

Design Methodology for Wireless Backhaul/Fronthaul Using Free Space Optics and Fibers

Atri Mukhopadhyay  and Marco Ruffini , *Senior Member, IEEE*

Abstract—The recent increase in data rates for Free space optics (FSO) transmission technology, means they could be used for designing the backhaul/fronthaul for 5G and beyond cellular networks. The flexibility and cost-effectiveness provided by FSO are the primary reasons for the mobile operators to investigate the potential of the technology as a mobile backhaul/fronthaul. Unfortunately, the reliability of FSO links is weather dependent, especially if the link covers considerable distance. Optical fibers, on the other hand, are expensive but more reliable. Hence, optimally designing a hybrid network consisting of both fiber and FSO connections can bring in the cost-effectiveness of FSO as well as the robustness of fibers. In such a design, the more important links are connected using fibers while the links with higher tolerance towards failure are designed using FSO. Therefore, in this paper, we propose a hybrid FSO/fiber backhaul/fronthauling methodology for connecting wireless base stations (BSs) to the network core. We first formulate a mixed integer non-linear program (MINLP) for determining the number of splitter/FSO distribution points required in the network that optimally provide connectivity to the BSs. The MINLP is designed to identify the locations of the splitters/FSO distribution points as well. Thereafter, we solve the MINLP with the help of particle swarm optimisation (PSO) and mixed integer linear programming (MILP) techniques. We also derive a heuristic for solving the MILP. Finally, we propose another method for determining the number of splitters required in a relatively shorter time: K-means cluster based method. The results verify that the hybrid network is cost-effective while conforming to the data rate and reliability requirements of the links. The proposal allows evaluation of a seamless design solution of the hybrid network with practical time-complexity.

Index Terms—Free space optics, passive optical networks, MINLP, MILP, particle swarm optimisation, K-means clustering, heuristic.

I. INTRODUCTION

THE telecommunication industry has witnessed a phenomenal increase in the use of smart devices over the last

Manuscript received 29 April 2022; revised 20 August 2022; accepted 23 September 2022. Date of publication 29 September 2022; date of current version 1 January 2023. This work was supported in part by Disruptive Technologies Innovation Fund (DTIF) under Grant DT 2019 0014B and in part by Science Foundation Ireland (SFI) under Grants SFI-17/CDA/4760, 12/RC/2276_P2, and 13/RC/2077_P2. (Corresponding author: Atri Mukhopadhyay.)

The authors are with the CONNECT Center, School of Computer Science and Statistics, Trinity College Dublin, Dublin 2, Ireland (e-mail: amukhopa@tcd.ie; marco.ruffini@tcd.ie).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JLT.2022.3210524>.

Digital Object Identifier 10.1109/JLT.2022.3210524

decade. As a result, the data rate requirement for catering to the communication demands of the end users has increased many-folds. The evolution of mobile networks into the fifth generation (5G) networks is a result of the extensive bandwidth requirements of the present age [1], [2]. 5G is expected to provide end users with ultra-low latency, high data rate support and ubiquitous access [3].

Network densification is a key strategy to meet the 5G requirements [2], as it enables higher spectrum reuse and therefore, higher capacity and reliability. The decrease in the size of the cells also signifies that larger number of base stations (BSs) are required to provide wireless coverage over a geographical area. As a result, network capital (CAPEX) and operational expenditure (OPEX) increase considerably, both due to the cost of BSs and high-speed connectivity (backhaul/fronthaul) [1].

The present day cellular backhaul rely mostly on two transport technologies: optical fibers and microwave radio links [4], [5], [6]. Optical fibers provide high capacity and reliability but need high initial investment [7]. Further, the fixed connections provided by optical links provide limited flexibility. On the other hand, microwave has high CAPEX, high OPEX due to spectrum licensing fees [8], lower data rate support and lower energy efficiency [9], [10]. In order to reduce the OPEX of licensed microwave links, unlicensed point to point microwave links have been employed as mobile backhaul in the recent years. Unfortunately, unlicensed microwave links are more prone to interference and are prone to unpredictable deployment issues [11]. Free space optics (FSO) is also being considered as an alternative to microwave that provides wireless connectivity. In contrast to microwave, FSO can provide high data rate connections between two points up to several kilometers [7]. FSO provides a high reuse factor, an inherent security, and robustness to electromagnetic interference. Since, FSO operates in frequencies in the unlicensed band above 300 GHz, the OPEX is low as well [7].

Unfortunately, FSO is sensitive to weather conditions. Its reachability degrades heavily in the event of fog and precipitation [12], [13]. Hence, reliability of the connection is of primary consideration when deploying a FSO link [14]. Consequently, a hybrid backhaul/fronthaul consisting of fiber and FSO can be a potent solution. The reliability of fiber over longer ranges can be complemented with the cost effectiveness and data carrying potential of FSO over shorter ranges. Further, the cost of point-to-point (P2P) fiber can be reduced by providing

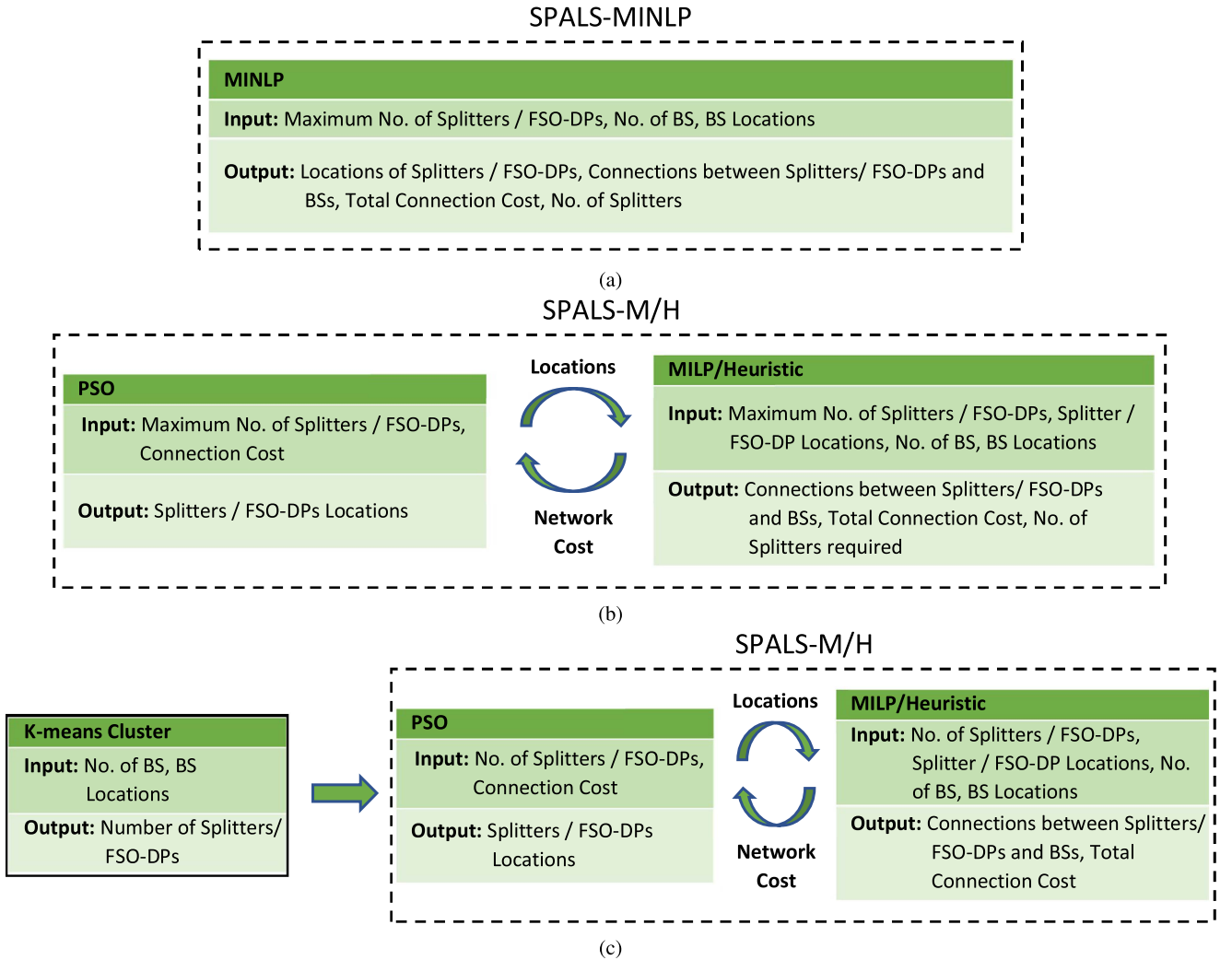


Fig. 2. High level description of the Proposal (a) SPALS-MINLP (b) SPALS-M/H (c) kMSPALS-M/H.

- 4) The splitter locations are identified using particle swarm optimisation (PSO). The PSO module accepts the number of splitters required as an input parameter. The PSO and the MILP/Heuristic algorithm work in tandem and form the SPALS module (see Fig. 2(b)).
- 5) Finally, we have proposed a K-means clustering based solution to identify number of splitters required for providing connectivity in a relatively shorter time. Note that we compromise on optimality for obtaining a faster solution in the the K-means Clustering method. We add the prefix “kM” whenever we use K-means Clustering algorithm for determining the number of splitters.

The paper is organized as follows. Section II summarises the state of the art. We present the system model in Section III. The SPALS-MINLP has been formulated in Section IV. Thereafter, we present the PSO, MILP and heuristic method for solving the optimisation problem in Section V. The K-means Clustering method for determining the number of splitters required is illustrated in Section VI. The simulation setup is elaborated in Section VII. The results are discussed in Section VIII followed by the conclusion.

