Capacity Enhancement of RF Wireless Mesh Networks Through FSO Links

Farshad Ahdi and Suresh Subramaniam

***Abstract*—RF-based technologies are often used to address the last-mile problem. However, RF technologies generally create a bottleneck for the ever-increasing demand for bandwidth, causing the under-utilization of core network resources. Enhancing the capacity of wireless mesh networks can be done efficiently and cost effectively through boosting the capacity of some strategically located links. In this paper, we propose link augmentation via free-space optics (FSO) links. The high capacity of optical links is the key in link capacity enhancement in the resulting hybrid network. Adopting a TDMA-based framework, we formulate the problem of transceiver placement and the scheduling of RF links to maximize network capacity. To avoid the complexity of the ILP, we also propose a randomized greedy heuristic algorithm. In addition, simulated annealing is used to provide a baseline for comparison and to show the efficiency of our heuristic. We show by simulations that our heuristic, which can be implemented efficiently, achieves a high fraction of the optimal network capacity.**

***Index Terms*—Capacity augmentation; Free space optics; ILP; Wireless mesh.**

I. Introduction

There is an ever-increasing demand for bandwidth due to emerging media-rich applications. Fiber-based networks address such demands; however, the cost of the infrastructure as well as infeasibility of deployment in some areas are the major barriers. Using wireless technologies to form wireless mesh networks (WMNs) addresses the last-mile problem. These networks are easy to deploy and incur little deployment cost, but at the price of lower data rates. Despite their popularity, RF-based WMNs degrade the overall network performance in terms of delay and throughput. In addition, the asymmetrical growth of bandwidth demands in the uplink and downlink directions leads to the under-utilization of the network. For example, the emergence of media-rich applications such as on-demand streaming of HD videos over the Internet has significantly tilted the growth rate in the downlink direction. Appropriate scheduling along with partial capacity enhancement of WMNs can greatly improve such inefficiencies.

Manuscript received July 30, 2015; revised April 14, 2016; accepted May 14, 2016; published June 22, 2016 (Doc. ID 246910).

The authors are with the Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA (e-mail: farshad.ahdi@gmail.com).

http://dx.doi.org/10.1364/JOCN.8.000495

Boosting the network capacity using RF-based solutions such as the deployment of multiple radios or MIMO antennas at some nodes is the standard practice. Besides requiring significant upgrades in the physical layer, these approaches are not scalable mainly due to the interference they introduce to the network. A wireless hybrid non-RF approach is pursued in this paper, as it can potentially enhance the network capacity significantly. In particular, free-space optics (FSO) as the link upgrade technology is investigated due to its high bandwidth, low interference, and quick setup time. Unlike their RF counterparts, FSO links are strongly dependent on weather conditions (e.g., fog, rain), so a network solely based on FSO technology may not be a reliable solution. Combining RF and FSO technologies, however, appears to be a promising approach.

Figure 1 shows a possible architecture that addresses the last-mile problem by using a WMN in the front end with a wired (fiber) network in the back end. In this architecture, adopted in our paper, the wireless gateways are the highest aggregation point in the WMN; they collect end users' traffic in a multihop manner. The aggregated traffic is routed from (to) the gateways to (from) the Internet. Due to ever-increasing need for bandwidth and long-term changes in traffic patterns, congestion in some links could lead to network bottlenecks. FSO transceivers can be used effectively to upgrade (create) a few links between nodes that are within line of sight of each other, as shown in Fig. 1.

Sarkar *et al.* proposed a similar tree-like fiber-based architecture and addressed the gateway placement problem as a cost minimization problem [1]. Among the existing technologies, a passive optical network (PON) may be used for the high-speed optical backbone, and LTE, WiMAX, or Wi-Fi can be used as a WMN in the front end. Narlikar *et al.* studied a general framework for link activation, routing, and scheduling of packets in wireless networks [2]. They showed that well-known scheduling techniques such as weighted fair queuing can be applied to the wireless multihop network under their so-called even–odd framework to guarantee an end-to-end delay. A delay-aware routing algorithm was proposed by Sarkar *et al.* [3]. Raez *et al.* propose a capacity and delay-aware routing scheme for a single-radio WMN [4].

Using multiple radios at bottleneck nodes is suggested for capacity enhancement of WMN in Ref. [5]. However, this may require a new channel assignment for a large part of the network and could incur a high cost. We proposed a

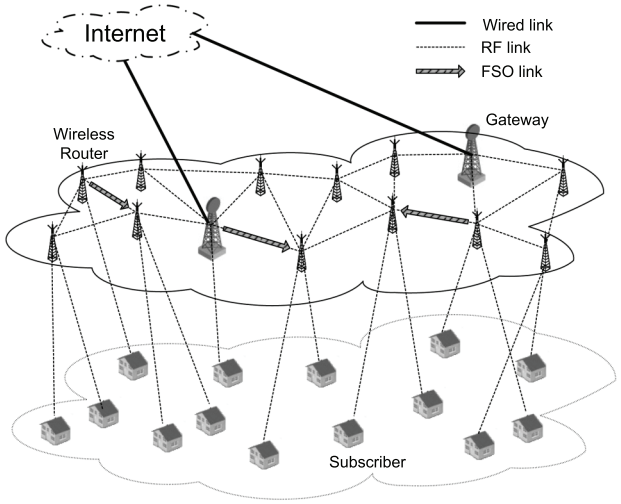

Fig. 1. Network architecture for hybrid wireless optical network.

MIMO-based solution in Ref. [6] to benefit spatial multiplexing through multiple antennas and to avoid new channel assignments to neighboring nodes imposed by using multiple radios. Mumey *et al.* suggested a joint-stream control and scheduling problem for MIMO-based wireless networks [7]. Such a solution requires every node to be upgraded with a MIMO antenna; this is not a cost-effective approach and is probably unnecessary, because the traffic distribution is typically non-uniform in hybrid wireless optical access networks. This is because there is more traffic to and from the gateways than from the wireless routers.

Krishnamurthy *et al.* investigated an FSO-based mesh network and proposed a routing and load-balancing algorithm to reduce the blocking probability [8]. Kashyap *et al.* proposed an integrated topology control and routing in FSO mesh backbone networks [9]. They emphasized the cost of FSO transceivers as the major constraint and showed that the corresponding problem is NP-hard. Smadi *et al.* proposed an algorithm for placing FSO gateways in WMNs [10].

Tang *et al.* consider the enhancement of RF WMN network throughput using FSO links in Ref. [11]. They consider both fading and non-fading communication channels for the RF portion of the network. There are some similarities between their work and our problem in Ref. [12] in that they propose a heuristic for the placement of FSO links. However, their physical layer model and formulation are different compared to ours in this paper. In their formulation, they maximize the total throughput of the network without considering fairness in network capacity allocation. In this paper, we define the network capacity such that the per-node traffic demand rates are maximized to guarantee fairness. We use a similar fading model for the FSO links as theirs in our formulation.

Rak *et al.* consider a hybrid FSO/RF approach with a special focus on the resistance of FSO paths against link outages due to adverse weather conditions [13]. They proposed a link-disjoint FSO paths algorithm to increase reliability. Their objective is different from this research in that their focus is mostly on reliable routing in an FSO-based network, while ours is about an existing RF-based network with some upgrades on the bottleneck links for increasing the capacity. In the context of FSO network reliability, we proposed an algorithm for improving hybrid FSO/RF network reliability through transceiver reconfiguration in Ref. [14]. We suggested using FSO relays for network disaster recovery in Ref. [15].

In this paper, we propose a novel solution for incorporating FSO technology into existing RF-based WMNs to greatly improve network capacity (throughput) at a fraction of the entire network upgrade cost. The contribution of this paper is two-fold. First, using a unified RF link scheduling framework and FSO link placements scheme, we find the optimal locations for a given number of FSO transceivers as well as the matching RF link schedules that maximize the network capacity. In other words, given the expected long-term uplink and downlink traffic demands at each node, we formulate an integer linear program (ILP) for the joint assignment of FSO transceivers and RF link scheduling to find the strategically located nodes which increase the overall network capacity the most if they get upgraded. Second, to avoid the complexity of the original ILP, we relax it and propose a simple probabilistic greedy algorithm, which is shown by simulation to compute a comparable solution at a significantly lower computational cost. In both schemes, the computed schedules provide all nodes with appropriate routes to and from the gateway(s).

