Next-Generation Wireless Backhaul Design for Rural Areas

Gabriele Gemmi
Northeastern University
g.gemmi@northeastern.edu

Llorenç Cerdà-Alabern
Universitat Politècnica de Catalunya
llorenc.cerda@upc.edu

Leonardo Maccari
Università Ca' Foscari Venezia
leonardo.maccari@unive.it

***Abstract*—Rural areas often face significant challenges in accessing reliable broadband Internet due to high infrastructure costs and low population density. To address this issue, we propose a model for evaluating the performance and the cost of a mesh-based, last- and middle-mile replacement for broadband connection in these underserved regions. We use open data from ten underserved municipalities to assess the demand, plan the mesh network, and estimate the allocated capacity per user. We consider two designs: a low-cost network using the classical 5 GHz unlicensed band, and a high-performance one using mmWave frequency. For both designs, we estimate the Operating Expenditure and the amortized Capital Expenditure using realistic device prices and operating cost estimations. We compare the price of the mesh-based solution with alternatives based on xDSL and satellite connectivity and show that it has competitive prices compared to existing offers, covering a larger portion of households than DSL. We open-source both the code and the elaborated data to reproduce, extend, and improve our results in different settings.**

***Index Terms*—WISP, WBN, Digital Divide, Wireless Backhaul, Resiliency, Economic Modeling**

I. Introduction

Access to broadband Internet is crucial in rural areas, where the need is often even greater than in urban centers due to the general lack of services. However, low population density significantly reduces the economic feasibility of telecommunications providers, delaying the deployment of fiber connections. A key challenge is the high cost of last- and middle-mile infrastructure required to connect homes to a fiber-enabled Point of Presence (PoP). In response, grassroots initiatives have emerged to bridge this gap, often relying on low-cost wireless equipment and community-driven efforts to establish connectivity. These initiatives include non-profit community networks directly operated by users [1], as well as Wireless Internet Service Providers (WISPs) [2] working with reduced profit margins.

In this paper, we investigate the deployment of networks by WISPs in rural areas, setting out a detailed economic model including the cost of the fixed initial assets, the so-called Capital Expenditure (CapEx), and the recurring expenses required to sustain the network, the Operating Expenditure (OpEx). From a fiber-connected location, we plan the deployment of aerial fiber (passing on power-lines) to a set of gateway nodes. From each gateway, we design a multi-hop Wireless Backhaul Network (WBN) made of relay nodes that bring connectivity close to the users, and finally, each user connects wirelessly to any relay. We consider two designs, the first uses the bands around 5 GHz with 802.11ac devices, and the second uses the bands around 70 GHz with 802.11ad devices. The first is a mature technology with a large market of relatively low-cost devices, it supports longer links, and has more separate channels compared to the second. With it, we design a network of minimal cost and minimal guaranteed performance. The second design has higher costs and higher performance and is more likely to be future-proof.

Our methodology uses open data from public administrations to plan the network and estimate the demand for connectivity in rural areas, which offers realistic constraints for network planning. We apply our model to ten digitally divided municipalities in central Italy, exploiting the availability of geographical and demographic open data in public repositories, and we compare the estimated cost with that of available connectivity options in these areas.

To the best of our knowledge, this is the first paper proposing:

- a methodology to design a reliable backhaul topology and node placement using geographical data;
- a detailed techno-economic model for network deployment in rural areas with a mixed wireless/wired approach;
- a comparison showing that in a real-world situation, the prices of a mesh backhaul are competitive with the available connectivity options.

For the sake of reproducibility and extension, all our data and source code are freely accessible[1].

II. Related Works

Since low-cost wireless equipment was made available at the beginning of the 2000s, wireless community networks [3] and WISPs [4] proliferated. Rapidly, these types of networks attracted the attention of the research community studying various aspects related to routing, scalability, security, measurements, testbeds, topologies, performance, usage patterns, evolution, and mobility. See e.g. the work of Neumann and Ben-David [5], [2] and the references therein for wireless community networks and WISPs, respectively. These technologies were considered good candidates to provide Internet access to developing and rural regions.

A hot research topic and a promising technology for closing the digital divide is given by low-orbit satellites. They have been considered as a backhaul option for local networks [6], however, their global performance, cost-effectiveness, and

[1] https://github.com/UniVe-NeDS-Lab/ODCM

manageability are still to be studied, especially in underserved areas with low income. In the near future, the integration of satellite networks and community networks is a promising option, and it is already part of the research agenda [7]. We will include satellite backhaul in our analysis.

A. *Techno-economic Analysis*

Some WISP techno-economic analysis related to ours can be found in the literature. Maccari et al. [8] propose a planning tool using open source resources about the physical terrain to investigate the economic feasibility of the growth of a wireless community network. Cameron et al. [9] focus on WISP with a backhaul powered by solar or wind energy systems and formulate an optimisation model to minimize energy consumption. In [10], the authors explain their experience deploying a wireless network on a previously-unconnected region in rural Northern California. The authors describe the challenges and give some planning guidelines. For instance, the authors explain that the network design must include site topography and line-of-sight. As in our study, the authors use land datasets for this purpose. In [11], the design and installation of a WISP in mountainous areas in Pakistan is described. In this paper, the authors focus on the technical details such as the location and height of the antennas, as well as the equipment installed on each tower. In [12], the authors use data from a real deployment in a project to provide Internet in a district of Kerala in India. The authors use a simple CapEx/OpEx model to investigate the lowest deployment costs using different technologies (WiFi, WiMax, CDMA450, and WipLL) and conclude that WiFi incurs the lowest costs for backhaul and CDMA450 for access. In [13], the authors develop a network planning tool for a WBN with mmWave IEEE 802.11ay devices. The tool is used to simulate the deployment of a WBN connecting 100 randomly selected houses in an urban and a rural area in Belgium. As in our work, the authors use realistic topographic data describing the location of buildings and streets. This information is included in a graph, and an optimisation problem with integer programming (MIP) is formulated to determine the optimal connections and the role of the nodes. Compared to our solution, the MIP solution in [13] does not scale to thousands of nodes. Moreover, there is no economic analysis. Finally, Prieto-Egido et al. [11] describe a techno-economic deployment of 3G cells with multi-hop WiFi backhaul to connect them to the operator's core network in 6 communities in the Peruvian Amazon. However, these studies differ from ours in several aspects. Some are very specific to certain deployments, and others use simple economic models that do not consider user demand, detailed CapEx/OpEx aspects, or network backhaul design.

In the context of 5G/6G cellular networks, there is a large number of works dealing with economics and topology planning [14]. Oughton et al. [15] propose an optimal network planning and cost assessment tool for 5G networks, and Yaghoubi et al. [16] formulate an optimization model for the backhaul design maximizing reliability for a channel model that includes rain attenuation. However, 5G has a specific focus on increasing user performance in dense areas, but no provision for underserved rural areas [17]. Some effort in this direction is ongoing for 6G [18], but it is far from being a viable present solution. In this regard Saarnisaari et al. [19] analyze the challenges of 6G development for remote area connectivity.

In a previous contribution [20], we introduced and evaluated the first version of our model. Here we extend it in two ways, first, we improve the previous model using a more realistic channel model, an improved centrality metric for topology planning, thus, all the results we present were obtained with a better-performing configuration than the conference paper. Second, we introduce a new design using mmWave links to have a higher-priced and higher-capacity option.

B. *On the Definition of Broadband*

Surprisingly, there is no accepted technical definition of a "broadband" service. The European Union policy goals in 2018 mention "*basic broadband*" as up to 30 Mb/s and *ultra-fast broadband*" as over 100 Mb/s [21]. However, there is no reference to the fact that this number must be considered as a guaranteed minimum or a target maximum speed per subscriber. Traditionally, most commercial offers provide a peak value subject to contention among subscribers sharing the same last-mile connection. Real contention ratios are normally not disclosed, and the guaranteed minimum is often a very low number that is used to assess the mere presence of a DSL service. The only official reference we can rely on is a 2018 document by the UK regulator that sets a minimum for the contention ratio to 50:1 [22]. Due to this ambiguity, we always refer to the "minimum capacity" as the effective capacity that is available for the user at peak time. We compare this with the "advertised capacity" offered by the operators, which should be scaled by an unknown factor. We target four profiles:

- Profile 1 (P1), minimal service/lowest cost: minimum capacity 5 Mb/s, 802.11ac
- Profile 2 (P2), base service, low cost, limited upgrade possibility: minimum capacity 20 Mb/s, 802.11ac
- Profile 3 (P3), advanced service, high cost, high upgrade possibility: minimum capacity 40 Mb/s, 802.11ad
- Profile 4 (P4), high-end service, high cost, limited upgrade possibility: minimum capacity 80 Mb/s, 802.11ad

III. Designing and Evaluating a Wireless Backbone

The network architecture we consider is depicted in Fig. 1a, and comprises separate clusters. In each cluster, we have one gateway connected to the global Internet via optical fiber, and a set of buildings where customers reside. The buildings are connected using wireless links and form a mesh network. Our goal is to design a realistic network with a repeatable and sound process, then verify its per-subscriber cost and compare it with the other available options.

