 Report*. Accessed: 2024-01-14. 2023. URL: https://onestore.nokia.com/asset/f/213660.

[Ela+20] Hadeel Elayan, Osama Amin, Basem Shihada, Raed M. Shubair, and Mohamed-Slim Alouini. "Terahertz Band: The Last Piece of RF Spectrum Puzzle for Communication Systems". In: *IEEE Open Journal of the Communications Society* 1 (2020), pp. 1–32. DOI: 10.1109/OJCOMS.2019.2953633.

[Sha+23] Akram Shafie, Nan Yang, Chong Han, Josep Miquel Jornet, Markku Juntti, and Thomas Kürner. "Terahertz Communications for 6G and Beyond Wireless Networks: Challenges, Key Advancements, and Opportunities". In: *IEEE Network* 37.3 (2023), pp. 162–169. DOI: 10.1109/MNET.118.2200057.

[Haa18] Harald Haas. "LiFi is a paradigm-shifting 5G technology". In: *Reviews in Physics* 3 (2018), pp. 26–31. ISSN: 2405-4283. DOI: 10.1016/j.revip.2017.10.001.

[BTH19] Rui Bian, Iman Tavakkolnia, and Harald Haas. "15.73 Gb/s Visible Light Communication With Off-the-Shelf LEDs". In: *Journal of Lightwave Technology* 37.10 (2019), pp. 2418–2424. DOI: 10.1109/JLT.2019.2906464.

[Haa+16] Harald Haas, Liang Yin, Yunlu Wang, and Cheng Chen. "What is LiFi?" In: *Journal of Lightwave Technology* 34.6 (2016), pp. 1533–1544. DOI: 10.1109/JLT.2015.2510021.

[Ayy+16] Moussa Ayyash et al. "Coexistence of WiFi and LiFi toward 5G: concepts, opportunities, and challenges". In: *IEEE Communications Magazine* 54.2 (2016), pp. 64–71. DOI: 10.1109/MCOM.2016.7402263.

[DBS18] Mohammad Amir Dastgheib, Hamzeh Beyranvand, and Jawad A. Salehi. "Optimal Visible Light Communication Access Point Placement Under Stationary Distribution of Users' Mobility". In: *2018 9th International Symposium on Telecommunications (IST)*. 2018, pp. 96–101. DOI: 10.1109/ISTEL.2018.8661152.

[DBS20] Mohammad Amir Dastgheib, Hamzeh Beyranvand, and Jawad A. Salehi. "Optimal Placement of Access Points in Cellular Visible Light Communication Networks: An Adaptive Gradient Projection Method". In: *IEEE Transactions on Wireless Communications* 19.10 (2020), pp. 6813–6825. DOI: 10.1109/TWC.2020.3006204.

[VB19] Anna Maria Vegni and Mauro Biagi. "Optimal LED placement in indoor VLC networks". In: *Opt. Express* 27.6 (Mar. 2019), pp. 8504–8519. DOI: 10.1364/OE.27.008504.

[MS23] Mahmoud Mohammadi and Seyed Mohammad Sajad Sadough. "Improved LED arrangement through outage probability minimization in LiFi communication systems". In: *IET Communications* (2023). DOI: 10.1049/cmu2.12598.

[Yan+20] Yang Yang, Zhiyu Zhu, Caili Guo, and Chunyan Feng. "Power efficient LED placement algorithm for indoor visible light communication". In: *Opt. Express* 28.24 (Nov. 2020), pp. 36389–36402. DOI: 10.1364/OE.410502.

[Gop+22] Govind R. Gopal, Elina Nayebi, Gabriel Porto Villardi, and Bhaskar D. Rao. "Modified Vector Quantization for Small-Cell Access Point Placement With Inter-Cell Interference". In: *IEEE Transactions on Wireless Communications* 21.8 (2022), pp. 6387–6401. DOI: 10.1109/TWC.2022.3148996.

[WH15] Yunlu Wang and Harald Haas. "Dynamic Load Balancing With Handover in Hybrid Li-Fi and Wi-Fi Networks". In: *Journal of Lightwave Technology* 33.22 (2015), pp. 4671–4682. DOI: 10.1109/JLT.2015.2480969.

