

# QoS-Driven Load Balancing in Hybrid LiFi and WiFi Networks

Xiping Wu, *Senior Member, IEEE*, and Dominic C. O'Brien, *Member, IEEE*

**Abstract**—This work studies the quality of service (QoS) performance of wireless networks that integrate light fidelity (LiFi) and wireless fidelity (WiFi). While the hybrid network is potential for improving network capacity, load balancing becomes essential and challenging due to the nature of heterogeneous access points (APs). A number of studies have been conducted to address this issue, focusing on maximising the network capacity with user fairness constraints. However, in practice, QoS metrics including packet loss ratio and latency are important to network services. In this paper, QoS-driven load balancing is studied for hybrid LiFi and WiFi networks (HLWNets) in two scenarios: single-AP association (SA) and multi-AP association (MA). In each case, an optimisation problem is formulated to minimise the packet loss ratio and latency, and a low-complexity iterative algorithm is proposed to solve the problem. Results show that the novel methods, especially MA, can effectively balance the traffic loads among the APs and improve the QoS performance. In addition, the more subflows the better performance MA provides. Targeting the same level of QoS, MA can achieve a system throughput up to 160% higher than the signal strength strategy and 130% higher than the proportional fairness load balancing.

**Index Terms**—Light fidelity (LiFi), wireless fidelity (WiFi), hybrid network, load balancing, quality of service (QoS), visible light communication (VLC)

## I. INTRODUCTION

LIGHT fidelity (LiFi) is potentially a key technology towards the sixth generation (6G) wireless communications [1]. LiFi access points (APs) can be embedded into the existing light infrastructure, realising a dual-purpose system of illumination and communication. In comparison with its radio frequency (RF) counterpart wireless fidelity (WiFi), LiFi provides a number of advantages including: i) availability in RF-restricted areas; ii) vast and licence-free visible optical spectrum of  $\sim 300$  THz; and iii) secure communication since light does not penetrate opaque structures. Experimental work demonstrates that LiFi is capable of supporting a very high link data rate, in the range of hundreds of Gbps [2]. However, LiFi is susceptible to blockages caused by opaque objects, such as furniture and human bodies. Also, LiFi covers a relatively small area with a single AP, usually a few meters in diameter. To combine the high data rate of LiFi and the ubiquitous coverage of WiFi, hybrid LiFi and WiFi networks (HLWNets) are proposed in recent years [3]. In fact, the

concept of hybrid infrared (IR) and RF can be traced back in the 1990s [4]. While the hybrid networks are capable of enhancing network performance, the access point selection (APS) process becomes challenging. This is because the WiFi AP usually has a larger coverage area but a lower capacity than the LiFi AP [5]. As a result, using the signal strength strategy (SSS) would readily congest the WiFi AP. In other words, load balancing is essential for HLWNets.

A number of studies have been carried out to tackle the load balancing issue in HLWNets. These methods fall into three categories: optimisation-based [6]–[8], rule-based [9], [10], and learning-based [11]–[13]. In [6], an optimisation problem was formulated to maximise the network capacity with proportional fairness. In [7], the optimisation was extended to a joint implementation of load balancing and power allocation. The concept of mobility-aware load balancing was introduced in [8] to cope with handovers due to user mobility. Optimisation-based methods usually require an excessive amount of processing power. This motivates the development of rule-based approaches. In [9], game theory was applied to adapt the APS strategy of each user. An APS approach based on fuzzy logic was proposed in [10], which can achieve near-optimal performance at significantly reduced computational complexity. However, rule-based methods are less adaptive to changes in the network settings. To make load balancing more smart, researchers attempt to apply machine learning. In [11], learning techniques were proposed to update the probability distribution of each AP for making APS decisions. The authors in [12] employed reinforcement learning to balance the traffic loads, where the reward function is designed to maximize the long-term average network throughput as well as the users' satisfaction and fairness. Based on artificial neural network, a handover scheme was developed in [13] to adapt the selection preference between LiFi and WiFi.

The existing load balancing methods for HLWNets have three major limitations. First, these methods focus on maximising network throughput in bits, while most of them do not consider quality of service (QoS) metrics in the packet level. In practice, a certain QoS level is required to support network services, which may have different demands for data rate. Maximising network throughput does not guarantee an optimal solution for processing the packets. Second, they do not consider the impact of buffer size, which limits the queueing length of packets as well as affecting the achievable throughput. Third, most of the existing schemes only consider single-AP association, i.e., each user is served by a single AP. This association manner is due to the restriction of the conventional transmission control protocol (TCP). The Internet

This work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 825651 (ELIOT).

X. Wu is with the School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland (e-mail: xiping.wu@ucd.ie).

D. C. O'Brien is with Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K. (e-mail: dominic.obrien@eng.ox.ac.uk).

Engineering Task Force (IETF) has been putting forward multipath TCP (MPTCP) [14] since 2013, so that the user can be served by multiple APs at the same time. This enables multi-AP association, which can provide a more flexible way of resource allocation and better balance the traffic loads. However, this point has not been well addressed in the current literature.

A QoS-based gate way selection was proposed in [15] for IEEE 802.11 based vehicular ad hoc networks. Fuzzy logic was used to solve a multi-criteria optimisation problem, which considers the received signal strength (RSS) and traffic load. However, this method does not provide a solution to directly optimising the QoS metrics. Apart from that, the fuzzy rules that are specially tailored for homogeneous networks are not suitable for hybrid networks. In this paper, QoS-driven load balancing is investigated in HLWNets for the first time, to the best of our knowledge. The main contributions of the paper are summarised below:

- The load balancing issue is addressed in both scenarios of single-AP association and multi-AP association. For each scenario, a multi-objective optimisation problem is formulated to minimise the packet loss ratio and latency, based on deriving the closed-form expressions of these metrics.
- Novel iterative algorithms are proposed to solve the optimisation problems, with analysis of the optimality and computational complexity. These algorithms can achieve near-optimal solutions with a significantly reduced amount of runtime, in the range of 1-10 ms.
- QoS-driven load balancing and throughput-driven load balancing are compared through analysing the QoS performance of HLWNets, with different packet arrival rates and numbers of users. The impacts of buffer size and light-path blockage are also studied.

The remainder of this paper is organised as follows. Section II introduces the system model of HLWNet, including the network structure, channel models and link capacity. The QoS metrics used in this paper are derived in Section III. QoS-driven load balancing is studied for the scenarios of single-AP association and multi-AP association in Section IV and V, respectively. Simulation results are presented in Section VI. Finally, conclusions are drawn in Section VII.

## II. SYSTEM MODEL

Consider an indoor HLWNet that consists of one WiFi AP and a number of LiFi APs. The WiFi AP is placed in the centre of the room, providing coverage for the entire room. Each LiFi AP is embedded into a ceiling light-emitting diode (LED) lamp, covering a confined area. The LiFi APs share the optical spectra with a reuse factor of 4 [16]. Fig. 1 shows the network topology of HLWNet, which adopts a symmetrical wiring structure for LiFi [17]. The users are uniformly distributed in the room area, and time-division multiple access (TDMA) is employed for each AP to serve multiple users.

### A. LiFi Channel Model

The LiFi channel is comprised of line-of-sight (LoS) and non line-of-sight (NLoS) paths, as shown in Fig. 2. Between

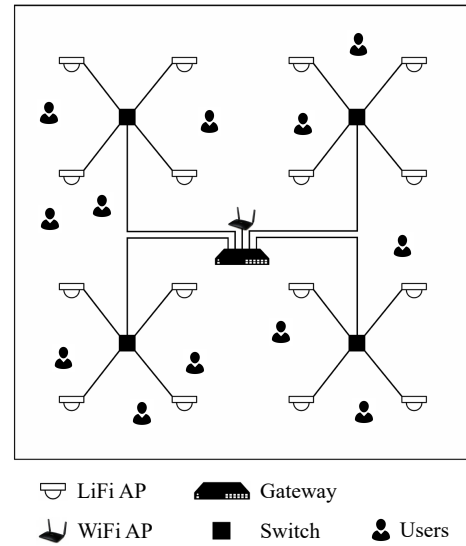


Fig. 1. Network topology of an indoor HLWNet.