Our methodology follows four phases, the first is the technical characterization of the building blocks: nodes and links. The second is the estimation of the demand of the population in the area, using geographical and societal open data. The third is backhaul planning using graph theory, and the last one is cost

Symb.	Description	Symb.	Description
$\mathcal{V}$	Set of buildings (nodes)	u	Total num. of subscribers
$\mathcal{E}$	Potential links in $\mathcal{V}$	$\mathcal{B} \subset \mathcal{V}$	Set of gateway nodes
$\mathcal{R} \subset \mathcal{V}$	Set of relay nodes	$\mathcal{L} \subset \mathcal{V}$	Set of leaf nodes
$n \in \mathcal{V}$	A building (node)	$s(n)$	Num. of subscribers in n
$\delta(n)$	Degree of node n	k	Num. of clusters
$\mathcal{V}_i$	Set of nodes in cluster i	$\mathcal{E}_i$	Set of links in cluster i
$r \in \mathcal{R}$	Relay node	$g \in \mathcal{B}$	Gateway node
$\phi(r)$	Angle covered by r	$sp(r)$	Num. of shortest paths to r
c_{sub}	Min. guaranteed cap.	$N(n)$	Num. of dev. in $n \in \{\mathcal{B} \cup \mathcal{R}\}$
c_s	Cluster size	s_r	Fraction of served households

(a) Topology model variables.

Symb.	Description	Symb.	Description
C	Total CapEx	O	Yearly OpEx
S_c	Monthly cost per subs.	$C_f(\mathcal{B})$	CapEx fiber cost of $\mathcal{B}$
C_g	CapEx of $g \in \mathcal{B}$	C_r	CapEx of $r \in \mathcal{R}$
O_w	OpEx IXP leasing cost	O_t	OpEx transit cost
O_m	OpEx maintenance cost	O_p	OpEx powering cost

(b) Cost model variables.

Symb.	Values		Description
	ac	ad	
β	120°	90°	Beamwidth of relay antenna
λ	360	1200	Channel capacity (Mb/s) at Median MCS
η	0.84	0.84	802.11ac MAC efficiency [23]
ρ	30	55	Maximum EIRP (dBm) due to regulations
f	5.8	70	Transmission Center Frequency (GHz)
γ_r	19	20	Relay antenna gain (dBi)
γ_l	27	46	Leaf antenna gain (dBi)
Γ	40	1080	Channel width (MHz)

(c) Parameters used in the experiments.

Symb.	Value	Description	Source
d_l	300 €	Deploy. cost of a leaf node	*
a_l	60-100 €	Cost range for an 11ac leaf radio	[24]
	180-300 €	Cost range for an 11ad leaf radio	[24]
d_r	1000 €	Deploy. cost of a relay node	*
a_p	120-200 €	Cost range for an 11ac PtMP radio	[24]
	276-461 €	Cost range for an 11ad PtMP radio	[24]
a_r	300-500 €	Cost range for a relay router	[25]
d_g	10 000 €	Deploy. cost of a gateway node	*
a_g	5000 €	Cost range for a gateway router	[25]
d_f	8946-18 512 €	Deploy. cost of aerial fiber (per km)	[26]

* Values obtained by interviews, similar to those in [27].

(d) CapEx costs. Ranges are explained in Appendix C

Symb.	Value	Description	Source
m_u	200 €/h	Unplanned maintenance cost	*
m_p	50 €/h	Planned maintenance cost	*
mttf_r	22.8 y	Mean time to failure of a router	[28] [25]
mttf_a	11.4 y	Mean time to failure of a radio	[24]
mttr_r	2 h	Mean time to repair of a router	[28]
mttr_a	4 h	Mean time to repair of a radio	*
o_w	1260-1680 €y	1 Gb/s wholesale at the local IXP	[27]
o_t^{10}	27 456 €/y	10 Gb/s of transport to local IXP	[29]
o_t^{100}	55 200 €/y	100 Gb/s of transport to local IXP	[29]
p_c	0.18-0.33 €	Cost range for a kWh of Electricity	[30]
p_g	200 W	Power consump. of a gateway router	[25]
p_r	20 W	Power consump. of a relay router	[24]
p_d	25 W	Power consump. of a wireless device	[24]

* Values obtained by interviews.

(e) OpEx costs. Ranges are explained in Appendix C

TABLE I: Notation table.

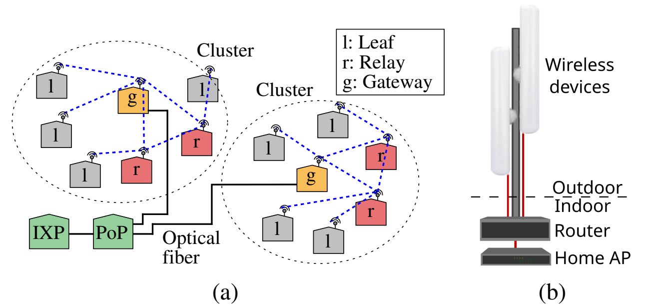

Fig. 1: a) Example topology made of two clusters. b) Graphical depiction of a relay node with two radios and one router.

estimation. In this section, we describe the first two steps, while in Sect. IV and Sect. V we will give an in-depth description of steps three and four. Some algorithms needed to reproduce the results are in the appendix, not to break the readability of the paper. In Tab. I we have grouped the symbols used in the paper to make them easier to find.

A. Modeling Network Nodes and Links

A network node is depicted in Fig. 1b and is made of an indoor and an outdoor part. Outdoors, there is one pole on which wireless devices (*devices*, from now on) are mounted, these are ISP-grade integrated radio and antenna devices that create point-to-point or point-to-multipoint links. This configuration has been used in real mesh networks made of hundreds of nodes studied in the literature [8], [31]. The router takes care of packet routing with some standard routing protocol, and inside the house of the subscriber there is a simple 802.11 Access Point. If only one wireless device is on the roof, the router is not present. When assessing the presence of line-of-sight for the link, we assume a 2-meter pole on the building roof.

We analyzed the available market products for Point to MultiPoint (PtMP) long-range links in the 802.11ac and 802.11ad standards and restricted our choice to two products of two well-known manufacturers, Mikrotik and Ubiquiti. We consider two kinds of devices: a sectorial antenna with beamwidth β, gain γ_r, and cost a_p for the relay nodes; and a more directive antenna with gain γ_l and cost a_l for the leaf nodes. For both devices, the maximum transmission power (including the antenna gain) has been set to ρ according to local regulations. The values extracted from the datasheets are in Tab. Ic, together with other parameters we need and found in the literature.

Given a link of a certain length, we compute the received signal power applying two different path loss models, one that was proposed for point-to-point suburban 802.11ac links at 5 GHz [32] and another for the mmWave spectrum [33]. For the mmWave spectrum, we evaluated 60 GHz in which Gas Absorption Loss (GAL) is present, and 70 GHz, in which GAL is absent. We then extracted from the datasheets [24] and the available online planning tools[2] the mapping between the received signal power and the Modulation and Coding Scheme

[2]https://ispdesign.ui.com/

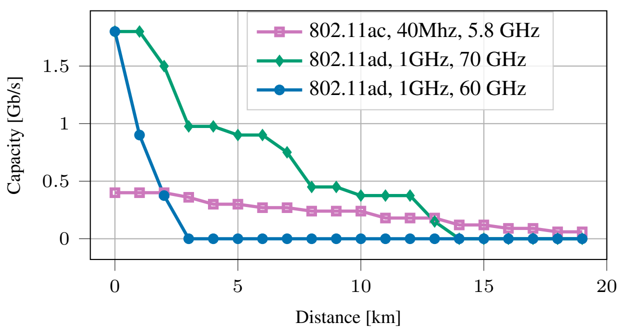

Fig. 2: Link capacity estimated on the data of real devices using different technologies and channel models.

(MCS), which we finally use for link capacity estimation. Fig. 2 shows the capacity versus the link length using the three frequencies.

As long as we operate outside of the gas absorption frequencies, we obtain performing links up to 10 km. At 60 GHz instead, 802.11ad becomes unusable for a distance larger than 2 km. Considering that in our visibility graphs we have as little as 1 % of the Line of Sight (LoS) links longer than 10 km but 36 % longer than 2 km, we take two design decisions: we use the mmWave bandwidth around 70 GHz and we don't use links that are longer than 10 km. This means we can use up to 12 independent channels for 802.11ac but only 4 for 802.11ad, and given the chosen frequencies, we consider only links in LoS.