II. RELATED WORKS

In this section, we briefly discuss the current state of the art on FSO networks. Excellent surveys on FSO networks can be found in [7] and [12], where detailed information about the advantages and applications of FSO, FSO channel modelling, FSO transceiver design, modulation techniques, channel coding and spatial diversity used in FSO, and information theoretical limits of an FSO system are available.

FSO has been an interesting candidate for supporting wireless networks for more than a decade now. However, the primary obstacle towards the adoption of FSO as wireless backhaul/fronthaul is the lack of reliability in longer links; especially when the weather conditions are not conducive. As a result, several innovative solutions have been proposed that enhance the link reliability. We can observe such attempts in [14], [16], [17], [18], [19], [20], [21], [22], [23], [24].

The authors of [14] introduce a linear programming model to establish a cellular backhaul using K -disjoint paths. They place mirrors in strategic locations to ensure line-of-sight (LoS). However, the mirrors merely act as reflectors, thereby increasing

the effective link path length. Hence, the reliability of the link decreases. K -disjoint paths are set up to enhance link reliability. We can find a proposal regarding setting up of backhaul links by choosing either fiber or FSO/Radio Frequency (RF) links in [18].

A Radio over FSO based fronthaul is presented in [19], where the signal from the FSO is coupled to optical fiber for further propagation towards the network core. In [20], a framework for cost optimal deployment for 5G network in a dense urban scenario is investigated. Both the number of radio resource heads (RRHs) and cost of the connections to the RRHs (using either FSO or fiber) is minimized. Two optimal FSO transceiver placement and resource allocation schemes that facilitate cooperative dynamic FSO networks are presented in [21]. The paper proposes the use of longer FSO links in clear weather. On the other hand, the FSO transceivers switch to using shorter links in case of bad weather conditions. Another fronthaul design using RF/FSO can be found in [22] where rate maximization is achieved by using efficient quantization schemes. Similar to [21], we find another network topology reconfiguration scheme in [24], where the FSO links are rearranged depending upon traffic demands or weather conditions.

We find the use of FSO in providing backhaul to wireless mesh networks (WMNs). In [16], an already operational WMN is considered that requires additional gateways for optimal operation. The additional gateways are connected to the network core using FSO/RF links. The capacity of WMNs is improved using strategically placed FSO links in [17].

In most of the works available in the literature, FSO is primarily used as an upgrade in network capacity. Further, most of the works consider that the fiber network is already present in the geographical area when the decision of using FSO links is undertaken. To the best of our knowledge, the joint problem of deciding the number of FSO transceivers required, identification of the optimum locations of the FSO transceivers and selecting the connection from the BS to the FSO-DP/splitter has not been researched upon. Therefore, in this paper, we proceed to solve this joint problem.

III. SYSTEM MODEL

In this section, we elaborate the networking environment where the SPALS will be operated. In order to provide an insight on the benefit of using SPALS, we have focused on a grid area ($20 \text{ km} \times 20 \text{ km}$) with random uniform distribution of active users (orange dots) as shown in Fig. 3. We have marked a separate area at the centre of the grid and have deployed users with higher density in that region as shown in Fig. 3 [25]. This distribution mimics the user distributions commonly found in the city centre and the suburban areas of a typical city. We have identified the positions of the BSs (black “*” signs in Fig. 3) using the K-means clustering algorithm with the restriction that a BS can provide coverage to a maximum of 10 active users, leading to an uncontended data rate of 100 Mbps per user (although different assumptions on user data rate can be easily considered). We also assume that the central distribution point “o” (shown as Splitter 1 in Fig. 1) is situated exactly at the middle of the grid, i.e., at co-ordinates (10000,10000).

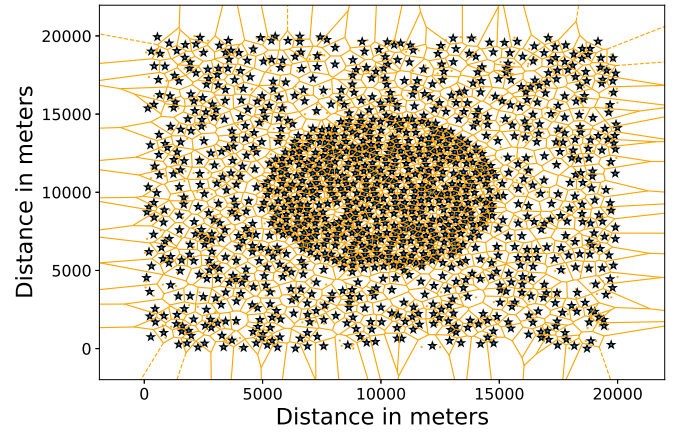


Fig. 3. User and Base Station Locations.

Thereafter, our proposal focuses on deciding the number and positions of the splitters and whether the connections are provided to the BSs through fiber or FSO links (see Fig. 1). For the small-cell fiber connectivity, we have considered passive optical networks (PONs), as this is being considered as a cost-effective alternative to point-to-point fiber links [26]. Hence, depending on the reliability and the data rate requirements of the connected BSs, we select the ideal technology (FSO or PON) for providing the connections. The overall objective is to minimize the cost of the network.

We next introduce the data rate, the reliability and the cost models used in the SPALS optimisation formulation. Please note that we have resorted to a generic model for our evaluation in this work. However, more complex and specialised models can easily be added to the SPALS optimisation formulation by including the data rate and reliability w.r.t. distance in a look-up table and thereafter, using the values as input parameters to the SPALS problem.

A. Data Rate Model

We adopt the model from [18] to determine the data rate ($D_{u,s}^{FSO}$) obtained by a FSO link given the distance of separation between the splitter s and the BS u . The relation between obtained data rate $D_{u,s}^{FSO}$ and distance $\delta_{u,s}$ is shown in (1).

$$D_{u,s}^{FSO} = \begin{cases} D_t, & \text{if } \delta_{u,s} \leq d_D \\ D_t e^{-(\delta_{u,s} - d_D)}, & \text{otherwise} \end{cases} \quad (1)$$

where, D_t is the maximum data rate provided by the FSO link and d_D is a threshold distance in kilometers.

B. Reliability Model

Similar to the data rate condition shown in (1), the reliability of the FSO link given the distance of separation between the splitter s and the BS u is calculated using (2) [18].

$$R_{u,s}^{FSO} = \begin{cases} 1, & \text{if } \delta_{u,s} \leq d_R \\ e^{-(\delta_{u,s} - d_R)}, & \text{otherwise} \end{cases} \quad (2)$$

where, d_R is a threshold distance in kilometers.

C. Cost Model

We consider FSO-DP setup cost, fiber cost and fiber installation cost for the evaluations. Setting up an FSO link requires two FSO devices (one at the FSO-DP and the other at the BS). The FSO-DPs are connected to the central distribution point via fibers. Therefore, we also need to consider the cost of connecting the FSO-DP to the central distribution point.

Fiber installation cost varies depending upon the scenario. In rural areas, trenching is required for installing the fiber. In urban areas, sewer ducts might be used for placing the fibers; thus reducing the trenching expenditure by up to 95% [27]. In our study we take in consideration the two extreme and opposite cases of 100% trenching requirement and 100% duct availability.

IV. SPLITTER PLACEMENT AND LINK SELECTION PROBLEM

In this section, we present the optimisation problem formulation for determining the number and locations of the splitters/FSO-DPs and selecting the connection technology between the BS and the splitter. For the discussions in this section and Section V, we assume that the maximum number of splitters available is equal to the number of BSs available in the scenario. The optimisation chooses a subset of the available splitters for the network deployment. We shall discuss the K-means Cluster method for a quick identification of the required number of splitters in Section VI.

Our target is to minimize the cost of the network by providing connections that are most cost effective given the circumstances. At the same time, we also need to ensure that the data-rate and the reliability provided by the links identified (both FSO and PON) are above the acceptable limits.

The cost (Ω) of the network is shown by (3).

$$\Omega = \sum_{\substack{u \in U \\ s \in S}} \theta_{u,s}^{FSO} C_{u,s}^{FSO} + \sum_{\substack{u \in U \\ s \in S}} \theta_{u,s}^{PON} C_{u,s}^{PON} + \sum_{s \in S} \beta_s \delta_{s,o} F \quad (3)$$

where, $\theta_{u,s}^{FSO}$ and $\theta_{u,s}^{PON}$ are binary variables that indicate if the link between BS u and splitter s is set up using FSO ($\theta_{u,s}^{FSO} = 1$) and fiber ($\theta_{u,s}^{PON} = 1$) respectively. $\theta_{u,s}^{FSO}$ and $\theta_{u,s}^{PON}$ are equal to zero otherwise. Similarly, $C_{u,s}^{FSO}$ and $C_{u,s}^{PON}$ are the expenditures incurred if the link between BS u and splitter s is set up using FSO and fiber respectively. β_s is a binary variable that marks whether the splitter (s) is used in the network ($\beta_s = 1$) or not ($\beta_s = 0$). $\delta_{s,o}$ is the distance between the splitter s and central distribution point o and F is the cost of fiber per meter (including fiber laying costs). The cost of fiber installation is considered only when the splitter is used, i.e. $\beta_s = 1$.

The first two terms of (3) aim at minimizing the overall cost of the network by considering the cost of the equipment used to set up a link between BS u and the splitter. Here, we have binary variables instead of a continuous variables because for setting up a link with a particular technology, we need to procure all

the necessary equipments² required to set up the connection. The third term of (3) evaluates the fiber cost for setting up a connection between the splitter s and the central distribution point (o).