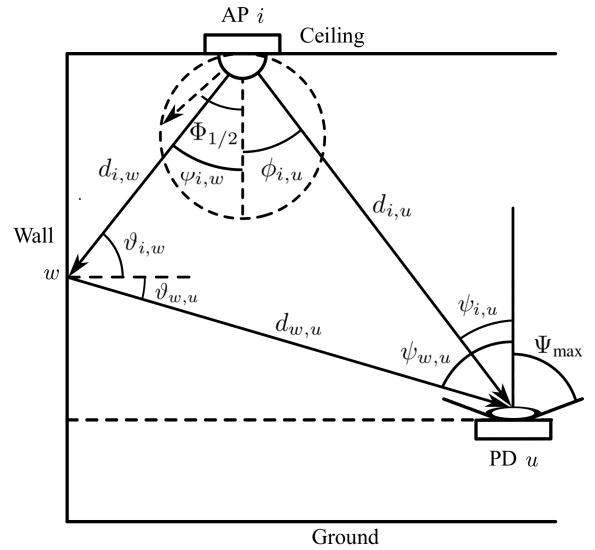


Fig. 2. The LoS and first-order NLoS paths of the LiFi channel.

AP  $i$  and user  $u$ , the LoS path is modelled as [18, eq. (10)]:

$$H_{\text{LoS}}^{i,u} = \frac{(m+1)A_{\text{pd}}}{2\pi d_{i,u}^2} \cos^m(\phi_{i,u}) g_f g_c(\psi_{i,u}) \cos(\psi_{i,u}), \quad (1)$$

where  $m = -\ln 2 / \ln(\cos \Phi_{1/2})$  is the Lambertian emission order, and  $\Phi_{1/2}$  is the half-intensity angle;  $A_{\text{pd}}$  is the physical area of the photodiode (PD);  $d_{i,u}$  is the Euclidean distance between the AP and the user;  $\phi_{i,u}$  and  $\psi_{i,u}$  denote the angles of irradiance and incidence;  $g_f$  is the transmittance of the optical filter; the optical concentrator gain  $g_c(\psi_{i,u})$  is given by:

$$g_c(\psi_{i,u}) = \begin{cases} \frac{n^2}{\sin^2(\Psi_{\text{max}})}, & 0 \leq \psi_{i,u} \leq \Psi_{\text{max}} \\ 0, & \psi_{i,u} > \Psi_{\text{max}} \end{cases}, \quad (2)$$

where  $n$  denotes the refractive index, and  $\Psi_{\text{max}}$  is the semi-angle of the field of view (FoV) of the PD.

With respect to the NLoS path, only first-order reflections are considered since higher-order reflections typically contribute little [18]. The channel gain of NLoS is given by (3) [16, eq. (3)], where  $d_{i,w}$  and  $d_{w,u}$  denote the Euclidean distances of the segment from the AP to a reflection area  $w$  and the segment from  $w$  to the user;  $\phi_{i,w}$  and  $\vartheta_{i,w}$  are the angles of radiance and incidence with respect to the first segment, and for the second segment they are  $\vartheta_{w,u}$  and  $\psi_{w,u}$ ;  $A_w$  denotes the area of  $w$ ;  $\rho_w$  is the wall reflectivity. The LiFi channel gain can be expressed as  $H_{\text{LiFi}}^{i,u} = H_{\text{LoS}}^{i,u} + H_{\text{NLoS}}^{i,u}$ .

At the receiver, photons are gathered and converted into an electric current:

$$I = R_{\text{pd}} H_{\text{LiFi}}^{i,u} P_{\text{mod}}, \quad (4)$$

where  $R_{\text{pd}}$  denotes the detector responsivity, and  $P_{\text{mod}}$  is the average modulated optical power. The signal-to-interference-plus-noise ratio (SINR) of a LiFi user can be written as:

$$\gamma_{\text{LiFi}}^{i,u} = \frac{(R_{\text{pd}} H_{\text{LiFi}}^{i,u} P_{\text{mod}})^2}{N_{\text{LiFi}} B_{\text{LiFi}} + \sum_{j \in \mathcal{I}_i, j \neq i} (R_{\text{pd}} H_{\text{LiFi}}^{j,u} P_{\text{mod}})^2}, \quad (5)$$

where  $B_{\text{LiFi}}$  is the bandwidth of the LiFi AP;  $N_{\text{LiFi}}$  is the power spectral density (PSD) of noise, which is assumed to be signal independent;  $\mathcal{I}_i$  represents the set of LiFi APs that employ the same optical spectrum as AP  $i$ .

### B. WiFi Channel Model

The WiFi channel gain can be expressed as [19]:

$$G_{\text{WiFi}}^{i,u} = \left| H_{\text{WiFi}}^{i,u} \right|^2 10^{\frac{-L(d_{i,u}) + X_\sigma}{10}}, \quad (6)$$

where  $H_{\text{WiFi}}^{i,u}$  is the channel transfer function, which follows a standard Rayleigh distribution;  $X_\sigma$ , which represents the fading due to shadowing, is a zero-mean Gaussian random variable with a standard deviation of 10 dB; and  $L(\cdot)$  denotes the free-space path loss:

$$L(d_{i,u}) = 20 \log_{10}(f_c d_{i,u}) - 147.5, \quad (7)$$

where  $f_c$  is the central carrier frequency. The signal-to-noise ratio (SNR) of a WiFi user is written as follows:

$$\gamma_{\text{WiFi}}^{i,u} = \frac{G_{\text{WiFi}}^{i,u} P_{\text{WiFi}}}{N_{\text{WiFi}} B_{\text{WiFi}}}, \quad (8)$$

where  $B_{\text{WiFi}}$  is the bandwidth of the WiFi AP;  $P_{\text{WiFi}}$  is the transmit power; and  $N_{\text{WiFi}}$  denotes the PSD of noise.

### C. Link Capacity

The capacity of a WiFi link can be bounded by the Shannon capacity. As for LiFi,  $\gamma_{\text{LiFi}}^{i,u}$  is an electrical SINR for non-negative signals, for which a tighter bound can be found [20]. Let  $B_i$  denote the bandwidth of AP  $i$ . The link capacity can be expressed as:

$$r_{u,i} = \begin{cases} B_i \log_2 \left( 1 + \gamma_{\text{WiFi}}^{i,u} \right), & \text{for WiFi users} \\ \frac{B_i}{2} \log_2 \left( 1 + \frac{e}{2\pi} \gamma_{\text{LiFi}}^{i,u} \right), & \text{for LiFi users} \end{cases}. \quad (9)$$

## III. QUALITY OF SERVICE METRICS

In this section the packet loss ratio and latency are derived, given the packet arrival rate of each user. The impact of buffer size is also taken into account, which restricts the maximum queue length. The derived expressions will be used to formulate the optimisation problems in Sections IV and V.

### A. Packet Loss Ratio

Packet loss is either caused by errors in data transmission or network congestion. Since the capacity bounds in (9) define the maximum error-free data rates, here we focus on network congestion. In an M/M/1 queueing system, the packet arrival follows a Poisson process with rate  $\lambda$ , and the time for serving one packet is exponentially distributed with mean  $1/\mu$ . The average queue length, i.e., the average number of packets waiting for service, is given by [21]:

$$N_W = \frac{\lambda}{\mu - \lambda}. \quad (10)$$

The parameters  $\lambda$  and  $\mu$  are referred to as the arrival rate and the service rate, respectively. Regarding the gateway and switches, the service rate is usually much greater than that of the wireless network. Hence, we focus on analysing network congestion in the wireless domain. Let  $\mathbf{U}_i$  denote the set of users that are served by AP  $i$ . Let  $R_{u,i}$  denote the arrival data rate, i.e., the data rate that AP  $i$  needs to provide user  $u$ . Denoting the average packet size by  $L_P$ , the arrival rate at AP  $i$  can be written as:

$$\lambda_i = \frac{1}{L_P} \sum_{u \in \mathbf{U}_i} R_{u,i}. \quad (11)$$

To send out  $R_{u,i}$  bits to user  $u$ , the amount of time spent by AP  $i$  is  $R_{u,i}/r_{u,i}$ . Thus, the service rate of AP  $i$  can be computed by:

$$\mu_i = \frac{\lambda_i}{\sum_{u \in \mathbf{U}_i} \frac{R_{u,i}}{r_{u,i}}}. \quad (12)$$