[WWH17] Yunlu Wang, Xiping Wu, and Harald Haas. "Load Balancing Game With Shadowing Effect for Indoor Hybrid LiFi/RF Networks". In: *IEEE Transactions on Wireless Communications* 16.4 (2017), pp. 2366–2378. DOI: 10.1109/TWC.2017.2664821.

[WH19a] Xiping Wu and Harald Haas. "Mobility-aware load balancing for hybrid LiFi and WiFi networks". In: *J. Opt. Commun. Netw.* 11.12 (Dec. 2019), pp. 588–597. DOI: 10.1364/JOCN.11.000588.

[3GP22a] 3GPP. *5G; Service requirements for cyber-physical control applications in vertical domains*. Technical Specification 22.104. 3GPP, May 2022.

[3GP22b] 3GPP. *5G; Study on scenarios and requirements for next generation access technologies*. Technical Report 38.913. 3GPP, May 2022.

[Wu+21a] Xiping Wu, Mohammad Dehghani Soltani, Lai Zhou, Majid Safari, and Harald Haas. "Hybrid LiFi and WiFi Networks: A Survey". In: *IEEE Communications Surveys & Tutorials* 23.2 (2021), pp. 1398–1420. DOI: 10.1109/COMST.2021.3058296.

[DH15] Svilen Dimitrov and Harald Haas. *Principles of LED Light Communications: Towards Networked Li-Fi*. Cambridge University Press, 2015. DOI: 10.1017/CBO9781107278929.

[CBH18] Cheng Chen, Rui Bian, and Harald Haas. "Omnidirectional Transmitter and Receiver Design for Wireless Infrared Uplink Transmission in LiFi". In: *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*. 2018, pp. 1–6. DOI: 10.1109/ICCW.2018.8403727.

[Wu+21b] Xiping Wu, Mohammad Dehghani Soltani, Lai Zhou, Majid Safari, and Harald Haas. "Hybrid LiFi and WiFi Networks: A Survey". In: *IEEE Communications Surveys & Tutorials* 23.2 (2021), pp. 1398–1420. DOI: 10.1109/COMST.2021.3058296.

[IEE23] IEEE. "IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks–Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Light Communications". In: *IEEE Std 802.11bb-2023* (2023), pp. 1–37. DOI: 10.1109/IEEESTD.2024.10315104.

[Gmb] aeroLiFi GmbH. *Our Products*. Accessed: 2024-07-21. URL: https://aerolifi.com/products/.

[WH20] Xiping Wu and Harald Haas. "Load Balancing for Hybrid LiFi and WiFi Networks: To Tackle User Mobility and Light-Path Blockage". In: *IEEE Transactions on Communications* 68.3 (2020), pp. 1675–1683. DOI: 10.1109/TCOMM.2019.2962434.

[Sol+19] Mohammad Dehghani Soltani, Ardimas Andi Purwita, Zhihong Zeng, Harald Haas, and Majid Safari. "Modeling the Random Orientation of Mobile Devices: Measurement, Analysis and LiFi Use Case". In: *IEEE Transactions on Communications* 67.3 (2019), pp. 2157–2172. DOI: 10.1109/TCOMM.2018.2882213.

[Fir+21] Fadhil Firyaguna, Jacek Kibilda, Carlo Galiotto, and Nicola Marchetti. "Performance Analysis of Indoor mmWave Networks With Ceiling-Mounted Access Points". In: *IEEE Transactions on Mobile Computing* 20.5 (2021), pp. 1940–1950. DOI: 10.1109/TMC.2020.2972282.

[IAY20] Sifat Ibne Mushfique, Ahmad Alsharoa, and Murat Yuksel. "Optimization of SINR and Illumination Uniformity in Multi-LED Multi-Datastream VLC Networks". In: *IEEE Transactions on Cognitive Communications and Networking* 6.3 (2020), pp. 1108–1121. DOI: 10.1109/TCCN.2020.2972310.

[Wu+20] Xiping Wu, Dominic C. O'Brien, Xiong Deng, and Jean-Paul M. G. Linnartz. "Smart Handover for Hybrid LiFi and WiFi Networks". In: *IEEE Transactions on Wireless Communications* 19.12 (2020), pp. 8211–8219. DOI: 10.1109/TWC.2020.3020160.