We must ensure that the link between BS u and splitter s is able to satisfy the data rate requirements of BS u , which is given by (4).

$$\sum_{s \in S} \theta_{u,s}^{FSO} D_{u,s}^{FSO} + \sum_{s \in S} \alpha_{u,s}^{PON} D_{u,s}^{PON} \geq D_u^{th}, \forall u \quad (4)$$

where $0 \leq \alpha_{u,s}^{PON} \leq 1$ is a continuous variable that indicates that the link between BS u and splitter s is set up using fiber. The value of $\alpha_{u,s}^{PON}$ indicates the load that BS u brings in to the PON. $D_{u,s}^{FSO}$ and $D_{u,s}^{PON}$ captures the data rate supported if the link between BS u and splitter s is set up using FSO and fiber respectively. D_u^{th} is the data rate requirement of BS u .

Next, we ensure that the PON has enough capacity to support the traffic forwarded by the created FSO links between the BS u and splitter s (via an ONU with FSOD installed).

$$\theta_{u,s}^{FSO} D_u^{th} \leq \alpha_{u,s}^{FSO} D_{u,s}^{PON}, \forall u, s \quad (5)$$

where $0 \leq \alpha_{u,s}^{FSO} \leq 1$ is a continuous variable that indicates the load that the established FSO link between BS u and splitter s has introduced in the PON.

The total load carried by the PON should be within the capacity limits as shown in (6).

$$\sum_{\substack{u \in U \\ s \in S}} \alpha_{u,s}^{FSO} + \sum_{\substack{u \in U \\ s \in S}} \alpha_{u,s}^{PON} \leq 1 \quad (6)$$

We also ensure that either a FSO or a fiber connectivity is used to establish a link between BS u and splitter s .

$$\sum_{s \in S} \theta_{u,s}^{FSO} + \sum_{s \in S} \theta_{u,s}^{PON} \leq 1, \forall u, s \quad (7)$$

Thereafter, we need to guarantee that if a FSO or a fiber link is connected to a splitter, then the splitter needs to be marked as active.

$$\theta_{u,s}^{FSO} + \theta_{u,s}^{PON} \leq \beta_s, \forall u, s \quad (8)$$

(9) and (10) provide the relation between the continuous decision variables $\alpha_{u,s}^{FSO}$, $\alpha_{u,s}^{PON}$ and binary decision variables $\theta_{u,s}^{FSO}$, $\theta_{u,s}^{PON}$ respectively. $\alpha_{u,s}^{FSO} > 0$ and $\alpha_{u,s}^{PON} > 0$ signifies that the link between BS u and splitter s is set up using FSO and fiber respectively.

$$0 \leq \alpha_{u,s}^{FSO} \leq \theta_{u,s}^{FSO} \leq 1, \forall u, s \quad (9)$$

$$0 \leq \alpha_{u,s}^{PON} \leq \theta_{u,s}^{PON} \leq 1, \forall u, s \quad (10)$$

The reliability constraint is shown by (11).

$$\sum_{s \in S} \theta_{u,s}^{FSO} R_{u,s}^{FSO} + \sum_{s \in S} \theta_{u,s}^{PON} \geq R_u^{th}, \forall u \quad (11)$$

²The necessary equipment for a fiber link are optical transceivers and processing units, ONU cabinet, fiber links. On the other hand, the necessary equipment for setting up a FSO link are FSODs on both the FSOD housing ONU and the BS. The general ONU equipment are also required for the ONU connected to the FSOD.

where, $R_{u,s}^{FSO}$ is the reliability of the FSO link between BS u and splitter s while R_u^{th} is the reliability requirement threshold of BS u . Here, we assume that the fiber links are 100% reliable. Note that fibers have a failure probability which is directly proportional to the fiber length. However, the time between fiber-cuts are quite large and the downtime is long (several hours or days). Since, here we are interested in the short-term availability of links (i.e., due to atmospheric conditions), we are neglecting fiber link failures.

The distance between the BS u and splitter s is given by (12).

$$\delta_{u,s} = \sqrt{(L_s^x - L_u^x)^2 + (L_s^y - L_u^y)^2}, \forall u, s \quad (12)$$

where, L_u^x and L_s^x are the x-coordinates and L_u^y and L_s^y are the y-coordinates of BS u and splitter s respectively.

Similarly, the distance between splitter s and the central distribution point (o) (shown as Splitter 1 in Fig. 1) is given by (13).

$$\delta_{s,o} = \sqrt{(L_s^x - L_o^x)^2 + (L_s^y - L_o^y)^2}, \forall s \quad (13)$$

where, L_o^x and L_o^y are the x-coordinates and y-coordinates central distribution point o .

The cost of a fiber connection between BS u and splitter s is given by (14).

$$C_{u,s}^{PON} = \delta_{u,s}F, \forall u, s \quad (14)$$

The distance dependent datarate model shown in (1) can be represented by (15)–(17).

$$\delta_{u,s} \geq d_D + 0.001 - M(1 - \gamma_{u,s}), \forall u, s \quad (15)$$

$$D_t e^{-(\delta_{u,s} - d_D)} - M(1 - \gamma_{u,s}) \leq D_{u,s}^{FSO}, \forall u, s \quad (16)$$

$$D_{u,s}^{FSO} \leq D_t e^{-(\delta_{u,s} - d_D)} + M(1 - \gamma_{u,s}), \forall u, s \quad (17)$$

where, M is a large integer and $0 \leq \gamma_{u,s} \leq 1$.

Similarly, the distance dependent reliability model introduced in (2) can be represented by (18) - (20).

$$\delta_{u,s} \geq d_R + 0.001 - N(1 - \psi_{u,s}), \forall u, s \quad (18)$$

$$e^{-(\delta_{u,s} - d_R)} - N(1 - \psi_{u,s}) \leq R_{u,s}^{FSO}, \forall u, s \quad (19)$$

$$R_{u,s}^{FSO} \leq e^{-(\delta_{u,s} - d_R)} + N(1 - \psi_{u,s}), \forall u, s \quad (20)$$

where, N is a large integer and $0 \leq \psi_{u,s} \leq 1$.

Thus, we formulate the SPALS optimization problem in (21).

$$\text{minimize } \Omega \quad (21a)$$

$$\text{subject to (4) - (20)} \quad (21b)$$

The meanings of the symbols used in (3) - (20) are summarised in Table I.

Solving (21) returns the the number of splitters/FSO-DPs required, the co-ordinates of the splitters/FSO-DPs and the connection technology required to set up the BS links from the splitters/FSO-DPs.

TABLE I
SYMBOLS USED IN THE OPTIMISATION FORMULATION

Symbol	Description
S	Set of Splitters
U	Set of Base Stations
Ω	Total Cost of the Network
$\theta_{u,s}^{FSO}$	Binary decision variable where $\theta_{u,s}^{FSO} = 1$ denotes that BS u is connected to splitter s using a FSO link
$\theta_{u,s}^{PON}$	Binary decision variable where $\theta_{u,s}^{PON} = 1$ denotes that BS u is connected to splitter s using a fiber link
β_s	Binary decision variable where $\beta_s = 1$ signifies the usage of a splitter in the network
$\alpha_{u,s}^{FSO}$	Continuous decision variable where $\alpha_{u,s}^{FSO} > 0$ indicates the load on PON due to BS u connected to splitter s using a FSO link
$\alpha_{u,s}^{PON}$	Continuous decision variable where $\alpha_{u,s}^{PON} > 0$ denotes that BS u is connected to splitter s using a fiber link
$C_{u,s}^{FSO}$	Cost of FSO link for BS u and splitter s
$C_{u,s}^{PON}$	Cost of fiber link for BS u and splitter s
$D_{u,s}^{FSO}$	Data rate supported by FSO link for BS u and splitter s
$D_{u,s}^{PON}$	Data rate supported by fiber link for BS u and splitter s
D_u^{th}	Minimum data rate requirement for BS u
$R_{u,s}^{FSO}$	Reliability of FSO link for BS u and splitter s
R_u^{th}	Reliability requirement threshold for BS u
L_j^x	The x co-ordinate for the j^{th} node ($j \in \{U, S, O\}$)
L_j^y	The y co-ordinate for the j^{th} node ($j \in \{U, S, O\}$)
$\delta_{u,s}$	Distance of BS u from splitter s
$\delta_{s,o}$	Distance of splitter s from central distribution point o
F	Fiber cost per meter
D_t	Maximum data rate provided by FSO link
d_D	Maximum distance upto which a FSO link provides full data rate ($D_t^{FSO} = D_t$)
d_R	Maximum distance upto which a FSO link provides 100% reliability ($R_u^{FSO} = 1$)
M, N	Large positive number

V. SOLVING THE OPTIMISATION PROBLEM

In this section, we provide the procedure for solving the SPALS optimisation problem formulated in Section IV, which belongs to the class of Mixed Integer Non-Linear Program (MINLP). Further, the selection of the optimum number of splitters to be deployed out of the set of possible splitters makes the basic objective of the MINLP similar to a bin-packing problem [15]. Therefore, the MINLP is inherently *np-hard*. A global MINLP solver can be employed to solve it. Hence, to check the effectiveness of the formulated MINLP, we have used Python based Pyomo package [28] in conjunction with COUENNE [29], a global MINLP solver. As expected, due to the hardness of the problem, the evaluation of the formulated optimisation problem takes excessively long time; even for a (2 km \times 2 km) grid, the computation time is of the order of days. This indicates that we need an efficient heuristic algorithm if we want to solve the same problem for our desired (20 km \times 20 km) or even larger grids. Hence, we split the problem into the following two sub-parts:

- 1) Determination of splitter locations.
- 2) Identification of connection technology and optimum number of splitters.

