Middle-mile Network Optimization in Rural Wireless Meshes

Yung-Fu Chen
Dept. of Computer Science and Engineering
The Ohio State University
Columbus, OH, USA
chen.6655@osu.edu

Anish Arora
Dept. of Computer Science and Engineering
The Ohio State University
Columbus, OH, USA
anish@cse.ohio-state.edu

Abstract—The status quo of limited broadband connectivity in rural areas motivates the need for fielding alternatives such as long-distance wireless mesh networks. A key aspect of fielding wireless meshes cost-effectively is planning how to connect the last-mile networks to the core network service providers (i.e., the network between the edge access terminals and the landline / optical fiber terminals) with minimal infrastructure cost and throughput constraints. This so-called middle-mile network optimization, which includes topology construction, tower height assignment, antenna and orientation selection, as well as transmit power assignment, is known to be a computationally hard problem.

In this paper, we provide the first polynomial time approximation solution for a generalized version of the middle-mile network optimization problem, wherein point-to-point (i.e., WiFi p2p) links are deployed to bridge last-mile networks. Our solution has a cost performance ratio of $O(\ln |A| + \frac{|B|}{|A|} + \frac{|A|+|B|}{\gamma})$, where A and B respectively denote the number of terminals and non-terminals and γ is the ratio of $\frac{link\ capacity}{terminal\ demand}$. Furthermore, our solution extends to hybrid networks, i.e., point-to-multipoint (i.e., WiFi p2mp) or omnidirectional (i.e., TV White Space) can serve as hyperlinks in addition to point-to-point links, to further reduce the cost of wireless links. We provide a complementary heuristic for our middle-mile network optimization solution that adds hyperlinks if and only if they reduce the cost.

Index Terms—Network planning, Rural connectivity, Low cost networking, Wireless mesh networks, 802.11/.22, WiFi, TVWS

I. Introduction

Even as the demand for broadband escalates worldwide to support real-time Internet services such as streaming, distance learning and telehealth, rural areas tend to be underserved if not unserved. In the US alone, of the 24 million Americans who lack access to fixed broadband services of at least 25 Mbps/3 Mbps as of 2016, the vast majority live in rural areas: 31.4% of the population in rural areas versus 2.1% in urban areas. And while fiber optic cable deployment continues, its high cost implies that wireless backbones remain a necessary part of the addressing the broadband digital divide for the foreseeable future.

Among the wireless approaches, wireless mesh networks offer a resilient alternative for spanning tens to a hundred kilometers in rural areas. They offer the potential of being low-cost, in part because they use unlicensed spectrum, unlike cellular network. IEEE 802.11-based long-distance wireless mesh networks have therefore been widely deployed to fill rural broadband gaps worldwide, [1]–[3]. Directional antennas of small beamwidth with WiFi radio are mounted on the top of towers of height sufficient to line-of-sight propagation over the links. TV White Space (TVWS) meshes are also starting to proliferate [4], [5] as a low-cost alternative since their spectrum offers non-line-of-sight propagation over foliage and obstructions, as compared to the 2.4GHz and 5GHz spectrum used in many IEEE 802.11 meshes. This allows TVWS to use lower height towers than WiFi, which is particularly cost-efficient for hilly rural areas. Also, the IEEE 802.22 TVWS range is competitive with that of 802.11 long-distance wireless links: a base station can provide the transmission radius up to 100 kilometers [6], [7].

A major challenge for wireless meshes is sustainability. The growth of demand for their services versus their profitability is a chicken-and-egg problem, which is made harder by a lack of networking planning tools that support incremental deployment. Part of the challenge is to contain the deployment cost, which can be dominated by tower costs. Towers are typically located at or near population hubs that need network connectivity, but suitable towers/structures are not easily available or afforded to providers. Tower cost is height dependent (say $100-$5000 for towers from 10-45m). Communication equipment cost is also a factor: while WiFi equipment is comparatively cheap (say $50 for WiFi p2p antennas), TVWS equipment is yet to enjoy economy of scale.

Anecdotally, our experience in dealing with rural broadband wireless providers corroborate the challenge in cost, throughput assurance, and flexibility for network growth. In a project [8] on infrastructure supporting rural counties in southern Ohio for data-driven services and analytics related to the opioid epidemic, we interacted with providers to deploy WiFi, TVWS, or hybrid meshes that would be leveraged by project stakeholders among other users. Providers point to high costs incurred for tower access, socio-political difficulty of leveraging community-shared tower structures for cost reduction, and networks planning that tends to err on the side of wasted capacity or does not efficiently support incremental growth of the networks.

Wireless Mesh Planning and MNO. Network planning in wireless meshes can be divided into the *last-mile network optimization (LNO)* problem and the *middle-mile network*

optimization (MNO) problem. Both are NP-hard [9], [10]. The former problem involves placing a minimum number of edge access terminals (i.e., gateways) to collect traffic flows from home APs while satisfying their QoS requirements. A polynomial time LNO approximation algorithm [9] exists that recursively calculates minimum weighted Dominating Sets (DS) while guaranteeing QoS requirements in each iteration.

MNO seeks to connect the edge-access terminals to core network service providers with minimal infrastructure cost, while satisfying the throughput constraints of each edge-access terminal. The MNO problem was first formulated [10] as:

> *Given a set of terminals to be connected to a given landline (and a set of non-terminals with existing towers to be used as relay vertices), determine the network topology, the tower heights of terminals, antenna types and orientations, transmit powers, and the route for each terminal to the landline, such that throughput constraints of terminals are satisfied and the total cost of towers and antennas is minimized.*

MNO problem is not yet well solved: Its solution in [10] has exponential time complexity, due to the formulation of height assignment and transmit power assignment as Linear Programming (LP) problems. Although a greedy approach has been proposed [11] for the NP-hard subproblem of topology construction along with tower height assignment, which has an approximation ratio of $O(\log n)$, we are not aware of approximation algorithms for MNO problem per se.

Contributions of the paper. In this paper, we design an approximation algorithm to solve MNO problem. To that end, we introduce the problem of Topology Construction of Minimum Steiner Tree (SteinerTC) problem.[1] Then, we solve MNO problem in two parts: solving first the SteinerTC problem and solving thereafter the Capacitated Network Design (CND) problem [12]. Next, we analytically evaluate the worst-case cost bound of the solution. Lastly, we extend the algorithm for solving MNO problem in hybrid networks, that allow incorporation of point-to-multipoint (i.e., WiFi p2mp) and omnidirectional links (i.e., TVWS) as hyperlinks. We give a complementary heuristic that replaces p2p links by p2mp or omnidirectional hyperlinks if and only if they reduce the total cost.

The contribution of the paper is four-fold: First, ours is the first polynomial-time approximation algorithm for MNO problem. Second, we show the cost performance ratio of our MNO solution is $O(\ln |A| + \frac{|B|}{|A|} + \frac{|A|+|B|}{\gamma})$, where $|A|$ is the number of terminals, $|B|$ is the number of non-terminal vertices and γ is the ratio of $\frac{link\ capacity}{terminal\ demand}$. Third, for the subproblem of SteinerTC, which we show is NP-hard (Theorem 1) by reducing from TC problem, we give a greedy solution with an approximation ratio of $O(\ln |A|)$ for the tower cost (Theorem 2). Finally, our MNO solution for hybrid networks achieves Capacity Flexibility, i.e., the added hyperlinks yield

---

[1]The SteinerTC problem—which is derived from the Topology Construction (TC) problem [11]—incorporates topology construction and tower height assignment in MNO problem.

residual capacity that can be leveraged to add additional edge access terminals in the future without incurring additional tower cost.

Organization. In Section II, we discuss the related work on TC and CND problems. In Section III, we formalize the MNO problem. We then present our approximation solution for a homogeneous case (p2p links only) of MNO problem and, in Section IV, analyze the cost bound. We describe, in Section V, our complementary heuristic that further reduces the infrastructure cost by introducing hyperlinks, and make concluding remarks, in Section VI.

II. RELATED WORK

A. Topology Construction (TC)

The TC problem is to determine a network topology that spans all the given terminals as well as a minimal height of antenna tower of each terminal such that line-of-sight propagation using a WiFi p2p link exists for each edge. TC was formalized as follows [11]:

> *Given an undirected graph G = (V, E), obstruction locations and their heights on each edge in G, and a cost function of tower height, determine a set of tower heights at vertices in V such that the subgraph induced by the edges covered by the set of tower heights is a subgraph (i.e., a spanning tree) of G spanning all vertices such that the total cost of the towers is minimized.*

The TC problem is known to be NP-hard and solvable by iteratively choosing the height increments of a vertex and a set of its neighbors with the best cost-to-benefit ratio until a spanning tree is determined.