[Rhe+11] Injong Rhee, Minsu Shin, Seongik Hong, Kyunghan Lee, Seong Joon Kim, and Song Chong. "On the Levy-Walk Nature of Human Mobility". In: *IEEE/ACM Transactions on Networking* 19.3 (2011), pp. 630–643. DOI: 10.1109/TNET.2011.2120618.

[Lee+09] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong. "SLAW: A New Mobility Model for Human Walks". In: *IEEE INFOCOM 2009*. 2009, pp. 855–863. DOI: 10.1109/INFCOM.2009.5061995.

[Hon+99] Xiaoyan Hong, Mario Gerla, Guangyu Pei, and Ching-Chuan Chiang. "A group mobility model for ad hoc wireless networks". In: *Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems*. 1999, pp. 53–60. DOI: 10.1145/313237.313248.

[MW47] H. B. Mann and D. R. Whitney. "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other". In: *The Annals of Mathematical Statistics* 18.1 (1947), pp. 50–60. DOI: 10.1214/aoms/1177730491.

[BH95] Yoav Benjamini and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal statistical society: series B (Methodological)* 57.1 (1995), pp. 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.

[Pan+19] Chunyu Pan, Jirong Yi, Changchuan Yin, Jian Yu, and Xuehua Li. "Joint 3D UAV Placement and Resource Allocation in Software-Defined Cellular Networks With Wireless Backhaul". In: *IEEE Access* 7 (2019), pp. 104279–104293. DOI: 10.1109/ACCESS.2019.2927521.

[UUC19] Ismail Uluturk, Ismail Uysal, and Kwang-Cheng Chen. "Efficient 3D Placement of Access Points in an Aerial Wireless Network". In: *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. 2019, pp. 1–7. DOI: 10.1109/CCNC.2019.8651769.

[MA03] RT Marler and JS Arora. *Review of multi-objective optimization concepts and methods for engineering*. Tech. rep. ODL-01.01. Optimal Design Laboratory, University of Iowa, Iowa City, IA, 2003.

[Deb+02] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. "A fast and elitist multiobjective genetic algorithm: NSGA-II". In: *IEEE Transactions on Evolutionary Computation* 6.2 (2002), pp. 182–197. DOI: 10.1109/4235.996017.

[DA+95] Kalyanmoy Deb, Ram Bhushan Agrawal, et al. "Simulated binary crossover for continuous search space". In: *Complex systems* 9.2 (1995), pp. 115–148.

[Cha+22] Florian Charlier et al. *Statannotations*. Version v0.6. Oct. 2022. DOI: 10.5281/zenodo.7213391.

[Pas+17] S Thomas Valerrian Pasca, R Adharsh Srivats, A Antony Franklin, and Bheemarjuna Reddy Tamma. "Optimal placement of colocated and non-colocated LWA nodes in dense deployments". In: *2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*. 2017, pp. 1–6. DOI: 10.1109/ANTS.2017.8384088.

[EAE19] Noha A. Elmosilhy, Ahmed M. Abd El-Haleem, and Mahmoud M. Elmesalawy. "Optimal Deployment of Heterogeneous Wireless Nodes in Integrated LTE|Wi-Fi Networks". In: *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. 2019, pp. 1035–1041. DOI: 10.1109/UEMCON47517.2019.8993033.

[MM22] Ahmad Mazaherifar and Seyedakbar Mostafavi. "UAV Placement and Trajectory Design Optimization: A Survey". In: *Wireless Personal Communications* 124.3 (June 2022), pp. 2191–2210. ISSN: 0929-6212. DOI: 10.1007/s11277-021-09451-7.

[Gha+18] Rozhina Ghanavi, Elham Kalantari, Maryam Sabbaghian, Halim Yanikomeroglu, and Abbas Yongacoglu. "Efficient 3D aerial base station placement considering users mobility by reinforcement learning". In: *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. 2018, pp. 1–6. DOI: 10.1109/WCNC.2018.8377340.

[WZZ18] Qingqing Wu, Yong Zeng, and Rui Zhang. "Joint Trajectory and Communication Design for Multi-UAV Enabled Wireless Networks". In: *IEEE Transactions on Wireless Communications* 17.3 (2018), pp. 2109–2121. DOI: 10.1109/TWC.2017.2789293.