B. Demand Estimation

A demand model is the number and the positions of subscribers spread around a region. Given a target number u of subscribers, we need to first enumerate all the buildings in the area and then assign a probability w_i to building i of having a subscriber. We then perform random extractions and repeat Monte Carlo simulations with statistical robustness. Here we provide an overview of the process; details are in Appendix A.

We consider ten rural municipalities in central Italy that have been chosen among those for which the morphological data were available and for their degree of digital division. In fact, for all the areas, the average maximum advertised download speed of traditional (xDSL) broadband connectivity was below 30 Mb/s, which normally corresponds to a much lower average speed. For each of the geographical regions, we retrieved the population and household census from the Italian Institute of Statistics (ISTAT). The municipalities have on average an area of 83 km^2, 1558 households, and 3110 buildings. ISTAT provides the number of households and inhabitants down to each census section, that is a polygon of variable dimension. In the areas under analysis, a census has an average surface of 4.52 km^2 and contains 70 buildings. Based on this data-set, and the position and size of the buildings extracted from OpenStreetMap (OSM), we assign to each building a probability w_i of having a household that could be a subscriber of the WISP. We call the subscriber ratio (s_r) the fraction of subscribers over the total households. This is a parameter that we set to 25%, 50%, and 100%, and we set the number u of subscribers consequently. Then we perform random extraction with reinsertion over the buildings to obtain the set of served buildings $\mathcal{V}$. Note that the same building n can be extracted multiple times, so we define a weight function $s(n)$ that is the number of subscribers in building $n \in \mathcal{V}$, so that $u = \sum_{n \in \mathcal{V}} s(n)$.

IV. Network Backhaul Design

Given $\mathcal{V}$ and the weight $s(n)$ $\forall n \in \mathcal{V}$, we can begin the network design. As an initial step, we need to assess the presence of LoS between any couple of buildings in $\mathcal{V}$. For this goal, we exploit the availability of open data for the terrain and the building shapes, and efficient GPU-based ray-tracing techniques. We use the same process used in past works [34] [35] to create a *visibility graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ made of all the nodes and all the potential links. For scalability reasons, we can not connect every node to a single gateway, we need to design a clustered network with one gateway per cluster.

As shown in Fig. 1a, we then have three different types of nodes. The first type is the gateway node, which is fiber-connected to some ISP core network. We call $\mathcal{B}$ the set of gateway nodes. The second type is the relay node, which is used to create a wireless backhaul that brings connectivity close to the users. Relay nodes route the user traffic, and we call $\mathcal{R}$ their set:

$$\mathcal{R} = \{n \in \mathcal{V} \mid \delta(n) > 1 \wedge n \notin \mathcal{B}\} \qquad (1)$$

where $\delta(n)$ is the degree of node n in the resulting network graph (i.e. the number of neighbors). Relays also provide connectivity to the subscribers that are assigned to the buildings where they are placed. The last type is leaf nodes, which are connected to the backhaul with a single edge. These are defined by the set $\mathcal{L}$:

$$\mathcal{L} = \{n \in \mathcal{V} \mid \delta(n) = 1 \wedge n \notin \mathcal{B}\} \qquad (2)$$

Given $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the wireless backhaul design problem consists of finding a set $\mathcal{B} \subset \mathcal{V}$ of gateway nodes and a set of edges $\mathcal{E}' \subseteq \mathcal{E}$ that interconnect all the nodes in $\mathcal{V}$ to a gateway in $\mathcal{B}$ with a multi-hop path, respecting some given performance indicators. We propose a heuristic *divide-et-impera* approach in five steps, each step is detailed in a dedicated subsection:

A) *Graph Partitioning*: We estimate the maximum number of subscribers per gateway and partition $\mathcal{V}$ into k clusters;
B) *Gateway Selection*: For each cluster, we optimize the position of the gateway;
C) *Distribution Tree Design*: We elaborate a strategy to select only the minimal subset of edges $\mathcal{E}_i^* \subset \mathcal{E}$ that guarantees connectivity;
D) *Graph Augmentation*: We enrich the internal topology of each cluster to make it robust to the failure of one edge;
E) *Assigning Devices*: Given a certain profile, we define the sufficient number of devices for each node to provide the target capacity per user.

Ideally, we could formulate an optimization problem with all the constraints and find an optimal solution. However, this approach can scale up to tens of nodes (like in previous works [13], [16], [36]) while in our setting we have thousands of nodes and tens of thousands of potential edges. As we target a usable solution, the next sub-sections will use state-of-the-art algorithms for solving each step. We customize algorithms to fit our specific problem and the real data we have, and we make data and source code available for further research that will improve and refine the algorithmic aspects.

A. Graph Partitioning

Since the cost of the gateway and its fiber connection is a large part of the CapEx, we need to minimize the number of clusters and try to assign to each cluster the maximum number of subscribers. However, we have a limited number of independent channels that limit the number of devices on the same node, including the gateway, to avoid cross-link interference. The total wireless capacity of the gateway is then a bottleneck, which is the same for each cluster and limits its maximum size.

Thus, given the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and $s(n)$ subscribers per node n, we want to solve the k-way graph partitioning problem:

$$\begin{aligned} &\min_{k} \{\mathcal{V}_1, \cdots \mathcal{V}_k\} \\ &\text{s.t.} \quad \cup_i \mathcal{V}_i = \mathcal{V};\ \mathcal{V}_i \cap \mathcal{V}_j = \emptyset,\ \forall i \neq j \\ &\qquad \textstyle\sum_{n \in \mathcal{V}_i} s(n) \simeq \frac{1}{k} \sum_{n \in \mathcal{V}} s(n),\ i \in \{1, \cdots k\} \end{aligned} \tag{3}$$

The well-known *METIS* graph partitioner [37] is one of the few algorithms that enable us to have clusters of the same (or at least similar) size. It uses the Kernighan–Lin (KL) algorithm [38] with complexity $O(|\mathcal{E}|)$.

After computing all the sets $\mathcal{V}_i$ we extract their associated subgraph $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$, where $\mathcal{E}_i$ is defined as:

$$\mathcal{E}_i = \{(n_j, n_k) \in \mathcal{E} \mid n_j, n_k \in \mathcal{V}_i\} \tag{4}$$

To solve problem given by eq. 3, one could start with a reasonable value of c_s and k, do the partition and later on check the minimum capacity condition. This is the process we adopted and that we discuss in Sect. VI-A to investigate the effect of c_s on the KPIs.

B. Gateway Selection

Given a certain $\mathcal{V}_i$ we need to identify the node that is in the best position for being the gateway of the cluster, minimizing the average hop-count to the gateway. We call $d(n_x, n_y)$ the hop-count from a node n_x to a node n_y in the same cluster, then the gateway g_i for cluster i is the node that maximizes the weighted closeness centrality [39]: the inverse of the sum of the weighted distance to all the other nodes of the network, as follows:

$$g_i = \text{argmax}_{n_x \in \mathcal{V}_i} \frac{1}{\sum_{n_y \in \mathcal{V}_i} s(n_y)\, d(n_x, n_y)}$$

Computing centrality requires computing the shortest path between any couple of nodes, so it has polynomial complexity with the size of $\mathcal{V}_i$, and it takes negligible time with thousands of nodes.

C. Distribution Tree Design

To connect each node of a cluster to its gateway, a subset of the available edges of $\mathcal{G}_i$ is sufficient. Recall that we are using directive antennas and that each node has a limited number of devices, so economic and technical constraints lead us to minimize the number of edges. The minimal edges tree is the Shortest Path Tree (SPT) computed with the classical Dijkstra's algorithm, using the link length as a weight. Actually, path loss is non-linear with the distance but after testing several weights we noticed only minimal differences, so we maintained the linear distance. At the end of this step, we have a set of graphs $\bar{\mathcal{G}}_i = (\mathcal{V}_i, \bar{\mathcal{E}}_i)$ so that the subset of edges $\bar{\mathcal{E}}_i \subseteq \mathcal{E}_i$ creates a SPT allowing each node in $\mathcal{V}_i$ to reach a gateway.

D. Graph Augmentation

The previous step produces a backhaul network in which every node has a path to one gateway using a minimal number of edges (a tree). The outcome is a topology in which the failure of a single link close to the gateway could disconnect large portions of a cluster. To increase the reliability of the network, additional edges must be added. This is a *Graph Augmentation Problem*, which can be formalized as follows. Given $\bar{\mathcal{G}}_i$ we want a set of edges $\mathcal{E}_i^* \subseteq \mathcal{E}_i$ such that:

- $|\mathcal{E}_i^*|$ is minimal and $\bar{\mathcal{E}}_i \subseteq \mathcal{E}_i^*$
- $\mathcal{G}_i^* = (\mathcal{V}_i, \mathcal{E}_i^*)$ is 2-edge-connected

A 2-edge-connected graph tolerates the failure of 1 edge without disconnecting any node. Like other graph combinatorial problems, this problem has been proven NP-hard. For this reason, a heuristic is needed to solve it [40]. This heuristic finds a plausible solution $\mathcal{E}_i^*$ with log-linear complexity.