Note that this TC formalization only considers topologies that span all vertices, i.e., each vertex is a terminal. One way of reducing the cost is to leverage pre-existing towers in rural areas to host non-terminal vertices (i.e., relays), thus eschewing additional tower cost while reducing the tower cost of some terminals. With this approach, the problem of determining the minimum cost subgraph of G that connects all terminals and may include some non-terminals becomes that of computing the Minimum Steiner Tree. This leads us to pose the Topology Construction problem for Minimum Steiner Tree (SteinerTC) problem of determining a set of tower height of terminals forming a Steiner tree with minimized cost. In Section III we show that SteinerTC is at least NP-hard and design an algorithmic solution for it.

B. Capacitated Network Design (CND)

Also related to our work is the CND problem of installing adequate capacity for meeting the bandwidth needs between a landline and edge access terminals. CND was formalized as follows [12], [13]:

> *Given an undirected graph G = (V, E), a set of terminals with their traffic demands, a landline, and uniform link capacity to be installed on edges, we are going to determine a minimum cost of capacity*

installation in G so that the installed capacity can route all traffics from terminals to the landline.

The authors in [12] propose an approximation algorithm by introducing a relation between the CND problem and minimum Steiner tree problem. First, they take advantage of known solutions to Steiner tree problem [14] to obtain a Steiner tree with minimum cost. Second, their solution divides terminals into groups whose internal traffic flows are equal to or less than the cable capacity, and assign a hub in each group to forward the aggregated traffic from the corresponding terminals to the landline through the additionally installed cable capacity[2], if necessary, on the shortest path between the hub and landline. Since the CND problem satisfies throughput constraints by finding a minimum cost of wired cable installation in a given graph instead of considering a wireless case where line-of-sight propagation is necessary to support the transmission of WiFi spectrum, the CND problem is only a particular instance of MNO problem. A solution to CND problem does not comprehensively solve MNO problem.

III. PROBLEM STATEMENT

In this section we describe the network model and state the middle-mile network optimization problem, whose goal, in addition to satisfying throughput constraints of terminals, is to minimize the cost infrastructure deployment consisting of towers and antennas, in a general way. The main notations are summarized in Table I. But first, we introduce the problem of SteinerTC.

A. Topology Construction of Minimum Steiner Tree (SteinerTC) Problem and Hardness

Unlike the TC problem, SteinerTC introduces non-terminals provided with towers as relays so as to minimize tower heights of terminals. We formulate the SteinerTC problem as follows:

> Given an undirected graph $G' = (V', E')$, a set of terminals $A \subset V'$, a set of non-terminals $B \subset V'$, $A \cup B = V'$, obstruction locations and their heights on each edge in G', and a cost function of tower height, we are going to determine a set of tower heights at vertices in A such that the subgraph induced by the edges covered by the set of tower heights in V' is a subgraph (i.e., a Steiner tree) of G' spanning all vertices in A such that the total cost of the towers at terminals is minimized.

Compared to the output graph of TC problem, which is a minimum cost spanning tree where the induced vertices are all terminals, the output graph of SteinerTC is a minimum cost Steiner tree where the spanned vertices consist of all terminals and some of the non-terminals.

We give a brief proof of the following theorem by reducing TC to SteinerTC.

Theorem 1. *SteinerTC is at least as hard as TC, which is NP-hard.*

---

[2]An edge in a graph can be installed more than one copy of cable.

Proof. For any two neighbor vertices u and v in the graph G in TC problem, assume we are allowed to create at least one intermediate vertex with fixed tower height between u and v but zero cost if there exists an edge (u, v) of line-of-sight propagation associated with height increments for u and v, denoted as $h^+(u)$ and $h^+(v)$ respectively. Accordingly, based on $h^+(u)$ and $h^+(v)$ which make edge (u, v) with cost increment $cTower^+(h^+(u))$ and $cTower^+(h^+(v))$ in TC problem, we could create one or more intermediate vertices $i_1, \ldots, i_r$ with fixed heights $h(i_1), \ldots, h(i_r)$ which make the same cost increment, $cTower^+(h^+(u)') + cTower^+(h^+(v)') = cTower^+(h^+(u)) + cTower^+(h^+(v))$, to make edges (u, i_1) and (i_r, v) of line-of-sight propagation with heights $h^+(u)'$ and $h^+(v)'$.

Note that $h(i_1), \ldots, h(i_r)$ are sufficient to make line-of-sight propagation over edges $(i_1, i_2), \ldots, (i_{r-1}, i_r)$. Thus we can say that for any new paths between (u, v), there exist a path with the minimal cost equal to the cost of $cTower^+(h(u)) + cTower^+(h(v))$. Some intermediate vertices with fixed tower heights could be reused as relays in a new path between any two neighbor vertices in G, thus the degree of intermediate vertex is ≥ 2.

Hence, for an instance of SteinerTC problem with the graph $G' = (V', E')$, there exists an instance of TC problem with the graph $G = (V, E)$, which can be transformed to G', where all the intermediate vertices are non-terminals in G'. □

B. Reformulation of MNO

Recall that the parameters to be determined in the MNO problem are network topology, tower heights, antenna type and orientations, transmit powers, and traffic routes. The SteinerTC problem—finding a minimum cost Steiner tree by identifying a set of tower heights—determines the first two of these parameters. Given this Steiner tree as an input graph, the CND problem is able to find a minimum capacity installation routing traffics to landline. This satisfies throughput constraints in a consideration of line-of-sight propagation over WiFi spectrum. Note that, in a homogeneous case (p2p links only) of the MNO problem, the parameters of antenna type and orientations and transmit powers can be simply determined on the basis of the edges in the Steiner tree. (In hybrid networks, the schemes of assigning antenna orientations of WiFi p2mp and transmit power are presented in Section V.) Hence, solving SteinerTC and CND problems in sequence suffices to determine the parameters in the MNO problem.

The input and output of MNO problem are reformulated as follows.

> ***Input***. An undirected graph $G = (V, E)$, a set $A \subset V$ of terminals and a set $B \subset V$ of non-terminals, $A \cup B = V$, a landline $LN \in A$, a cost function of tower height, a cost function of link installation, and throughput demand $dem(v) \leq U$ for each terminal.

Note that for any two vertices u and v, $(u, v) \in E$, $dist(u, v) \leq R$. For each edge $(u, v) \in E$, there exists a normalized

Symbol	Meaning
G	input graph of MNO problem $G = (V, E), V = A \cup B$
A	set of terminals
B	set of non-terminals
$LN \in A$	landline (core network service provider)
$U \in \mathbb{N}$	capacity of wireless p2p link
$cTower$	tower cost function $cTower : [HTMIN, HTMAX] \to \mathbb{R}$
$HTMIN \in \mathbb{R}$	minimum tower height
$HTMAX \in \mathbb{R}$	maximum tower height
$cLink$	link cost function $cLink : \mathbb{N} \to \mathbb{R}$
dem	throughput demand function $dem : V \to [0, U], \forall v \in B, dem(v) = 0$
$R \in \mathbb{R}$	maximal transmit distance of WiFi p2p links
ob	normalized obstruction height function $ob : E \to \mathbb{R}$
$G_{SteinerTC}$	output graph of MNO problem $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$ $A \subseteq V_{SteinerTC} \subseteq V$
h	tower height function $h : V \to [HTMIN, HTMAX]$ $\forall v \in B, h(v)$ is given and fixed
rt	route function from terminals to LN $rt : V \to P$ P: set of paths composed of edges in $E_{SteinerTC}$
f	traffic flow function $f : E_{SteinerTC} \to \mathbb{R}$
$linkset_{MP}$	hyperlink function at vertices with MP antennas $linkset_{MP} : V_{SteinerTC} \to CONFS$ $CONFS$: set of subset of MP antenna configurations (see Algorithm 3)
$linkset_{Omni}$	hyperlink function at vertices with $Omni$ antennas $linkset_{Omni} : V_{SteinerTC} \to VS$ VS: set of subsets of $V_{SteinerTC}$

obstruction height $ob(u, v)$ at the middle of edge (u, v). Line-of-sight propagation is available over a wireless p2p link connecting u and v if the sum of the tower heights at u and v is not less than twice of $ob(u, v)$.

> ***Output***. A connected graph $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, along with a height function h for each terminal in A, the route function rt for each terminal in V indicating the routing path from terminals to the landline, the traffic flow function f for each edge in $E_{SteinerTC}$, and the hyperlink functions $linkset_{MP}$ and $linkset_{Omni}$ if a hybrid network is constructed, such that throughput demands of terminals are satisfied and the cost of tower and wireless link installation $\sum_{v \in A} cTower(h(v)) + \sum_{e \in E_{SteinerTC}} cLink(\lceil \frac{f(e)}{U} \rceil)$ is minimized.