This approach is different from the existing work in that it takes RF link-scheduling into account in determining the optimal placement of the FSO links. The scheduling of RF links, which affects the traffic routing indirectly, influences the capacity of the network. Therefore, the optimal placement of FSO links cannot be addressed without considering the RF links schedule. To the best of our knowledge, this problem has not been investigated. We note that, typically, only a small set of links needs to be upgraded with FSO, and therefore, our approach requires minimal changes to the network infrastructure.

The rest of this paper is organized as follows. In Section II, we present the network model used to formulate the problem. The problem is stated and formulated as an ILP in Section III. We present a novel heuristic algorithm based on ILP relaxation and probabilistic scheduling in Section IV. In Section V, extensive simulation results are presented to show the performance of the optimal algorithm and the proposed heuristic on some representative network topologies. Results using a simulated annealing approach are also provided for comparison. Section VI concludes the paper.

II. Network Model

The multihop WMN is denoted by a directed graph $\mathcal{G} = (V, E)$, where V and E represent the sets of wireless routers and the directed links between them, respectively. It is assumed that the set of gateways (nodes connected to

the Internet), which is denoted by G, is a subset of V. A goal of ours is to find a subset of V to be equipped with FSO transceivers. It should be noted that while FSO technology needs line of sight, for RF transceivers, it is not a necessity. Therefore, not all links in E are eligible to be upgraded by FSO technology.

We consider only the WMN and the FSO network in our network model and ignore the user nodes. The end users are assumed to be connected to their respective access points (i.e., wireless routers) using separate frequency bands that do not interfere with the wireless mesh network. We assume that the destination of the aggregated uplink traffic to be one or more gateways and that of the downlink traffic to be the respective wireless router. Thus, a wireless router, v, has an aggregated uplink and downlink traffic demand denoted by $\mathbf{r}_u(v)$ and $\mathbf{r}_d(v)$, respectively.

A. Physical Layer and Interference Model

We assume that the RF antennas in the WMN are omnidirectional; however, our approach can be easily generalized for sectorized antennas. The nodes communicate in a half-duplex manner, i.e., simultaneous transmission and reception are not possible. It is assumed that the nominal capacity of each link is given and that fading is insignificant due to the wireless routers being static. However, the shadowing effect is considered, and this might cause two nodes that are physically proximate to not have a link between them, while two nodes that are separated by a longer distance will have a link between them. Such a model is used, for example, in Ref. [2]. Interference may occur at an RF node because of its neighboring nodes transmitting at the same time. All directed links that interfere with each link can be pre-calculated, and we assume that the set of interfering links for link $e \in E$ is given and is denoted by $I_e \subseteq E$. We assume the protocol model of interference in this paper [2].

FSO links are assumed to provide interference-free channels for communication. The average capacity of an FSO link can be calculated using the model in Ref. [16]. In this paper, we assume that FSO transceivers use narrow laser beams that are able to provide a capacity of up to $C_F = 2.5$ Gbps up to a distance of 4 km in clear weather conditions [9]. In our simulation, we conservatively consider a lower rate of about 1.2 Gbps [17]. We assume that FSO links may experience fading that may result in temporary link outages. To characterize the effective capacity of a link, we take the link availability into account. This parameter, which is denoted by π, is the probability that an FSO link successfully operates at the data rate of C_F, resulting in an effective capacity of $C_F\pi$ [11]. In fact, in order to obtain this effective capacity, the relay node must have enough memory to buffer the data during the fade and transmit it once the link becomes available again. The details of such a scheme are out of the scope of this paper. Finally, we assume that such possible link outages due to fading are statistically independent.

B. Link Scheduling Framework

We assume that all RF transmissions are done on a single common channel that is accessed according to a TDMA-based MAC layer protocol.[1] For this purpose, time is framed into non-overlapping intervals within which a set of link schedules is applied. These frames are periodically repeated, resulting in the periodic activation of the links according to their schedules. Our goal is to determine the set of schedules within a frame such that the overall network capacity is maximized. The schedules for link $e \in E$ are characterized by a *schedule vector* $\mathbf{S}(e)$. We assume that time within a frame is slotted in T equal-sized intervals. Therefore, the schedule vector $\mathbf{S}(e)$ consists of T binary elements that determine the status of link e at any time slot within a frame: $\mathbf{S}(e,t) = 1$ implies that link e is active in time slot t.

III. Problem Definition and Formulation

In this section, we formulate the problem of capacity enhancement of an existing RF-based wireless mesh network by upgrading some of the transceivers using FSO technology. Network capacity is defined as the total per-node traffic demand rates (both uplink and downlink) that are admissible within our proposed scheduling framework. More precisely, the network capacity, $\mathcal{R}$, is defined as the largest factor by which all traffic demands can be scaled with a feasible routing. In other words, if $\mathbf{r}_u(v)$ and $\mathbf{r}_d(v)$ are the uplink and downlink demands at node $v \in V$, the network capacity is defined as the largest $\mathcal{R}$ for which demands $\mathcal{R} \times \mathbf{r}_u(v)$ and $\mathcal{R} \times \mathbf{r}_d(v)$ can be successfully routed using our proposed framework.

In what follows, a matrix notation is adopted to simplify the equations. For example, dot operator $(\cdot)$ implies matrix multiplication and/stands for element-by-element matrix (vector) division. When no sign is used between two symbols, it represents the element-by-element product of matrices (vectors). We use $\mathbf{1}$ to show a matrix (vector) whose elements are all ones. As an example, the summation of all elements of a vector denoted by $\mathbf{q}$ can be simply shown by $\mathbf{q} \cdot \mathbf{1}$. In addition, when comparative operators such as $=$ or $\leq$ are used, the operators apply to every pair of corresponding elements in both matrices.

A. Problem Definition

Given:

- potential RF network topology, $\mathcal{G} = (V, E)$,
- subset of E for potential FSO upgrades,
- availability of the FSO links, π,
- number of FSO upgrades/transceivers, M,
- set of interfering links, I_e, for each link e,

[1]This assumption can be generalized by enabling subchannelization in neighboring nodes and thus allowing simultaneous transmission in non-overlapping frequency bands.

- number of time slots in a frame, T,
- uplink and downlink demand vectors, $\mathbf{r}_u$ and $\mathbf{r}_d$, and
- nominal capacity of RF and FSO links, C_R and C_F

Find:

- schedule vectors of the RF links for each link e, $\mathbf{S}_R(e)$,
- placement of FSO links, $\mathbf{S}_F$,
- uplink and downlink traffic, $\mathbf{\Lambda}_u$ and $\mathbf{\Lambda}_d$ such that the network capacity $\mathcal{R}$ is maximized.

In the following, we formulate the problem as an ILP, where the objective is to maximize $\mathcal{R}$.

B. ILP Formulation

Please refer to Table I for notations.