Augmenting the whole network, however, would be too costly, as it would approximately double the number of edges. Moreover, no commercial Internet Service Provider (ISP) guarantees a fault-tolerant connection to its subscribers. For this reason, we augment only the core of the cluster (the subgraph made of the relays and the gateway) to make it 2-edge-connected. This means that when a device in a relay fails, the leaf nodes connected to that device will be disconnected until the relay is repaired, but all the other nodes in the network will remain connected to the gateway using an alternative path.

E. Assigning Devices to Nodes

Given the topology, we need to assign a number of wireless devices per relay satisfying two constraints: the coverage of the neighbor nodes and the overall capacity required to route the traffic. The exact process is in Algorithm 1, described hereafter.

Let us call $sp(r)$ the number of shortest paths that go from any subscriber to a gateway and pass through relay r. This is computed using the $path\_to\_gateway$ function, which returns a path from the building of a subscriber to the gateway in the cluster, excluding the starting and the arrival node. We call c_{sub} the minimum guaranteed capacity (in Mb/s) per subscriber so that $sp(r) \times c_{sub}$ is the required minimum downstream capacity at r. The capacity of radio links depend on the MCS and on how many edges are shared on the same device. The distribution of the MCS is extremely skewed, with the majority of the links

Algorithm 1: Assign devices to relays

Input: graph G; subscribers $\mathcal{S}$; relays $\mathcal{R}$; gateways $\mathcal{B}$
Output: Per-node radios $N(v)$

1 **foreach** $s \in \mathcal{S}$ **do**
2   path ← path_to_gateway(s);
3   **foreach** $r \in$ *path* **do**
4         $sp(r) \mathrel{+}= 1$
5 **foreach** $r \in \mathcal{R} \cup \mathcal{B}$ **do**
6     $n_c \leftarrow \left\lceil \frac{sp(r)\, c_{\text{sub}}}{\lambda} \right\rceil$; $n_\phi \leftarrow \left\lceil \frac{\phi(r)}{\beta} \right\rceil$;
7   **if** $r \in \mathcal{B}$ **then** $N(r) \leftarrow \max(n_c, n_\phi)$;
8   **else** $N(r) \leftarrow \max(1 + n_c, n_\phi)$;
9 **return** $N(\cdot)$

negotiating the maximum MCS (83 % for 802.11ad and 91 % for 802.11ac). For this reason, we assume that our links have a capacity λ corresponding to the maximum MCS, ad then the minimum number of devices for the needed capacity is $\left\lceil \frac{sp(r)\, c_{\text{sub}}}{\lambda} \right\rceil$. However, if the total angle that must be covered by relay r is $\phi(r)$ then we need at least $\left\lceil \frac{\phi(r)}{\beta} \right\rceil$ devices. In relays the outgoing capacity is guaranteed by a dedicated upstream wireless device, while gateways are fiber connected. This is reflected in line 6-8 of Algorithm 1. Finally, we call $N = \sum_v N(v)$ the total number of devices in the network.

Evaluating KPIs: It must be noted that $N(r)$ is an approximation of the minimum number of required radios given its downstream capacity, but our planning does not include device orientation, so the approximation holds on average. Moreover, $N(r)$ is always capped by the maximum number of independent channels, so Algorithm 1 does not guarantee the capacity per subscriber. Only after we assign devices to nodes will we assign real capacity to links and links to devices, and then we can verify that the minimum capacity for a single subscriber matches c_{sub}. This is done by identifying the bottleneck on the path from the gateway to the subscriber, with a process detailed in Appendix B. We then average the minimum and maximum capacity of all subscribers and obtain $\bar{c}_{min}$ and $\bar{c}_{max}$, which are two key performance indicators of our network.

V. Cost Model

The other key performance indicator is the monthly cost per subscriber, given by the sum of two components, the monthly amortization of the CapEx, and the monthly OpEx:

$$S_c = \underbrace{\frac{1}{u}\left(\frac{C_f(\mathcal{B})}{180} + \frac{C - C_f(\mathcal{B})}{60}\right)}_{\text{Monthly CapEx}} + \underbrace{\frac{1}{u}\left(\frac{O}{12}\right)}_{\text{Monthly OpEx}} \tag{5}$$

where u is the number of subscribers, $C_f(\mathcal{B})$ is the cost of deploying fiber to the set of gateways $\mathcal{B}$ (that we amortize in 180 months, 15 years), O is the yearly OpEx, C is the total CapEx, so $C - C_f(\mathcal{B})$ is the CapEx for the wireless infrastructure (that we amortize in 5 years).

We assume that the fiber infrastructure is amortized over 15 years, as it serves multiple purposes and can be utilized for other business applications. Even in the case that the network business is interrupted for 15 years, it is an asset that can be monetized. In contrast, the entire wireless infrastructure is amortized over 5 years, reflecting the expectation of renewing it within this period. This assumption is based on the shorter lifespan of outdoor devices compared to indoor ones. Even if the mean time to fail is longer than 5 years, assuming a 5-year lifespan allows a complete topology redesign every five years to integrate new technologies and enhance the network. This approach aligns with other studies that consider a 10-year overall return period for both wired and wireless infrastructure, as referenced in Oughton [41]. In our previous work, we adopted a 5-year period, which represents a short return timeframe and results in a high net margin after 60 months. Below, we detail the estimation of each term in Eq. (5). Whenever possible instead of using one price tag, we use a range of prices derived from the literature, details are explained in Appendix C.

A. Estimating the CapEx

We assume that leaf nodes are equipped with a single device, whose cost is a_l while relay nodes and gateways have multiple devices and a router, whose costs are a_r and a_g respectively. Every node has a fixed cost for the physical installation and a home access point. Tab. Id reports the costs, with the source from which we extract the costs, which are data sheets, works in the literature, and interviews with members of the guifi.net community network based in Catalonia, which offers connectivity in rural areas.

Given a gateway $g \in \mathcal{B}$ its cost C_g depends on the number of the radio devices $N(g)$, times the cost of a radio a_p, plus the cost of the router a_g and the deployment cost d_g:

$$C_g = d_g + a_g + a_p \cdot N(g) \tag{6}$$

Similarly to the gateway node, the cost C_r of a relay node $r \in \mathcal{R}$ depends on the number of radio devices $N(r)$ times the cost of a radio a_p plus some fixed costs for the router and the deployment (a_r and d_r):

$$C_r = d_r + a_r + a_p \cdot N(r) \tag{7}$$

The cost of a leaf node is given simply by the sum of the cost of one radio a_l and the physical deployment d_l.

Finally, we estimate the cost for the deployment of aerial fiber to the gateways $C_f(\mathcal{B})$ from the closest PoP of some operator. Since traditional operators are present in these areas, we assume that the PoP is a point p_0 in the center of the municipality. Given the street graph, we compute the Steiner tree connecting all the gateways to the PoP along existing public streets, we sum the length (in km) of every arch of the tree, and we multiply it by the cost per km of aerial fiber d_f. Eq. (8) shows the composition of the CapEx of the network, which is the sum of the cost of gateways, relays, and leaf nodes, plus the cost of the aerial fiber backhaul:

$$C = \sum_{g\in\mathcal{B}} C_g + \sum_{r\in\mathcal{R}} C_r + (d_l + a_l)|\mathcal{L}| + C_f(\mathcal{B}) \tag{8}$$

B. Estimating the OpEx

As shown in Eq. (9), the yearly OpEx is made of four different parts: O_w is the cost of leasing the needed wholesale capacity at the closest Internet Exchange Point (IXP) (see Waites et al. [42] for a description on the role of IXPs in rural connectivity); O_t is the cost of the transit from the closest PoP of some operator to the IXP; O_m is the cost of maintenance of the backhaul; O_p is the cost of powering the whole infrastructure:

$$O = O_w + O_t + O_m + O_p \tag{9}$$

The basic costs we consider for the calculation are reported in Tab. Ie. The total capacity that the WISP needs to contract is given by the minimum guaranteed capacity provisioned to each subscriber (c_{sub}, in Mb/s), times the number of subscribers (u). We consider a yearly price for wholesale connectivity given by o_w (see Tab. Ie) with a minimum unit of 1 Gb/s.

$$O_w = \left\lceil \frac{c_{sub}}{1000} \cdot u \right\rceil \cdot o_w \tag{10}$$

The cost for the transport of the connectivity from the PoP to the regional IXP equals o_t^{10} if the transport is up to 10 Gb/s or o_t^{100} if it is between 10 and 100 Gb/s. We then have:

$$O_t = \begin{cases} o_t^{10} & \text{if } \frac{c_{sub}}{1000} \cdot u < 10\,\text{Gb/s} \\ o_t^{100} & \text{otherwise} \end{cases} \tag{11}$$

To estimate the yearly maintenance cost of the network, we take into account the failures of both routers and radio devices deployed in the wireless backhaul. For both, we have found realistic mean-time-to-failure (mttf) and mean-time-to-repair (mttr) values, which respectively express the average life of a device and the average time needed to repair/replace it after a failure.