Note that $linkset_{MP}$ and $linkset_{Omni}$ denote the set of hyperlinks. For vertices v where p2mp antenna(s) are installed, $linkset_{MP}$ specifies the p2mp antenna configurations at v that include the covered vertices and effective direction/range. For vertices v where an omnidirectional antenna is installed, $linkset_{Omni}$ returns vertices in v's neighborhood that are connected to the omnidirectional antenna at v.

IV. POLYNOMIAL TIME APPROXIMATION ALGORITHM

A. Solution to SteinerTC Problem

A greedy algorithm, TC-ALGO, is proposed in [11] to solve TC with a logarithmic cost bound. In each iteration of TC-ALGO, all terminals are considered with different height increments associated with cost increments, using a doubling search, to find the one with the lowest cost-to-benefit ratio. For each combination of a terminal and a certain height increment considered, TC-ALGO considers its neighbor terminals with the minimum height increments to achieve line-of-sight links by calling the subroutine of STAR-TC-ALGO. Its evaluation metric is the cost-to-benefit ratio defined as the minimum increment in cost for the maximum reduction in the number of isolated components. A component of a graph denotes a maximal connected subgraph.

Given a terminal v with a height increment $h^+(v)$, STAR-TC-ALGO sorts the terminal's neighbors $u_1, \ldots, u_s$ by the cost of the neighbors' minimum height increments $cTower^+(h(u_1)), \ldots, cTower^+(h(u_s))$, that cover edges $(v, u_1), \ldots, (v, u_s)$. Then it returns the list of selected neighbors achieving lowest cost-to-benefit ratio associated with v and $h^+(v)$. In each iteration of TC-ALGO, the terminal with the height increment and some of its neighbors with minimum sufficient height increments achieving the lowest cost-to-benefit ratio over all combinations are selected to add to h. TC-ALGO iteratively choose the height increments until the number of components is reduced to one.[3] Moreover, components covered by the height function h are denoted by $cover(h)$. At the beginning, $cover(h)$ stands for all isolated terminals due to zero height increment at all terminals.

In the TC problem all vertices are considered terminals and assigned tower heights. However, in SteinerTC problem some of the vertices considered non-terminals with fixed tower heights can work as relays between two terminals to further reduce the tower heights at terminals if applicable. To find the lowest cost-to-benefit ratio in SteinerTC problem, for a given terminal v with a height increment, we search not only the adjacent terminals of v but the terminal u that can be connected through some non-terminals $v_1, \ldots, v_r$ by the path $(v, v_1, \ldots, v_r, u)$. We use the term of *logical neighbor* of a given vertex v to denote a terminal u that can be connected by a single edge (u, v) or a path $(v, v_1, \ldots, v_r, u)$ where only v and u are terminals.

STAR-SteinerTC-ALGO, presented in Algorithm 1, is adapted from STAR-TC-ALGO. (Note that the main greedy routine, SteinerTC-ALGO, to SteinerTC problem remains the same as TC-ALGO.[4]) In STAR-SteinerTC-ALGO, line 4 collects the logical neighbors of v that are not in the same component as v in $cover(h)$, since the benefit of v cannot be improved by increasing the height of the vertices in the same

---

[3]The benefit, standing for the number of reduced components, is at least one. This guarantees the progress of component reduction in each iteration.

[4]In each iteration, the greedy algorithm of SteinerTC-ALGO considers different height increments only for terminals, since the tower heights of non-terminals are fixed.

component of v. The for loop from line 5 to 13 determines the smallest height increment of each vertex u in nbr_{logic} that makes u connected to v. If there is a single edge (u,v) between u and v, we directly assign the smallest height increment at u to cover this edge. Otherwise, we search a minimum cost path[5] from v to u including only non-terminals, $v_1, \ldots, v_s$, except u and v and then assign the smallest height increment at u to cover (v_s, u). Note that the tower heights at $v_1, \ldots, v_s$ must be high enough to cover edges $(v, v_1), \ldots, (v_{s-1}, v_s)$. A vertex with the lowest cost increment (and well as the lowest height increment) is chosen from the same component and its vertices are sorted in ascending order of the cost increment from line 14 to 17. Note also that each element in L reduces the number of components in $cover(h)$ by exact one. Then we can find the lowest cost-to-benefit ratio by selecting of first k elements in L. The for loop from line 19 to 24 determines the size of k to obtain the lowest cost-to-benefit ratio r'_{best}. The remaining lines add the corresponding height increments for v and the chosen logical neighbors of v to the height increment function $incr$ and then return it.

The performance bound of our greedy algorithm is shown in Theorem 2, whose proof is relegated to Appendix A.

Theorem 2. *The tower cost associated with the height function h determined by SteinerTC-ALGO is at most $2 \ln |A| OPT$.*

B. Solution to CND Problem

The approximation algorithm presented in [12] installs capacity per a given minimum cost Steiner tree to route the traffic flows from terminals to the landline. The algorithm first partitions terminals into groups according to their throughput demands. Note that the total aggregated traffic in a group is upperbounded by the cable capacity. For each group, it chooses the vertex with shortest path to the landline as a hub. The algorithm of the grouping guarantees that any internal aggregated traffic in a group does not exceed $U - dem(v_h)$, where U is the uniform cable capacity and v_h is the hub in this group. Thus, installing only one cable over every edge in the minimum cost Steiner tree can support the traffic flows from terminals to hubs. This capacity installation makes the connectivity in the Steiner tree.

Moreover, for each group, one cable is installed on per edge over the shortest path from its hub to the landline if additional capacity is needed. This capacity installation guarantees the bandwidth requirement from hubs to the landline. Hence, the capacity installation from terminals to hubs and from hubs to the landline satisfies the throughput constraints. Note that the cost of capacity installation from terminals to the landline is dominated by their hop distances to the landline. In our solution to CND problem, we take advantage of the grouping scheme in [12] on the minimum cost Steiner tree, $G_{SteinerTC}$, determined by SteinerTC-ALGO. For each group, the vertex with shortest hop distance to the landline LN is selected as a hub.

---

[5] The minimum cost path between v and u can be found in polynomial time since the heights of v and all non-terminals are given.

Algorithm 1: STAR-SteinerTC-ALGO

1 Input: $G = (V, E)$, terminal set A, non-terminal set B, height function h, vertex v, height increment δ at v
2 Output: cost-benefit-ratio r'_{best}, height increment function $incr$
3 $cTower^+(h(v)) := cTower(h(v) + \delta) - cTower(h(v))$;
4 $nbr_{logic} := \{u \mid v, u \in A$, u is not in the same component as v in $cover(h)$, there exists a path $p = (v, v_1, \ldots, v_r, u)$ from v to u in G s.t. $|p| \geq 2$, $v_1, \ldots, v_r \in B$, heights $(h(v) + \delta)$ at v and heights $h(v_1), \ldots, h(v_r)$ cover edges $(v, v_1), \ldots, (v_{r-1}, v_r)\}$;
5 **for** *each vertex $u \in nbr_{logic}$* **do**
6   **if** $(u, v) \in E$ **then**
7         $h^+(u) :=$ smallest β s.t. heights $(h(v) + \delta)$ at v and $(h(u) + \beta)$ at u cover edge (u, v);
8   **else**
9   search the path $p_{best} = (v, v_1, \ldots, v_s, u)$ from v to u with lowest cost increment;
10         $h^+(u) :=$ smallest β s.t. heights $(h(v) + \delta)$ at v and $(h(u) + \beta)$ at u cover path $p_{best} = (v, v_1, \ldots, v_s, u)$;
11   **end**
12     $cTower^+(h(u)) := cTower(h(u) + h^+(u)) - cTower(h(u))$;
13 **end**
14 $L :=$ list of vertices in nbr_{logic} in ascending order of $cTower^+$;
15 **for** *each component $D \in cover(h)$* **do**
16   remove from L all vertex $u \in D \cap nbr_{logic}$ except the one with lowest $cTower^+$;
17 **end**
18 $r'_{best} := \infty$; $k_{best} := 0$;
19 **for** $1 \leq k \leq |L|$ **do**
20     $r'_{tmp} := \frac{cTower^+(h(v)) + \sum_{i=1}^{k} cTower^+(h(L[i]))}{k}$;
21   **if** $r'_{tmp} < r'_{best}$ **then**
22         $k_{best} := k$; $r'_{best} := r'_{tmp}$;
23   **end**
24 **end**
25 **for** $u \in V$ **do**
26     $incr(u) := 0$;
27 **end**
28 **for** $u \in L[1 \ldots k_{best}]$ **do**
29     $incr(u) := h^+(u)$;
30 **end**
31 $incr(v) := \delta$;
32 **return** $(r'_{best}, incr)$;

For the capacity installation, we first install one wireless link per edge in $G_{SteinerTC}$. Second, to route the aggregated traffics of groups to LN, one additional wireless link, if necessary, is installed per edge on the paths from hubs to LN, which are formed by the corresponding edges in $G_{SteinerTC}$. In our solution, the route function rt and the traffic flow function f are determined by the grouping scheme and the terminals demands.