Objective: max $\mathcal{R}$

Flow constraints: A gateway as the interconnection node between the wireless and the wireline network (Internet) sinks the uplink traffic from the wireless routers and sources from the downlink traffic in the reverse direction. Therefore, at the downlink traffic origin and uplink traffic destination, we have

$$\forall\, e \in E_o(v) \quad \mathbf{\Lambda}_u(e) = 0, \quad \forall\, e \in E_i(v) \quad \mathbf{\Lambda}_d(e) = 0, \tag{1}$$

TABLE I
SYMBOL DESCRIPTION

Symbol	Description
$\mathcal{G}$, G, V, E	Graph, gateways, transceivers, links
$\mathbf{r}_u$, $\mathbf{r}_d$	Traffic demand: uplink, downlink
e_t, e_r	Transmitter, receiver node of link e
$\mathbf{S}$, $\mathbf{S}(e)$	Schedule: matrix, vector for link e
$\pi(e)$	Avaiabilty of FSO link e
T, M	Number of: time slots in a frame, FSO links
$\mathcal{R}$, $\mathbf{\Lambda}_u$, $\mathbf{\Lambda}_d$	Network capacity, traffic flow: uplink, downlink
$E_o(v)$, $E_i(v)$	Links from and to node v: outgoing, incoming
$\mathbf{V}_o$, $\mathbf{V}_i$, $\mathbf{V}_{oi}$	Link matrix: outgoing, incoming, outgoing–incoming
$\mathbf{c_F}$, $\bar{\mathbf{c}}_\mathbf{R}$	Link capacity: FSO, RF (effective time averaged)
$\mathbf{X}(v,t)$, $\mathbf{Y}(v,t)$	Node v status at time slot t: transmitting, receiving
I_e, $\mathbf{I}$	Interfering links set of e, interference matrix
C_R, C_F	Nominal link capacity: RF, FSO
$\mathbf{p}_t(v)$, $\mathbf{p}_r(v)$	Node v status probability: transmitting, receiving
$\boldsymbol{\sigma}$, $\hat{\boldsymbol{\sigma}}$	Time-averaged schedule vector: relaxed ILP, PGS
α, η	Projection of $\hat{\boldsymbol{\sigma}}$ on $\boldsymbol{\sigma}$, link probability update factor
$\bar{\mathbf{c}}_\mathbf{R}$, $\bar{\hat{\mathbf{c}}}_\mathbf{R}$	Time-averaged RF capacity: relaxed ILP, PGS
$\boldsymbol{\rho}(t)$	Link activation probability at time slot t
θ_0, θ_{fin}, θ, β	Initial, final, current temp., decay factor (SA)
$\mathcal{A}$, L	The set of active links, remaining links
OPT, PGS, SA	Optimal Alg., rand greedy Alg., simulated annealing

where $\mathbf{\Lambda}(e)$ denotes the traffic flow along link e. For the uplink traffic flow of any non-gateway node, we have

$$\sum_{e\in E_o(v)} \mathbf{\Lambda}_u(e) - \sum_{e\in E_i(v)} \mathbf{\Lambda}_u(e) = \mathcal{R} \times \mathbf{r}_u(v), \tag{2}$$

where $v \in V\backslash G$. The downlink traffic flow constraints can be obtained similarly. For the sake of simplification, we define two $|V| \times |E|$ matrices as follows:

$$\mathbf{V}_i(v,e) = \begin{cases} 1 & e \in E_i(v) \\ 0 & o.w. \end{cases} \quad \mathbf{V}_o(v,e) = \begin{cases} 1 & e \in E_o(v) \\ 0 & o.w. \end{cases}. \tag{3}$$

In addition, we define $\mathbf{V}_{oi} = \mathbf{V}_o - \mathbf{V}_i$. Using this notation, Eq. (2) can be simply written as $\mathbf{V}_{oi} \cdot \mathbf{\Lambda}_u = \mathbf{r}_u \mathcal{R}$, where $\mathbf{\Lambda}_u$ and $\mathbf{r}_u$ are the vectors of the uplink flow and demands on all links and nodes, respectively.

Link capacity constraints: Since for every link, the total uplink and downlink flows can be at most equal to the effective capacity of each link, we must have

$$\mathbf{\Lambda}_u + \mathbf{\Lambda}_d \le \mathbf{c_F} + \bar{\mathbf{c}}_\mathbf{R}. \tag{4}$$

Here, elements of vector $\mathbf{c_F}$ are either the nominal capacity of an FSO link if an upgrade happens or 0 otherwise. Also, $\bar{\mathbf{c}}_\mathbf{R}$ denotes a vector whose elements are the effective, time-averaged capacity for each RF link.[2]

Half-duplexing constraints: An RF transceiver may not transmit and receive simultaneously. Let the binary variables $\mathbf{X}(v,t)$ and $\mathbf{Y}(v,t)$ denote the transmitting and receiving statuses of the RF interface at node v in time slot t, respectively. Being in either mode makes the corresponding variable 1; otherwise, it is 0, and we must have $\mathbf{X}(v,t) + \mathbf{Y}(v,t) \le 1$ or in general matrix form $\mathbf{X} + \mathbf{Y} \le \mathbf{1}$.

Link activity constraints: No outgoing link is active unless the corresponding transmitter is in transmission mode. Therefore, we have

$$\mathbf{X}_v(t) = \sum_{e\in E_o(v)} \mathbf{S}_R(e,t). \tag{5}$$

The above constraint can be summarized by $\mathbf{V}_o \cdot \mathbf{S}_R = \mathbf{X}$. With minor modifications, a similar constraint can be derived for incoming links.

Interference constraints: Once link e is active in time slot t, i.e., $\mathbf{S}_R(e,t) = 1$, all its interfering links $e' \in I_e$ must be inactive in t. In other words, we must have

$$\sum_{e'\in I_e} \mathbf{S}_R(e',t) \le |E| \times (1 - \mathbf{S}_R(e,t)). \tag{6}$$

We summarize these constraints by $(\mathbf{I} + |E|) \cdot \mathbf{S}_R \le |E|$, where $\mathbf{I}$ denotes the interference matrix, which is an $|E| \times |E|$ binary matrix whose eth row is zero except for all elements of I_e.

[2]This capacity $\bar{\mathbf{c}}_\mathbf{R}$ is defined a bit later in the subsection "Average link capacity."

TABLE II
ILP FOR CAPACITY MAXIMIZATION

capacity = max $\mathcal{R}$		
Flow	$\mathbf{V}_i(v)\cdot\mathbf{\Lambda}_u = 0 \quad \mathbf{V}_o(v)\cdot\mathbf{\Lambda}_d = 0$	$v \in G$
	$\mathbf{V}_{oi}\cdot\frac{\mathbf{\Lambda}_u}{\mathbf{r}_u} = -\mathbf{V}_{oi}\cdot\frac{\mathbf{\Lambda}_d}{\mathbf{r}_d} = \mathcal{R}$	$v \in V\backslash G$
	$\mathbf{\Lambda}_u + \mathbf{\Lambda}_d \le C_F\mathbf{S}_F\pi_F + C_R(\frac{1}{T}\mathbf{S}_R\cdot\mathbf{1})$	
Schedule	$\mathbf{V}_i\cdot\mathbf{S}_R = \mathbf{Y} \quad \mathbf{V}_o\cdot\mathbf{S}_R = \mathbf{X}$	
	$\mathbf{X}+\mathbf{Y}\le\mathbf{1} \quad (\mathbf{I}+\lvert E\rvert)\cdot\mathbf{S}_R \le \lvert E\rvert$	
Location	$\mathbf{S}_F\cdot\mathbf{1}\le M$	

Average link capacity: Depending on the link activity in each time slot, the overall capacity of the link can be determined in a time frame. The time-averaged activity of link e in a time frame can be calculated by $\frac{1}{T}\sum_t \mathbf{S}_R(e,t)$. Therefore, we have

$$\bar{\mathbf{c}}_\mathbf{R} = C_R\left(\frac{1}{T}\mathbf{S}_R\cdot\mathbf{1}\right), \tag{7}$$

where C_R is the nominal capacity of an RF link.

Link capacity: The binary variable $\mathbf{S}_F(e)$ indicates whether a link is upgraded with an FSO or not. The capacity of such a link can be calculated as follows:

$$\mathbf{c}_\mathbf{F}(e) = C_F\mathbf{S}_F(e)\pi_F(e), \tag{8}$$

where C_F is the nominal capacity of an FSO link and $\pi_F(e)$ denotes the availability of link e.

Number of upgrades constraint: Considering the total number of upgrades, M, we have

$$\sum_e \mathbf{S}_F(e) \le M. \tag{9}$$

The resulting ILP is shown in Table II.