With Eq. (12) we calculate the yearly cost of maintenance as the number of devices divided by the mttf (which yields the number of yearly failures) times the cost of the intervention. Note that in Eq. (5) we already assume that every 60 months the wireless devices need to be replaced, so the cost of the failed hardware is already discounted in the CapEx. Eq. (12) then does not repeat the cost of wireless devices. Its main purpose is to account for the human intervention needed to repair a broken node. The equation is made of four terms:

$$\begin{aligned} O_m = {} & \frac{|\mathcal{B}|}{\text{mttf}_r}(\text{mttr}_r \cdot m_u) + \frac{|\mathcal{R}|}{\text{mttf}_r}(\text{mttr}_r \cdot m_u) \\ & + \sum_{r \in \mathcal{R}} N(r) \frac{1}{\text{mttf}_a}(\text{mttr}_a \cdot m_u) + \frac{|\mathcal{L}|}{\text{mttf}_a}(\text{mttr}_a \cdot m_p) \end{aligned} \tag{12}$$

The first term takes into account the failure of gateway routers, which are one per gateway; the following term takes into account the failure of routers on relay nodes; the third term takes into account the failure of radio devices on relays, and the last one considers the failure of leaf nodes. The reason we separate leaf nodes from relay nodes is that most user contracts allow to delay the technical repair until the next working day, while relay nodes need to be repaired as soon as possible as they can impact many users, moreover, even if we have a redundant backhaul network, correlated failures could disconnect large portions of the topology. Thus, we use m_u in Eq. (12) for relays and m_p for the leaf nodes.

To estimate the yearly energy consumption of the network (O_p) we use the maximum instantaneous power consumption values provided by the datasheets of the devices in the gateway, the routers, and the single radios, p_g, p_r and p_d, respectively. We rescale it by a factor of 0.7 as the network is not always at its maximum power, and multiply by the number of Gateways, Relays, and radio devices. We finally multiply by 24 (hours) and 365 (days) to obtain the yearly energy consumption in kWh. We use the average cost of the energy (p_c) obtained by the national regulator to obtain the yearly cost.

$$\begin{aligned} O_p = {} & \left(|\mathcal{B}|\, p_g + |\mathcal{R}|\, p_r + \Big(\sum_{r \in \mathcal{R}} N(r) + |\mathcal{L}| \Big) p_d \right) \\ & \cdot 365 \cdot 24 \cdot 0.7 \cdot p_c \end{aligned} \tag{13}$$

VI. Experiments and Results

In our experiments, we vary the minimum guaranteed capacity per subscriber c_{sub}, the cluster size $c_s \in \{50, 100, 200\}$, and the fraction of served households $s_r \in \{0.25, 0.5, 1\}$. Each combination of parameters has been run 50 times in ten different areas with a different random seed and thus a different set of subscribers. Images report the average, we don't report the 95 % confidence intervals as they are less than 3 % of the measure.

A. Defining the Cluster Size

The only parameter we need to pre-configure is the maximum size of each cluster c_s, which is limited by the maximum capacity we can offer with one single gateway, which is in turn capped by the number of devices we can mount on a gateway. Fig. 3 then shows on the x-axis the expected minimum capacity allocated per subscriber c_{sub} and on the y-axis the maximum number of devices placed on the gateway computed using Algorithm 1 in each run of the simulations (averaged on all the runs).

The first observation is that a lower penetration rate (so a geographically sparser network) does not impact the number of devices on the gateway, as varying s_r the curves remain clustered in groups with very close values. We notice that the curves can have initial horizontal plateau, followed by a linear growth. The plateau is more visible for 802.11ad and happens when the given number of devices already provides excess capacity, so c_s can be increased without requiring more devices. In the linear growth the more capacity we request, the more capacity (and thus the more devices) we need at the gateway. In general, the mesh backhaul is not a bottleneck; the bottleneck is the maximum number of devices we can mount on the gateway (the red line, showing the independent channels). If we consider 802.11ac we notice that we could achieve the maximum target performance (profile P2, $c_{sub} = 20$ Mb/s) with $c_s = 200$ with almost all penetration ratios. However, it would be a borderline choice. From a network management point of view, 12 devices

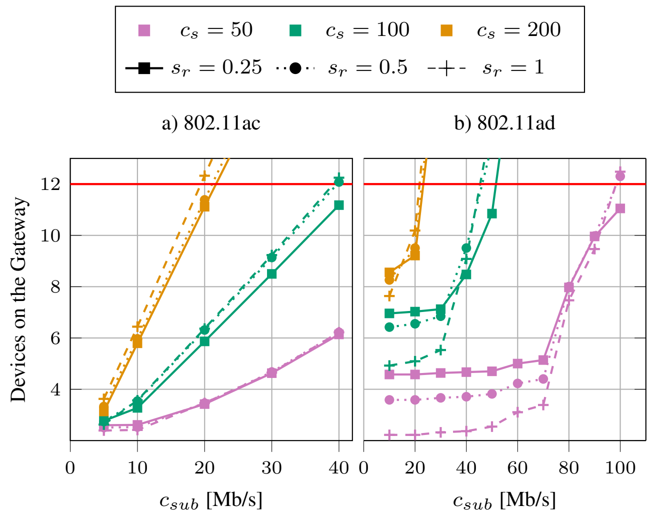

Fig. 3: The maximum number of devices mounted on the gateway, as a function of the expected minimum capacity per subscriber, c_{sub}; cluster size, c_s; penetration ratio, s_r; and technology used: 802.11ac (left), 802.11ad (right). The red line corresponds to the maximum number of independent channels for the standard used.

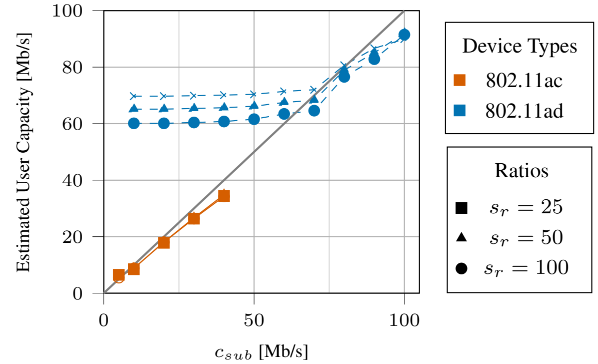

Fig. 4: Effective estimated throughput vs Minimum guaranteed bandwidth for 802.11ac ($c_s = 100$) and 802.11ad ($c_s = 50$).

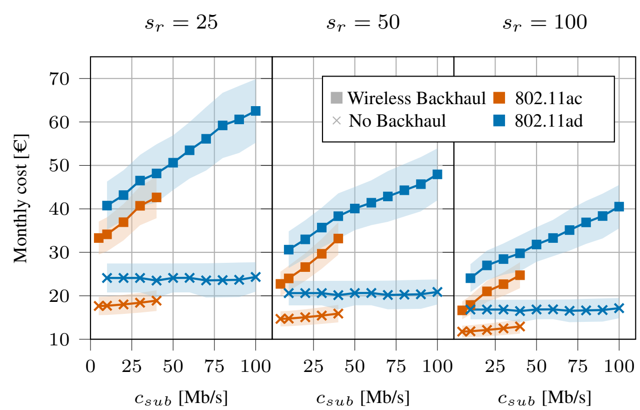

Fig. 5: Monthly cost per subscriber, S_c, for 802.11ac ($c_s = 100$) and 802.11ad ($c_s = 50$).

on a single node are hard to manage because they need physical space, power, a complex configuration, and more devices in cascade (switches, routers, etc.). Instead, for $c_s = 50$ we have excess capacity, and we would increase the total cost. If we consider 802.11ad instead, $c_s = 50$ is the only configuration that satisfies both P3 and P4. From now on, we will restrict our results to these two cluster sizes: $c_s = 100$ for 802.11ac and $c_s = 50$ for 802.11ad.