C. Performance Analysis

In this section, we will derive the cost performance ratio of our approximation algorithm to MNO problem by analyzing the total cost of tower and wireless link installation $(\sum_{v \in A} cTower(h(v)) + \sum_{e \in E_{SteinerTC}} cLink(\lceil \frac{f(e)}{U} \rceil))$. In the following, let $cTower(SteinerTC)$ and $cLink(CND)$ respectively denote the tower cost and wireless link cost in our solution.

First, we analyze the cost of tower deployment. Let $cTower(OPT)$ be the cost of towers in the optimal solution to SteinerTC problem. Suppose there exists a minimum cost spanning tree where all vertices are terminals. The tower cost

of this spanning tree is denoted by $cTower(MST)$. Since our solution to SteinerTC problem exploits non-terminals with existing towers to further minimize the cost of towers, $cTower(SteinerTC) \leq cTower(MST)$. In addition, since our solution to SteinerTC problem has the cost bound ratio of $2\ln|A|$ to the optimal solution, we get the equation:

$$cTower(OPT) \leq cTower(SteinerTC)$$
$$\leq 2\ln|A|cTower(OPT) \qquad (1)$$

Second, we analyze the cost of wireless link installation. Let $cLink(OPT)$ be the cost of all wireless links in the optimal solution to CND problem. Since $cLink(CND)$ consists of the cost of installed links from terminals to hubs and the cost of installed links from hubs to the landline, we use $cLink(CND_{T2H})$ and $cLink(CND_{H2L})$ to respectively denote the two costs. Let $cLink(MST_{T2H})$ be the cost of link installation from terminals to hubs in MST where each edge connects two terminals and the number of edges, $|A|-1$, is minimal. Thus, we know $cLink(MST_{T2H}) \leq cLink(OPT)$ and $cLink(MST_{T2H}) \leq cLink(CND_{T2H})$. Since the worst case of the number of edges in a Steiner tree is the number of all vertices minus one ($|A|+|B|-1$), we get the worst-case bound of $cLink(CND_{T2H})$:

$$cLink(MST_{T2H}) \leq cLink(CND_{T2H})$$
$$\leq \frac{|A|+|B|-1}{|A|-1}cLink(MST_{T2H})$$

Also, we infer the bound ratio of $cLink(CND_{T2H})$ to $cLink(OPT)$:

$$cLink(CND_{T2H}) \leq \frac{|A|+|B|-1}{|A|-1}cLink(MST_{T2H})$$
$$\leq \frac{|A|+|B|-1}{|A|-1}cLink(OPT) \leq \frac{|A|+|B|}{|A|}cLink(OPT) \qquad (2)$$
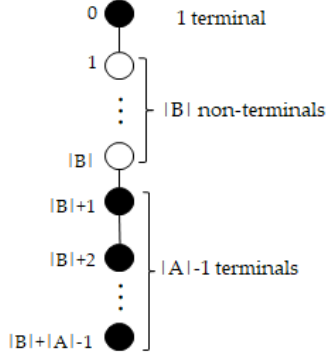$$= (1+\frac{|B|}{|A|})cLink(OPT)$$

Fig. 1. Worst-case topology of CND solution.

Moreover, the upper bound of $cLink(CND_{H2L})$ is obtained by considering the worst case of total hop distance from hubs to the landline. Let all terminals have the uniform demand and all wireless links have a uniform link capacity. Let γ denote the ratio of $\frac{link\ capacity}{terminal\ demand}$. To calculate the lower bound of $cLink(OPT)$, we consider the best case of the topology, where all the vertices have 1-hop distance to the landline. In that case, a wireless link is installed on every edge from terminals to the

landline. Thus, we have the best-case bound of $cLink(OPT)$, $link\ cost \times (|A|-1) \leq cLink(OPT)$. Now we calculate the worst case of the topology that is a long-chain topology where the 0th vertex is the landline (terminal), the 1st to the $(|B|)$-th vertices are non-terminals, and the $(|B|+1)$-th to $(|B|+|A|-1)$-th vertices are the remaining $|A|-1$ terminals. As Fig. 1 depicted, the numbers labeled at left stand for the hop distances for the vertices to the landline. Note that the grouping scheme to CND problem separates all terminals vertex into $\frac{|A|\times terminal\ demand}{link\ capacity} = \frac{|A|}{\gamma}$ groups. For convenience, $\frac{|A|}{\gamma}$ is denoted as m. For each group, we select a hub with minimum hop distance to send the aggregated traffic equal to (or less than) the link capacity to the landline. Thus, in the 1st group containing terminals with the labels, $0, |B|+1, \ldots, |B|+\gamma-1$, we select vertex 0 as a hub. In the 2nd group containing terminals with the labels, $|B|+\gamma, |B|+\gamma+1, \ldots, |B|+2\gamma-1$, we select vertex $|B|+\gamma$ as a hub. In the last group containing terminals with the labels, $B+(m-1)\gamma, B+(m-1)\gamma+1, \ldots, B+m\gamma-1)$, we select vertex $B+(m-1)\gamma$ as a hub. Now the total hop distance from hubs to the landline in the worst case can be presented as follows:

$$0 + (|B|+\gamma) + \cdots + (|B|+(m-1)\gamma)$$
$$= \gamma[(\frac{|B|}{\gamma}+1) + \cdots + (\frac{|B|}{\gamma}+m-1)]$$
$$= \gamma[(m-1)\frac{|B|}{\gamma} + (1+\cdots+(m-1))]$$
$$= \gamma[(m-1)\frac{|B|}{\gamma} + \frac{m(m-1)}{2}]$$
$$= (m-1)(|B|+\frac{m\gamma}{2})$$
$$= (\frac{|A|}{\gamma}-1)(|B|+\frac{|A|}{2}), m \leftarrow \frac{|A|}{\gamma}$$
$$\leq \frac{|A|}{\gamma}(|B|+\frac{|A|}{2}) = \frac{|A|^2+2|A||B|}{2\gamma}$$

According to the best-case bound of $cLink(OPT)$ and the worst-case bound of $cLink(CND_{H2L})$, we know $link\ cost \times (|A|-1) \leq cLink(OPT)$ and $link\ cost \times (|A|-1) \leq cLink(CND_{H2L}) \leq link\ cost \times \frac{|A|^2+2|A||B|}{2\gamma}$. Hence, the performance ratio of $cLink(CND_{H2L})$ to $cLink(OPT)$ can be stated as follows:

$$cLink(CND_{H2L}) \leq \frac{|A|^2+2|A||B|}{2\gamma(|A|-1)}cLink(OPT)$$
$$\leq \frac{|A|^2+2|A||B|}{\gamma|A|}cLink(OPT) \leq \frac{|A|+2|B|}{\gamma}cLink(OPT) \qquad (3)$$

Note that $|A|-1 \geq \frac{|A|}{2}$ due to $|A| \geq 2$.

Finally, we calculate the cost performance ratio of our approximation algorithm to MNO problem. Let C_{OPT} be the total infrastructure cost of the optimal solution to MNO problem. Since the sum of the costs of the respective optimal solutions in SteinerTC and CND problems is the cost lower bound for C_{OPT}, we know $cTower(OPT) + cLink(OPT) \leq C_{OPT}$. Let C_{ALGO} be the total infrastructure cost of our solution to MNO problem, where $C_{ALGO} = cTower(SteinerTC) + cLink(CND_{T2H}) + cLink(CND_{H2L})$. Now we are going to the discuss the three following cases of terminal demand and articulate their performance ratio.