[GRV21] Govind R. Gopal, Bhaskar D. Rao, and Gabriel Porto Villardi. "Access Point Placement for Hybrid UAV-Terrestrial Small-Cell Networks". In: *IEEE Open Journal of the Communications Society* 2 (2021), pp. 1826–1841. DOI: 10.1109/OJCOMS.2021.3100060.

[Gur23] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2023. URL: https://www.gurobi.com.

[Chu20] Tinkle Chugh. "Scalarizing Functions in Bayesian Multiobjective Optimization". In: *2020 IEEE Congress on Evolutionary Computation (CEC)*. 2020, pp. 1–8. DOI: 10.1109/CEC48606.2020.9185706.

[Wie82] Andrzej P Wierzbicki. "A mathematical basis for satisficing decision making". In: *Mathematical modelling* 3.5 (1982), pp. 391–405. DOI: 10.1016/0270-0255(82)90038-0.

[BD20] Julian Blank and Kalyanmoy Deb. "Pymoo: Multi-Objective Optimization in Python". In: *IEEE Access* 8 (2020), pp. 89497–89509. DOI: 10.1109/ACCESS.2020.2990567.

[Moc89] Jonas Mockus. *Bayesian Approach to Global Optimization*. Kluwer academic publishers, 1989. DOI: 10.1007/BF00940509.

[Bre01] Leo Breiman. "Random forests". In: *Machine learning* 45 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[GEW06] Pierre Geurts, Damien Ernst, and Louis Wehenkel. "Extremely randomized trees". In: *Machine learning* 63 (2006), pp. 3–42. DOI: 10.1007/s10994-006-6226-1.

[WO22] Xiping Wu and Dominic C. O'Brien. "QoS-Driven Load Balancing in Hybrid LiFi and WiFi Networks". In: *IEEE Transactions on Wireless Communications* 21.4 (2022), pp. 2136–2146. DOI: 10.1109/TWC.2021.3109716.

[Sol+23] Mohammad Dehghani Soltani et al. "Terabit Indoor Laser-Based Wireless Communications: Lifi 2.0 For 6G". In: *IEEE Wireless Communications* 30.5 (2023), pp. 36–43. DOI: 10.1109/MWC.007.2300121.

[Est+24] Rafael Estepa, Mark Davis, Vicente Mayor, and Antonio Estepa. "Supporting VoIP communication in IEEE 802.11 ax networks: A new admission control and scheduling resource allocation scheme". In: *Computer Communications* 224 (2024), pp. 225–241. ISSN: 0140-3664. DOI: 10.1016/j.comcom.2024.06.010.

[ASS23] Nurul Aini Amran, Mohammad Dehghani Soltani, and Majid Safari. "Link Blockage Analysis for Indoor Optical Wireless Communications". In: *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*. 2023, pp. 5568–5573. DOI: 10.1109/GLOBECOM54140.2023.10437747.

[Lan+10] Tian Lan, David Kao, Mung Chiang, and Ashutosh Sabharwal. "An Axiomatic Theory of Fairness in Network Resource Allocation". In: *2010 Proceedings IEEE INFOCOM*. 2010, pp. 1–9. DOI: 10.1109/INFCOM.2010.5461911.

[WSH17] Xiping Wu, Majid Safari, and Harald Haas. "Joint Optimisation of Load Balancing and Handover for Hybrid LiFi and WiFi Networks". In: *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. 2017, pp. 1–5. DOI: 10.1109/WCNC.2017.7925839.

[WWH16] Yunlu Wang, Xiping Wu, and Harald Haas. "Fuzzy logic based dynamic handover scheme for indoor Li-Fi and RF hybrid network". In: *2016 IEEE International Conference on Communications (ICC)*. 2016, pp. 1–6. DOI: 10.1109/ICC.2016.7510823.

[WN03] Dapeng Wu and R. Negi. "Effective capacity: a wireless link model for support of quality of service". In: *IEEE Transactions on Wireless Communications* 2.4 (2003), pp. 630–643. DOI: 10.1109/TWC.2003.814353.

[Jin+17] Yaqi Jin, Xiaodong Xu, Yunting Wang, and Xiaofeng Tao. "Multi-QoS mobile services guaranteed resource allocation with effective capacity". In: *2017 IEEE/CIC International Conference on Communications in China (ICCC)*. 2017, pp. 1–6. DOI: 10.1109/ICCChina.2017.8330461.