B. Capacity per Subscriber

Fig. 4 reports the effective average minium capacity per subscriber $\bar{c}_{min}$ estimated as explained in Appendix B as a function of the desired capacity per user c_{sub}. Note that this parameter is measured considering the bottleneck of the backhaul; it is not affected by the capacity leased at the IXP (we always lease enough capacity to sustain c_{sub}). In ideal conditions this would be a straight line on the diagonal of the graph, however, we see that we often have excess capacity, while in P4 (802.11ad, $c_{sub} = 80$ Mb/s) we are slightly below the required values when the penetration rate is lower than 100 %. 802.11ad has an excess minimum capacity of up to 60 Mb/s, so even at P3 the guaranteed performance is much higher than requested.

Overall, Fig. 3 and 4 show that when tested on realistic conditions (real maps, real subscriber sparsity, realistic link capacity estimation), our network design can provide the requested capacity per subscriber, which is our first key performance indicator.

VII. Cost Analysis

This section provides an analysis of the cost of the infrastructure, starting with Fig. 5 that shows the trend of S_c increasing c_{sub}, for the chosen cluster sizes. The image gives several insights that we anticipate here and discuss in depth in the next subsections.

The first is that the monthly cost increases linearly with c_{sub}, which depends mostly on the increase in the OpEx due to the cost of wholesale bandwidth at the IXP and the leasing cost of the fiber to the IXP itself. The CapEx only slightly increases with c_s, as some more devices need to be installed in some cases (as shown in Fig. 4, for 802.11ad we have excess capacity to the gateway up to $c_{sub} = 60$ Mb/s). However, the CapEx is affected by customer penetration: since the geographical distribution does not change, the number of clusters does not decrease with s_r, and it directly affects the cost of the fiber backhaul. Thus, reducing s_r does not reduce the CapEx proportionally, and the same CapEx must be amortized among less subscribers. The third insight is that the customer's monthly prices are extremely competitive both on the lower end and on the higher end.

1) Evaluating the Satellite Option: As we will see in the next section, the cost items that have the largest impact are the cost of fiber deployment and the cost of wholesale capacity. For this reason, we report in Fig. 5 the cost of the mesh infrastructure, excluding the upfront cost to deploy the aerial fiber ($C_f(\mathcal{B})$) and the recurring costs of the Wholesale Capacity (O_w) and Fiber transit to the IXP (O_t). The rationale is that we could use the mesh network for the distribution of capacity that is then provided to the gateway by a satellite link. In fact Starlink also offers backhaul connectivity, the so-called community gateways[3]. The cheapest offer starts at \$75,000/Gb/s/month (completely out of scale compared with data in Tab. Ie) with a one-time upfront cost of \$1,250,000 (larger than the whole CapEx in any combination we consider). As a consequence, the subscriber cost would be much larger for any profile. Cheapest options for a satellite backhaul have been used in related works [6], but with non-LEO satellites and lower capacity. If other alternatives become available, our model can be easily used to make an updated comparison.

A. Cost Breakdown

Fig. 6 helps us analyze the monthly cost of the infrastructure, as defined in Eq. (5). We use monthly amortization for both CapEx (top row) and OpEx (bottom row), so we can compare them and identify how the single cost items impact the total. We also report for each cost item a bar that shows the variation between the minimum and maximum price, and the bar shows the average. The CapEx is due primarily to the deployment costs, which include both the deployment of the wireless nodes and the deployment of the optical fiber. When we increase c_{sub} using the same wireless technology, the growth of the CapEx is extremely small, as the deployment cost is essentially the same, and there is very little increase in the cost for the devices. When comparing 802.11ad to 802.11ac there is a substantial increase in the radio cost, due to the higher prices of the devices, together with an increase of the fiber and deployment cost, which is raised by the smaller size of the clusters.

The figure shows that the amortized OpEx is generally higher than the CapEx, especially when the capacity of the network increases (profiles 3 and 4). This is because the largest portion of the OpEx is due to the cost of the wholesale capacity at the IXP and the cost of leasing the fiber to reach the IXP, which increases at least linearly with the offered capacity. This reflects the difficulty of traditional operators to offer connectivity in rural areas, as the cost of bringing enough capacity close to sparsely populated areas is very high. The third voice is the energy cost, and the unplanned maintenance is instead a marginal operating cost.

The shaded bars in Fig. 5 (min–max for each point) and the whiskers in Fig. 6 (per–item ranges) take into account the price ranges we justify in Appendix C and reveal that the dispersion around the mean is modest. The average dispertion across all the parameters is roughly $\pm 12\%$. Likewise, for the cost-breakdown in Fig. 6 no individual CapEx or OpEx component varies by more than 30% across the extreme assumptions (bulk-purchase discounts, PPP currency conversion, or energy-price swings). This limited spread confirms that the overall conclusions are robust: even under the most pessimistic input values, the wireless backhaul remains cost-competitive, while under favourable conditions it delivers additional head-room for profit.

[3]See https://www.starlinkinternet.info/community-gateway

B. Comparison with Available Offers

We conclude our analysis by showing a price comparison with available commercial offers in the areas. Tab. II shows the comparison of the WBN approach for two values of s_r with xDSL offers [43] and Starlink satellite connectivity [44]. We compare two profiles (P1 and P3) in terms of minimum capacity, maximum capacity, monthly cost, monthly cost per maximum Mb/s achievable, and the fraction of unconnected households.

Speed measurements for the xDSL come from the Italian Regulator on Communications[4] that published sampled maximum speeds for all the areas, while for Starlink, they represent the maximum capacity advertised by Starlink itself. As discussed in Sect. II-B, we don't have a strict definition of the expected performance of commercial offers, so we compare the minimum computed capacity of a WBN with the advertised capacity of commercial offers. The minimum capacity offered by an xDSL is the minimum negotiated link performance, but in practical terms, it could be much lower. Similarly, the satellite Starlink connection has a maximum sponsored capacity, but the description of the offers explicitly states that traffic shaping applies to subscribers based on their contract and the generated traffic. It is impossible to know how much capacity the subscribers effectively have, based on their density per square km.

The comparison shows that with 25 % penetration, a wireless backhaul has prices that are competitive with commercial offers. When penetration reaches 50 %, the costs are significantly lower both in the low range (P1 compared to xDSL) and in the high range (P3 compared to SAT). Also, in terms of maximum capacity per subscriber per Mb/s our solution outperforms the other ones in all scenarios. The difference leaves a high margin to include the interests of a loan (if the WISP can not afford the initial upfront cost) and also profit if the WISP is not a community initiative but a for-profit one. It is also important to note that existing ISPs have been publicly financed in recent years to extend their infrastructure in remote areas, while we assume the WISP takes all the costs and the risk.

Finally, we report the number of households that we could not connect to the visibility graph compared to the number of households that are declared impossible to serve by the telecommunication minister with DSL. The latter is 2 to 3 times larger than the former. Note that our methodology does not assume any practical adjustment (high trellises, or nodes placed on strategic positions to increase coverage) that is applied in the real world to increase the penetration, so we expect that the number of unconnected subscribers can be lowered substantially.

VIII. Conclusions

Mesh networks have been used in the past to provide connectivity to disconnected areas, but there is a lack of evidence of their sustainability. This paper takes a techno-economic approach to evaluate the feasibility of a wireless mesh backhaul

[4]https://maps.agcom.it/

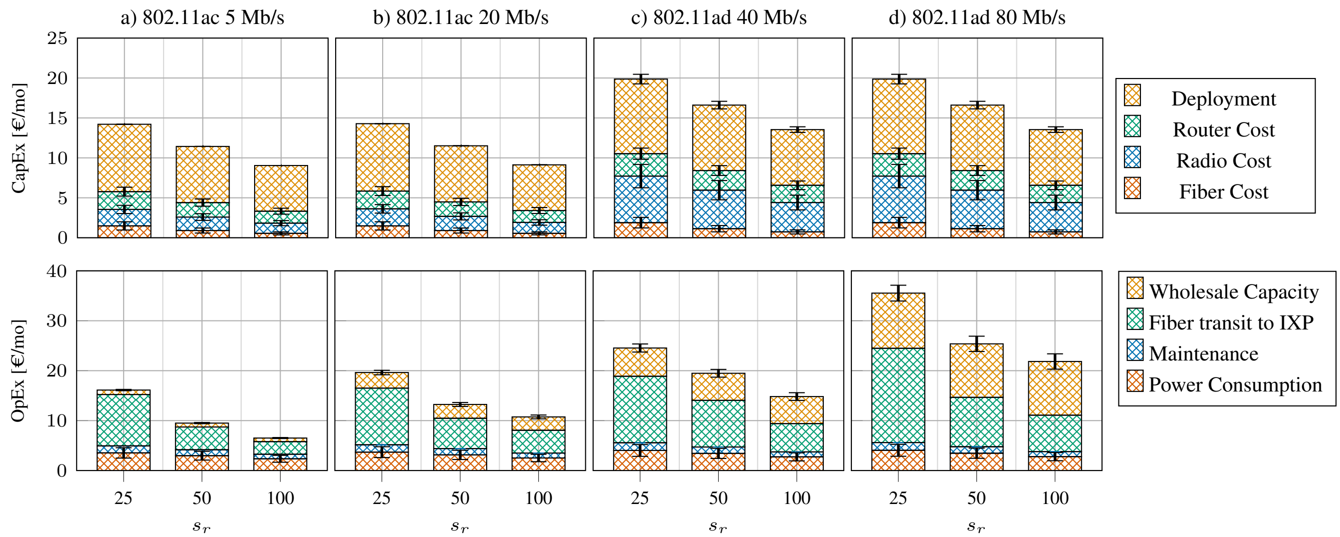

Fig. 6: CapEx (top) and OpEx (bottom) for 4 different configurations.