In the presented formulation, T plays a key role in terms of performance and complexity. For the same frame length, clearly, the achievable capacity monotonically increases for integer multiples of T. However, the complexity of the problem significantly increases with T. We will explore this trade-off later in the paper. Apart from the effect of T on the complexity and on the capacity, there is a practical consideration in the choice of T. A too-small value for T could possibly leave too few slot options for activating the links and may possibly make the resulting topology disconnected. On the other hand, when T is too large (without changing the duration of the frame), too-frequent switching of the transceivers is required, possibly increasing packet delays.[3]

IV. ILP RELAXATION AND HEURISTIC ALGORITHMS

The network capacity maximization ILP is computationally expensive. Some relaxation can be applied in order to reduce its complexity and obtain an upper bound on the network capacity. While relaxation is meant to reduce the complexity of the ILP in order to obtain a meaningful upper bound, deliberate relaxations are necessary to avoid ending up with a loose upper bound. Investigating the constraints of the ILP reveals that there are three main categories of constraints (as shown in Table II), all of which have different levels of complexity.

The *flow* constraints are mostly real-valued. These constraints specify the routing information by providing the amount of traffic flow on each link. The second batch of constraints is related to the RF links' *schedules*, which characterize the solution space for the schedule matrix. The variables involved in this part of the ILP are all integers and their numbers scale as $O(T)$. The third group of constraints is about the FSO upgrade *locations*. These constraints, although they involve integer variables, are far fewer and scale as $O(|E|+|V|)$. As a result, the relaxation of the latter constraints provides us with minimal improvement compared to the second group, namely, schedule constraints. In this paper, we opt to only focus on relaxing the scheduling constraints. In other words, the relaxed version of our ILP still provides the placements (locations) of the upgrades, but not the detailed RF link activation schedules. Later, we propose a randomized algorithm to approximate the optimal schedule matrix, but at a significantly lower computational cost.

Assume the optimal schedule matrix, $\mathbf{S}$, is found for a given T by solving the ILP. This $|E|\times T$ binary matrix determines the status of all links at any time slot t. Regardless of the sequence of activation, if all schedule (column) vectors are used to activate the links, the optimal capacity can be achieved. Consider an approach to randomly select the sequence of schedule vectors for a frame without replacement such that at the first round, the probability of selection for each schedule vector is $\frac{1}{T}$. After selecting the first schedule, the second one is selected from the remaining schedules with probability $\frac{1}{T-1}$. This procedure is continued until all schedules are used. At the end of this procedure, the expected number of times where link e is active, denoted by $\boldsymbol{\sigma}_e$, can be calculated as $\frac{1}{T}\sum_t \mathbf{S}_R(e,t)$. Obviously, the time-averaged vector $\boldsymbol{\sigma}$ whose entries are real-valued is the centroid[4] of the schedule matrix.

The time-averaged vector $\boldsymbol{\sigma}$ can be used to relax the scheduling constraints and still preserve meaningful information about the schedule matrix. Using this variable, the set of constraints that have the term $\mathbf{S}_R(e,t)$ can be condensed to convert the ILP to one that is much simpler, with far fewer variables and constraints. For example, we can take the summation over t on both sides of Eq. (5) to get

$$\frac{1}{T}\sum_t[\mathbf{X}_v(t)] = \frac{1}{T}\sum_t\left[\sum_{e\in E_o(v)}\mathbf{S}_R(e,t)\right]. \tag{10}$$

By replacing the term $\sum_t \mathbf{S}_R(e,t)$ with $T\boldsymbol{\sigma}_e$, the resulting constraint becomes

[3]Consideration of delays is outside the scope of this paper.

[4]The centroid is considered as the *average* position of all points (vectors) represented by the matrix.

TABLE III
RELAXED ILP FOR CAPACITY ENHANCEMENT

UB = max $\mathcal{R}$		
Flow	$\mathbf{V}_i(v)\cdot\mathbf{\Lambda}_u = 0 \quad \mathbf{V}_o(v)\cdot\mathbf{\Lambda}_d = 0$	$v \in G$
	$\mathbf{V}_{oi}\cdot\frac{\mathbf{\Lambda}_u}{\mathbf{r}_u} = -\mathbf{V}_{oi}\cdot\frac{\mathbf{\Lambda}_d}{\mathbf{r}_d} = \mathcal{R}$	$v \in V\backslash G$
	$\mathbf{\Lambda}_u + \mathbf{\Lambda}_d \le C_F\mathbf{S}_F\pi_F + C_R\boldsymbol{\sigma}$	
Schedule	$\mathbf{V}_i\cdot\boldsymbol{\sigma} = \mathbf{p}_r \quad \mathbf{V}_o\cdot\boldsymbol{\sigma} = \mathbf{p}_t$	
	$\mathbf{p}_r + \mathbf{p}_t \le \mathbf{1} \quad (\mathbf{I} + \lvert E\rvert)\cdot\boldsymbol{\sigma} \le \lvert E\rvert$	
Location	$\mathbf{S}_F\cdot\mathbf{1} \le M$	

$$\mathbf{p}_t(v) = \sum_{e\in E_o(v)} \boldsymbol{\sigma}_e, \tag{11}$$

where $\mathbf{p}_t(v)$ is the probability of node v being in transmission mode at time slot t. Using the above randomized approach, we obtain the relaxed ILP shown in Table III, which is much easier to solve.

It should be noted that the solution of the relaxed ILP just provides an upper bound on capacity, and it may not be achievable. For example, the relaxed constraint resulting from Eq. (6) does not guarantee interference prevention in each time slot, as Eq. (6) does. Moreover, the solution of the relaxed ILP does not provide detailed information about the schedules and activity of the links in each slot. In the next section, we propose a probabilistic algorithm that uses the relaxed ILP solution to activate the links. In fact, any algorithm (if it exists) that is able to realize the solution of the relaxed ILP is optimal.

A. Solution Space Characteristics

The optimal schedule matrix can be approximated by appropriate sampling of the schedule vector space. For this purpose, T *schedule vectors* must be found to determine the status of the links. From a geometrical perspective, each schedule vector represents a point in the solution space which is itself a subspace of an $|E|$-dimensional space. Each point in the solution space corresponds to a maximal set of non-interfering links that can be active in a half-duplex manner simultaneously.

Loosely speaking, the set of scheduling constraints in the ILP characterizes a subset of points in the $|E|$-dimensional space. The flow constraints along with the maximization objective select T points as schedule vectors to form the schedule matrix. The vector $\boldsymbol{\sigma}$ can be used to approximate the weights of different links' involvement in the suboptimal solution. Based on this, we propose a probabilistic greedy scheduling algorithm to construct the schedule matrix. Once the matrix is calculated, it can be plugged back into the original ILP to find the actual traffic flow on each link. The resulting program is an LP with much less complexity, as the integer variables no longer exist and the only unknowns are the traffic flows, which are real-valued.

In constructing the schedule matrix, all constraints that shape the solution space are taken into account. Based on its definition, $\boldsymbol{\sigma} = \frac{1}{T}\mathbf{S}_R\cdot\mathbf{1}$ is the centroid of the schedule matrix. In other words, the relaxed ILP suggests a direction that is a good approximation of the schedule matrix through $\boldsymbol{\sigma}$. We define a constructed schedule matrix, $\hat{\mathbf{S}}$, to be an α-approximate schedule matrix if the resulting $\hat{\mathcal{R}} = \alpha\cdot\mathcal{R}$, where $0 < \alpha \le 1$ and $\mathcal{R}$ is the solution of the relaxed ILP. In order to find a good approximation of the optimal solution, we may maximize the projection of $\hat{\boldsymbol{\sigma}}$ on $\boldsymbol{\sigma}$, which can be performed by maximizing the minimum element of $\hat{\boldsymbol{\sigma}}/\boldsymbol{\sigma}$ according to the following Lemma.

Lemma 1: $\hat{\mathbf{S}}$ is α_m-approximate if $\alpha_m = \min_e(\boldsymbol{\alpha}_e)$.