[Jin+16] Fan Jin, Xuan Li, Rong Zhang, Chen Dong, and Lajos Hanzo. "Resource Allocation Under Delay-Guarantee Constraints for Visible-Light Communication". In: *IEEE Access* 4 (2016), pp. 7301–7312. DOI: 10.1109/ACCESS.2016.2564298.

[JZH15] Fan Jin, Rong Zhang, and Lajos Hanzo. "Resource Allocation Under Delay-Guarantee Constraints for Heterogeneous Visible-Light and RF Femtocell". In: *IEEE Transactions on Wireless Communications* 14.2 (2015), pp. 1020–1034. DOI: 10.1109/TWC.2014.2363451.

[Luo17] Xiliang Luo. "Delay-Oriented QoS-Aware User Association and Resource Allocation in Heterogeneous Cellular Networks". In: *IEEE Transactions on Wireless Communications* 16.3 (2017), pp. 1809–1822. DOI: 10.1109/TWC.2017.2654458.

[Gol89] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. 1st. Addison-Wesley Longman Publishing, 1989.

[Sol+17] Mohammad Dehghani Soltani, Hossein Kazemi, Majid Safari, and Harald Haas. "Handover Modeling for Indoor Li-Fi Cellular Networks: The Effects of Receiver Mobility and Rotation". In: *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. 2017, pp. 1–6. DOI: 10.1109/WCNC.2017.7925676.

[LCW16] Yun Li, Bin Cao, and Chonggang Wang. "Handover schemes in heterogeneous LTE networks: challenges and opportunities". In: *IEEE Wireless Communications* 23.2 (2016), pp. 112–117. DOI: 10.1109/MWC.2016.7462492.

[DEA15] Ergin Dinc, Ozgur Ergul, and Ozgur B. Akan. "Soft Handover in OFDMA Based Visible Light Communication Networks". In: *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*. 2015, pp. 1–5. DOI: 10.1109/VTCFall.2015.7391146.

[VL12] Anna Maria Vegni and Thomas D. C. Little. "Handover in VLC systems with cooperating mobile devices". In: *2012 International Conference on Computing, Networking and Communications (ICNC)*. 2012, pp. 126–130. DOI: 10.1109/ICCNC.2012.6167395.

[ABG14] Atiq Ahmed, Leila Merghem Boulahia, and Dominique Gaïti. "Enabling Vertical Handover Decisions in Heterogeneous Wireless Networks: A State-of-the-Art and A Classification". In: *IEEE Communications Surveys & Tutorials* 16.2 (2014), pp. 776–811. DOI: 10.1109/SURV.2013.082713.00141.

[WH19b] Xiping Wu and Harald Haas. "Handover Skipping for LiFi". In: *IEEE Access* 7 (2019), pp. 38369–38378. DOI: 10.1109/ACCESS.2019.2903409.

[SHI+14] Huaizhou SHI, R. Venkatesha Prasad, Ertan Onur, and I.G.M.M. Niemegeers. "Fairness in Wireless Networks:Issues, Measures and Challenges". In: *IEEE Communications Surveys & Tutorials* 16.1 (2014), pp. 5–24. DOI: 10.1109/SURV.2013.050113.00015.

[JY22] Jonggyu Jang and Hyun Jong Yang. "α-Fairness-Maximizing User Association in Energy-Constrained Small Cell Networks". In: *IEEE Transactions on Wireless Communications* 21.9 (2022), pp. 7443–7459. DOI: 10.1109/TWC.2022.3158694.

[Xu+22] Yongjun Xu, Bowen Gu, Dong Li, Zhaohui Yang, Chongwen Huang, and Kai-Kit Wong. "Resource Allocation for Secure SWIPT-Enabled D2D Communications With α Fairness". In: *IEEE Transactions on Vehicular Technology* 71.1 (2022), pp. 1101–1106. DOI: 10.1109/TVT.2021.3129787.

[MW00] J. Mo and J. Walrand. "Fair end-to-end window-based congestion control". In: *IEEE/ACM Transactions on Networking* 8.5 (2000), pp. 556–567. DOI: 10.1109/90.879343.