The average size of data waiting for service is  $L_W = N_W L_P$ . Combining (10) and (12), we have:

$$L_W = \frac{L_P}{\max \left( \left( \sum_{u \in \mathbf{U}_i} \frac{R_{u,i}}{r_{u,i}} \right)^{-1} - 1, 0 \right)}. \quad (13)$$

The parameter  $L_W$  increases with  $R_{u,i}$ , and packet loss occurs when  $L_W > L_B$ , where  $L_B$  denotes the buffer size. Here the subscription  $i$  is omitted for  $L_W$  and  $L_B$  for the sake of simplicity. The maximum achievable data rate  $\hat{R}_{u,i}$  is achieved when  $L_W = L_B$ :

$$\sum_{u \in \mathbf{U}_i} \frac{\hat{R}_{u,i}}{r_{u,i}} = \frac{L_B}{L_B + L_P}. \quad (14)$$

For the link between AP  $i$  and user  $u$ , the packet loss ratio can be computed as follows:

$$\eta_{u,i} = \begin{cases} 1 - \frac{\hat{R}_{u,i}}{R_{u,i}}, & \text{if } R_{u,i} > \hat{R}_{u,i} \\ 0, & \text{if } R_{u,i} \leq \hat{R}_{u,i} \end{cases}. \quad (15)$$

$$H_{\text{NLoS}}^{i,u} = \int_{A_w} \frac{(m+1)A_{\text{pd}}}{2(\pi d_{i,w} d_{w,u})^2} \rho_w \cos^m(\phi_{i,w}) g_f g_c \cos(\psi_{w,u}) \cos(\vartheta_{i,w}) \cos(\vartheta_{w,u}) dA_w. \quad (3)$$

*Lemma 3.1:* In an M/M/1 queueing system, the users associated to the same AP have the same value of  $\eta_{u,i}$ .

*Proof:* With respect to user  $u$ , the number of dropped packets is  $N_D^{u,i} = \eta_{u,i} R_{u,i} / L_P$ . For an M/M/1 queueing system,  $N_D^{u,i}$  is proportional to the user's arrival data rate. This gives:

$$\frac{N_D^{u,i}}{\sum_{u \in \mathbf{U}_i} N_D^{u,i}} = \frac{R_{u,i}}{\sum_{u \in \mathbf{U}_i} R_{u,i}}, \quad \forall u \in \mathbf{U}_i. \quad (16)$$

Therefore, we have:

$$\sum_{u \in \mathbf{U}_i} \eta_{u,i} R_{u,i} = \eta_{u_1,i} \sum_{u \in \mathbf{U}_i} R_{u,i} = \dots = \eta_{u_n,i} \sum_{u \in \mathbf{U}_i} R_{u,i}. \quad (17)$$

From the above equation it can be derived that:

$$\eta_{u_1,i} = \eta_{u_2,i} = \dots = \eta_{u_n,i}. \quad (18)$$

Hence,  $\eta_{u,i}$  can be reduced to  $\eta_i$ . Substituting (14) into (15),  $\eta_i$  is computed by:

$$\eta_i = \max \left( 1 - \frac{L_B}{L_B + L_P} \left( \sum_{u \in \mathbf{U}_i} \frac{R_{u,i}}{r_{u,i}} \right)^{-1}, 0 \right). \quad (19)$$

The average packet loss ratio of the entire network can be calculated as follows:

$$\eta = \frac{\sum_i \eta_i \left( \sum_{u \in \mathbf{U}_i} R_{u,i} \right)}{\sum_i \sum_{u \in \mathbf{U}_i} R_{u,i}}. \quad (20)$$

### B. Latency

There are two main factors that affect network latency: packet processing and propagation. For an indoor network, the propagation time is on the order of 10 ns and thus is negligible. Let  $T_i$  denote the average time it takes to process one packet. According to Little's theorem,  $T_i$  equals the average queue length divided by the average service rate. Let  $L_i$  denote the average queue length of AP  $i$ . When the buffer is not fully used,  $L_i = L_W$ , which can be obtained from (13). When the buffer is used up,  $L_i$  is limited to  $L_B$ . Therefore, we have  $L_i = L_W$  with the constraint  $L_i \leq L_B$ . The average service rate is  $\sum_{u \in \mathbf{U}_i} r_{u,i} R_{u,i} / \sum_{u \in \mathbf{U}_i} R_{u,i}$ . As a result, the packet latency at AP  $i$  can be obtained by:

$$T_i = L_i \frac{\sum_{u \in \mathbf{U}_i} R_{u,i}}{\sum_{u \in \mathbf{U}_i} r_{u,i} R_{u,i}}. \quad (21)$$

The average latency of the entire network is computed by:

$$T = \frac{\sum_i T_i \left[ (1 - \eta_i) \sum_{u \in \mathbf{U}_i} R_{u,i} \right]}{\sum_i \left[ (1 - \eta_i) \sum_{u \in \mathbf{U}_i} R_{u,i} \right]}. \quad (22)$$

## IV. LOAD BALANCING FOR SINGLE-AP ASSOCIATION

With single-AP association (SA), each user is only served by one AP, and all the packets belonging to one user flow into a single host AP. As a result, the arrival data rate  $R_{u,i}$  is equivalent to  $R_u$ . In this scenario, the task is assigning an appropriate AP to each user.

### A. Optimisation Problem

The aim is to minimise the packet loss ratio first and then minimise the latency. Let  $\chi_{u,i}$  indicate the AP assignment:  $\chi_{u,i} = 1$  means that user  $u$  is connected to AP  $i$  and otherwise  $\chi_{u,i} = 0$ . The set of  $\chi_{u,i}$  is denoted by  $\chi$ . Replacing  $R_{u,i}$  by  $R_u$  in (20) and (22), the objective functions can be expressed as:

$$F_{\text{SA}}^\eta(\chi) = \frac{\sum_i \eta_i \left( \sum_u \chi_{u,i} R_u \right)}{\sum_i \sum_u \chi_{u,i} R_u}. \quad (23)$$

and:

$$F_{\text{SA}}^T(\chi) = \frac{\sum_i T_i \left[ (1 - \eta_i) \sum_u \chi_{u,i} R_u \right]}{\sum_i \left[ (1 - \eta_i) \sum_u \chi_{u,i} R_u \right]}. \quad (24)$$

The optimisation problem is formulated as follow:

$$\begin{aligned} \min \quad & \{ F_{\text{SA}}^T(\chi) \mid \chi = \min F_{\text{SA}}^\eta(\chi) \} \\ \text{s.t.} \quad & \chi_{u,i} \in \{0, 1\}, \quad \forall u, i; \\ & \sum_i \chi_{u,i} = 1, \quad \forall u; \\ & L_i \leq L_B. \end{aligned} \quad (25)$$

Three constraints are considered. The first constraint states that each link is either on or off, while the second one indicates that each user is connected to one and only one AP. The third constraint reflects that the queue length does not exceed the buffer size. In practice, there might be other constraints such as routing, jitter, etc. Here in the considered network, packets are forwarded from the gateway to the user through only one AP. Hence, there is no additional constraint on routing. Due to the difficulty of analysing the latency of each packet, it is difficult to implement a constraint on jitter. Instead, the jitter performance is numerically analysed in Section VI. The problem in (25) can be solved by exhaustively searching the possible solutions. However, this requires an excessive amount of computational complexity. Alternatively, we propose an iterative algorithm to reduce the processing power.