Proof: For matrix $\hat{\mathbf{S}}$, we must have $\bar{\hat{\mathbf{c}}}_{\mathbf{R}}(e) \ge \alpha_{\min}\cdot C_R\cdot\boldsymbol{\sigma}_e$, as $\alpha_{\min}$ is the minimum projection of $\hat{\boldsymbol{\sigma}}$ on $\boldsymbol{\sigma}$. In other words, we can have $\bar{\hat{\mathbf{c}}}_{\mathbf{R}}(e) \ge \alpha_{\min}\cdot\bar{\mathbf{c}}_{\mathbf{R}}(e)$ for all links. Therefore, we have $\hat{\Lambda}_e^u + \hat{\Lambda}_e^d \le \alpha_{\min}\cdot(\bar{\mathbf{c}}_{\mathbf{R}}(e) + \mathbf{c}_{\mathbf{F}}(e))$. As a result, the link with the minimum effective capacity (i.e., projection) becomes the bottleneck for the network flow that results in $\hat{\mathcal{R}} = \alpha_{\min}\cdot\mathcal{R}$. ∎

B. Probabilistic Greedy Scheduling (PGS) Algorithm

Our PGS algorithm essentially reverse-engineers the time-averaged schedule vector to approximate the schedule matrix. In other words, by knowing the constraints, we sample the vector space in a way that the centroid of the sampled points is in the direction of interest indicated by $\boldsymbol{\sigma}$. Ideally, a subset of points in the solution space must be selected such that their centroid is as close as possible to the time-averaged schedule vector provided by the relaxed ILP. The first step in PGS is to solve the relaxed ILP to calculate $\boldsymbol{\sigma}$. Using the solution of the relaxed ILP, the placements of the FSO upgrades are found. Then, an iterative procedure finds the T schedule vectors (columns), as the main objective of PGS is to find a schedule matrix that is α-approximate, and α is maximized.

Each schedule vector partitions the set of links into two subsets of active and idle links at each iteration. In order to maximize the projection ratio, namely, $\max(\min_e(\boldsymbol{\alpha}_e))$, the frequency of link activation must be done such that the selection probability is $\alpha\cdot\boldsymbol{\sigma}_e$. This is similar to the concept of weighted max–min fairness. We design a simple iterative procedure to activate links e in each time slot t. Suppose that the calculated schedule matrix is α-approximate; this means that the expected number of slots in which link e is active in one frame $\hat{\mathbf{S}}$ is $\alpha\boldsymbol{\sigma}_e T$. Therefore, we have

$$\mathbb{E}\left[\sum_{t\le T}\hat{\mathbf{S}}(e,t)\right] = \alpha\boldsymbol{\sigma}_e T$$
$$\frac{1}{T}\sum_{t\le T}\mathbb{E}[\hat{\mathbf{S}}(e,t)] = \frac{1}{T}\sum_{t\le T}\boldsymbol{\rho}_e(t) = \alpha\boldsymbol{\sigma}_e, \tag{12}$$

where $\boldsymbol{\rho}_e(t)$ is the probability of link e being active in time slot t. Equation (11) suggests that the time average of the selection probabilities must be equal to $\alpha\boldsymbol{\sigma}_e$. In order to take fairness into account, we cannot simply assume that $\boldsymbol{\rho}_e(t)$ remains fixed during the selection procedure. This is because there is strong dependence among the links due to

interference. In other words, a link with a high probability of selection will dominate those with smaller probabilities by putting them in the back-off set more frequently than their fair share. The resulting solution would be far from fair and would reduce the resulting network capacity significantly.

We consider a randomized round robin approach to apply fairness in the link selection procedure such that the selection weight is $\rho_e(t)$. In other words, for the links to be selected, we take their prior selection record into account. For example, suppose that we have already selected t schedule vectors and we know that link e has been active in h_t time slots so far. In order to be fair to other links, we propose updating the selection probability as follows:

$$\rho_e\left(t+1|\sum_{\tau\le t}\hat{\mathbf{S}}(e,\tau)=h_t\right)=\alpha\max\left(0,\boldsymbol{\sigma}_e-\frac{\eta h_t}{T}\right), \quad (13)$$

where $\alpha\boldsymbol{\sigma}_e T$ is the expected number of times link e is active in the entire time frame, and $\eta=\frac{1}{\alpha}$ and $\max(0,\ldots)$ is used to ensure the weight cannot be a negative number. Since the probability of links' selection is performed comparatively, we can simply factor α out and derive a recursive approach, as follows:

$$\rho_e(t)=\begin{cases}\rho_e(t-1) & \hat{\mathbf{S}}(e,t-1)=0\\ \max\left(0,\rho_e(t-1)-\frac{\eta}{T}\right) & \hat{\mathbf{S}}(e,t-1)\neq 0\end{cases}, \quad (14)$$

where Eq. (14) is initialized by $\rho_e(1)=\boldsymbol{\sigma}_e$. The iterative procedure selects the links to be active one by one at each time slot t probabilistically. Let us denote the set of remaining links to be processed by L at each iteration. At each iteration, this procedure is initialized with $L=E$. Each link e is randomly selected with a probability proportional to $\rho_e(t)$. Once a link is selected, a set of variables must be updated, as follows:

- The selected link in this stage, say e, is added to the active links' set, denoted by $\mathcal{A}$.
- All outgoing links from the receiver of e, namely e_r, and incoming links to its transmitter, namely e_t, are removed from L to comply with half-duplexing.
- Link e and its potential interfering links, namely, I_e, are removed from L.
- All remaining links e' in L where e is among their interfering links must be removed from L.
- Update schedule vector $\hat{\mathbf{S}}(e,t)=1$.
- The selection probability for e is reduced by $\frac{\eta}{T}$.

The procedure is repeated among the remaining links until L is empty. The first iteration determines the schedule vector $\hat{\mathbf{S}}(e,1)$. In order to get the remaining schedule vectors, we repeat the algorithm $T-1$ more times. The PGS algorithm is summarized in Algorithm 1. It should be noted that the above algorithm does not guarantee Eq. (12) and in practice, we often have $\frac{1}{T}\sum_{t\le T}\rho_e(t)\le\frac{\sigma_e}{\eta}$. In other words, the effective α always satisfies $\alpha\le\frac{1}{\eta}$.

Algorithm 1 The PGS algorithm

```
Data: V, E, I, p, T, η
Result: Ŝ
begin
    τ ← 1
    Ŝ ← 0
    while τ ≤ T do
        L ← E
        A ← Ø
        for e ∈ L do
            if τ = 1 then ρ_e(τ) = σ_e
            else
                ρ_e(τ) = ρ_e(τ − 1)
            end
        end
        while L ≠ Ø do
            e = RandomPick(p, L)
            A ← A ∪ {e}
            ρ_e(τ) = max(0, ρ_e(τ) − η/T)
            L ← L\{e ∪ I_e ∪ E_o(e_r) ∪ E_i(e_t)}
            for e' ∈ L do
                if e ∈ I_e' then L ← L\e'
            end
            Ŝ(e, τ) ← Ŝ(e, τ) + 1
        end
        τ ← τ + 1
    end
end
```

The output of this algorithm finds $\alpha=\min(\frac{1}{T}\hat{\mathbf{S}}_R\cdot\mathbf{1}/\boldsymbol{\sigma})$ for a given η. A good value needs to be found for η that maximizes the calculated α. Due to the ease of computation, this heuristic can be run multiple times for random η values that are uniformly selected in a given range. We later show that the distribution of good choices for η is concentrated around a certain value that can be found by trial and error. In our simulation, depending on the computational budget we run the heuristic for a certain number of times, say 1000, and keep the best result at the end of the procedure.

C. Geometrical Interpretation of PGS

The PGS algorithm may not realize the time-averaged schedule vector of the relaxed ILP, namely $\boldsymbol{\sigma}$. However, we will show later that it performs quite well when compared with the optimal solution. Considering the geometrical notion of the schedule matrix, the centroid of the schedule vectors, which determines the performance of PGS, must provide a balance between being along the $\boldsymbol{\sigma}$ vector (fairness) and having a large projection on the coordinate axes (individual link throughput). This balance is controlled by parameter η, for which a good value needs to be found. A simple analytical bound on the performance of the heuristic based on η can be derived due to its relationship with α, namely $\alpha\le\frac{1}{\eta}$.