Case 1: The total demand from terminals is not larger than link capacity. In this case, $C_{ALGO} = cTower(SteinerTC) + cLink(CND_{T2H})$ since installation of a single wireless link on each edge from terminals to hubs is sufficient to convey

the total traffic from all terminals. According to (1) and (2), we have

$$cTower(SteinerTC)$$
$$\leq 2\ln|A|cTower(OPT) \leq 2\ln|A|C_{OPT}$$

and

$$cLink(CND_{T2H})$$
$$\leq (1+\tfrac{|B|}{|A|})cLink(OPT) \leq (1+\tfrac{|B|}{|A|})C_{OPT}$$

Hence, the cost performance ratio of case 1 is

$$\frac{C_{ALGO}}{C_{OPT}} \leq (1+2\ln|A|+\tfrac{|B|}{|A|}) \qquad (4)$$

Case 2: The total demand from terminals is larger than link capacity and the terminal demands are uniform. In this case, $C_{ALGO} = cTower(SteinerTC) + cLink(CND_{T2H}) + cLink(CND_{H2L})$ According to (3), we know

$$cLink(CND_{H2L})$$
$$\leq \tfrac{|A|+2|B|}{\gamma}cLink(OPT) \leq \tfrac{|A|+2|B|}{\gamma}C_{OPT}$$

Hence, the cost performance ratio of case 2 is

$$\frac{C_{ALGO}}{C_{OPT}} \leq (3+\ln|A|+\tfrac{|B|}{|A|}+\tfrac{|A|+2|B|}{\gamma})$$
$$= O(\ln|A|+\tfrac{|B|}{|A|}+\tfrac{|A|+|B|}{\gamma}) \qquad (5)$$

Case 3: The total demand from terminals is larger than link capacity and the terminal demands are non-uniform. In this case, we can reuse the result from (5) by letting γ be $\frac{link\ capacity}{maximum\ demand}$.

V. COST MINIMIZATION HEURISTIC FOR HYBRID MESH NETWORKS

In this section, we present a heuristic for adding p2mp or omnidirectional hyperlinks, of say WiFi for p2mp or TVWS for omnidirectional, if and only if they reduce the total cost while preserving the satisfaction of throughput constraints. The algorithm iteratively finds a vertex that has some links with residual capacity and then replaces those links with a hyperlink while considering cost reduction. Below, we first articulate a concept of link replacement. Second, we present a replacement scheme for the case of p2mp hyperlinks that substitute p2p links without changing the topology $G_{SteinerTC}$, tower heights, and the traffic routing. Third, we introduce a replacement scheme for the case of omnidirectional hyperlink that substitute some p2p links and for altering the corresponding tower heights adapting to omnidirectional antennas and their subordinate directional antennas when that can offer cost reduction. Finally, we discuss the assignment of transit power in hybrid networks.

A. Hyperlink Replacement Strategy

The solution to the CND problem that we have adopted involves a partitioning wherein terminals are divided into groups whose aggregated traffics do not exceed the link capacity. This scheme results in links within groups that are not fully utilized. This underutilized capacity of wireless link is called *residual capacity*. For the traffic flow function f returned from MNO problem, it computes the required traffic for each edge

$e \in E_{SteinerTC}$. Let U be the link capacity over its spectrum. We assume that a p2p link (which has two p2p antennas) and a p2mp hyperlink (which as one p2mp antenna and one or more p2p antennas) have the same capacity U, and yet an omnidirectional hyperlink (which as one omnidirectional antenna and one or more its subordinate directional antennas) can have different capacity U_{Omni}. The residual capacity of edge e is thus defined as $U\lceil\frac{f(e)}{U}\rceil f(e)$.

We now describe our cost minimization heuristic that replaces underutilized edges with p2mp and omnidirectional hyperlinks. Note that one edge is installed with one or more p2p links. The related notations are listed in Table II. Recall that although the main goal of hyperlink replacement is to minimize the infrastructure cost in hybrid networks, the replacement must satisfy the throughput constraints. The heuristic does not change the MNO solution topology or the traffic routing, since changing the topology will alter the traffic flows and that may break the constraints that were guaranteed by solving CND. In other words, replacement with a p2mp / omnidirectional hyperlink only happens at "candidate" vertices with respect to some/all of their respective neighbors; it does not include any vertex which is at a distance of more than one hop from any candidate vertex.

More specifically, given a candidate vertex v, the replacement only involves the edges connected to its children instead of the edge $(v, v.parent)$ that conveys the aggregated traffic of all its children. That is, we choose to not deploy a hyperlink that covers both $(v, v.parent)$ and the edges from v to its children. The reasons for this are: First, for p2mp antennas, the replacement of $(v, v.parent)$ may undermine throughput performance because only one link in a p2mp hyperlink can transmit at a given time. Second, for omnidirectional links, such replacement occupies twice the capacity over the spectrum but does not effectively reduce the cost of tower height, which is the dominant factor for infrastructure cost. At v, we apply its height to *HTOmni*, which may increase cost ($h(v) < $ *HTOmni*) or reduce cost ($h(v) > $ *HTOmni*). Nevertheless, at vertex $v.parent$, we apply its height to $max(h(v.parent), HTOmniSD)$ because $v.parent$ still has other induced edges(s). Note that $h(v)$ is likely larger than *HTOmniSD* (10m) and the available channels in the omnidirectional link's spectrum are limited and location-dependent. Therefore, allowing replacement only at children vertices offers the chance to replace more p2p links as well as reduce the corresponding tower heights.

B. MP Deployment

For a p2mp hyperlink, we use one p2mp antenna to share the capacity U among its corresponding p2p antennas. A p2mp antenna is modeled by three variables: direction, beamwidth, and radius. We denote, for a p2mp antenna is deployed at v, the direction of the antenna pointed towards vertex u as $\overrightarrow{vu}$. We assume the p2mp antenna is flexible so as to configure its beamwidth and transmit radius so that one p2mp antenna can be set to different sector ranges; this is typical of today's p2mp antennas.

Symbol	Meaning
U	p2p link / p2mp hyperlink capacity
R_{MP}	maximal transmit distance of MP antenna, $R_{MP} \leq R$
BW	tunable beamwidth of MP antenna, $BW \leq BWMAX$
$BWMAX$	maximal beamwidth of MP antenna
RAD_{MP}	tunable radius of MP antenna, $RAD_{MP} \leq R_{MP}$
X	set of vertices covered by a hyperlink
$cAntenna$	antenna cost function $cAntenna : \{PP, MP, Omni, OmniSD\} \rightarrow \mathbb{R}$
rng_{MP}	MP range verification function $rng_{MP} : E_{SteinerTC} \times V_{SteinerTC} \times V_{SteinerTC} \times \mathbb{R} \times \mathbb{R} \rightarrow \{0,1\}$
U_{Omni}	omnidirectional hyperlink capacity
R_{Omni}	maximal transmit distance of omnidirectional antenna
RAD_{Omni}	radius of omnidirectional antenna, $RAD_{Omni} \leq R_{Omni}$
$HTOmni$	fixed tower height of omnidirectional antenna
$HTOmniSD$	fixed tower height of subordinate directional antenna of omnidirectional hyperlink
rng_{Omni}	omnidirectional range verification function $rng_{Omni} : V_{SteinerTC} \times \mathbb{R} \times V_{SteinerTC} \times \mathbb{R} \rightarrow \{0,1\}$
V_{Omni}	set of vertices with omnidirectional antennas
$V_{OmniSD}(v)$	set of vertices only covered by the omnidirectional hyperlink induced from v

For MP deployment, there are four constraints that must be guaranteed if and only if a p2mp hyperlink is deployed at v to replace p2p links associated with the edges $(v,x), x \in X$, where X denotes the set of selected children of v.

$Cost : cAntenna(MP) < cAntenna(PP)|X|$
$Capacity : \sum_{x \in X} f((v,x)) \leq U$
$Range : \prod_{(v,x) \in E_{SteinerTC}, x \in X} rng_{MP}((v,x), v, u, BW, RAD_{MP}) = 1$
$Interference : \sum_{e \in E_{SteinerTC} \setminus \{(v,x)|x \in X\}} rng_{MP}(e, v, u, BW, RAD_{MP}) = 0$

The cost and capacity constraints ensure that the replacement reduces the total cost and satisfies the throughput demand. The range verification function rng_{MP} examines if the edge segment (v,x) is covered by the p2mp antenna at v and pointed towards u, with the beamwidth BW and radius RAD_{MP}. The interference constraint guarantees that the deployed p2mp antenna does not affect any edge other than $(v,x), x \in X$.

We can now state the cost minimization problem of hybrid MP deployment as follows[6]:

> **Input.** $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, *the corresponding traffic flows on edges in $E_{SteinerTC}$.*
> **Output.** *A set of vertices where p2mp antennas are deployed, along with direction, beamwidth, and radius assignments such that all 4 constraints are satisfied and the cost of wireless links is minimized.*

The solution in Algorithm 2 considers candidate vertices (and their respective children) in a bottom-up fashion. For each candidate vertex it performs local MP deployment with the greedy Algorithm 3 that selects the antenna configuration

[6]Note that we can maximize the cost reduction by maximizing the number of replaced p2p links. The problem can be considered a maximum coverage problem [15].

(direction, beamwidth, radius) containing a largest number of uncovered vertices at each iteration.