---

**Algorithm 1** The SA Iterative Algorithm
 

---

**Input:**  $R_u, \gamma_{u,i}$ 
**Output:**  $\chi$ 

- 1: initialisation:
    - for** each user  $u$  **do**
    - find the AP  $i_u$  that maximises  $\gamma_{u,i}$
    - $\chi^{u,i_u} \leftarrow 1$
    - end for**
  - 2: compute  $\eta_i$  and  $T_i$  for each AP
  - 3: sort users according to the priority rule
  - 4: **for** each user  $u$  **do**
    - for** each AP  $i$  except  $i_u$  **do**
    - $\chi_{u,j} \leftarrow 0, \quad \forall j \neq i$
    - $\chi_{u,i} \leftarrow 1$
    - compute  $\eta_i$  and  $T_i$
    - end for**
    - find the set of APs that minimise  $\eta_i$
    - find the AP  $i$  that minimises  $T_i$  in the above set
    - if**  $i \neq i_u$  **then**
    - $i_u \leftarrow i$
    - break loop
    - end if**
    - end for**
  - 5: repeat step 2 until no user changes its host AP
  - 6: **return**  $\chi$
- 

### B. Proposed Iterative Algorithm

The principle is letting the users select an AP in turn. Initially, each user is associated with the AP that offers the best channel quality, following the signal strength strategy (SSS). At this stage, multiple users might choose the same AP, making that AP overloaded. These users are allowed to look for an AP that can enhance the QoS, until no one can find a better solution. Among these users, the one with a worse QoS performance is more likely to change its AP association. As a result, the user with a higher packet loss ratio is given priority to change its host AP. When several users have the same level of packet loss ratio, the one with a higher latency has priority. As for the users with the same packet loss ratio and latency, the user with a lower SINR benefits more from changing its host AP and thus is given priority. This method is referred to as the SA iterative algorithm. Its pseudocode is shown in Algorithm 1.

### C. Analysis of Optimality and Complexity

It is difficult to theoretically analyse the optimality and complexity of the iterative algorithm due to its nature. Alternatively, numerical analysis is carried out by comparing the SA iterative algorithm with exhaustive search. Taking 4 LiFi APs as an example, Fig. 3 shows the outcomes for two cases  $N_u = 10$  and  $N_u = 20$ , where  $N_u$  denotes the number of users. The results of the iterative algorithm are shown in solid lines, whereas the solutions of exhaustive search are represented by dashed lines. It is observed that the proposed algorithm converges quickly and approaches the optimal solution closely.

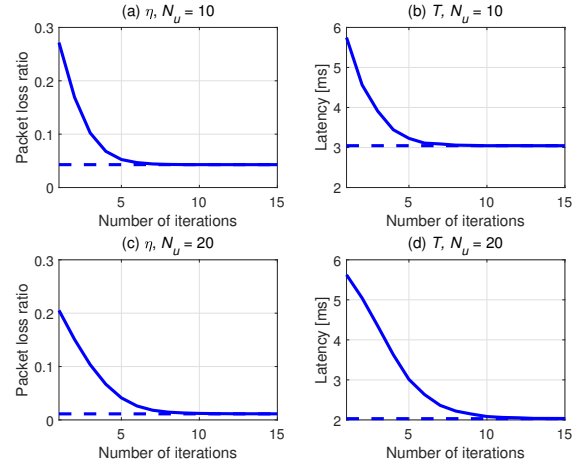


Fig. 3. Optimality versus the number of iterations (solid line: SA; dashed line: exhaustive search).

In average, only 6 iterations are required in the case of 10 users, and 10 iterations are needed when  $N_u$  increases to 20.

Now we analyse the computational complexity. Regarding the exhaustive search, the number of candidates (i.e., possible solutions) is  $N_{AP}^{N_u}$ , where  $N_{AP}$  denotes the number of LiFi APs. For each candidate, the objective function is computed through multiplying the matrix of AP selection and the matrix of channel capacity. Both matrices have a dimension of  $N_u$ -by- $N_{AP}$ . The standard way of multiplying an  $m$ -by- $n$  matrix by an  $n$ -by- $p$  matrix has complexity  $O(mnp)$ . Therefore, each candidate has complexity between  $O(N_{AP}^2 N_u)$  and  $O(N_{AP} N_u^2)$ . The overall complexity of exhaustive search is between  $O(N_{AP}^2 N_u N_{AP}^{N_u})$  and  $O(N_{AP} N_u^2 N_{AP}^{N_u})$ . Let  $T_{ES}(N_u)$  denote the runtime of exhaustive search in the case of  $N_u$  users. Taking  $T_{ES}(2)$  as a reference point,  $T_{ES}(N_u)$  can be estimated with a lower bound and an upper bound:

$$\begin{cases} T_{ES}^{\text{lower}}(N_u) = T_{ES}(2) \times \frac{N_u N_{AP}^{N_u-2}}{2} \\ T_{ES}^{\text{upper}}(N_u) = T_{ES}(2) \times \frac{N_u^2 N_{AP}^{N_u-2}}{4} \end{cases} \quad (26)$$

As for the SA algorithm, it has complexity  $O(N_{\text{iter}} N_c)$ , where  $N_{\text{iter}}$  can be obtained through counting the average number of required iterations and  $N_c$  is the number of users that switch the host AP per iteration. The value of  $N_c$  is between 1 and  $N_u$ . Thus, the runtime of the SA algorithm can also be estimated with a lower bound and an upper bound:

$$\begin{cases} T_{SA}^{\text{lower}}(N_u) = T_{SA}(2) \times \frac{N_{\text{iter}}(N_u)}{N_{\text{iter}}(2)} \\ T_{SA}^{\text{upper}}(N_u) = T_{SA}(2) \times \frac{N_u N_{\text{iter}}(N_u)}{2 N_{\text{iter}}(2)} \end{cases} \quad (27)$$

Fig. 4 shows the runtime of the SA algorithm in comparison with exhaustive search, which is referred to as ES. As shown, for both methods, the measured runtime is well between the derived lower and upper bounds. The two methods cost a similar amount of runtime when  $N_u = 2$ . However, as  $N_u$  becomes larger, the runtime of ES increases exponentially, while for SA it has a linear growth. When  $N_u$  increases to

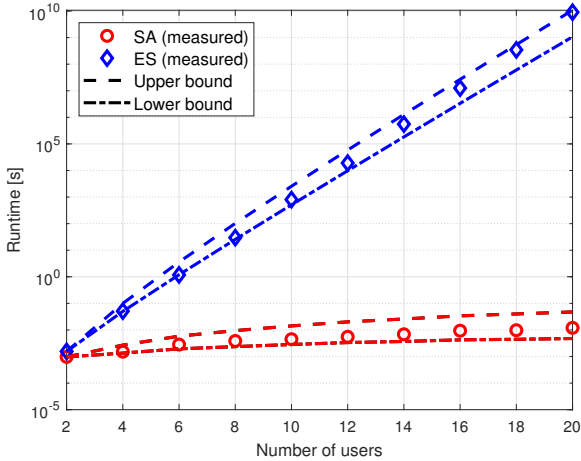


Fig. 4. Runtime versus the number of users for SA and ES.

20, SA can reduce the computation complexity by  $7.5 \times 10^{11}$  over ES.

## V. LOAD BALANCING FOR MULTI-AP ASSOCIATION

With multi-AP association (MA), each user is served by multiple APs at the same time, with each AP providing a subflow. The user monitors the SINR situations of nearby APs through dedicated channels. With  $N_f$  subflows, the user is assigned to  $N_f$  APs which offer the highest values of SINR. The task is distributing the packets requested by each user among its subflows.