[NH09] Dusit Niyato and Ekram Hossain. "Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach". In: *IEEE Transactions on Vehicular Technology* 58.4 (2009), pp. 2008–2017. DOI: 10.1109/TVT.2008.2004588.

[JCH84] Rajendra K Jain, Dah-Ming W Chiu, and William R Hawe. *A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems*. Tech. rep. DEC-TR-301. Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, 1984.

[BPB11] Sébastien Barré, Christoph Paasch, and Olivier Bonaventure. "MultiPath TCP: From Theory to Practice". In: *NETWORKING 2011*. Ed. by Jordi Domingo-Pascual, Pietro Manzoni, Sergio Palazzo, Ana Pont, and Caterina Scoglio. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 444–457. ISBN: 978-3-642-20757-0. DOI: 10.1007/978-3-642-20757-0_35.

[Wir] Wireshark. *Protocol analyzer*. (accessed July 16, 2024). URL: https://www.wireshark.org/.

[Wan+17] Yunlu Wang, Dushyantha A. Basnayaka, Xiping Wu, and Harald Haas. "Optimization of Load Balancing in Hybrid LiFi/RF Networks". In: *IEEE Transactions on Communications* 65.4 (2017), pp. 1708–1720. DOI: 10.1109/TCOMM.2017.2654249.

[Ahm+22] Rizwana Ahmad, Mohammad Dehghani Soltani, Majid Safari, and Anand Srivastava. "Reinforcement Learning-Based Near-Optimal Load Balancing for Heterogeneous LiFi WiFi Network". In: *IEEE Systems Journal* 16.2 (2022), pp. 3084–3095. DOI: 10.1109/JSYST.2021.3088302.

[Arf+21] Mohamed Amine Arfaoui et al. "Invoking Deep Learning for Joint Estimation of Indoor LiFi User Position and Orientation". In: *IEEE Journal on Selected Areas in Communications* 39.9 (2021), pp. 2890–2905. DOI: 10.1109/JSAC.2021.3064637.

[Bui+17] Nicola Bui, Matteo Cesana, S. Amir Hosseini, Qi Liao, Ilaria Malanchini, and Joerg Widmer. "A Survey of Anticipatory Mobile Networking: Context-Based Classification, Prediction Methodologies, and Optimization Techniques". In: *IEEE Communications Surveys & Tutorials* 19.3 (2017), pp. 1790–1821. DOI: 10.1109/COMST.2017.2694140.

[Arf+24] Mohamed Amine Arfaoui, Ali Ghrayeb, Chadi Assi, and Marwa Qaraqe. "Deep Learning Based Proactive Optimization for Mobile LiFi Systems With Channel Aging". In: *IEEE Transactions on Communications* 72.6 (2024), pp. 3543–3557. DOI: 10.1109/TCOMM.2024.3366405.

[Zha+18] Rong Zhang, Ying Cui, Holger Claussen, Harald Haas, and Lajos Hanzo. "Anticipatory Association for Indoor Visible Light Communications: Light, Follow Me!" In: *IEEE Transactions on Wireless Communications* 17.4 (2018), pp. 2499–2510. DOI: 10.1109/TWC.2018.2797182.

[Das+18] Mohammad Amir Dastgheib, Hamzeh Beyranvand, Jawad A. Salehi, and Martin Maier. "Mobility-Aware Resource Allocation in VLC Networks Using T-Step Look-Ahead Policy". In: *Journal of Lightwave Technology* 36.23 (2018), pp. 5358–5370. DOI: 10.1109/JLT.2018.2872869.

[CDS18] Ursula Challita, Li Dong, and Walid Saad. "Proactive Resource Management for LTE in Unlicensed Spectrum: A Deep Learning Perspective". In: *IEEE Transactions on Wireless Communications* 17.7 (2018), pp. 4674–4689. DOI: 10.1109/TWC.2018.2829773.

[Li+20] Jing Li, Xing Zhang, Jiaxin Zhang, Jie Wu, Qi Sun, and Yuxuan Xie. "Deep Reinforcement Learning-Based Mobility-Aware Robust Proactive Resource Allocation in Heterogeneous Networks". In: *IEEE Transactions on Cognitive Communications and Networking* 6.1 (2020), pp. 408–421. DOI: 10.1109/TCCN.2019.2954396.