Algorithm 2: MP-DEPLOY-ALGO

1 Input: $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, traffic flow function f
2 Output: p2mp hyperlink function $linkset_{MP}$
3 **for** $v \in V_{SteinerTC}$ **do**
4    $linkset_{MP} := \{\}$;
5 **end**
6 $R := V_{SteinerTC} \setminus \{leaves\}$;
7 **while** $R \neq \{\}$ **do**
8   pick a vertex $v \in R$ such that $v.depth$ is maximum;
9    $linkset_{MP}(v) :=$ MP-ANT-REPLACE$(G_{SteinerTC}, f, v)$;
10    $R := R \setminus \{v\}$;
11 **end**
12 **return** $linkset_{MP}$;

In MP-ANT-REPLACE, the while loop from line 4 to 32 finds the p2mp antenna with a configuration (direction, beamwidth, radius) covering a maximum number of uncovered vertices at each iteration. The for loop from line 6 to 17 tries to assign the antenna direction to every vertex u in the available children and then gradually decreases its beamwidth and radius until the constraints (cost, capacity, and interference) are satisfied. The tuning of antenna configuration is presented in the while loop from line 8 to 12. Let L be the candidate list of the available antenna configurations to be deployed in each iteration, as assigned from line 13 to 16. After L is computed, from line 18 to 31 we choose the one that covers the maximum vertices. If there is no available candidate found, from line 18 to 19, we end the replacement at v. Otherwise, in the for loop from line 22 to 28 we find the candidate with the best coverage. Then we adopt this antenna configuration and remove the vertices covered by it. Note that this iteration (loop from line 4 to 32) ends at line 8, when there is no configuration satisfying the constraints. Finally, the set of selected antenna configurations is returned as $linkset_{MP}(v)$.

C. Omnidirectional Antenna Deployment

For an omnidirectional hyperlink, we use one omnidirectional antenna to share its capacity, U_{Omni}, with its corresponding subordinate directional antennas. Unlike p2mp antennas, an omnidirectional antenna covers its subordinate directional antennas according to its omnidirectional coverage shaped by R_{Omni}[7]. We assume that the spectrum of the omnidirectional hyperlinks does not overlap with that of the p2p/p2mp links, and also that these links are capable of non-line-of-sight propagation; these assumptions are again common; i,e, if we use TVWS omnidirectional hyperlinks with WiFi p2p/p2mp links. The implication of the latter assumption in particular is that omnidirectional antenna and their subordinate directional antennas are mounted at a fixed height ($HTOmni$ and $HTOmniSD$) for non-line-of-sight propagation so the needed tower heights to omnidirectional hyperlinks are considerably shorter.

[7]R_{Omni} varies with the coding schemes (i.e., 64QAM, 16QAM, QPSK) [7]. To support broadband services, we select the one that achieves the best capacity of U_{Omni}.

Algorithm 3: MP-ANT-REPLACE

1. Input: $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, traffic flow function f, candidate vertex v
2. Output: deployed p2mp antenna set of v, $linkset_{MP}(v) = \{(u_1, BW_1, RAD_{MP1}), \dots\}$
3. $S := \{u | u \in v.children\}$; $end := false$;
4. **while** $S \neq \{\} \wedge end \neq true$ **do**
5. $index := 0$;
6.   **for** $u \in S$ **do**
7. $BW := BWMAX$; $RAD_{MP} := R_{MP}$; $X := \{x | x \in v.children \wedge rng_{MP}((v,x), v, u, BW, RAD_{MP}) = 1\}$;
8.   **while** $\left(\frac{cAntenna(MP)}{cAntenna(PP)} < |X| \wedge (\sum_{x \in X} f((v,x)) \leq U) \wedge\right.$ $\left.(\sum_{e \in (E_{SteinerTC} \setminus \{(v,x)|x \in X\})} rng_{MP}(e, v, u, BW, RAD_{MP}) = 0)\right)$ $\neq true \wedge |X| > 1$ **do**
9.   decrease BW to the closest vertex $x \in X$ (ignore any vertex at the boundary);
10. $X := \{x | x \in v.children \wedge rng_{MP}((v,x), v, u, BW, RAD_{MP}) = 1\}$;
11.   decrease RAD_{MP} to the vertex $x \in X$ s.t. $dist(v,x)$ is maximum;
12.   **end**
13.   **if** $|X| > 1$ **then**
14. $L[index] := (u, BW, RAD_{MP})$;
15. $index := index + 1$;
16.   **end**
17.   **end**
18.   **if** $index = 0$ **then**
19. $end := true$;
20.   **else**
21. $cover_{best} := 0$;
22.   **for** $i \leftarrow 0, \dots, index - 1$ **do**
23. $cover := \#\{x | x \in v.children \wedge rng_{MP}((v,x), v, L[i].direction, L[i].BW, L[i].RAD_{MP}) = 1\}$;
24.   **if** $cover > cover_{best}$ **then**
25. $cover_{best} := cover$;
26. $select := i$;
27.   **end**
28.   **end**
29. $linkset_{MP}(v) := linkset_{MP}(v) \cup L[select]$;
30.   remove the set of vertices $\{x | x \in v.children \wedge rng_{MP}((v,x), v, L[select].direction, L[select].BW, L[select].RAD_{MP}) = 1\}$ from S;
31.   **end**
32. **end**
33. **return** $linkset_{MP}(v)$;

For Omnidirectional deployment, four constraints must be guaranteed if and only if an omnidirectional hyperlink at a candidate vertex v is going to replace WiFi p2p links associated with the edges $(v,x), x \in X$:

$$Cost : cAntenna(Omni) + cAntenna(OmniSD)|X|$$
$$+ cTower(\max(HTOmni, h(v)))$$
$$+ \sum_{u \in V_{OmniSD}} cTower(HTOmniSD)$$
$$< 2 \times cAntenna(PP)|X| + cTower(h(v))$$
$$+ \sum_{u \in V_{OmniSD}} cTower(h(u))$$
$$Capacity : \sum_{x \in X} f((v,x)) \leq U_{Omni}$$
$$Range : dist(v,x) \leq R_{Omni}, (v,x) \in E_{SteinerTC}, x \in X$$
$$Interference : \sum_{u \in V_{Omni}} rng_{Omni}(u, u.RAD_{Omni}, v,$$
$$RAD_{Omni}) = 0$$

Since Omnidirectional deployment as formulated not only affects the antenna cost but also changes the tower heights associated with the hyperlink, the cost constraint guarantees that the total infrastructure cost consisting of link installation and tower deployment must be reduced. The capacity and range constraints ensure that the terminal demands are still satisfied and the corresponding vertices are covered by the omnidirectional antenna. The interference constraint guarantees that the deployed omnidirectional antenna at the candidate vertex v does not affect any other omnidirectional hyperlink. The range verification function rng_{Omni} examines if the candidate vertex v with radius RAD_{Omni} overlaps any other omnidirectional antenna at u with its radius $u.RAD_{Omni}$, $u \in V_{Omni}$.

We can now state the cost minimization problem of hybrid Omnidirectional deployment as follows:

> **Input.** $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, *the corresponding traffic flows on edges in $E_{SteinerTC}$, the tower height at vertices in $V_{SteinerTC}$.*
> **Output.** *A set of omnidirectional antennas with the covered vertices deployed subordinate directional antennas such that the four constraints are satisfied and the cost of towers and wireless links is minimized.*

The solution in Algorithm 4 also considers each candidate vertex in a bottom up manner and tries to deploy an omnidirectional antenna for its children using Algorithm 5. The algorithm maximizes the cost reduction by maximizing the covered vertices whose tower heights are changed to *HTOmniSD*. Recall that it is profitable to change a tower height to *HTOmniSD* due to low cost (i.e., only \$100 for 10m).