### A. Optimisation Problem

With respect to user  $u$ , the ratio between the number of packets that flows into AP  $i$  and the number of overall packets is denoted by  $\rho_{u,i}$ . The set of  $\rho_{u,i}$  is denoted by  $\rho$ . The objective functions are written as:

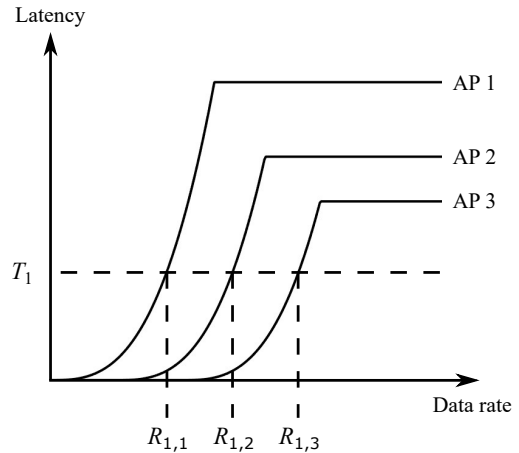
$$F_{MA}^{\eta}(\rho) = \frac{\sum_i \eta_i \left( \sum_u \rho_{u,i} R_u \right)}{\sum_i \sum_u \rho_{u,i} R_u}. \quad (28)$$

and:

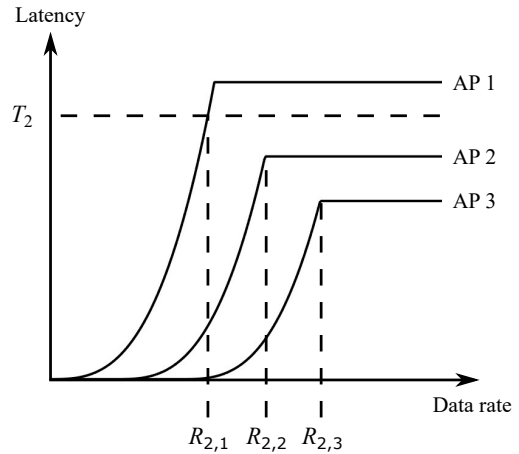
$$F_{MA}^T(\rho) = \frac{\sum_i T_i \left[ (1 - \eta_i) \sum_u \rho_{u,i} R_u \right]}{\sum_i \left[ (1 - \eta_i) \sum_u \rho_{u,i} R_u \right]}. \quad (29)$$

The traffic load of a user is split among the selected APs, whereas  $\rho_{u,i} = 0$  for the remaining APs. Thus, the optimisation problem can be formulated as follows:

$$\begin{aligned} \min \quad & \{ F_{MA}^T(\rho) \mid \rho = \min F_{MA}^{\eta}(\rho) \} \\ \text{s.t.} \quad & 0 \leq \rho_{u,i} \leq 1, \quad \forall u; \\ & \rho_{u,i} = 0, \quad u \notin \mathbf{U}_i; \\ & \sum_i \rho_{u,i} = 1, \quad \forall u; \\ & L_i \leq L_B. \end{aligned} \quad (30)$$



(a) Below saturation



(b) Above saturation

Fig. 5. Water filling algorithm for splitting traffic load among subflows.

The above problem differs from (25) in two aspects: i) unlike  $\chi_{u,i}$ ,  $\rho_{u,i}$  is a continuous variable; and ii) the AP association is already determined. An iterative algorithm is also proposed to solve this problem.

### B. Proposed Iterative Algorithm

The concept of water filling is adopted to divide the user's traffic load among its subflows. Fig. 5 exemplifies this process, where the latency of each subflow is presented as a function of data rate. In general, the required data rate is split for all subflows to reach the same level of latency. For instance, in Fig. 5(a),  $R_1$  is split into  $R_{1,1}$ ,  $R_{1,2}$  and  $R_{1,3}$  so that the 3 APs provides the same latency  $T_1$ . As the data rate keeps increasing, the latency of each AP reaches a saturation due to the limited buffer size. This saturation point indicates the start of packet loss. As a result, the data rate that an AP can support is restricted. In Fig. 5(b), for example,  $R_2$  is split into  $R_{2,1}$ ,  $R_{2,2}$  and  $R_{2,3}$ , where  $R_{2,2}$  and  $R_{2,3}$  are the maximum data rates that AP 2 and AP 3 can provide without packet loss. In this case, the latency values are different among the APs.

Let  $\hat{T}_i$  denote the saturation value of latency for the  $i$ -th subflow. Without loss of generality, we assume:

$$\hat{T}_1 \leq \hat{T}_2 \leq \dots \leq \hat{T}_i \leq \dots \leq \hat{T}_{N_f}. \quad (31)$$

---

**Algorithm 2** The Water Filling Algorithm for Distributing  $R_u$ 


---

**Input:**  $R_u, f_{u,i}(R)$ 
**Output:**  $\rho_{u,i}$ 

- 1: compute  $\hat{T}_{u,i}, \hat{R}_{u,i}$  and  $\hat{R}_u^i$  for each AP
  - 2: sort APs with  $\hat{T}_{u,i}$  in ascending order
  - 3: **if**  $R_u < \hat{R}_u^1$  **then**  
     compute  $\rho_{u,i}$  following (33)  
   **else if**  $R_u < \hat{R}_u^{N_f}$  **then**  
     compute  $\rho_{u,i}$  following (34)  
   **else**  
     compute  $\rho_{u,i}$  following (35)  
   **end if**
  - 4: **return**  $\rho_{u,i}$
- 

The data rate that the user can achieve at the saturation latency of the  $j$ -th subflow is denoted by  $\hat{R}_u^j$ , which can be computed by:

$$\hat{R}_u^j = \sum_{i=1}^j \hat{R}_{u,i} + \sum_{i=j+1}^{N_f} f_{u,i}^{-1}(\hat{T}_j), \quad (32)$$

where  $\hat{R}_{u,i}$  is the maximum data rate that AP  $i$  can provide user  $u$  without packet loss;  $R = f_{u,i}^{-1}(T)$  is the inverse function of  $T = f_{u,i}(R)$ , which can be acquired from (21). This function describes the relation between data rate and latency with respect to the  $i$ -th subflow. When the data rate requirement is below  $\hat{R}_u^1$ ,  $\rho_{u,i}$  can be readily obtained by solving the following equation set:

$$\begin{cases} f_{u,1}(\rho_{u,1}R_u) = \dots = f_{u,N_f}(\rho_{u,N_f}R_u) \\ \sum_{i=1}^{N_f} \rho_{u,i} = 1 \end{cases}. \quad (33)$$

For a required data rate between  $\hat{R}_u^j$  and  $\hat{R}_u^{j+1}$ , we have  $R_{u,i} = \hat{R}_{u,i}$  for  $i \leq j$ . The corresponding  $\rho_{u,i}$  equals  $\hat{R}_{u,i}/R_u$ . For  $i > j$ ,  $\rho_{u,i}$  can be calculated by:

$$\begin{cases} f_{u,j+1}(\rho_{u,j+1}R_u) = \dots = f_{u,N_f}(\rho_{u,N_f}R_u) \\ \sum_{i=j+1}^{N_f} \rho_{u,i} = 1 - \sum_{i=1}^j \frac{\hat{R}_{u,i}}{R_u} \end{cases}. \quad (34)$$

When larger than  $\hat{R}_u^{N_f}$ , the required data rate is split across the APs in proportion to  $\hat{R}_{u,i}$ . As a result,  $\rho_{u,i}$  is computed as follows:

$$\rho_{u,i} = \hat{R}_{u,i} \left( \sum_{i=1}^{N_f} \hat{R}_{u,i} \right)^{-1}. \quad (35)$$

The pseudocode of the water filling process is given in Algorithm 2. With this data stream segmentation approach, an iterative algorithm is constructed to process the users in turn. This approach is named the MA iterative algorithm, and its pseudocode is shown in Algorithm 3.

### C. Analysis of Optimality and Complexity

An experimental study is conducted to evaluate the optimality and complexity of the MA iterative algorithm. Taking

---

**Algorithm 3** The MA Iterative Algorithm
 

---

**Input:**  $R_u, \gamma_{u,i}$ 
**Output:**  $\rho$ 

- 1: initialisation:  
   **for** each user  $u$  **do**  
     find the AP  $i$  that maximises  $\gamma_{u,i}$   
      $\rho_{u,i} \leftarrow 1$   
   **end for**
  - 2: compute  $\eta_i$  and  $T_i$  for each AP
  - 3: sort users according to the priority rule
  - 4: **for** each user  $u$  **do**  
   compute  $\rho_{u,i}$  via the water filling algorithm  
   **if**  $\rho_{u,i}$  changes **then**  
     break loop  
   **end if**  
   **end for**
  - 5: repeat step 2 until no user changes its  $\rho_{u,i}$
  - 6: **return**  $\rho$
- 

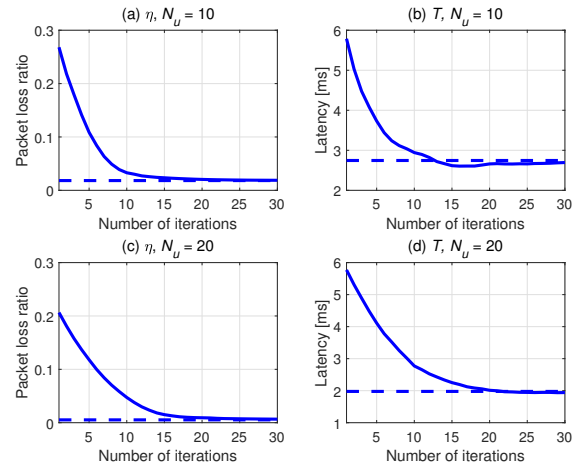


Fig. 6. Optimality versus the number of iterations (solid line: MA; dashed line: OPTI).