	WBN/P1		WBN/P3			
	0.25	0.5	0.25	0.5	xDSL	SAT
Upfront cost (€)	524.2 ±30%	372.5 ±29%	857.9 ±29%	641.1 ±28%	480	369
Monthly cost (S_c)	33.3 ±11%	22.7 ±12%	48.2 ±14%	38.3 ±14%	34	46.2
Max Speed (Mb/s)	325	326	1373	1377	12.6	250
Min Speed (Mb/s)	6.5	6.0	70.1	65.7	7.2	-
Monthly Max (€/Mb/s)	0.10	0.07	0.04	0.05	2.6	0.4
Unconnected Households (%)	12	6	7	7	18	-

TABLE II: Comparison of WBN Profile 1 and 3 ($s_r = [0.25, 0.5]$) with available commercial offers (xDSL and SAT).

that replaces the last and middle mile of a traditional fiber-connected communication network. Our model uses the highest realism, leveraging open demographic and geographic data, a market analysis of the current device prices, and state-of-the-art algorithms for network planning. We show that a wireless backhaul can be used to offer both minimal service at low cost and professional services at a higher target cost, and that it is not only technically but also economically viable and competitive with existing offers. The model is fully configurable and reusable, we share all the code and the generated data so that the same approach can be repeated in different locations and can be updated with new prices and new technologies, and become the starting point for future analyses.

Acknowledgment

This work was funded by the Spanish grants PDC2023-145809-I00/AEI/10.13039/501100011033 and PID2023-146066OB-I00, and Generalitat de Catalunya through 2021-SGR-01059.

References

[1] P. Micholia, M. Karaliopoulos *et al.*, "Community networks and sustainability: a survey of perceptions, practices, and proposed solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3581–3606, 2018.

[2] Y. Ben-David, "Connecting the last billion," Ph.D. dissertation, UC Berkeley, 2015. [Online]. Available: https://digitalassets.lib.berkeley.edu/etd/ucb/text/BenDavid_berkeley_0028E_15758.pdf

[3] P. A. Frangoudis, G. C. Polyzos, and V. P. Kemerlis, "Wireless community networks: an alternative approach for nomadic broadband network access," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 206–213, 2011.

[4] S. Hasan, Y. Ben-David *et al.*, "The challenges of scaling WISPs," in *Annual symposium on computing for development (DEV)*, 2015.

[5] A. Neumann, "Cooperation in open, decentralized, and heterogeneous computer networks," Ph.D. dissertation, Universitat Politècnica de Catalunya (UPC), 2017. [Online]. Available: http://upcommons.upc.edu/handle/2117/114450

[6] A. Chiha, M. Van der Wee *et al.*, "Techno-economic viability of integrating satellite communication in 4g networks to bridge the broadband digital divide," *Telecommunications Policy*, vol. 44, no. 3, 2020.

[7] P. Hu, "Closing the management gap for satellite-integrated community networks: A hierarchical approach to self-maintenance," *IEEE Communications Magazine*, vol. 59, no. 12, 2021.

[8] L. Maccari, G. Gemmi *et al.*, "Towards scalable community networks topologies," *Ad Hoc Networks*, vol. 94, 2019.

[9] D. Cameron, A. Valera, and W. K. Seah, "Elasticwisp: Energy-proportional wisp networks," in *Network Operations and Management Symposium (NOMS)*. IEEE/IFIP, 2020.

[10] T. Pötsch, P. Schmitt *et al.*, "Helping the lone operator in the vast frontier," in *Workshop on Hot Topics in Networks (HotNets)*. ACM, 2016.

[11] A. Hameed, A. N. Mian, and J. Qadir, "Low-cost sustainable wireless internet service for rural areas," *Wireless Networks*, vol. 24, no. 5, pp. 1439–1450, 2018.

[12] S. Mubaraq, J. Hwang *et al.*, "Economic analysis of networking technologies for rural developing regions," in *Workshop on Internet Economics (WIE)*, vol. 10. ACM, 2005.

[13] G. Castellanos, B. De Beelde *et al.*, "Evaluating 60 ghz fwa deployments for urban and rural environments in belgium," *Sensors*, vol. 23, no. 3, p. 1056, 2023.

[14] T. Sharma, A. Chehri, and P. Fortier, "Review of optical and wireless backhaul networks and emerging trends of next generation 5G and 6G technologies," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 3, p. e4155, 2021.

[15] E. J. Oughton, K. Katsaros *et al.*, "An open-source techno-economic assessment framework for 5G deployment," *IEEE Access*, vol. 7, pp. 155 930–155 940, 2019.

[16] F. Yaghoubi, M. Furdek *et al.*, "Reliable topology design of wireless networks under correlated failures," in *International Conference on Communications (ICC)*. IEEE, 2018.

[17] L. Maccari, M. Karaliopoulos *et al.*, "5g and the internet of everyone: Mo-

tivation, enablers, and research agenda," in *2018 European Conference on Networks and Communications (EuCNC)*, 2018.

[18] A. Chaoub, M. Giordani *et al.*, "6g for bridging the digital divide: Wireless connectivity to remote areas," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 160–168, 2022.

[19] H. Saarnisaari, A. Chaoub *et al.*, "Wireless terrestrial backhaul for 6g remote access: Challenges and low power solutions," *Frontiers in Communications and Networks*, vol. 2, p. 710781, 2021.

[20] G. Gemmi, L. Cerdà-Alabern, and L. Maccari, "A realistic open-data-based cost model for wireless backhaul networks in rural areas," in *18th International Conference on Network and Service Management (CNSM)*, 2022.

[21] E. S. Report, "Broadband in the eu member states: despite progress, not all the europe 2020 targets will be met."

[22] M. UK Department of Digital, Culture and Sport, "A new broadband universal service obligation: Government's response to consultation on design."

[23] E. H. Ong, J. Kneckt *et al.*, "Ieee 802.11 ac: Enhancements for very high throughput wlans," in *International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2011.

[24] "Mikrotik wireless devices' prices and datasheets," accessed: 2022-04-28. [Online]. Available: https://mikrotik.com/products/group/wireless-systems

[25] "Mikrotik routers' prices and datasheets," accessed: 2022-04-28. [Online]. Available: https://mikrotik.com/products/group/ethernet-routers

[26] Fiber Broadband Association, "2023 fiber deployment cost annual report," https://fiberbroadband.org/wp-content/uploads/2024/01/Fiber-Deployment-Annual-Report-2023_FBA-and-Cartesian.pdf, 2023, accessed May 2025.

[27] L. . T. Cerdà-Alabern, "On the Guifi.net community network economics," *Computer Networks*, vol. 168, no. 4, pp. 3581–3606, 2018.

[28] S. Verbrugge, D. Colle *et al.*, "Methodology and input availability parameters for calculating OpEx and CapEx costs for realistic network scenarios," *Journal of Optical Networking*, vol. 5, no. 6, pp. 509–520, 2006.

[29] "Xarxa oberta fiber transport service," accessed: 2022-05-03. [Online]. Available: https://web.archive.org/web/20220503174738/https://xarxaoberta.cat/services/transparent

[30] Eurostat, "Electricity price statistics – first half 2024," https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics, 2024, accessed May 2025.

[31] G. Gemmi, L. Cerdà-Alabern *et al.*, "Toward Smart Community Networks," *IEEE Network*, vol. 37, no. 2, 2023.

[32] X. Zhao, J. Kivinen *et al.*, "Propagation characteristics for wideband outdoor mobile communications at 5.3 ghz," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 507–514, 2002.

[33] E. T. S. I. (ETSI), "ETSI TR 138 901 V14.0.0 (2017-03): 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 14.0.0 Release 14)," Tech. Rep. TR 138 901 V14.0.0, March 2017, available online: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/14.00.00_60/tr_138901v140000p.pdf.

[34] G. Gemmi, R. Lo Cigno, and L. Maccari, "On the properties of next generation wireless backhaul," *IEEE Transactions on Network Science and Engineering*, 2022.

[35] ——, "On cost-effective, reliable coverage for los communications in urban areas," *IEEE Transactions on Network and Service Management*, 2022.