In order to clarify this bound, recall that the link selection algorithm is a contention-based procedure because of interference. In other words, interference does not allow some links to be active simultaneously. Hypothetically, assume that there is no interference and the time-averaged activation for link e is $\boldsymbol{\sigma}_e$. Without interference, this link is expected to be active in $\boldsymbol{\sigma}_e T$ time slots, which translates to $\eta = 1$. In reality, due to interference, this link can only be active in a fraction of this many time slots. Based on the probability update formula Eq. (14), the total number of time slots in which link e can be selected is $\lceil \frac{\boldsymbol{\sigma}}{\eta/T} \rceil$, providing that the probability tends to zero after T time slots. Therefore, the link activation ratio is at most around $\frac{\boldsymbol{\sigma}}{\eta}$, decided by η.

Parameter η plays a key role in providing a fairness level by limiting the total number of active time slots for links with higher activation probability and providing the low probability links a chance of activation. For example, in the case that η is considerably large, the fairness level is high, as after a few times of being active, the probability of selection becomes zero even for highly probable links, thereby resulting in all links having the same fractional chance of activation. Geometrically, this corresponds to the case where several zero vectors (a vector with all elements equal to zero) are included as schedule vectors in the schedule matrix. These vectors correspond to points located at the origin of the coordinate system, and make it easier for the centroid to be along the $\boldsymbol{\sigma}_e$ vector. Such values, however, provide a higher fairness level to limit the performance of the heuristic by $\alpha \leq \frac{1}{\eta}$. When η is zero or a small number, which is close to the case with no probability update, the links with the higher probability of selection will dominate the other links, reducing the overall performance. This case geometrically refers to the state where the centroid has a very small or no projection for some of the links (with low probabilities) and large projection for some others. As a final note, good performance of the heuristic, which will be shown later, is due to updating the probability of link selection at each iteration (e.g., $\eta > 0$). As mentioned, the main reason for updating this probability is to provide fairness in the contention process. Therefore, depending on the network topology, a good value for η must be selected to provide a balance between fairness and throughput.

D. Simulated Annealing

In order to provide a baseline to compare the performance of PGS, simulated annealing (SA) is employed. This probabilistic algorithm can provide a good approximation of the globally optimal solution [18]. SA is analogous to the thermodynamics process involved in the formation of crystals such that a temperature parameter, θ, is incorporated into the optimization procedure, and it is decayed geometrically according to $\theta \cdot \beta$ for the next iteration, where $\beta \leq 1$ is a decay factor. SA is initialized with a trial point based on current estimates, which might not even be close to optimal, and it iteratively evaluates the objective function in a new proposed location. The new value is accepted if it improves the solution. Otherwise, it is accepted with a certain probability which is calculated based on the current temperature and the difference between the cost function at the new value and the last found solution. It is known that if the cooling procedure is sufficiently slow, the globally optimal solution will be reached. The SA algorithm is summarized in Algorithm 2.

Algorithm 2 Simulated Annealing

Data: $V, E, I, p, \theta_{fin}, \beta$
Result: $\hat{\mathbf{S}}$
begin
  $\theta = \theta_0 \leftarrow 1$
  while $\theta \geq \theta_{fin}$ **do**
    $\alpha_{min} \leftarrow \min_e(\alpha_e)$
    $e_{min} \leftarrow Arg = \min_e(\alpha_e)$
    $\tau \leftarrow \{t | t \in \hat{\mathbf{S}}(t, e_{min}) = 0\}$
    $\hat{\mathbf{S}}' \leftarrow \hat{\mathbf{S}}$
    $\hat{\mathbf{S}}'(E, \tau) \leftarrow \mathbf{0}$
    $L \leftarrow E$
    $\mathcal{A} \leftarrow \emptyset$
    while $L \neq \emptyset$ **do**
      if $\mathcal{A} = \emptyset$ **then** $\mathcal{A} \leftarrow e_{min}$
      else
        $e = RandomPick(p, L)$
        $\mathcal{A} \leftarrow \mathcal{A} \cup \{e\}$
      end
      $L \leftarrow L \backslash \{e \cup I_e \cup E_o(e_r) \cup E_i(e_t)\}$
      for $e' \in L$ **do**
        if $e \in I_{e'}$ **then** $L \leftarrow L \backslash e'$
      end
      $\hat{\mathbf{S}}'_e(\tau, e) \leftarrow \hat{\mathbf{S}}'_e(\tau, e) + 1$
    end
    $\alpha'_{min} \leftarrow \min_e(\alpha'_e)$
    if $\alpha_{min} < \alpha'_{min}$ **then** $\hat{\mathbf{S}} \leftarrow \hat{\mathbf{S}}'$
    else if $rand < \exp(-\frac{\alpha_{min} - \alpha'_{min}}{\theta})$ **then**
      $\hat{\mathbf{S}} \leftarrow \hat{\mathbf{S}}'$
    end
    $\theta \leftarrow \theta \cdot \beta$
  end
end

V. Simulation Results

The RF physical layer in our experiment is WiMAX (802.16e), where with 20 MHz of bandwidth, the link capacity can be as high as 75 Mbps [19]. We consider FSO links using narrow laser beams that provide a capacity of up to 1.2 Gbps within a range of around 4 km [17]. For the first part of the simulation, we assume the FSO links are available 100% of the time. In Subsection V.C, we consider various availabilities of the FSO links and analyze the performance of the algorithm. CPLEX 12.2 is used to solve the integer linear programs. We assume that there is only one gateway in the wireless mesh network. In fact, by adding more gateways or relocating them appropriately, higher gains can be achieved, but we do not investigate that in this paper. Such approaches can be found in Ref. [1]. We compare the performance of the proposed heuristics with that of the optimal algorithm.

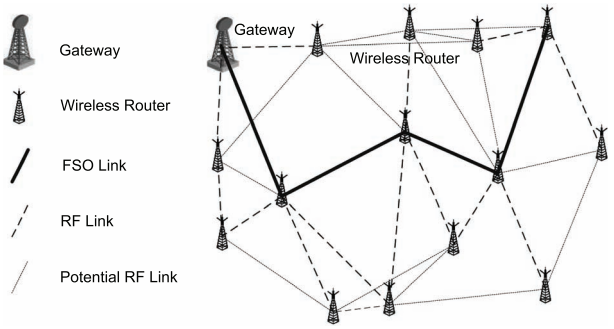

Fig. 2. Random wireless mesh network topology.

For the simulation, we generate 20 random graphs consisting of 15 nodes each, one of which is a gateway. The nodes are randomly placed in a square area of side 10 km, and the maximum range of RF transmission is 4 km. Figure 2 shows one such graph. We assume that all nodes have equal uplink and downlink demands and the capacities of all RF links are the same and equal to 75 Mbps. It should be noted that actual capacities do not affect the formulation and different link capacities can be considered. We employ the standard interference model, i.e., interference is restricted to nodes within one hop from each other. For example, link e_1 is in the interfering set of e_2, i.e., $e_1 \in I_{e_2}$, if its transmitter has a link to the receiver of e_2.

The solution of the ILP in Table II provides the optimal placements of FSO links. Figure 2 shows an example of optimal solution for $M = 4$, i.e., 4 links are upgraded. As can be seen, the FSO links provide a high-capacity backbone for the WMN through which the uplink and downlink traffic for several nodes are routed to and from the gateway. This figure shows the general topology of the underlying RF wireless mesh network. Among the RF links, some are active in at least one time slot within a frame, and these are shown using dashed lines. The other RF links that are never used in a frame are illustrated using dotted lines.