4 LiFi APs and  $N_f = 3$  as an example, Fig. 6 shows the packet loss ratio and latency of the MA algorithm as a function of the number of iterations in solid lines. Here we use the MATLAB function ‘fmincon’ as a baseline to solve the optimisation problem in (30). The corresponding results are referred to as OPTI and presented in dashed lines. As shown, the steady performance achieved by the MA algorithm method closely approaches the optimal solution. With 10 and 20 users, the proposed method requires 15 and 20 iterations, which are slightly higher than the numbers of iterations needed by the SA algorithm. Fig. 7 shows the runtime of the MA algorithm in comparison with OPTI. As  $N_u$  increases, the runtime of the MA algorithm increases modestly, while OPTI remains the same level of runtime. Despite this, the MA algorithm consumes about 13% runtime of OPTI when  $N_u = 20$ . As for  $N_u = 2$ , the MA algorithm can save more time, costing only 0.6% runtime of OPTI.

In Section IV-C, it is shown that the runtime of SA is between 1 ms and 10 ms. Here the runtime of MA reaches

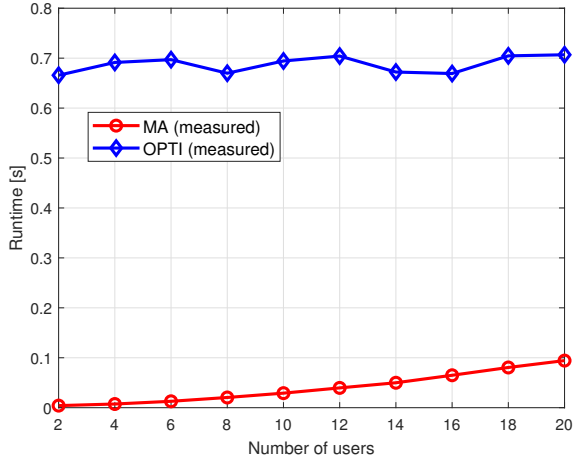


Fig. 7. Runtime versus the number of users for MA and OPTI.

up to 100 ms, though it has already been greatly reduced over OPTI. To further decrease the amount of runtime, artificial intelligence could be a feasible approach. A number of studies were carried out to apply learning methods in software-defined networking, e.g., using reinforcement learning to realise adaptive routing [22] and using deep neural network to implement adaptive channel assignment [23]. In the future, we will study the application of learning methods to perform QoS-aware network management for HLWNets.

## VI. SIMULATION RESULTS

In this section, Monte Carlo simulations are carried out through MATLAB to evaluate the QoS performance of the proposed methods. Two baseline approaches, the SSS method and the proportional fairness (PF) method [6], are adopted. We consider 1 WiFi AP and 16 LiFi APs, with a 2.5 m separation between two adjacent LiFi APs. The vertical distance between the user and the LiFi AP is fixed to be 3 m. Packets are generated under the assumption of an M/M/1 system, and the average packet size is assumed to be 1 KB. The buffer size is set to be 128 KB, except when analysing its impact on the network performance. Other parameters used for the simulations are listed in Table I.

### A. Impact of Arrival Data Rate

Taking 5 users as an example, Fig. 8 presents the packet loss ratio and latency as a function of the overall arrival data rate, which is denoted by  $R$ . As shown, MA can greatly reduce the packet loss ratio compared with SA, which marginally outperforms PF, while SSS performs the worst. This is because SA can only balance the traffic loads when multiple users share the same AP which provides the best channel quality. However, with a small number of users, it is more likely that each user has a different AP that offers the best channel quality. Unlike SA, MA distributes the traffic loads of one user to multiple APs. This allows MA to balance the traffic loads even when there are only a few users. Also, MA performs better with a larger number of subflows. When  $R = 500$  Mbps,

TABLE I  
SIMULATION PARAMETERS

Parameter	Value
Gain of the optical filter, $g_f$	1
Refractive index, $n$	1.5
Half-intensity radiation angle, $\Phi_{1/2}$	60°
FoV semi-angle of the PD, $\Psi_{\max}$	90°
Physical area of the PD, $A_{\text{pd}}$	1 cm <sup>2</sup>
Wall reflectivity, $\rho_w$	0.8
Detector responsivity, $R_{\text{pd}}$	0.53 A/W
Modulated optical power, $P_{\text{mod}}$	1 Watt
LiFi bandwidth, $B_{\text{LiFi}}$	20 MHz
LiFi PSD of noise, $N_{\text{LiFi}}$	10 <sup>-21</sup> A <sup>2</sup> /Hz
WiFi transmit power, $P_{\text{WiFi}}$	20 dBm
WiFi bandwidth, $B_{\text{WiFi}}$	20 MHz
WiFi PSD of noise, $N_{\text{WiFi}}$	-174 dBm/Hz

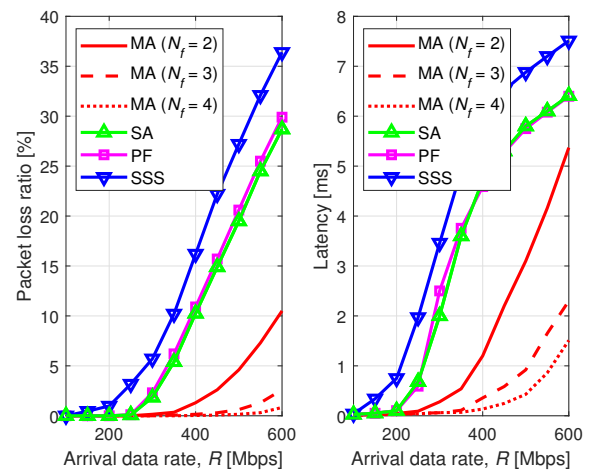


Fig. 8. QoS versus arrival data rate ( $N_u = 5$ ).

for example, MA with  $N_f = 2$  achieves an  $\eta$  of 4%, which is much lower than that of SA ( $\eta = 20\%$ ). When  $N_f = 3$ , MA reduces  $\eta$  to about 1%. As  $N_f$  is further increased to 4, MA decreases  $\eta$  to almost zero. A similar trend is found in the results with respect to latency. At  $R = 500$  Mbps, SA and SSS experience a latency of 5.8 and 6.8 ms, respectively. Meanwhile, the latency offered by MA is 3.1 ms for  $N_f = 2$ , and for  $N_f = 4$  it reduces to 0.4 ms.

### B. Impact of the Number of Users

In Fig. 9, the achievable throughput when  $\eta$  is below 1% is shown for different values of  $N_u$ . Three outcomes are observed. First, SA outperforms PF and SSS more significantly as  $N_u$  increases, as explained. Specifically, SA achieves almost the same throughput as SSS when  $N_u = 2$  but can double the throughput when  $N_u = 16$ . Second, MA improves the throughput over SA, but this improvement decreases as  $N_u$  increases, especially for  $N_f = 2$ . The reason for this trend is that with a larger  $N_u$ , the resource granularity of each user becomes smaller for SA. As a result, SA tends to achieve a performance closer to MA. Third, MA acquires a substantial increase in throughput with more subflows. While MA with  $N_f = 2$  achieves almost the same throughput as SA when



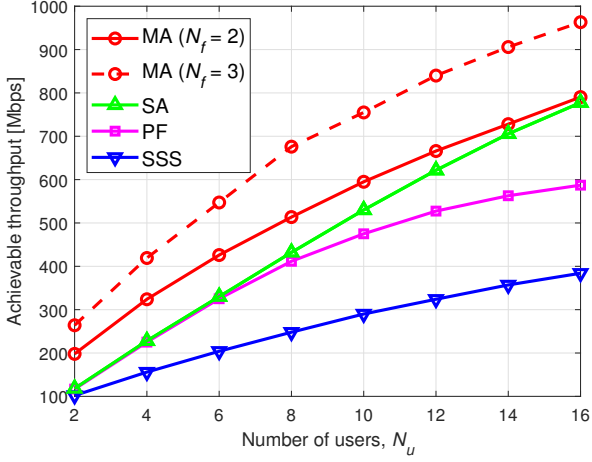


Fig. 9. Achievable throughput ( $\eta \leq 1\%$ ) versus  $N_u$ .