[Fró+24] Dayrene Frómeta Fonseca, Borja Genovés Guzmán, Giovanni Luca Mertena, Rui Bian, Harald Haas, Domenico Giustiniano, et al. "Prediction-model-assisted reinforcement learning algorithm for handover decision-making in hybrid LiFi and WiFi networks". In: *Journal of Optical Communications and Networking* 16.2 (Feb. 2024), pp. 159–170. DOI: 10.1364/JOCN.495234.

[TQS23] Rongshun Tang, Chenhao Qi, and Yan Sun. "Blockage Prediction and Fast Handover of Base Station for Millimeter Wave Communications". In: *IEEE Communications Letters* 27.8 (2023), pp. 2142–2146. DOI: 10.1109/LCOMM.2023.3289581.

[LD10] Ailsa H Land and Alison G Doig. "An Automatic Method for Solving Discrete Programming Problems". In: *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*. Springer Berlin Heidelberg, 2010, pp. 105–132. DOI: 10.1007/978-3-540-68279-0_5.

[Fan+22] Wenhao Fan et al. "Joint Task Offloading and Service Caching for Multi-Access Edge Computing in WiFi-Cellular Heterogeneous Networks". In: *IEEE Transactions on Wireless Communications* 21.11 (2022), pp. 9653–9667. DOI: 10.1109/TWC.2022.3178541.

[XS23] Xiangyang Xu and Yu Song. "A Deep Reinforcement Learning-Based Optimal Computation Offloading Scheme for VR Video Transmission in Mobile Edge Networks". In: *IEEE Access* 11 (2023), pp. 122772–122781. DOI: 10.1109/ACCESS.2023.3327921.

[Yun+22] Jusik Yun, Yunyeong Goh, Wonsuk Yoo, and Jong-Moon Chung. "5G Multi-RAT URLLC and eMBB Dynamic Task Offloading With MEC Resource Allocation Using Distributed Deep Reinforcement Learning". In: *IEEE Internet of Things Journal* 9.20 (2022), pp. 20733–20749. DOI: 10.1109/JIOT.2022.3177425.

[Zha+20] Yufeng Zhan, Song Guo, Peng Li, and Jiang Zhang. "A Deep Reinforcement Learning Based Offloading Game in Edge Computing". In: *IEEE Transactions on Computers* 69.6 (2020), pp. 883–893. DOI: 10.1109/TC.2020.2969148.

[SKO23] Akito Suzuki, Masahiro Kobayashi, and Eiji Oki. "Multi-Agent Deep Reinforcement Learning for Cooperative Computing Offloading and Route Optimization in Multi Cloud-Edge Networks". In: *IEEE Transactions on Network and Service Management* 20.4 (2023), pp. 4416–4434. DOI: 10.1109/TNSM.2023.3267809.

[Fen+21] Guangsheng Feng, Xin Li, Zihan Gao, Chengbo Wang, Hongwu Lv, and Qian Zhao. "Multi-Path and Multi-Hop Task Offloading in Mobile Ad Hoc Networks". In: *IEEE Transactions on Vehicular Technology* 70.6 (2021), pp. 5347–5361. DOI: 10.1109/TVT.2021.3077691.

[Den+24] Yiqin Deng, Haixia Zhang, Xianhao Chen, and Yuguang Fang. "Multi-Hop Task Routing in Vehicle-Assisted Collaborative Edge Computing". In: *IEEE Transactions on Vehicular Technology* 73.2 (2024), pp. 2444–2455. DOI: 10.1109/TVT.2023.3312142.

[Fen+22] Chuan Feng, Pengchao Han, Xu Zhang, Bowen Yang, Yejun Liu, and Lei Guo. "Computation offloading in mobile edge computing networks: A survey". In: *Journal of Network and Computer Applications* 202 (2022), p. 103366. DOI: 10.1016/j.jnca.2022.103366.

[Sal+21] Umber Saleem, Yu Liu, Sobia Jangsher, Yong Li, and Tao Jiang. "Mobility-Aware Joint Task Scheduling and Resource Allocation for Cooperative Mobile Edge Computing". In: *IEEE Transactions on Wireless Communications* 20.1 (2021), pp. 360–374. DOI: 10.1109/TWC.2020.3024538.