A. Performance Comparison of PGS and SA With the Optimal Performance

The optimal capacity of the network is calculated for all representative graphs using FSO upgrades. For the sake of simplicity, we consider $M \in \{1, 2, 3, 4\}$ for $T = 10$ and 20, as larger values result in very complex and computationally expensive solutions. We consider larger values for T in our heuristics. Figure 3 shows the average of optimal network capacities for the random graphs. As expected, the capacities increase with the number of FSO links. It is noteworthy that the performance increases more than linearly as M increases.

The error bars show one standard deviation of the computed capacities in different cases. For $T = 20$, the achievable network capacity is slightly higher than that for $T = 10$, as expected. This is because the former provides a larger solution space due to more relaxed constraints.

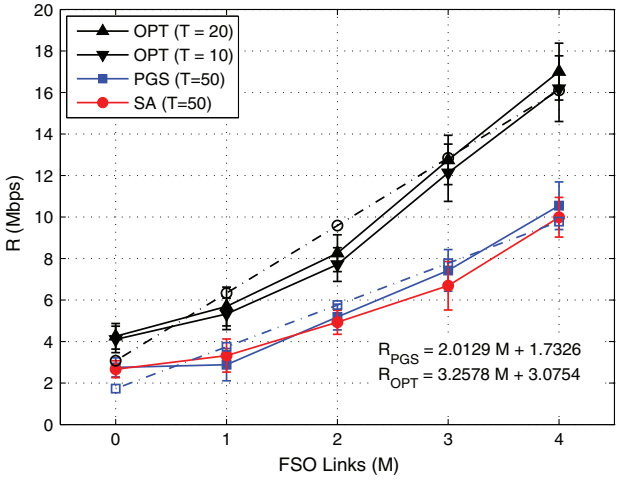

Fig. 3. Performances of PGS and SA compared with the optimal performance.

As can be seen in this figure, the interpolating line for the optimal solution has a slope of 3.2578, which implies about 76% of potential network capacity enhancement per each additional FSO link. However, the gain is lower at smaller values of M and higher for larger values.

Figure 3 also shows results for PGS and SA for $T = 20$ and 50. These approaches do not provide a competitive solution for $T = 10$, as the corresponding solution space is limited in size, resulting in low performance (not shown here). The general trend of capacity enhancement is similar to the optimal solution, but with a lower efficiency for both approaches. Based on the calculation, the interpolating line for PGS has a slope of 2.0129, which is about 50% of network capacity improvement per each additional FSO link. In other words, the performance of these approaches is around 62% of the optimal solutions. SA and PGS have a similar performance, with SA performing slightly better for $M = 1$, and PGS outperforming SA for larger values of M. PGS, however, has lower complexity than SA.

Table IV shows runtimes of all approaches.[5] According to this table, the average runtimes for the optimal algorithm are about two orders of magnitude larger than those for PGS and SA; for larger network topologies, the ILP solution is computationally prohibitive and was unable to obtain results in a reasonable amount of time. Further, the runtime of the algorithm scales almost linearly with T for the same algorithms. As can be seen, PGS runs generally faster than SA.

B. Comparison of PGS and SA

The performances of the proposed PGS algorithm and the simulated annealing approach are generally close to each other. This is mainly because the core of both algorithms is based on the solution of the relaxed ILP shown

[5]Multiple machines equipped with Quad Core Intel Xenon processors (8M cache, 2.13 GHz) and 8 GB RAM were used.

TABLE IV
ALGORITHM RUNTIME IN SECONDS

Method	OPT	PGS	SA
$T = 10$	3.536×10^4	NA	NA
$T = 20$	8.121×10^4	1.142×10^2	1.519×10^2
$T = 50$	NA	2.345×10^2	3.211×10^2

in Table III. The relaxed ILP solution, which is characterized by $\boldsymbol{\sigma}$, governs the general direction of the schedule matrix regardless of the approach. That direction itself may not necessarily be optimal, but it suggests a suboptimal solution. Therefore, the performance deficiency incurred by the limitations of the relaxed solution is justified through the trade-off with the extensive complexity of the original ILP.

The quality of the PGS solution compared with that of SA depends on different factors. For example, it appears that when the size of the solution space is small (e.g., $T = 20$), SA performs better, while for larger solution spaces (e.g., $T = 50$), that is not necessarily true. The reason is that SA generally performs a randomized exhaustive search depending on the computational budget assigned to it through the cooling temperature. For a given computational budget, when the space size is smaller, the algorithm has a better chance to obtain a globally optimal solution. However, when the space size is large, for the same computational budget, it is less likely to find a globally optimal solution. Even when the budget is higher, there is a chance that SA loses its performance due to the concavity of the huge solution space. On the other hand, PGS has a generally random behavior as opposed to the exhaustive search behavior of SA. In other words, depending on the input variable η, the solution space is sampled at random, especially when η is in a range that can potentially provide a good solution.

Figure 4(a) shows α for both SA and PGS calculated for $T = 50$. Clearly, except for the case of $M = 1$, PGS has a generally better performance. It must be mentioned that PGS incurs much less computational complexity as opposed to SA, as can be seen in Table IV.

C. Effect of FSO Link Availability on Network Capacity

In order to analyze the effect of FSO link availability on the network capacity, we consider three low availabilities for the FSO links, namely, $\pi \in \{0.1, 0.2, 0.3\}$, which result in an effective capacity of $\mathbf{c_F} \in \{160, 320, 480\}$ Mbps, where the nominal capacity of an FSO link is 1.2 Gbps. We assume the availability to be the same for all FSO links and solve the problem using PGS for the cases where $M \in \{1, 2, 3, 4, 5\}$ FSO links can be added to the network. Figure 4(b) shows the solutions of the heuristic averaged over the generated random graphs.

As can be seen in Fig. 4(b), when the availability of the FSO links is very low, i.e., $\pi = 0.1$, adding more than 2 links does not provide a significant improvement. The reason is that the effective capacity of FSO links is comparable to that of RF links. In other words, adding more links does not solve the bottleneck issue that might exist due to some congested RF links. Especially, no improvement is observed for $M > 3$. When the FSO links are available with a higher probability, namely $\pi = 0.2$, more improvement is observed and the network capacity curve is saturated at a higher value. The same trend is observed for $\pi = 0.3$ where $M > 5$.

In the latter scenario, the achievable throughput is close to the case where the FSO links are 100% available and $M \leq 4$, as shown in Fig. 3. In conclusion, in order to achieve the optimal network capacity for a certain number of FSO link upgrades, these links do not need to have 100% availability. This observation shows that upgrading a limited number of network links with FSO technology is worthwhile even if the links are not highly available. In addition, obtaining almost the same network capacity as $\pi = 1$ for $\pi = 0.3$ availability of the FSO link confirms that the PGS algorithm performs well at lower link availabilities.

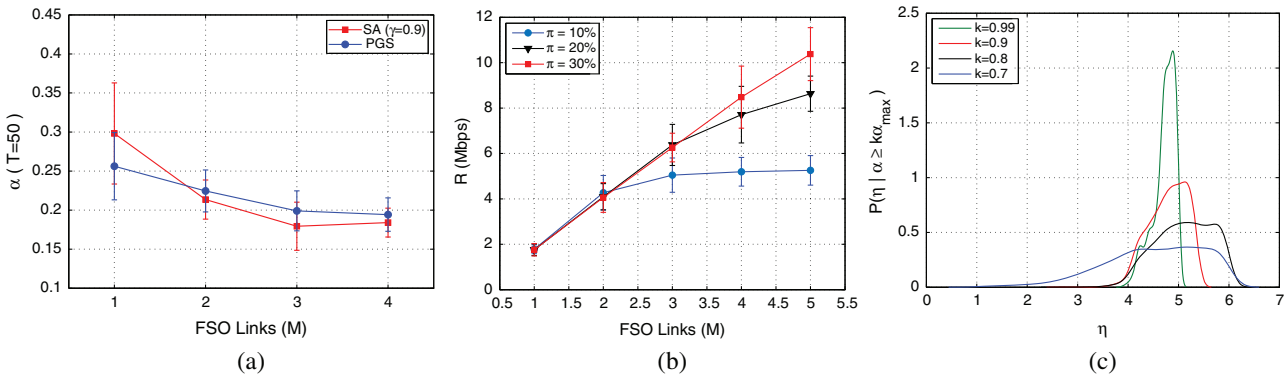

Fig. 4. (a) Comparison between achieved α for PGS and SA for $T = 50$. (b) Effect of link availability on the network capacity for $T = 50$. (c) Estimated probability density function of α versus η.