A. Determination of Splitter Locations

In this section, we provide the procedure to identify the optimal locations of the splitters (provided that the maximum

number of splitters is known and is equal to K) so that the connections to the BSs can be performed in a cost-effective manner.

Since, we are working with multiple splitter locations on a large grid area ($20 \text{ km} \times 20 \text{ km}$), the number of combinations for the splitter locations is extremely high. Therefore, we resort to Particle Swarm Optimisation (PSO) for the purpose of identifying splitter locations [30]. PSO is a meta-heuristic that is inspired by swarm intelligence, social behavior, and food searching of a bird flock or a fish school. This algorithm has been widely used in the literature to solve non-linear optimisation problems.

In our problem, we generate L particles that form the initial population \mathcal{P} . Each of the L particles contains randomly assigned x and y coordinates of the K splitter locations.

$$\mathbf{W}^{(l)} = \begin{bmatrix} x^{(l)} \\ y^{(l)} \end{bmatrix} \quad (22)$$

where, $\mathbf{W}^{(l)}$ is a matrix that holds K random splitter locations for the particle ($l = 1, \dots, L$).

Given the splitter locations, the utility (U) is the cost returned by Section V-B after selecting the optimum splitter count and optimum connection technology. In each iteration, the PSO records the best solution among the solutions obtained by all the particles ($\mathbf{W}^{(global)}$), i.e., the splitter location combination that returns the lowest connection cost among all the particles. Additionally, each particle also records the position combination of its best performance ($\mathbf{W}^{(l,local)}$). Thereafter, the PSO calculates the velocity term $V_k^{(l)} \forall k = 1, 2, \dots, K$ at each iteration t .

$$\begin{aligned} V_k^{(l)}(t+1) &= \psi V_k^{(l)}(t) + c_1 \phi_1 \left(\mathbf{W}_k^{(l,local)}(t) - \mathbf{W}_k^{(l)}(t) \right) \\ &+ c_2 \phi_2 \left(\mathbf{W}_k^{(global)}(t) - \mathbf{W}_k^{(l)}(t) \right) \end{aligned} \quad (23)$$

where, ψ is the inertia weight that controls the convergence speed. c_1 and c_2 represent the size of the step that the particle takes toward its best individual local candidate solution $\mathbf{W}^{(l,local)}$ and the global best solution $\mathbf{W}^{(global)}$ respectively. The parameters ϕ_1 and ϕ_2 are two random positive numbers generated for each k (i.e, for each element of $\mathbf{W}^{(l)}$). The PSO updates each element k of the particle $\mathbf{W}^{(l)}$ by the following equation.

$$\mathbf{W}_k^{(l)}(t+1) = \mathbf{W}_k^{(l)}(t) + V_k^{(l)}(t+1) \quad (24)$$

The process is repeated till the PSO reaches maximum number of iterations (I_{max}). The value of I_{max} and the number of particles (L) influence the accuracy of the PSO algorithm. Using larger values of I_{max} and L return solutions that are closer to the global optimum. However, larger values of L and I_{max} require more time to execute the PSO algorithm. Hence, a trade-off between speed and accuracy is required to solve our problem. Since, we are solving a network deployment problem, the splitter locations are determined only during the network deployment phase. Hence, we have the luxury to invest sufficient time for solving the problem. Therefore, we recommend that sufficiently

large values of I_{max} and L should be used. Thus, we shall be able to achieve a solution that is very close to the global optimum.

B. Identification of Connection Technology and Optimum Number of Splitters

In this sub-section, we discuss how separating the optimisation problem into two sub-problems, i.e, the splitter location identification and the connection technology selection, reduces the MINLP introduced in Section IV into an Mixed Integer Linear Program (MILP).

1) *Converting the MINLP Into an MILP*: In the reduced optimisation problem, the distance of the splitter from the central distribution point ($\delta_{s,o}$) is provided as a constant value because the positions of the splitters are assumed to be already known (the splitter locations are obtained from the PSO module described in Section V-A). Further, the constraints (12) - (20) are no longer required because the distance of the BS u from the splitter s ($\delta_{u,s}$) is provided as a parameter. Note, when the splitter positions are already fixed, $\delta_{u,s}$, $\delta_{s,o}$, $C_{u,s}^{PON}$, $D_{u,s}^{FSO}$ and $R_{u,s}^{FSO}$ can be pre-calculated using the equations (12), (13), (14), (1) and (2) respectively. Therefore, the corresponding constraints are no longer required to be a part of the optimisation problem. Further, the variables $\gamma_{u,s}$ and $\psi_{u,s}$ are no longer required and as a result, are removed when designing the reduced optimisation problem. Pre-calculating listed set of values converts the MINLP into an MILP as -

- The objective function becomes linear as $C_{u,s}^{PON}$ and $\delta_{s,o}$ are no longer optimisation variables.
- Constraint (4) becomes linear as $D_{u,s}^{FSO}$ and $D_{u,s}^{PON}$ are pre-calculated parameters.
- Constraint (11) also becomes linear as $R_{u,s}^{FSO}$ is no longer an optimisation variable.

The resulting optimisation problem looks like (25), where $C_{u,s}^{PON}$, $D_{u,s}^{FSO}$, $D_{u,s}^{PON}$, $R_{u,s}^{FSO}$ and $\delta_{s,o}$ are pre-calculated and supplied as parameters.

$$\text{minimize} \sum_{\substack{u \in U \\ s \in S}} \theta_{u,s}^{FSO} C_{u,s}^{FSO} + \sum_{\substack{u \in U \\ s \in S}} \theta_{u,s}^{PON} C_{u,s}^{PON} + \sum_{s \in S} \beta_s \delta_{s,o} F \quad (25a)$$

$$\text{s.t. (4) - (11)} \quad (25b)$$

Hence, the optimisation framework presented in (25) is an MILP. Hence, it can be solved by a linear solver like CPLEX [31] or Gurobi [32]. We summarise the steps of pre-processing that converts the MINLP to MILP in Algorithm 1.

The value of the objective function of the MILP (25) act as the cost function of the PSO described in Section V-A. Working in tandem with the PSO algorithm, the MILP obtains the locations of the splitters and also the connection technologies to the BSs that return the lowest cost (see Fig. 2). The connections are identified by the values of $\theta_{u,s}^{FSO}$ and $\theta_{u,s}^{PON}$.

Thus, we solve the SPALS problem. Since, Optimisation (25) is a MILP, the combination of PSO and MILP results in a much quicker method to find a near optimal splitter location. We call the combination of PSO and MILP solution method of SPALS as SPALS-M.

Algorithm 1: Pre-processing for Conversion from MINLP to MILP.

Data: $S, U, L_s^x, L_s^y, L_u^x, L_u^y, L_o^x, L_o^y, F, d_D, d_R, D_u^{th}, D_t$
Result: $\delta_{u,s}, C_{u,s}^{PON}, D_{u,s}^{FSO}, R_{u,s}^{FSO}, C_{u,o}^{PON}$

```

while  $s \in S$  do
   $\delta_{s,o} = \sqrt{(L_s^x - L_o^x)^2 + (L_s^y - L_o^y)^2}$ ;
   $C_{s,o}^{PON} = \delta_{s,o}F$ ;
  while  $u \in U$  do
     $\delta_{u,s} = \sqrt{(L_s^x - L_u^x)^2 + (L_s^y - L_u^y)^2}$ ;
     $C_{u,s}^{PON} = \delta_{u,s}F$ ;
    if  $\delta_{u,s} \leq d_D$  then
       $D_{FSO}^{u,s} = D_t$ ;
    else
       $D_{FSO}^{u,s} = D_t e^{-(\delta_{u,s} - d_D)}$ ;
    if  $\delta_{u,s} \leq d_R$  then
       $R_{FSO}^{u,s} = 1$ ;
    else
       $R_{FSO}^{u,s} = e^{-(\delta_{u,s} - d_R)}$ ;

```

2) *Moving Towards a Heuristic:* Even though the solution times of the MILPs are shorter than MINLPs, our designed MILP, being a bin packing problem, is also inherently *np-hard*. Therefore, as the size of the data set increases, the solution time of the MILP may no longer be practical. As a result, we look into alternatives to solve our MILP (25) with polynomial time complexity.

Since, the MILP (25) is trying to establish a FSO or a fiber link between the BS $u \in U$ and the splitter $s \in S$ after identifying the optimum number of splitters required, we can infer the following

- If $R_{u,s}^{FSO} < R_u^{th}$ and/or $D_{u,s}^{FSO} < D_u^{th}$ then it is not possible to set up a FSO link for connecting u and s because the reliability and/or the data rate constraints are no longer satisfied by the best available FSO link.
- If $C_{u,s}^{PON} < C_{u,s}^{FSO}$, then the optimisation will always select a fiber link for connecting s and u because the target of the optimisation is to minimise cost. Note that PON always meets the data rate and the reliability requirements.
- BS u will always be connected to the splitter s that has the lowest fiber connection cost to the central distribution point while satisfying the reliability and the data rate constraints. Since we are going towards a greedy heuristic, we shall be selecting the BSs one by one for connection. Therefore, if a previously chosen BS (u') is already connected to a certain splitter (s'), the connection cost of that particular splitter (s') to the central distribution point will be deemed as zero for the current BS (u).

Therefore, we can perform the following operation without changing the solution of the problem (the process is summarised in Algorithm 2).

- If $R_{u,s}^{FSO} < R_u^{th}$ or $D_{u,s}^{FSO} < D_u^{th}$ or $C_{u,s}^{PON} < C_{u,s}^{FSO}$ then we set the cost of the FSO link ($C_{u,s}^{FSO}$) between s and u as very high ($C_{u,s}^{FSO} \gg C_{u,s}^{PON}$). This measure ensures that

Algorithm 2: Heuristic Pre-processing.