Algorithm 4: OMNI-DEPLOY-ALGO

1. Input: $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, traffic flow function f, height function h
2. Output: omnidirectional hyperlink function $linkset_{Omni}$
3. $V_{Omni} := \{\}$;
4. **for** $v \in V_{SteinerTC}$ **do**
5. $linkset_{Omni} := \{\}$;
6. **end**
7. $R := V_{SteinerTC} \setminus \{leaves\}$;
8. **while** $R \neq \{\}$ **do**
9.   pick a vertex $v \in R$ such that $v.depth$ is maximum;
10. $linkset_{Omni}(v) := $ OMNI-ANT-REPLACE$(G_{SteinerTC}, f, h, V_{Omni}, v)$;
11.   **if** $linkset_{Omni}(v) \neq \{\}$ **then**
12. $V_{Omni} := V_{Omni} \cup v$;
13.   **end**
14. $R := R \setminus \{v\}$;
15. **end**
16. **return** $linkset_{Omni}$;

In OMNI-ANT-REPLACE, the while loop from line 5 to 17 finds a maximum set of vertices where we are going to substitute with subordinate directional antennas. A way to trim the set to satisfy the cost and capacity constraints is shown from line 6 to 15. First of all, line 6 verifies if the configuration of the candidate vertex v and radius RAD_{Omni} has no overlap with any other deployed omnidirectional antenna. If there is no overlap, then line 7 examines whether any vertex not in V_{OmniSD} exists. We then remove the one with maximum flow at line 8. Otherwise, at line 10 we remove the one in V_{OmniSD} with maximum flow. This enhances the chance to satisfy capacity constraint as well as remain more residual

Algorithm 5: OMNI-ANT-REPLACE

1. Input: $G_{SteinerTC} = (V_{SteinerTC}, E_{SteinerTC})$, traffic flow function f, height function h, omnidirectional antenna set V_{Omni}, candidate vertex v
2. Output: deployed subordinate directional antennas of v, $linkset_{Omni}(v) = \{u_1, \ldots, u_{|X|}\}$
3. $X := \{x | x \in v.children \wedge dist(v, x) \le R_{Omni}\}$;
4. $V_{OmniSD} := \{x | x \in X, (v, x) \text{ is the only induced edge from } x\}$;
5. **while** $(cAntenna(Omni) + cAntenna(OmniSD)|X| + cTower(\max(HTOmni, h(v))) + \sum_{u \in V_{OmniSD}} cTower(HTOmniSD) < 2 \times cAntenna(PP)|X| + cTower(h(v)) + \sum_{u \in V_{OmniSD}} cTower(h(u)) \wedge \sum_{x \in X} f((v, x)) \le U_{Omni}) \ne true \wedge |X| > 0$ **do**
6.   **if** $\sum_{u \in V_{Omni}} rng_{Omni}(u, u.RAD_{Omni}, v, RAD_{Omni}) = 0$ **then**
7.   **if** $|X \setminus V_{OmniSD}| > 0$ **then**
8.   remove $x \in X \setminus V_{OmniSD}$ s.t. $f(v, x)$ is maximum, from X;
9.   **else**
10.   remove $x \in X$ s.t. $f(v, x)$ is maximum, from X;
11.   **end**
12.   **else**
13.   remove $x \in X$ s.t. $dist(v, x)$ is maximum, from X;
14.   decrease RAD_{Omni} to the vertex $x \in X$ s.t. $dist(v, x)$ is maximum;
15.   **end**
16. $V_{OmniSD} := \{x | x \in X, (v, x) \text{ is the only induced edge from } x\}$;
17. **end**
18. **if** $|X| > 0$ **then**
19. $linkset_{Omni}(v) = X$;
20. **else**
21. $linkset_{Omni}(v) = \{\}$;
22. **end**
23. **return** $linkset_{Omni}(v)$;

capacity. If there exists overlap between the candidate vertex v and any other omnidirectional antenna, from line 13 to 14, we remove the vertex in X with maximum distance to v and then retune the radius RAD_{Omni}. After that, we determine from line 18 to 19, the covered vertices if any set X found to satisfy the constraints. The set of covered vertices is then returned as $linkset_{Omni}(v)$.

D. Power Assignment and Interference Avoidance

In a homogeneous network (i.e., with only p2p links), the antenna transmit power varies according to the link distance. To mitigate the interference between any two p2p links, power selection is possible using, say, the 2P MAC protocol [16] that solves interference issues more efficiently than CSMA/CA for long-distance links. In hybrid network, in addition, our algorithm for MP / Omnidirectional deployment considers the interference constraint by tuning the transmit radius (and beamwidth of p2mp antennas) to the farthest covered vertex. Hence, the transit power of antennas is assigned during the replacement and cause no additional interference.

VI. CONCLUSIONS AND FUTURE WORK

Terrestrial wireless meshes of long-distance links offer an important alternative to address the digital divide for rural areas. Adoption of new technologies for high throughput, low latency modalities, such as optical wireless, can further improve their viability. We have focused on the middle-mile network optimization aspect of their sustainable development. We have proposed the first MNO solution that works in polynomial time. Our solution has an approximation ratio of $O(\ln|A| + \frac{|B|}{|A|} + \frac{|A|+|B|}{\gamma})$ in tower and antenna cost. We also introduced the SteinerTC problem to ease the MNO problem by determining a network topology covered by minimal tower heights that assure line-of-sight links. Our MNO solution applies to hybrid networks, allowing lower-cost hyperlinks to share capacity and eliminate underutilized capacity.

Our future work includes evaluation of our solution in targeted rural areas of southern Ohio in partnership with local wireless network providers. This will include cost-analysis for likely scenarios of network growth that build on capacity flexibility offered by our solution.

REFERENCES

[1] K. Chebrolu, B. Raman, and S. Sen, "Long-distance 802.11 b links: performance measurements and experience," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, 2006, pp. 74–85.
[2] Airjaldi networks. [Online]. Available: https://airjaldi.com
[3] K. Matthee, G. Mweemba, A. V. Pais, G. Van Stam, and M. Rijken, "Bringing internet connectivity to rural zambia using a collaborative approach," in *2007 International Conference on Information and Communication Technologies and Development*. IEEE, 2007, pp. 1–12.
[4] Microsoft airband initiative. [Online]. Available: https://www.microsoft.com/en-us/corporate-responsibility/airband
[5] Mircosoft partnerships. [Online]. Available: http://whitespaces.microsoftspectrum.com
[6] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "Ieee 802.22: the first worldwide wireless standard based on cognitive radios," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*. IEEE, 2005, pp. 328–337.
[7] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell, "Ieee 802.22: The first cognitive radio wireless regional area network standard," *IEEE Communications magazine*, vol. 47, no. 1, pp. 130–138, 2009.
[8] Spokes: Medium: Midwest: Collaborative: Community-driven data engineering for substance abuse prevention in the rural midwest. [Online]. Available: https://nsf.gov/awardsearch/showAward?AWD_ID=1761969
[9] B. Aoun, R. Boutaba, Y. Iraqi, and G. Kenward, "Gateway placement optimization in wireless mesh networks with qos constraints," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2127–2136, 2006.
[10] S. Sen and B. Raman, "Long distance wireless mesh network planning: problem formulation and solution," in *Proceedings of the 16th international conference on World Wide Web*. ACM, 2007, pp. 893–902.
[11] D. Panigrahi, P. Dutta, S. Jaiswal, K. Naidu, and R. Rastogi, "Minimum cost topology construction for rural wireless mesh networks," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 771–779.
[12] R. Hassin, R. Ravi, and F. S. Salman, "Approximation algorithms for a capacitated network design problem," *Algorithmica*, vol. 38, no. 3, pp. 417–431, 2004.
[13] F. S. Salman, J. Cheriyan, R. Ravi, and S. Subramanian, "Approximating the single-sink link-installation problem in network design," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 595–610, 2001.
[14] H. Takahashi, "An approximate solution for the steiner problem in graphs," *Math. Japonica.*, vol. 6, pp. 573–577, 1990.
[15] R. Cohen and L. Katzir, "The generalized maximum coverage problem," *Information Processing Letters*, vol. 108, no. 1, pp. 15–22, 2008.
[16] B. Raman and K. Chebrolu, "Design and evaluation of a new mac protocol for long-distance 802.11 mesh networks," in *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking*, 2005, pp. 156–169.
[17] P. N. Klein and R. Ravi, "A nearly best-possible approximation algorithm for node-weighted steiner trees," *J. Algorithms*, vol. 19, no. 1, pp. 104–115, 1995.

Lemma 1. *Given a height increment δ at v, STAR-SteinerTC-ALGO returns height increments of v's logical neighbors such that the cost-to-benefit ratio of v is minimized.*

Proof. Let $incr_{OPT}$ be the increment function with height increment δ at v and the height increments of some of the logical neighbors of v. Let J be the set of chosen logical neighbors of v in $incr_{OPT}$ such that the cost-to-benefit ratio at v is optimal. Note that the benefit of $incr_{OPT}$ at v is $|J|$ that minimized the cost-to-benefit ratio. If there exists more than one vertex in J in the same component D_i, the benefit of $incr_{OPT}$ at v is less than J. Then we can remove one vertex in $J \cap D_i$ to obtain a lower cost-to-benefit ratio at v. This contradicts the definition of $incr_{OPT}$.

Let $incr_{ALGO}$ be the increment function returned by STAR-SteinerTC-ALGO and K be the set of chosen logical neighbors of v in $incr_{ALGO}$. The benefit of $incr_{OPT}$ at v is $|K|$ since line 16 in Algorithm 1 remains only one vertex in each component. Now we will compare the selection of J and K. First, suppose $|K|$ is equal to $|J|$. Note that, for each component having at least one logical neighbor of v, STAR-SteinerTC-ALGO only remains one vertex with lowest cost increment. $incr_{ALGO}$ also considers the same set of vertices, denoted as L, because choosing another vertex in the component makes higher cost increment but identical benefit. Since the for loop from line 19 to 24 chooses first $|K|$ elements in L with ascending order of cost increment, the sum of cost increments with size $|K|$ is minimized. Thus, $incr_{ALGO}$, equal to $incr_{OPT}$, is optimal.