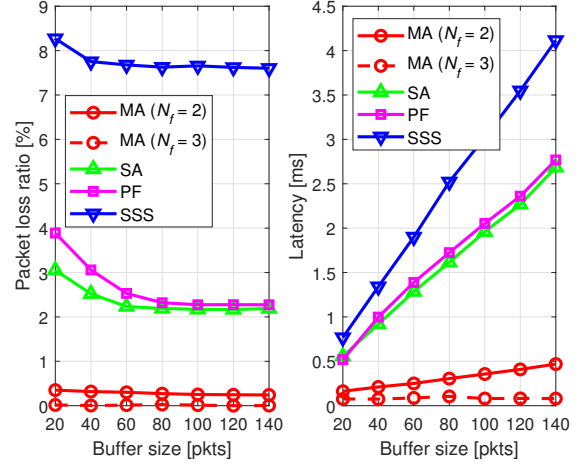


Fig. 11. QoS versus buffer size ( $N_u = 5$ ,  $R = 300$  Mbps).

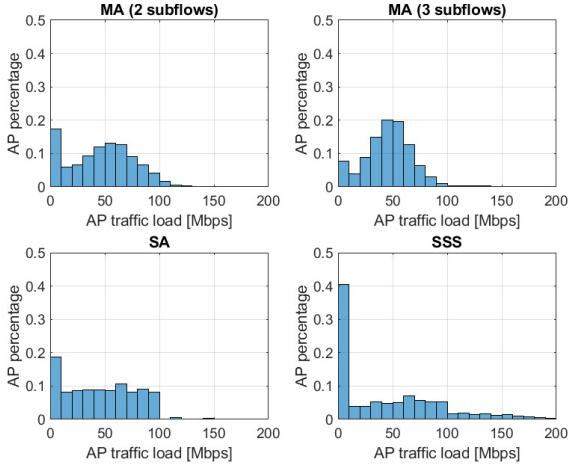


Fig. 10. Traffic load distribution ( $N_u = 16$ ,  $R = 800$  Mbps).

$N_u = 16$ , MA with  $N_f = 3$  can provide an increase of 185 Mbps (24%). However, the improvement percentage offered by MA with  $N_f = 3$  still decreases as  $N_u$  increases. For instance, MA with  $N_f = 3$  improves the throughput over SA by 123% when  $N_u = 2$ . This value reduces to 24% when  $N_u = 16$ .

### C. Traffic Load Distribution

Taking  $N_u = 16$  with  $R = 800$  Mbps as an example, Fig. 10 presents the traffic load distribution among the APs. Regarding SSS, the percentage of APs that have an extremely low load is much higher than other load situations. The reason is that with this method, a large portion of LiFi APs are unused. Compared with SSS, SA can substantially reduce this percentage, from over 40% to less than 20%. Also, SSS has a noticeable AP percentage above 100 Mbps, while almost none appears in SA. In general, the traffic load distribution tends to be uniform in SA. When MA is used, the distribution tends to be Gaussian. With 2 subflows, MA slightly decreases the portion of extremely low loads over SA. Meanwhile, the AP percentage around 50 Mbps becomes higher in MA than in

SA. When  $N_f$  increases to 3, MA shows a more pronounced Gaussian distribution of the traffic loads, with the portion of extremely low loads reduced to about 8%. This signifies that MA (especially with a larger  $N_f$ ) offers a better capability of load balancing than SA, since the ideal outcome is that each AP can undertake the same traffic load.

### D. Impact of Buffer Size

The choice of buffer size is a trade-off between packet loss ratio and latency. A larger buffer size can reduce the packet loss ratio, while a smaller buffer size offers a lower latency. Fig. 11 shows the packet loss ratio and latency as a function of the buffer size  $L_B$  in the unit of packets. As  $L_B$  increases, the packet loss ratio with respect to SA and SSS first decreases and then remains a certain value when  $L_B$  is above 60 packets. This trend is consistent with (19), which indicates that  $\eta$  decrease as  $L_B/(L_B + L_P)$  increases. When  $L_B$  is sufficiently large,  $L_B/(L_B + L_P)$  approaches 1 and  $\eta$  reaches a lower limit. Meanwhile, MA requires a much smaller  $L_B$  to achieve its lower limit of  $\eta$ . It is worth noting that the packet loss ratio of PF noticeably falls behind that of SA when the buffer size is below 80 packets. This signifies the importance of considering the impact of buffer size on achievable data rates. As for latency, it increases linearly with  $L_B$  for all the methods. Thus, an oversized buffer can hardly improve the packet loss ratio but significantly compromise the latency. When  $L_B = 60$  packets, SSS provides  $\eta = 7.6\%$  with a latency of 1.9 ms. For SA,  $\eta = 2.2\%$  and the latency is 1.3 ms. Compared to SA and SSS, MA with  $N_f = 2$  offers a much lower  $\eta$  and latency, which are 0.4% and 0.16 ms respectively when  $L_B = 20$  packets. With  $N_f = 3$ , MA can further reduce the packet loss ratio and latency to 0.02% and 0.07 ms.

### E. Impact of Light-path Blockage

In practice, the LiFi channel might be blocked by opaque objects such as human bodies or furniture. The blockage probability can be calculated by considering the human body as a cylinder object [24]. With a higher blockage probability,

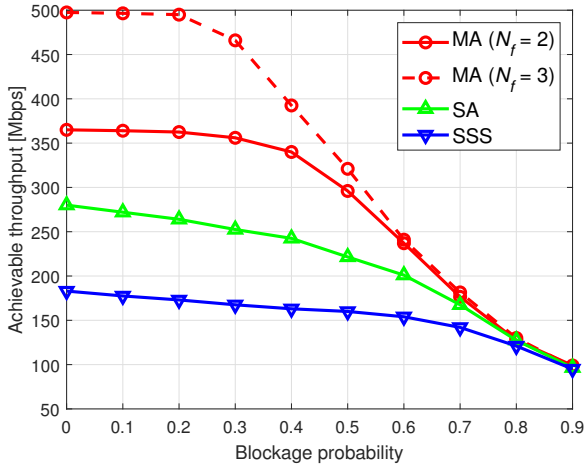


Fig. 12. Achievable throughput versus blockage probability ( $N_u = 5$ ).

the user is less likely to be served by the AP that provides the best channel quality. As a result, the achievable throughput becomes lower. Here we study the network performance at different blockage probabilities. With a small number of users and a sufficient buffer size, PF has a similar performance to SA. In this case, only SSS is presented as a baseline. Fig. 12 shows that as the blockage probability increases, the achievable throughput of MA decreases faster than that of SA, while SSS has the slowest decreasing slope. When the blockage probability reaches 0.9, all the methods offer almost the same achievable throughput. This is because with a high blockage probability, the users would mostly rely on WiFi, and the load balancing abilities of SA and MA would lose effectiveness. However, when the blockage probability is below 0.2, the achievable throughput of MA is nearly unchanged. As for SA and SSS, the achievable throughput noticeably decreases when the blockage probability increases from 0 to 0.2. This signifies that MA is more robust than SA against occasional light-path blockages.

#### F. User Fairness

Jain's fairness index is used to measure the user fairness, which is computed by [10, eq. (14)]:

$$\zeta = \frac{\left( \sum_{u=1}^{N_u} (1 - \eta_u) \right)^2}{N_u \sum_{u=1}^{N_u} (1 - \eta_u)^2}. \quad (36)$$

Taking  $N_u = 5$  as an example, Fig. 13 presents the user fairness as a function of the arrival data rate. As shown, MA is able to achieve a higher user fairness than SA and SSS, especially for a higher arrival data rate. This is attributed to the multiple subflows provided by MA. At  $R = 1$  Gbps, for example, SSS and SA can only offer a user fairness of 0.81 and 0.84, respectively. Using 3 subflows, MA can improve the user fairness to 0.94, which is 16% higher than SSS and 12% higher than SA. A similar but less pronounced trend is found in the cases of larger numbers of users.