Data: $S, U, L_s^x, L_s^y, L_u^x, L_u^y, F, d_D, d_R, D_u^{th}, R_u^{th}, D_t$,
 $D_{u,s}^{FSO}, R_{u,s}^{FSO}, C_{u,s}^{FSO}, C_{u,s}^{PON}, C_{u,o}^{FSO}$
Result: $C_{u,s}^{FSO}, U_{sorted}^{Distance}$

```

while  $u \in U$  do
  while  $s \in S$  do
    if  $D_{u,s}^{FSO} < D_u^{th}$  or  $R_{u,s}^{FSO} < R_u^{th}$  or
        $C_{u,s}^{FSO} > C_{u,s}^{PON}$  then
       $C_{u,s}^{FSO} = C_{u,s}^{FSO}_{max}$ ;
   $U_{sorted}^{Distance} = \text{sortDescending}(U)$ ;

```

Algorithm 3: Heuristic.

Data: $D_{u,s}^{FSO}, D_{u,s}^{PON}, C_{u,s}^{FSO}, C_{u,s}^{PON}, \delta_{u,s}, U_{sorted}^{Distance}$
Result: $Obj, \theta_{u,s}^{FSO}, \theta_{u,s}^{PON}, \beta_s$

```

totCap $_s = 1, \forall s$ ;
Obj = 0;
while  $s \in S$  do
   $\beta_s = 0$ ;
  while  $u \in U_{sorted}^{Distance}$  do
    while  $s \in S$  do
      temp $C_s^{FSO} = C_{u,s}^{FSO} + (1 - \beta_s)C_{s,o}^{PON}$ ;
      temp $C_s^{PON} = C_{u,s}^{PON} + (1 - \beta_s)C_{s,o}^{PON}$ ;
      min $C_s^{FSO} = \min(tempC_s^{FSO})$ ;
      minLoc $^{FSO} = \arg \min(tempC_s^{FSO})$ ;
      min $C_s^{PON} = \min(tempC_s^{PON})$ ;
      minLoc $^{PON} = \arg \min(tempC_s^{PON})$ ;
      if min $C_s^{PON} > minC_s^{FSO}$  then
         $\theta_{u,minLoc}^{FSO} = 1$ ;
        Obj = Obj +  $C_{u,minLoc}^{FSO}$ ;
         $\beta_{minLoc}^{FSO} = 1$ ;
      else
         $\theta_{u,minLoc}^{PON} = 1$ ;
        Obj = Obj +  $C_{u,minLoc}^{PON}$ ;
         $\beta_{minLoc}^{PON} = 1$ ;
    while  $s \in S$  do
      Obj = Obj +  $\beta_s C_{u,o}^{PON}$ ;

```

a fiber link is always selected over FSO link for connecting s and u .

The reduced problem can be solved using a heuristic approximation algorithm shown in Algorithm 3. We reduce the number of candidate splitters for a certain BS using the fact that the least expensive splitter in the current iteration will always be used for providing connection to a BS. Hence, effectively, for every BS, there is only a single candidate splitter with two possible connections (FSO and PON). The fiber link capacities can be ignored for the optimisation as the problem is modeled to have PON capacity suitable to transport the BS data rate requirements. Even if the aggregated BS data rate requirement becomes high, PON capacity can be enhanced by adding higher channel rates (100 Gbps and above) and multiple wavelength channels, possibly even dedicating one WDM channel to just one BS (i.e., through point-to-point wavelength overlay). Thus,

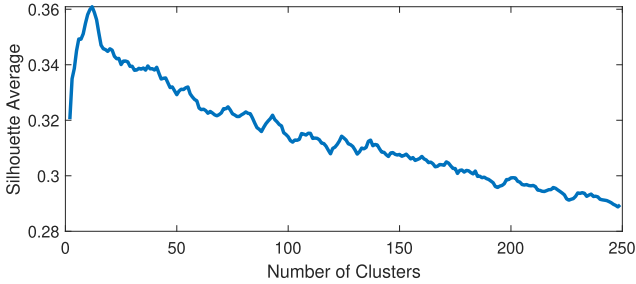


Fig. 4. K-means Cluster Silhouette Average for the Example Distribution of Base Stations.

TABLE II
SIMULATION PARAMETERS

Parameter	Values
Fiber cost	1 Unit per meter [20]
Trenching cost	1300 Units per meter [20]
Cost (Ducts already present)	65 Units per meter
Single FSO device cost	5000 Units [20]
d_D	3 km
d_R	2 km
Grid size	20 km \times 20 km
Central distribution point coordinate	(10000m, 10000m)

the problem becomes equivalent to a choosing between FSO and fiber links. We name the combination of PSO and Heuristic solution method of SPALS as SPALS-H.

Note that the worst case time complexity of the heuristic pre-processing is $O(|S||U|)$, where $|S|$ and $|U|$ are the number of splitters and BSs respectively. Similarly, the worst case time complexity of the Heuristic is $O(|S||U| + |S|)$.

VI. K-MEANS CLUSTERING BASED SOLUTION

In Section V, we have presented a method of solving the SPALS-MINLP (21) with reduced complexity. However, identifying the number of splitters to be placed from a higher number of candidate splitters, is computationally expensive. Therefore, in this section, we develop a K-means Clustering based solution for SPALS (kMSPALS) for quickly determining the number of splitters. Note that in kMSPALS, we sacrifice optimality for faster results.

We observe that the cost involved in providing connectivity either with FSO or fiber links is directly proportional to the distance of separation between the splitter location and the BS. Therefore, we infer that the K-means clustering algorithm can identify the number of cluster centroids required to group the number of BSs. Thereafter, we can use the method given in Section V-A to identify the exact splitter locations.

From the literature, we can find that the Silhouette method is one of the most efficient methods of determining the number of clusters required to optimally group a set of data-points [33], [34]. The Silhouette method uses a collection of *proximities* to construct silhouettes. The proximities include *similarities* and *dissimilarities* between the objects of the clusters [33].

Let us assume that there are K clusters and a sample is assigned to cluster $C(i)$. Let $|C(i)|$ denote the number of

TABLE III
SIMULATION PARAMETERS FOR SCALED DOWN NETWORK

Parameter	Values
Fiber cost	1 Unit per meter
d_D	0.4 km
d_R	0.2 km
Single FSO device cost	1000 Units
Grid size	2 km \times 2 km
Central distribution point coordinate	(1000m, 1000m)

samples in cluster $C(i)$. Then, we calculate two metrics $a(i)$ and $b(i)$ for the sample i using (26). Formally, $a(i)$ is the average dissimilarity of sample i to its cluster and $b(i)$ is the average dissimilarity of sample i to the closest cluster apart from cluster $C(i)$.

$$a(i) = \frac{1}{|C(i)| - 1} \sum_{\substack{i, j \in C(i) \\ i \neq j}} d(i, j)$$

$$b(i) = \min_{i \neq j} \left(\frac{1}{|C(j)|} \sum_{j \in C(j)} d(i, j) \right) \quad (26)$$

After calculating $a(i)$ and $b(i)$, the Silhouette of sample i is calculated as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$$= \begin{cases} 1 - \frac{a(i)}{b(i)}, & a(i) < b(i) \\ 0, & a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, & a(i) > b(i) \end{cases} \quad (27)$$

where, $s(i)$ is the Silhouette value of sample i , and $-1 \leq s(i) \leq 1$. Note that higher $s(i)$, denotes better clustering [33].

Finally, the average of all the Silhouette values is calculated.

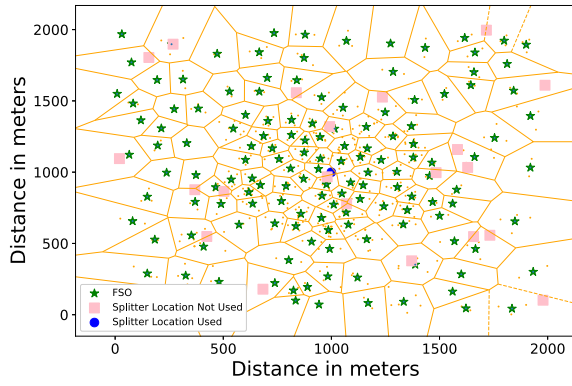
$$\overline{S}_K = \frac{\sum_i s(i)}{N} \quad (28)$$

where, N is the total number of samples and K is the number of clusters.

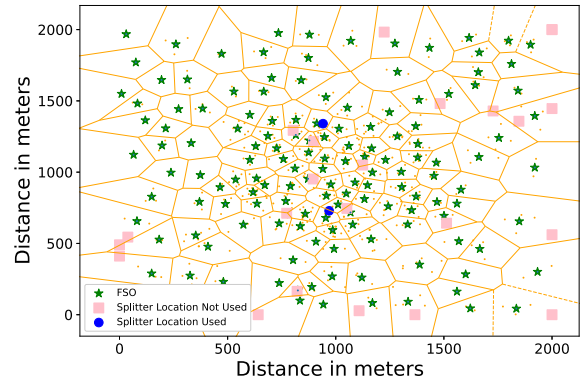
In kMSPALS, the number of clusters (K) is varied from 2 to the total number of samples present in the data set and the value of \overline{S}_K is noted in each case. The value of K that returns the highest \overline{S}_K is determined as the optimum number of clusters. In Fig. 4, we show the silhouette average of the BS distribution illustrated in Fig. 3. From Fig. 4, we can see that the maximum value of \overline{S}_K is obtained for $K = 12$. Therefore, we select $K = 12$ as the number of splitters that are required to be placed for kMSPALS in the example grid area.