[36] G. Gemmi, M. Polese *et al.*, "Optimizing and Managing Wireless Backhaul for Resilient Next-Generation Cellular Networks," in *International Conference on Network and Service Management (CNMS)*, 2024.

[37] G. Karypis and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," *SIAM Journal on scientific Computing*, vol. 20, no. 1, pp. 359–392, 1998.

[38] B. W. Kernighan and S. Lin, "An efficient heuristic procedure for partitioning graphs," *The Bell System Technical Journal*, vol. 49, no. 2, pp. 291–307, 1970.

[39] N. Mark, *Networks: an introduction*. Oxford University Press New York, 2010.

[40] S. Khuller and R. Thurimella, "Approximation Algorithms for Graph Augmentation," *Journal of Algorithms*, vol. 14, no. 2, pp. 214–225, 1993.

[41] E. J. Oughton, "Policy options for broadband infrastructure strategies: A simulation model for affordable universal broadband in africa," *Telematics and Informatics*, vol. 76, 2023.

[42] W. Waites, J. Sweet *et al.*, "RemIX: A Distributed Internet Exchange for Remote and Rural Networks," in *Workshop on Global Access to the Internet for All (GAIA)*. ACM, 2016.

[43] "Trasparenza tariffaria delle offerte di linea fissa," accessed: 2022-08-29. [Online]. Available: https://web.archive.org/web/20220829180318/https://www.tim.it/assistenza/trasparenza-tariffaria/trasparenza-tariffaria-delle-offerte-di-linea-fissa

[44] "Order starlink," accessed: 2024-11-13. [Online]. Available: https://www.starlink.com/

[45] "Lidar datasets released by the italian ministry of environment," accessed: 2022-04-28. [Online]. Available: https://www.mite.gov.it

APPENDIX

A. Computing the Demand

Algorithm 2: Compute a random subscriber set $\mathcal{V}$

Input: $\mathcal{C} = \{c_j\}$ census polygons; h_j households per census; $\mathcal{P} = \{n_i\}$ building polygons; v_i building volumes; u number of subscribters

Output: sampled set $\mathcal{V}$ of u subscribers

foreach $n_i \in \mathcal{P}$ **do** // Step 1
    if $v_i < 100\,\mathrm{m}^3$ **then continue**;
    foreach $c_j \in \mathcal{C}$ **do**
        $V_i^j \leftarrow \dfrac{\mathrm{area}(c_j \cap n_i)}{\mathrm{area}(n_i)} v_i$
foreach $c_j \in \mathcal{C}$ **do** // Step 2
    $S_j \leftarrow \sum_{n_k} V_k^j$
    foreach n_i *with* $V_i^j > 0$ **do**
        $H_i^j \leftarrow V_i^j h_j / S_j$
foreach $n_i \in \mathcal{P}$ **do** // Step 3
    $w_i \leftarrow \dfrac{\sum_j H_i^j}{\sum_k h_k}$
$\mathcal{V} \leftarrow$ sample_with_replacement($\mathcal{P}$, w_i, u) // Step 4
return $\mathcal{V}$

Let $\mathcal{C} = \{c_j\}$ be a set of census section polygons for the municipality under analysis, and let h_j be the number of households in the polygon, both obtained by ISTAT. Let $\mathcal{P} = \{n_i\}$ be a set of polygons representing buildings, obtained from the OSM data-set. We use highly precise morphological open-data [45] to compute the volume v_i of each building n_i, we remove buildings smaller than $100\,\mathrm{m}^3$, as they have a low probability of being inhabited. The process is split in 4 steps which we describe here, the pseudo-code with precise formulas is in Algorithm 2:

1) We compute V_i^j, that is the volume of the building n_i in the section c_j. V_i^j takes into account the fact that a building may lie across two different census sections.
2) We then normalize V_i^j over the volume of all buildings in the census area, and multiply it by the number of households in the area, obtaining H_i^j: the number of households we expect to live in building n_i that pertain to the census area c_j
3) We sum H_i^j over all the possible census sections, and then normalize again over the total number of households in all census sections. We obtain a set of dimensionless weights $\{w_i\}$

4) We sample u subscribers on $\mathcal{P}$ with probabilities $\{w_i\}$.

B. Capacity Computation

Let $n_0, n_1, n_2 \dots n_t$ be a sequence of nodes on the shortest path from a gateway ($g = n_0$) to a subscriber (s hosted on node n_t). We call $\Delta_d(n)$ the set of the neighbors downstream of n with $|\Delta_d(n)| = \delta_d(n)$, similarly $\Delta_u(n)$ is the set of upstream neighbors, which by design has cardinality one, as the topology is a tree. We will denote the only node in $\Delta_u(n)$ as n_u. Let then $s(n)$ be the number of subscribers in node n and $s_d(n)$ the number of

Gabriele: is $s_d(n)$ the same as $sp(n)$ in algorithm 1? then change it

subscribers passing through n (downstream). We call $b(n_i, n_j)$ the negotiated bit-rate on the link $n_i \to n_j$. Such bit-rate is computed by leveraging mappings between MCS and received signal power, as explained in Sect. III-A. Such bit-rate is multiplied by a MAC efficiency parameter $\eta = 0.84$ [23]. Then we derive the average downstream capacity per neighbor of node n:

$$\hat{\lambda}_d(n) = \frac{\eta}{\delta_d(n)} \sum_{n_j \in \Delta_d(n)} b(n, n_j) \tag{14}$$

And the average upstream capacity:

$$\hat{\lambda}_u(n) = \eta \, b(n_u, n), \tag{15}$$

The minimum bit-rate per shortest path on node n based on the number of devices and the number of paths passing through the node is the minimum of its upstream and downstream capacities:

$$\lambda_d(n) = \min\left(\frac{\hat{\lambda}_d(n) \cdot N(n)}{s_d(n)}, \frac{\hat{\lambda}_u(n)}{s(n) + s_d(n)}\right) \tag{16}$$

Note that in our model, we do not assign a specific orientation to the devices, so we can not effectively compute the number of edges per device, and we must rely on an average per node. On the path from g to s, if $n = g$ then Eq. (16) the upstream capacity is not taken into account, as it runs on optical fiber. The minimum capacity per subscriber s is the bottleneck on the path from the gateway:

$$c_{\min}(s) = \min_{n_i \in \{n_0 \dots n_t\}} \lambda_d(n_i) \tag{17}$$

and $\bar{c}_{min}$ is the average on all subscribers. To compute $c_{max}(s)$ we use:

$$c_{\max}(s) = \min_{n_i \in \{n_1 \dots n_t\}} b(n_{i-1}, n_i) \tag{18}$$

where $\bar{c}_{\max}(s)$ is the average on all subscribers.

C. Price Ranges

To capture cost variability in our techno-economic model, we define lower and upper bounds for each parameter using multiple sources, including vendor data, empirical studies, and public reports.

CPEs and Routers: For CPEs and routers, we use the retail price listed in Mikrotik's catalog [24] as the upper bound. The lower bound applies a 40% discount to reflect typical reductions from bulk purchasing and supplier-level agreements.

Aerial Fiber: Fiber deployment costs are based on the 25th and 75th percentiles reported in the Fiber Broadband Association's 2023 Annual Report [26]. Original values in USD per foot were converted to Euros per kilometer using OECD purchasing power parity (PPP) factors for Italy, yielding a realistic cost range of approximately €8946–€18512 per kilometer.

Transit at IXP: For Internet transit at the regional IXP, we use €1680/Gbps/year from Cerdà-Alabern et al. [27] as an upper bound. To represent cost savings from peering, the lower bound accounts for a 25% reduction due to locally exchanged traffic, resulting in an effective cost of €1260/Gbps/year.

Energy Cost: Energy costs are based on Eurostat's average 2024 electricity prices for Italy [30]. We use €0.1867/kWh (non-household consumers) as the lower bound and €0.3274/kWh (household consumers) as the upper bound.

Gabriele Gemmi received a double Ph.D. in computer science at the Ca' Foscari University of Venice (IT) and in computer architecture at Polytechnic University of Catalonia (ES) in 2023. Currently, he is a Research Scientist at the Institute for Wireless Internet of Things at Northeastern University, USA. His research interests include wireless communication networks, distributed systems, and network economics.

Llorenç Cerdà-Alabern is an associate professor at the Computer Architecture Department, Universitat Politècnica de Catalunya, since 2000. He has published a substantial number of papers and participated in national and EU-funded research projects. His current areas of research include wireless 802.11, mesh networks, and community networks.

Leonardo Maccari is an associate professor at Ca' Foscari University of Venice, Italy. He co-authored more than 80 publications in refereed conferences, journals, book chapters, and patents. Among his research interests are network protocols for large-scale wireless mesh networks, wireless backhaul networks, interdisciplinary studies, and network security.