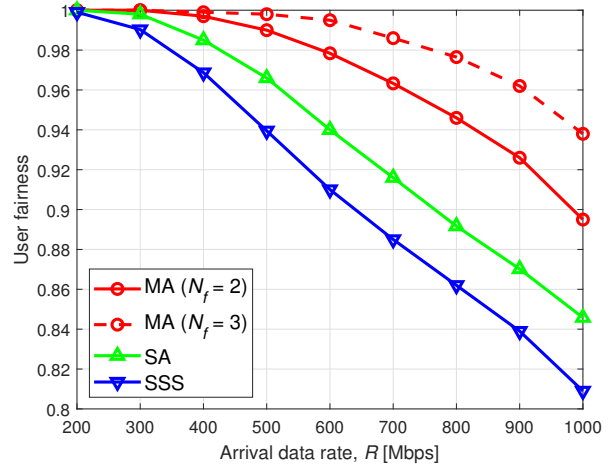


Fig. 13. User fairness versus arrival data rate ( $N_u = 5$ ).

## VII. CONCLUSION

QoS-driven load balancing was studied for HLWNets in two scenarios: single-AP association and multi-AP association. With SA, the requested packets of each user flow to a single AP. As for MA, these packets are distributed among multiple subflows. In each scenario, an iterative algorithm was proposed to solve the optimisation problem, which aims to minimise the packet loss ratio and latency. These algorithms are proven to achieve near-optimal solutions with significantly reduced complexity. Results show that SA outperforms SSS and PF more significantly as the number of users increases. Compared with SA, MA can further enhance the QoS performance, especially when a small number of users are presented. The more subflows the better performance MA achieves, though the improvement becomes less pronounced. For practical use, 3 subflows would be a recommended choice. In this case, MA can improve the achievable throughput over SSS and PF by up to about 160% and 130%, respectively. Moreover, MA is robust against occasional light-path blockages. The convergence of different wireless technologies is gathering a significant momentum in the future wireless communication systems. Our study signifies that an adequate QoS-driven load balancing is of vital importance for exploring the potential of network convergence.

## REFERENCES

- [1] Z. Zhang *et al.*, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, Sep. 2019.
- [2] A. Gomez *et al.*, "Design and demonstration of a 400 Gb/s indoor optical wireless communications link," *J. Lightw. Technol.*, vol. 34, no. 22, pp. 5332–5339, Nov. 2016.
- [3] X. Wu, M. D. Soltani, L. Zhou, M. Safari, and H. Haas, "Hybrid LiFi and WiFi networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 1398–1420, 2nd Quart. 2021.
- [4] H. Zeino and M. Misson, "A pico-cellular concept for a wireless hybrid LAN using infrared modulation," in *Proceedings of 2nd IEEE Int. Conf. on Universal Personal Communications*, vol. 2, Ottawa, 1993, pp. 776–780.
- [5] C. Chen, D. A. Basnayaka, and H. Haas, "Downlink performance of optical attocell networks," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 137–156, Jan. 2016.

- [6] X. Li, R. Zhang, and L. Hanzo, "Cooperative load balancing in hybrid visible light communications and WiFi," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1319–1329, Apr. 2015.
- [7] M. Obeed, A. M. Salhab, S. A. Zummo, and M. S. Alouini, "Joint optimization of power allocation and load balancing for hybrid VLC/RF networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 5, pp. 553–562, May 2018.
- [8] X. Wu and H. Haas, "Mobility-aware load balancing for hybrid LiFi and WiFi networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 11, no. 12, pp. 588–597, Dec. 2019.
- [9] Y. Wang, X. Wu, and H. Haas, "Load balancing game with shadowing effect for indoor hybrid LiFi/RF networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2366–2378, Apr. 2017.
- [10] X. Wu, M. Safari, and H. Haas, "Access point selection for hybrid LiFi and Wi-Fi networks," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5375–5385, Dec. 2017.
- [11] J. Wang, C. Jiang, H. Zhang, X. Zhang, V. C. M. Leung, and L. Hanzo, "Learning-aided network association for hybrid indoor LiFi-WiFi systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3561–3574, Apr. 2018.
- [12] R. Ahmad, M. D. Soltani, M. Safari, A. Srivastava, and A. Das, "Reinforcement learning based load balancing for hybrid LiFi WiFi networks," *IEEE Access*, vol. 8, pp. 132 273–132 284, 2020.
- [13] X. Wu and D. C. O'Brien, "A novel machine learning-based handover scheme for hybrid LiFi and WiFi networks," in *2020 IEEE Globecom Workshops (GC Wkshps)*, Taipei, 2020, pp. 1–5.
- [14] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP extensions for multipath operation with multiple addresses," RFC 6824, Fremont, CA, USA, Tech. Rep., Jan. 2013.
- [15] G. e. m. Zhioua, N. Tabbane, H. Labiod, and S. Tabbane, "A fuzzy multi-metric QoS-balancing gateway selection algorithm in a clustered VANET to LTE advanced hybrid cellular network," *IEEE Trans. Veh. Technol.*, vol. 64, no. 2, pp. 804–817, Feb. 2015.
- [16] X. Wu and D. C. O'Brien, "Parallel transmission LiFi," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6268–6276, Oct. 2020.
- [17] O. Narmanlioglu, R. C. Kizilirmak, F. Miramirkhani, S. Safaraliev, S. M. Sait, and M. Uysal, "Effect of wiring and cabling topologies on the performance of distributed MIMO OFDM VLC systems," *IEEE Access*, vol. 7, pp. 52 743–52 754, 2019.
- [18] J. Kahn and J. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
- [19] E. Perahia and R. Stacey, *Next Generation Wireless LAN: 802.11n and 802.11ac*. Cambridge, U.K.: Cambridge University Press, 2013.
- [20] J. Wang, Q. Hu, J. Wang, M. Chen, and J. Wang, "Tight bounds on channel capacity for dimmable visible light communications," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3771–3779, Dec. 2013.
- [21] "Queueing systems," [http://www.cs.toronto.edu/~marbach/COURSES/CSC358\\_F19/delay.pdf](http://www.cs.toronto.edu/~marbach/COURSES/CSC358_F19/delay.pdf), accessed: 2020-01-10.
- [22] S.-C. Lin, I. F. Akyildiz, P. Wang, and M. Luo, "QoS-aware adaptive routing in multi-layer hierarchical software defined networks: A reinforcement learning approach," in *2016 IEEE Int. Conf. on Services Computing (SCC)*, San Francisco, CA, 2016, pp. 25–33.
- [23] F. Tang, Z. M. Fadlullah, B. Mao, and N. Kato, "An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5141–5154, Dec. 2018.
- [24] S. Jivkova and M. Kavehrad, "Shadowing and blockage in indoor optical wireless communications," in *2003 IEEE Global Telecommunications Conf. (GLOBECOM)*, vol. 6, San Francisco, CA, 2003, pp. 3269–3273.



**Xiping Wu** (Senior Member, IEEE) received the Ph.D. degree from the University of Edinburgh, U.K., in 2015. He is an Assistant Professor in the School of Electrical and Electronic Engineering, University College Dublin (UCD), Ireland. Prior to joining UCD, he was a Postdoctoral Research Associate at the University of Oxford. His main research interests are in the areas of 6G mobile communications, optical wireless communications (OWC), and the Internet of Things. A particular focus is on developing hybrid wireless networks that integrate OWC and radio frequency, empowered by software-defined networking and artificial intelligence. He has authored or co-authored over 50 publications in these areas.



**Dominic O'Brien** (Member, IEEE) is a Professor of Engineering Science at Oxford and leads the optical communications group. He is Director of the UK Quantum technology Hub in Quantum Computing and Simulation. He has a range of experience in optoelectronics research in industry (BT labs) and academia, and has authored or co-authored approximately 200 publications in this area. The group has a wide range of experience in free space optical communications and optical wireless, demonstrating integrated optical wireless transceivers, full room-scale systems at 300Mbit/s, and integrated visible light communications systems.