VII. SIMULATION SETUP

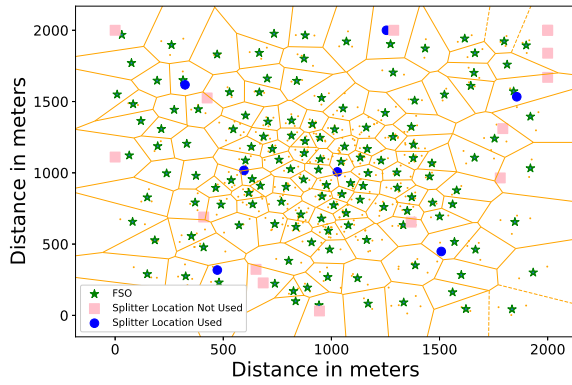
The evaluation framework for the proposal was generated using Python and the optimisation framework was coded using the Pyomo package. Initially, COUENNE, a non-linear solver was used for solving the SPALS-MINLP in a scaled down scenario (2 km \times 2 km grid). The Gurobi solver was used for solving the MILP (both in the scaled down and the original



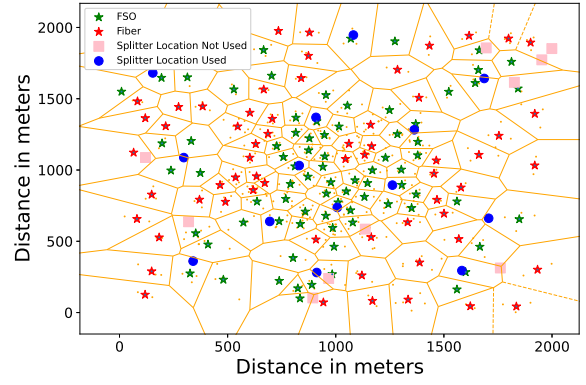
(a) Reliability Threshold = 0.1; Fiber = 0%, FSO = 100%



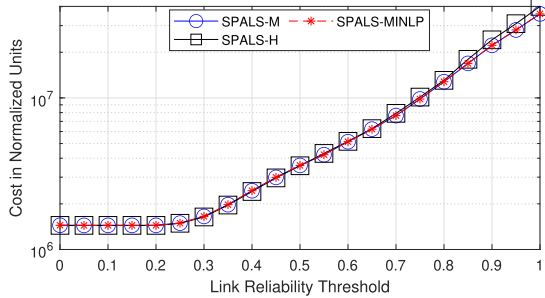
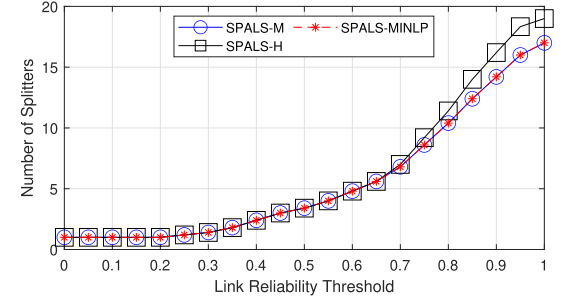
(b) Reliability Threshold = 0.4; Fiber = 0%, FSO = 100%



(c) Reliability Threshold = 0.7; Fiber = 0%, FSO = 100%



(d) Reliability Threshold = 0.9; Fiber = 33%, FSO = 67%

(e) Cost of a $(2\text{km} \times 2\text{km})$ network w.r.t the BS reliability requirement threshold.(f) Number of splitters required in a $(2\text{km} \times 2\text{km})$ network w.r.t the BS reliability requirement threshold.Fig. 5. Illustration of Base Station connection using FSO or Fiber applying SPALS on a $(2\text{km} \times 2\text{km})$ grid.

scenario). The parameters used in the study are summarised in Table II. The cost of fiber is taken as 1 unit per meter. If trenching is required for installing the fibers, an additional cost of 1300 units per meter has been assumed. On the other hand, if ducts are already present in a location (e. g. sewer ducts), trenching is no longer required. Hence, we have used 65 units per meter for installation cost when ducts are already present.

The cost of a hybrid RF/FSO link is taken to be independent of the distance (up to the system maximum reach). Since two FSODs are required to set up an FSO link (one at the ONU and the other at the BS), 10 k units has been adopted as the cost of a FSO link. It should be noted that the threshold distances for data rate (d_D) and reliability (d_R) were considered following the work in [18]. In particular, the reliability threshold distance

refers to a light fog scenario as we wanted to benchmark our proposal against a fairly challenging scenario for FSO communication. It is envisaged that such values would need to be adapted to specific locations, i.e., using statistical weather data, in order to obtain results that are accurate to given locations.

In order to benchmark our simplified proposals, SPALS-M and SPALS-H, we have first solved SPALS-MINLP using COUENNE. Unfortunately, the solution search space for the $20\text{km} \times 20\text{km}$ grid is very large and as a result, was unsolvable in a laboratory computer. Therefore, we scaled down the network to a $2\text{km} \times 2\text{km}$ grid. The parameters used for this case are shown in Table III. Notice that here some parameters were artificially scaled down to maintain the proportion to the network size and obtain insightful solution in the scaled down network.

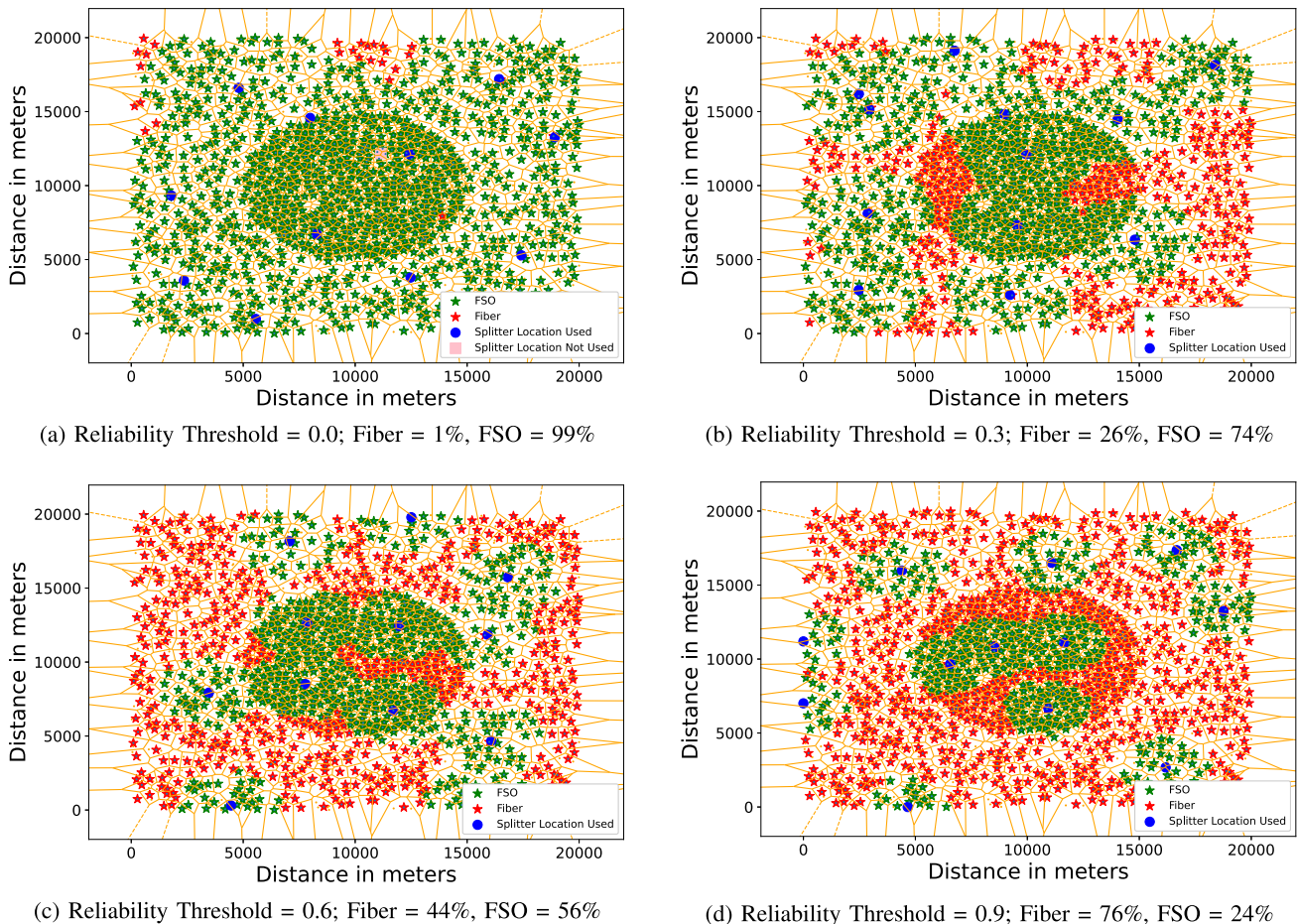


Fig. 6. Illustration of Base Station connection using FSO or Fiber applying kMSPALS; 12 s level splitters used.

Thus, while these are useful for the purpose of algorithm comparison, the parameters are not representative of actual system specifications.

VIII. RESULTS

In this section, we provide results that gauge the performance of the proposal.

A. Comparison With Performance of COUENNE

We start the evaluation of our proposal by first solving the SPALS-MINLP (21) with the help of COUENNE. For the experiment, we set the maximum number of splitters to be placed as 25. The MINLP selects the optimum number of splitters to be used for cost minimisation. We also take a reduced grid area of $2\text{ km} \times 2\text{ km}$. We resort to such a reduced system as COUENNE takes a long time to converge. While the area is reduced, the random assignment of end users still provides sufficient generality to carry out the comparison. We also execute SPALS-M and SPALS-H algorithms on the same grid.