Second, suppose $|K|$ is less than $|J|$. Since K is the first $|K|$ elements in L and K is the set with minimal cost increment in size $|K|$, $J \cap K = K$ and $J \setminus K$ contains at least one vertex u not in the components of vertices in K. Note that vertex u is the one with minimal cost increment in its component, thus u is in L. However, the for loop from line 19 to 24 in Algorithm 1 considers all sizes from 1 to $|L|$ to determine the lowest cost-to-benefit ratio. u is selected in $incr_{ALGO}$ if and only if u and the vertices with lower cost increments than u in L make lowest cost-to-benefit ratio. This conflicts with the assumption. Hence, $|K|$ must be identical to $|J|$. According to the first statement, $incr_{ALGO}$ is optimal.

Finally, suppose $|K|$ is larger than $|J|$. Since $J \cap K = J$, we can obtain a lower cost-to-benefit ratio by removing $L[|J| + 1], \ldots, L[|K|]$ from K. However, the set $L[1 \ldots |J|]$ is considered by STAR-SteinerTC-ALGO and chosen as $incr_{ALGO}$ if and only if $L[1 \ldots |J|]$ makes lowest cost-to-benefit ratio. This conflicted with the assumption. Hence, $|K|$ must be identical to $|J|$ and thus $incr_{ALGO}$ is optimal. $\square$

Lemma 2. *The cost-to-benefit ratio of the height increment selected in each iteration of SteinerTC-ALGO is at most twice the cost-to-benefit ratio of any height increment centered at v.*

Proof. Since SteinerTC-ALGO remains the same as TC-ALGO and STAR-SteinerTC-ALGO returns local optimum, it also achieves a 2-approximation of the optimal height increment centered at v as TC-ALGO does, which is proved in [11]. $\square$

Definition 1. *A spider is a tree containing at least two leaves and having disjoint paths from the root to all leaves. There is at most one vertex with degree larger than 2 in a spider. A root of a spider is a vertex with edge-disjoint paths to all leaves of the spider. The root is unique if and only if a spider contains more than two leaves.*

Definition 2. *Let $G = (V, E)$ be an undirected connected graph and let S be a subset of V. A spider decomposition $SD(G, S)$ is a set of vertex-disjoint spiders in G such that the union of leaves of spiders in $SD(G, S)$ contains all vertices in S.*

Lemma 3. *[17]. Let $G = (V, E)$ be an undirected connected graph and let S be a subset of V, where $|S| \geq 2$. There exist a spider decomposition $SD(G, S)$ in G.*

Lemma 4. *Consider iteration i in the greedy algorithm SteinerTC-ALGO. Let the height function at the beginning of iteration i be h_{i-1}. Let OPT be the cost of optimal height function h_{OPT}. The ratio of $\frac{cost}{benefit-1}$ associated with the height increment selected by iteration i is at most $\frac{2OPT}{|cover(h_{i-1})|}$.*

Proof. Let c_i be the cost increment and b_i be the benefit associated with the height increment determined in iteration i. Note that edges in $cover(h_{OPT})$ associated with the corresponding height increments can be used to connect $|cover(h_{i-1})|$ components into a tree. Let $cover(h_{i-1})$ have ϕ_{i-1} components $D_1, \ldots, D_{\phi_{i-1}}$. Let T_i be the graph obtained from $cover(h_{OPT})$ by contracting $D_j \cap cover(h_{OPT}), j = 1, \ldots, \phi_{i-1}$ to a supervertex. That is to say, starting with $cover(h_{OPT})$, for each component D_j, $D_j \cap cover(h_{OPT})$ is merged into a supervertex. All the supervertices will be connected in one single component after the final iteration is done. Note that the vertices in $D_j \setminus cover(h_{OPT})$, that are non-terminals, have already connected by the supervertex contracted from $D_j \cap cover(h_{OPT})$. Connecting all supervertices yields a Steiner tree spanning all terminals.

Let S denote the set of the ϕ_{i-1} supervertices. According to Lemma 3, T_i contains a spider decomposition $SD(T_i, S)$. Each spider in $SD(T_i, S)$ is associated with height increments (derived from h_{OPT}) at its leaves (i.e., supervertices). The cost increment from $cover(h_{i-1})$ to T_i is at most OPT. Let $r_1, \ldots, r_w$ be the roots of the spiders in $SD(T_i, S)$ and $m_1, \ldots, m_w$ be the number of supervertices in S in each of these spiders. Let sc_j denote the cost increment to form the spider with root r_j, where $m_j \geq 2$. Note that a spider with root r_j spans a subset of components $D_1, \ldots, D_{\phi_{i-1}}$, which is the m_j components corresponding to m_j supervertices in this spider. Accordingly, the cost-to-benefit ratio at r_j is $\frac{sc_j}{m_j-1}$. Note also that the cost-to-benefit ratio at r_j, a non-terminal

with fixed height, can be considered the cost-to-benefit ratio at a terminal v in the supervertices of the same spider, that is induced by an edge (v, u), $u \in B$, in the spider associated with the height increment $h^+(v)$. Therefore, the cost-to-benefit ratio at v in the spider is at least the cost-to-benefit ratio determined by STAR-SteinerTC-ALGO in a consideration of the same terminal v and height increment $h^+(v)$.

Since SteinerTC-ALGO chooses a vertex with lowest cost-to-benefit ratio in each iteration, for each spider in $SD(T_i, S)$, $\frac{c_i}{b_i} \le \frac{sc_j}{m_j-1}, j = 1, \ldots, w$. Moreover, $\frac{c_i}{b_i}$ is at most the average cost-to-benefit ratio over spiders in $SD(T_i, S)$, thus we have:

$$\frac{c_i}{b_i} \le \min_j \frac{sc_j}{m_j-1} \le \frac{\sum_{j=1}^{w} sc_j}{\sum_{j=1}^{w} (m_j-1)}$$
$$\implies \frac{c_i}{b_i+1} \le \min_j \frac{sc_j}{m_j} \le \frac{\sum_{j=1}^{w} sc_j}{\sum_{j=1}^{w} m_j}$$
$$\implies \frac{c_i}{b_i+1} \sum_{j=1}^{w} m_j \le \sum_{j=1}^{w} sc_j \le OPT$$

Due to $\sum_{j=1}^{w} m_j = \phi_{i-1}$,

$$\frac{c_i}{b_i+1} \le \frac{OPT}{\phi_{i-1}} = \frac{OPT}{|cover(h_{i-1})|} \tag{6}$$

Note that $c_i \le sc_j$. According to Lemma 2, a single iteration of the greedy algorithm SteinerTC-ALGO finds a 2-approximation of the optimal height increment, the minimum ratio of $\frac{cost}{benefit-1}$ for the height increment selected by SteinerTC-ALGO is thus at most $2 \times \frac{OPT}{|cover(h_{i-1})|} = \frac{2OPT}{|cover(h_{i-1})|}$. □

From Lemma 4, the approximation ratio of $2 \ln |A| OPT$ in Theorem 2 is shown as follows. According to the definition of benefit, $\phi_i = \phi_{i-1} - b_i$ in iteration i. Substituting Eq. 6 into this equation, we obtain

$$\phi_i = \phi_{i-1} - b_i \le \phi_{i-1} - \frac{b_i+1}{2} \le \phi_{i-1}\left(1 - \frac{c_i}{2OPT}\right)$$

Let the total number of iterations taken in SteinerTC-ALGO be p and $\phi_p = 1$. We have

$$\phi_p = \phi_0 \prod_{j=1}^{p} \left(1 - \frac{c_j}{2OPT}\right)$$
$$\implies 0 = \ln \frac{\phi_0}{\phi_p} + \ln \prod_{j=1}^{p} \left(1 - \frac{c_j}{2OPT}\right)$$

By using $\ln (1 + x) \le \ln x$, we get

$$0 = \ln \frac{\phi_0}{\phi_p} + \ln \prod_{j=1}^{p} \left(1 - \frac{c_j}{2OPT}\right)$$
$$\le \ln \frac{\phi_0}{\phi_p} + \sum_{j=1}^{p} \frac{-c_j}{2OPT}$$
$$\implies \frac{1}{2OPT} \sum_{j=1}^{p} c_j \le \ln \frac{\phi_0}{\phi_p}$$
$$\implies \sum_{j=1}^{p} c_j \le 2 \ln |A| OPT$$

This proves that the total cost for the height function h returned by SteinerTC-ALGO is at most $2 \ln |A| OPT$.