It must be noted that although in theory, the optimal network capacity can be achieved at a fraction of the full availability of FSO links, in practice, more memory is required in the relay nodes to be able to buffer data in the case that a channel fade happens. This is likely to increase the average end-to-end delay.

D. Analysis of PGS

As mentioned above, PGS is less complex than SA since parameter η splits the solution space into several subspaces and randomly samples the subspace with the highest likelihood of providing a better solution. In this section, using a Monte Carlo approach, the density function of α versus η is calculated to analyze the key role played by η for one representative graph with $M = 3$. Figure 4(c) shows the estimated probability density functions of α given that its value is greater than $k \times \alpha_{\max}$ in 10,000 trials performed. As can be seen in this figure, for $k \in \{0.7, 0.8, 0.9, 0.99\}$, the conditional density functions are concentrated around some value of η in the range of [1,2]. In other words, the better the performance, i.e., the larger the value of k, the sharper and narrower the density function becomes. This implies that for a reasonable quality of the solution, one may not need to search the entire space. Especially, it can be seen that when $k = 0.99$, only the range of Refs. [1,2] needs to be searched. It should be noted that the range depends on the network topology and the number of upgrades. A simple binary search algorithm may be used along with the upper bound $\alpha \le \frac{1}{\eta}$ to explore a good range for η.

VI. Conclusion

In this paper, we proposed a new method for network capacity enhancement using FSO transceivers in hybrid wireless-optical networks. We formulated the problem of joint FSO placement and link scheduling in the form of an integer linear program. It was shown that upgrading a few strategically located nodes with such transceivers can greatly improve the capacity of the network. We also proposed a probabilistic heuristic and a simulated annealing scheme that can perform competitively compared to the optimal approach at significantly lower computation costs. Considering switching delays imposed by link scheduling and the reliability of the FSO links is part of our future work.

References

[1] S. Sarkar, H.-H. Yen, S. Dixit, and B. Mukherjee, "Hybrid wireless-optical broadband access network (WOBAN): Network planning using Lagrangean relaxation," *IEEE/ACM Trans. Netw.*, vol. 17, no. 4, pp. 1094–1105, 2009.

[2] G. J. Narlikar, G. T. Wilfong, and L. Zhang, "Designing multihop wireless backhaul networks with delay guarantees," in *Proc. of IEEE INFOCOM*, Barcelona, Spain, April 2006.

[3] S. Sarkar, H.-H. Yen, S. S. Dixit, and B. Mukherjee, "DARA: Delay-aware routing algorithm in a hybrid wireless-optical broadband access network (WOBAN)," in *IEEE Int. Conf. on Communications (ICC)*, Glasgow, June 24–28, 2007, pp. 2480–2484.

[4] A. A. Reaz, V. Ramamurthi, S. Sarkar, D. Ghosal, S. S. Dixit, and B. Mukherjee, "CaDAR: An efficient routing algorithm for wireless-optical broadband access network," in *IEEE Int. Conf. on Communications (ICC)*, Beijing, May 19–23, 2008, pp. 5191–5195.

[5] A. S. Reaz, V. Ramamurthi, S. Sarkar, D. Ghosal, and B. Mukherjee, "Hybrid wireless-optical broadband access network (WOBAN): Capacity enhancement for wireless access," in *Global Telecommunications Conf. (GLOBECOM)*, 2008, pp. 2812–2816.

[6] F. Ahdi and S. Subramaniam, "Capacity enhancement of hybrid wireless optical networks using MIMO links," in *Global Telecommunications Conf. (GLOBECOM)*, 2011, pp. 1–6.

[7] B. Mumey, J. Tang, and T. Hahn, "Joint stream control and scheduling in multihop wireless networks with MIMO links," in *IEEE Int. Conf. on Communications (ICC)*, Beijing, May 19–23, 2008.

[8] S. Krishnamurthy and A. Acampora, "Capacity of a multihop mesh arrangement of radio cells connected by free-space optical links," in *12th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications*, San Diego, CA, Sept. 2001, vol. 2, pp. G-49–G-54.

[9] A. Kashyap, K. Lee, M. Kalantari, S. Khuller, and M. Shayman, "Integrated backup topology control and routing of obscured traffic in hybrid RF/FSO networks," *Comput. Netw.*, vol. 51, pp. 4237–4251, Oct. 2007.

[10] M. Smadi, S. Ghosh, A. Farid, T. Todd, and S. Hranilovic, "Free-space optical gateway placement in hybrid wireless mesh networks," *J. Lightwave Technol.*, vol. 27, no. 14, pp. 2688–2697, July 2009.

[11] Y. Tang and M. Brandt-Pearce, "Link allocation, routing, and scheduling for hybrid FSO/RF wireless mesh networks," *J. Opt. Commun. Netw.*, vol. 6, no. 1, pp. 86–95, 2014.

[12] F. Ahdi and S. Subramaniam, "Optimal placement of FSO links in hybrid wireless optical networks," in *Global Telecommunications Conf. (GLOBECOM)*, 2011, pp. 1–6.

[13] J. Rak and W. Molisz, "Reliable routing and resource allocation scheme for hybrid RF/FSO networks," in *16th Int. Conf. on Transparent Optical Networks (ICTON)*, 2014, pp. 1–4.

[14] F. Ahdi and S. Subramaniam, "Improving hybrid FSO/RF network reliability through transceiver reconfiguration," in *Global Telecommunications Conf. (GLOBECOM)*, 2012, pp. 2947–2952.

[15] F. Ahdi and S. Subramaniam, "Optimal placement of FSO relays for network disaster recovery," in *IEEE Int. Conf. on Communications (ICC)*, 2013, pp. 3921–3926.

[16] H. E. Nistazakis, E. A. Karagianni, A. D. Tsigopoulos, M. E. Fafalios, and G. Tombras, "Average capacity of optical wireless communication systems over atmospheric turbulence channels," *J. Lightwave Technol.*, vol. 27, no. 8, pp. 974–979, 2009.

[17] B. Epple and H. Henniger, "Discussion on design aspects for free-space optical communication terminals," *IEEE Commun. Mag.*, pp. 62–69, Oct. 2007.

[18] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[19] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*, Prentice Hall Communications Engineering and Emerging Technologies Series. Upper Saddle River, NJ: Prentice Hall, 2007.

Farshad Ahdi (M'06) earned his B.Sc. and M.Sc. in electrical engineering from Sharif University of Technology in 2003 and 2005, respectively. In 2010, he joined the University of Maryland, College Park, as a faculty research assistant, where he designed and developed wireless sensors hardware and software for traffic engineering applications. He earned his Ph.D. in electrical engineering from George Washington University in 2016, where he studied capacity and reliability planning of hybrid optical and wireless networks. He has published in several journals and conferences broadly in the area of optical, wireless, and sensor networks.

Suresh Subramaniam (S'95-M'97-SM'07-F'15) received his Ph.D. in electrical engineering from the University of Washington, Seattle, in 1997. He is a Professor in the Department of Electrical and Computer Engineering at George Washington University, Washington, DC. His research interests are in the architectural, algorithmic, and performance aspects of communication networks, with current emphasis on optical networks, cloud computing, and data center networks. He has published over 150 peer-reviewed papers in these areas.

Dr. Subramaniam is a co-editor of three books on optical networking. He is or has been on the editorial boards of six journals, including *IEEE/ACM Transactions on Networking* and IEEE/OSA *Journal of Optical Communications and Networking*, and has chaired several conferences, including IEEE INFOCOM 2013. He is a co-recipient of the Best Paper Awards at ICC 2006 and at the 1997 SPIE Conference on All-Optical Communication Systems. He is a Fellow of the IEEE.