In Fig. 5(e), we find that the final network deployment costs are similar for both SPALS-MINLP and SPALS-M/H (both MILP and Heuristic) based solutions. As expected, COUENNE, being a global optimiser provides better solution. On the other

hand, SPALS-H being a greedy heuristic, it provides slightly sub-optimal solution compared to SPALS-MINLP and SPALS-M. The difference in the cost returned by SPALS-MINLP and SPALS-H is approximately within 1.5% up to a reliability threshold of 0.9 and then about 12% at reliability threshold approximately equal to 1. SPALS-M, on the other hand, approximately reaches the global minimum. The slight difference between SPALS-MINLP and SPALS-M is mainly due to the stochastic nature of the PSO algorithm. Hence, we extrapolate that it is reliable to use the SPALS-M for larger grid areas and obtain sufficiently accurate results. Further, SPALS-H might be also used to obtain a fast solution with reasonable accuracy.

Further, it is observed in Fig. 5(e) the cost of the network starts increasing for values of the reliability threshold above 0.2. The primary reason behind the escalation in cost is the increase in the number of splitters required to meet the data rate and reliability thresholds of the network (as seen in Fig. 5(f)). Further, at higher reliability thresholds, a large number of BSs are connected using fiber links (as seen in Fig. 5(d)) as the FSO links become inadequate. Hence, the cost of the network increases even further.

Looking at the time of execution, SPALS-MINLP, SPALS-M and SPALS-H require (> 5 days), (≈ 2 hrs) and (≈ 15 mins) respectively for solving the optimization problem for a single

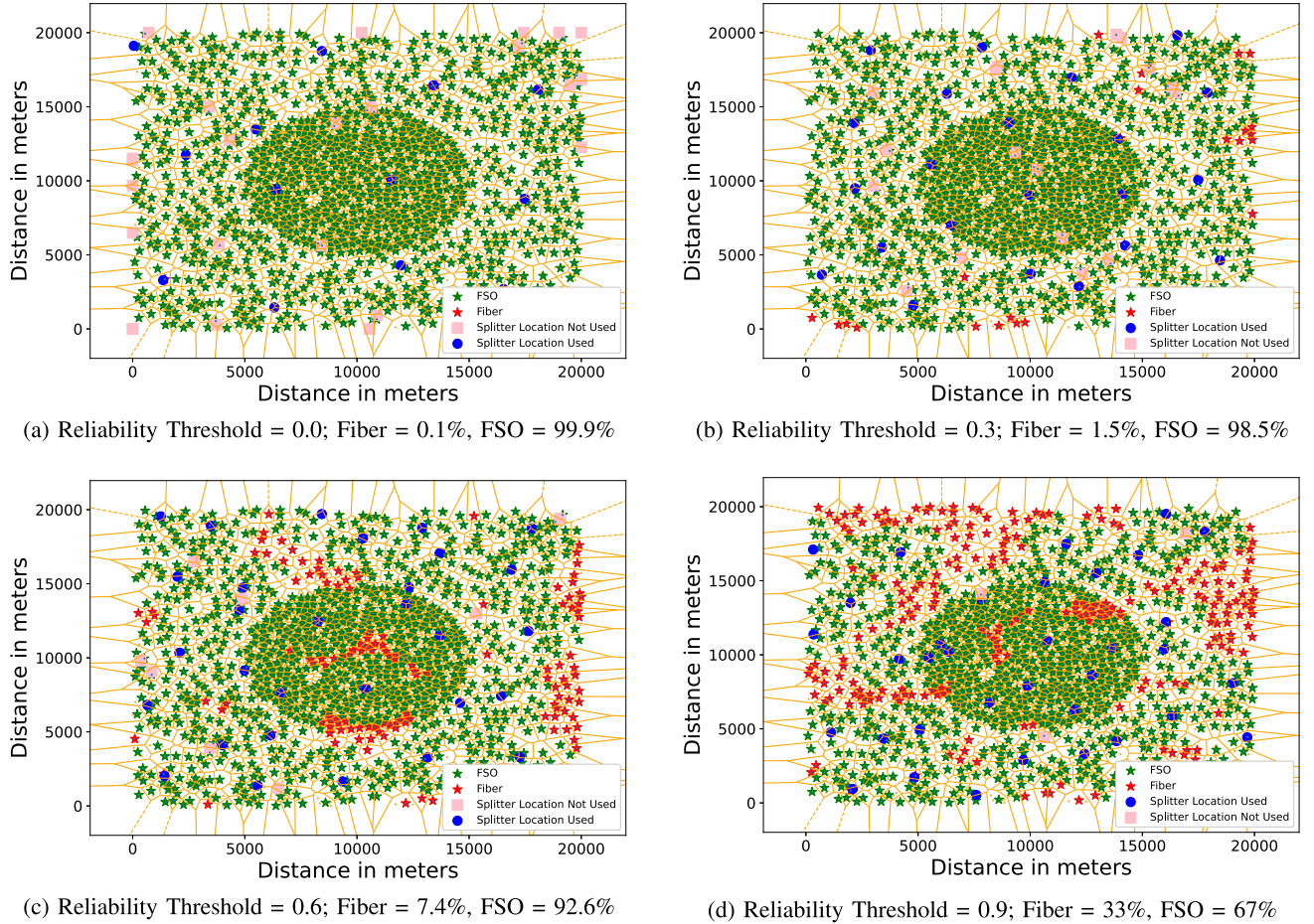


Fig. 7. Illustration of Base Station connection using FSO or Fiber applying SPALS-H.

reliability threshold value shown in Fig. 5(e). Hence, SPALS-M/H is more suitable for executing SPALS over larger grid areas.

B. Evaluation of SPALS in Larger Grids

1) *kMSPALS Evaluation*: We present the evaluation of kMSPALS first as the intuition with the fixed number of splitters returned by the K-Means Clustering algorithm helps in the explanation of the SPALS performance later on. Once the maximum number of splitters to be used is identified, we employ SPALS-H to obtain the positions of the splitters and optimum connections from the splitters to the BSs. The maximum number of splitters to be used for the network deployment is obtained using the K-Means Clustering algorithm. Hence, the number of splitters required cannot exceed the number of splitters returned by the K-Means Clustering algorithm.

In Fig. 6, the splitter locations that are used are represented by *blue dots*. On the other hand, the unused splitter locations are denoted by *pink squares*. The BSs connected using FSO to the splitter is marked by *green stars* and the BSs connected to the splitter using fiber is marked by *red stars*. We can clearly observe that as the desired reliability of the links increases, higher number of BSs are being connected using fiber links. However, as the reliability expectation of the links decreases,

slowly the majority of the connections shift towards FSO. Note that the PSO module of the SPALS-H algorithm optimises the splitter locations so that the cost of the network can be minimized while satisfying the operational constraints.

2) *SPALS Evaluation*: In Fig. 7, we obtain the number of splitters required to be deployed in the 20 km \times 20 km grid by using SPALS. The primary difference between Fig. 6 and Fig. 7 is in the difference in the maximum number of splitters that can be used for network deployment. The ideal value to be used for the maximum number of splitters is equal to the number of BSs in the area. However, a sufficiently large value for which we can observe unused splitter locations even at the highest desired link reliability threshold is an acceptable value for the maximum number of splitters.

The increase in the number of splitters in the network deployment has the following two contrasting effects on the network deployment cost -

- *The BSs are now closer to the splitters* - As a result, higher number of BSs can be connected using comparatively inexpensive FSO links. Alternatively, if the BS needs to be connected to the splitter using a fiber connection, the fiber laying cost is lower due to the lower distance between the splitter and the BS. Thus, network deployment cost decreases.

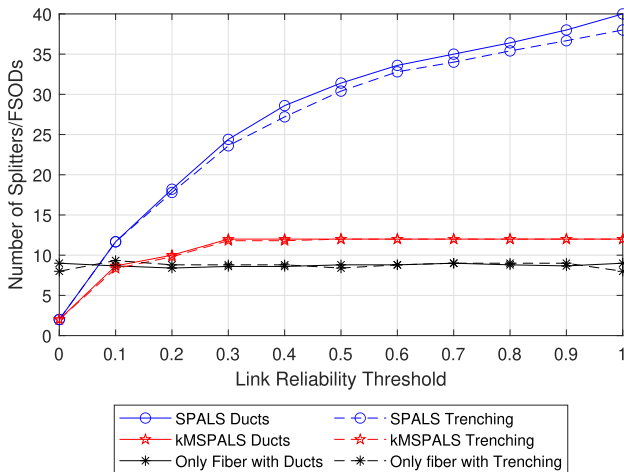


Fig. 8. Number of Splitters required for the 20 km \times 20 km grid w.r.t the link reliability threshold.

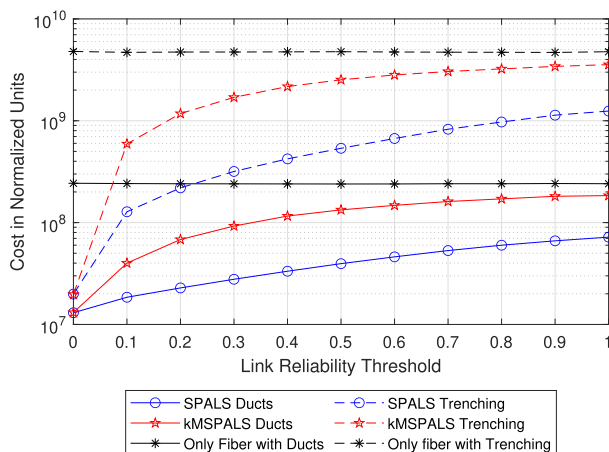


Fig. 9. The cost of a network w.r.t the link reliability threshold.