[Liu+24a] Yu Liu, Yingling Mao, Zhenhua Liu, and Yuanyuan Yang. "Deep Learning-Assisted Online Task Offloading for Latency Minimization in Heterogeneous Mobile Edge". In: *IEEE Transactions on Mobile Computing* 23.5 (2024), pp. 4062–4075. DOI: 10.1109/TMC.2023.3285882.

[CL21] Xing Chen and Guizhong Liu. "Energy-Efficient Task Offloading and Resource Allocation via Deep Reinforcement Learning for Augmented Reality in Mobile Edge Networks". In: *IEEE Internet of Things Journal* 8.13 (2021), pp. 10843–10856. DOI: 10.1109/JIOT.2021.3050804.

[Bra+20] Tristan Braud, Pengyuan Zhou, Jussi Kangasharju, and Pan Hui. "Multipath Computation Offloading for Mobile Augmented Reality". In: *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 2020, pp. 1–10. DOI: 10.1109/PerCom45495.2020.9127360.

[Zha+17] Heli Zhang, Jun Guo, Lichao Yang, Xi Li, and Hong Ji. "Computation offloading considering fronthaul and backhaul in small-cell networks integrated with MEC". In: *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 2017, pp. 115–120. DOI: 10.1109/INFCOMW.2017.8116362.

[Jan+23] Alba Jano, Mehmet Mert Bese, Nitinder Mohan, Wolfgang Kellerer, and Jörg Ott. "next GSIM: Towards Simulating Network Resource Management for Beyond 5G Networks". In: *2023 IEEE Future Networks World Forum (FNWF)*. 2023, pp. 1–7. DOI: 10.1109/FNWF58287.2023.10520638.

[Liu+24b] Yu Liu, Yingling Mao, Zhenhua Liu, Fan Ye, and Yuanyuan Yang. "Joint Task Offloading and Resource Allocation in Heterogeneous Edge Environments". In: *IEEE Transactions on Mobile Computing* 23.6 (2024), pp. 7318–7334. DOI: 10.1109/TMC.2023.3335198.

[PSL05] Kenneth Price, Rainer M Storn, and Jouni A Lampinen. *Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series)*. Springer-Verlag, 2005. DOI: 10.1007/3-540-31306-0.

[HJ61] Robert Hooke and T. A. Jeeves. "" Direct Search" Solution of Numerical and Statistical Problems". In: *J. ACM* 8.2 (Apr. 1961), pp. 212–229. ISSN: 0004-5411. DOI: 10.1145/321062.321069.

[KE95] J. Kennedy and R. Eberhart. "Particle swarm optimization". In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. Vol. 4. 1995, 1942–1948 vol.4. DOI: 10.1109/ICNN.1995.488968.

[RY00] T.P. Runarsson and Xin Yao. "Stochastic ranking for constrained evolutionary optimization". In: *IEEE Transactions on Evolutionary Computation* 4.3 (2000), pp. 284–294. DOI: 10.1109/4235.873238.

[Sch+17] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal Policy Optimization Algorithms". In: *arXiv preprint arXiv:1707.06347* (2017). DOI: 10.48550/arXiv.1707.06347.

[Raf+21] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. "Stable-Baselines3: Reliable Reinforcement Learning Implementations". In: *Journal of Machine Learning Research* 22.268 (2021), pp. 1–8. URL: http://jmlr.org/papers/v22/20-1364.html.

[Aki+19] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. "Optuna: A Next-generation Hyperparameter Optimization Framework". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2019. DOI: 10.1145/3292500.3330701.

[Haa+20] Harald Haas et al. "Introduction to indoor networking concepts and challenges in LiFi". In: *Journal of Optical Communications and Networking* 12.2 (2020), A190–A203. DOI: 10.1364/JOCN.12.00A190.

[Guz+23] Borja Genoves Guzman, Muhammad Sarmad Mir, Dayrene Frometa Fonseca, Ander Galisteo, Qing Wang, and Domenico Giustiniano. "Prototyping Visible Light Communication for the Internet of Things Using OpenVLC". In: *IEEE Communications Magazine* 61.5 (2023), pp. 122–128. DOI: 10.1109/MCOM.001.2200642.