- *Higher number of splitters needs to be connected to the PON using fibers* - As the splitters/FSO-DPs are providing connectivity to a set of BSs, the splitter connectivity must be absolutely reliable and therefore, fiber connections are used. Therefore, the connection to the splitters are expensive. Moreover, installing more splitters/FSO-DPs also increases the overall cost of the network. Thus, we observe an increase in network deployment cost.

The above two conflicting effects are weighed by SPALS to reduce the overall cost of the network by selecting the optimum number of splitters required. Hence, the cost returned by SPALS is lower than kMSPALS as observed from Fig. 9.

C. Comparison of Network Deployment Costs

Finally, the cost of the network is a direct consequence of the connection technology used in the network. FSO links are more cost effective over longer distances when compared to fiber links. However, an FSO link's reliability decreases with distance. Therefore, as the desired network reliability increases, the cost of the network increases as well (either due to the usage

of more splitters or more fiber connections). We can observe the effect of desired reliability on network cost in Fig. 9. Further, the cost of fiber deployment is heavily dependent on whether or not ducts are already available in the region or not. We show the extreme cases with full or zero availability of ducts in Fig. 9. Any practical deployment would cost somewhere between the extreme cases.

In SPALS, we evaluate the cost of network deployment with the maximum allowable splitter set cardinalities to be 50 (as we found it is a sufficiently large number for a 20 km \times 20 km grid). Thereafter, we choose the best splitter set that returns the minimum network deployment cost. kMSPALS, on the other hand, identifies the splitter set cardinality without considering the reliability and data rate constraints and hence provides suboptimal solution. As a result, the SPALS minimises the network deployment cost more efficiently as compared to kMSPALS. The observation from Fig. 9 reinforces our claim; even when the link reliability requirement is 1, the SPALS algorithm provides a cost reduction of about 73% and 70% w.r.t to a scenario with all fiber connections and where ducts are absent or present respectively. On the other hand, kMSPALS provides 25% and 23% improvement when ducts are absent and present respectively.

IX. CONCLUSION

FSO is a promising solution for providing backhaul/fronthauling services to 5G and beyond BSs. Unfortunately, the weather dependency of FSO links can be an obstacle towards its large scale adoption. However, FSO links are sufficiently reliable over shorter distances. As a result, their benefits can be availed in a hybrid network with FSO and optical fibers. In this paper, we have proposed an efficient and near optimal method for designing a hybrid FSO/fiber network with the objective of minimizing network deployment cost. The proposed method carefully considers the reliability and data rate requirements of the connected wireless BSs while selecting the link technology. The results reflect the dependency of network cost on the reliability requirements. The cost reduction is especially evident (up to 73% even for high desired link reliability) when ducts are not present in the scenario and trenching is required for fiber installation. The proposed network deployment method produces results in practical time limits and therefore, is suitable for adoption in network planning. As a future work, we intend to extend the deployed network to a multihop FSO scenario.

REFERENCES

- [1] J. G. Andrews et al., "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
- [2] M. Shafi et al., "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.
- [3] I. Chih-Lin, H. Li, J. Korhonen, J. Huang, and L. Han, "RAN revolution with NGFI (xhaul) for 5G," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 541–550, 2018.
- [4] O. Tipmangkolsilp, S. Zaghloul, and A. Jukan, "The evolution of cellular backhaul technologies: Current issues and future trends," *IEEE Commun. Surv. Tut.*, vol. 13, no. 1, pp. 97–113, Jan.–Mar. 2011.

- [5] A. Mukhopadhyay and G. Das, "A ring-based wireless optical network to reduce the handover latency," *J. Lightw. Technol.*, vol. 33, no. 17, pp. 3687–3697, Sep. 2015.
- [6] A. Mukhopadhyay and G. Das, "Design and analysis of a resilient packet ring-based backhaul for LTE networks that reduces handover latency," *Photonic Netw. Commun.*, vol. 32, no. 3, pp. 438–452, 2016.
- [7] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Commun. Surv. Tut.*, vol. 16, no. 4, pp. 2231–2258, Oct.–Dec. 2014.
- [8] M. Paolini, "Crucial economics for mobile data backhaul," 2011. [Online]. Available: https://www.alliancecorporation.ca/images/documents/broadband-documents/Cambridge_Broadband/Cambridge-Broadband-Whitepaper-Crucial-economics-for-mobile-data-backhaul.pdf
- [9] S. Chia, M. Gasparroni, and P. Brick, "The next challenge for cellular networks: Backhaul," *IEEE Microw. Mag.*, vol. 10, no. 5, pp. 54–66, Aug. 2009.
- [10] P. Monti, S. Tombaz, L. Wosinska, and J. Zander, "Mobile backhaul in heterogeneous network deployments: Technology options and power consumption," in *Proc. IEEE 14th Int. Conf. Transparent Opt. Netw.*, 2012, pp. 1–7.
- [11] M. Coldrey, J.-E. Berg, L. Manholm, C. Larsson, and J. Hansryd, "Non-line-of-sight small cell backhauling using microwave technology," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 78–84, Sep. 2013.
- [12] S. A. Al-Gailani et al., "A survey of free space optics (FSO) communication systems, links, and networks," *IEEE Access*, vol. 9, pp. 7353–7373, 2020.
- [13] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, *Optical Wireless Communications: System and Channel Modelling With Matlab*. Boca Raton, FL, USA: CRC Press, 2019.
- [14] Y. Li, N. Pappas, V. Angelakis, M. Pióro, and D. Yuan, "Optimization of free space optical wireless network for cellular backhauling," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1841–1854, Sep. 2015.
- [15] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*. Hoboken, NJ, USA: Wiley, 1990.
- [16] M. N. Smadi, S. C. Ghosh, A. A. Farid, T. D. Todd, and S. Hranilovic, "Free-space optical gateway placement in hybrid wireless mesh networks," *J. Lightw. Technol.*, vol. 27, no. 14, pp. 2688–2697, Jul. 2009.
- [17] F. Ahdi and S. Subramaniam, "Capacity enhancement of RF wireless mesh networks through FSO links," *J. Opt. Commun. Netw.*, vol. 8, no. 7, pp. 495–506, Jul. 2016.
- [18] A. Douik, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "Hybrid radio/free-space optical design for next generation backhaul systems," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2563–2577, Jun. 2016.
- [19] K. Ahmed, A. E. Morra, and S. Hranilovic, "Network planning of uplink all-optical passive FSO/OF C-RAN fronthaul," *J. Opt. Commun. Netw.*, vol. 11, no. 12, pp. 600–612, Dec. 2019.
- [20] F. Tonini, C. Raffaelli, L. Wosinska, and P. Monti, "Cost-optimal deployment of a C-RAN with hybrid fiber/FSO fronthaul," *J. Opt. Commun. Netw.*, vol. 11, no. 7, pp. 397–408, Jul. 2019.
- [21] A. S. Ghazy, M. A. Hasabelnaby, H. A. Selmy, and H. M. Shalaby, "Optimal transceiver placement and resource allocation schemes in cooperative dynamic FSO networks," *J. Opt. Commun. Netw.*, vol. 11, no. 11, pp. 512–524, Nov. 2019.
- [22] M. Najafi, V. Jamali, D. W. K. Ng, and R. Schober, "C-RAN with hybrid RF/FSO fronthaul links: Joint optimization of fronthaul compression and RF time allocation," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8678–8695, Dec. 2019.
- [23] B. Bag, A. Das, I. S. Ansari, A. Prokeš, C. Bose, and A. Chandra, "Performance analysis of hybrid FSO systems using FSO/RF-FSO link adaptation," *IEEE Photon. J.*, vol. 10, no. 3, pp. 1–17, Jun. 2018.
- [24] Z. Gu, J. Zhang, Y. Ji, L. Bai, and X. Sun, "Network topology reconfiguration for fso-based fronthaul/backhaul in 5G wireless networks," *IEEE Access*, vol. 6, pp. 69426–69437, Nov. 2018.
- [25] H. Ghazzai, E. Yaacoub, M.-S. Alouini, Z. Dawy, and A. Abu-Dayya, "Optimized LTE cell planning with varying spatial and temporal user densities," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1575–1589, Mar. 2016.
- [26] H. S. Abbas and M. A. Gregory, "The next generation of passive optical networks: A review," *J. Netw. Comput. Appl.*, vol. 67, pp. 53–74, 2016.
- [27] Huawei Technologies Co., Ltd., "Gigaband network development white paper," 2016. [Online]. Available: <https://www-file.huawei.com//media/CORPORATE/PDF/white%20paper/Gigaband-Network-EN.pdf>
- [28] W. E. Hart, J.-P. Watson, and D. L. Woodruff, "Pyomo: Modeling and solving mathematical programs in python," *Math. Program. Comput.*, vol. 3, no. 3, pp. 219–260, 2011.
- [29] COIN-OR (Project Manager P. Belotti), "Convex over and under envelopes for nonlinear estimation," [Online]. Available: <http://www.coin-or.org/Couenne/>
- [30] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, 1995, vol. 4, pp. 1942–1948.
- [31] Cplex, IBM ILOG, "V12, 1: User's manual for cplex," [Online]. Available: https://www.ibm.com/docs/en/SSSA5P_12.8.0/ilog.odms.studio.help/pdf/usrcplex.pdf
- [32] Gurobi Optimization, LLC, "Gurobi optimizer reference manual," 2022. [Online]. Available: <https://www.gurobi.com>
- [33] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, 1987.
- [34] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, vol. 344. Hoboken, NJ, USA: Wiley, 2009.