WhiteHaul: An Efficient Spectrum Aggregation System for Low-Cost and High Capacity Backhaul over White Spaces

Citation for published version:
Kassem, M, Kheirkhah, M, Marina, MK & Buneman, P 2020, WhiteHaul: An Efficient Spectrum Aggregation System for Low-Cost and High Capacity Backhaul over White Spaces. in *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services.* Association for Computing Machinery (ACM), pp. 338–351, 18th ACM International Conference on Mobile Systems, Applications, and Services, Toronto, Canada, 15/06/20. https://doi.org/10.1145/3386901.3388950

Digital Object Identifier (DOI):
10.1145/3386901.3388950

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services

WhiteHaul: An Efficient Spectrum Aggregation System for Low-Cost and High Capacity Backhaul over White Spaces

Mohamed M. Kassem
University of Edinburgh
m.kassem@sms.ed.ac.uk

Morteza Kheirkhah
University College London
m.kheirkhah@ucl.ac.uk

Mahesh K. Marina
University of Edinburgh
mahesh@ed.ac.uk

Peter Buneman
University of Edinburgh
opb@inf.ed.ac.uk

ABSTRACT

We address the challenge of backhaul connectivity for rural and developing regions, which is essential for universal fixed/mobile Internet access. To this end, we propose to exploit the TV white space (TVWS) spectrum for its attractive properties: low cost, abundance in underserved regions and favorable propagation characteristics. Specifically, we propose a system called WhiteHaul for the efficient aggregation of the TVWS spectrum tailored for the backhaul use case. At the core of WhiteHaul are two key innovations: (i) a TVWS conversion substrate that can efficiently handle multiple non-contiguous chunks of TVWS spectrum using multiple low cost 802.11n/ac cards but with a single antenna; (ii) novel use of MPTCP as a link-level tunnel abstraction and its use for efficiently aggregating multiple chunks of the TVWS spectrum via a novel uncoupled, cross-layer congestion control algorithm. Through extensive evaluations using a prototype implementation of WhiteHaul, we show that: (a) WhiteHaul can aggregate almost the whole of TV band with 3 interfaces and achieve nearly 600Mbps TCP throughput; (b) the WhiteHaul MPTCP congestion control algorithm provides an order of magnitude improvement over state of the art algorithms for typical TVWS backhaul links. We also present additional measurement and simulation based results to evaluate other aspects of the WhiteHaul design.

CCS CONCEPTS

• **Networks** → **Wireless access points, base stations and infrastructure**; **Cross-layer protocols**; **Mobile networks**.

KEYWORDS

Universal Internet access, Rural connectivity, Backhaul, TV white space spectrum, Spectrum aggregation, Multipath TCP

ACM Reference Format:

Mohamed M. Kassem, Morteza Kheirkhah, Mahesh K. Marina, and Peter Buneman. 2020. WhiteHaul: An Efficient Spectrum Aggregation System for Low-Cost and High Capacity Backhaul over White Spaces. In *The 18th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '20), June 15–19, 2020, Toronto, ON, Canada.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3386901.3388950

MobiSys '20, June 15–19, 2020, Toronto, ON, Canada

ACM ISBN 978-1-4503-7954-0/20/06…$15.00
https://doi.org/10.1145/3386901.3388950

1 INTRODUCTION

The beneficial impacts with Internet connectivity have been well documented [73]. However, almost half the world's population is still unconnected [13]. And, according to 2018 World Bank estimates, almost half the world's population is rural who make up the large fraction of the unconnected. This is, for example, apparent from the global mobile Internet penetration figures. In some regions like North America 4G/LTE amounts to nearly 90% of the overall mobile subscriptions and fast moving to 5G; however it is just around 7% in Sub-Saharan Africa with higher percentages of rural population [27]. In the past decade, there have been a series of efforts to remedy this through community cellular networks, initially focused on voice and SMS services [15, 41, 91] and more recently on LTE based mobile broadband Internet service [51, 80] leveraging the emergence of open-source software platforms, etc. Recent regulatory developments [67] promote such *local* wireless access networks by allowing the use of "unused" licensed spectrum at nominal cost.

Despite these developments, the infrastructure to connect the sparsely populated or low-income rural and developing regions is limited and remains a major roadblock [20, 54, 80, 83]. A key challenge is to create economical backhaul networks that connect access networks to the wider Internet [50, 56, 65]. Traditional approaches for backhaul connectivity rely on fiber, licensed microwave or satellite solutions that have high CAPEX or OPEX costs [88].

In this paper, we propose to exploit spectrum white spaces towards low cost backhaul for underserved regions. In particular, we focus on the TV white space (TVWS) spectrum — the portions of UHF TV bands unused by TV transmitters and wireless microphone users (the primary users of this spectrum). Led by the U.S. in 2008, several countries have made the TVWS spectrum unlicensed subject to interference protection for primary users (e.g., TV receivers) that requires consulting a geolocation database for available spectrum at a given location and time. TVWS spectrum is attractive for backhaul connectivity in rural and developing regions for multiple reasons. First, TVWS spectrum costs little (just database access fees) and there is an ample amount of it available in rural areas (in the region of 200+ MHz) with fewer TV transmitters and rare wireless microphone use, as also shown in Fig. 1 for our case study in Scotland. In developing countries, almost all of the UHF band is available as white space due to non-existent or limited presence of over-the-air TV [50].

Second, UHF TV spectrum has superior propagation characteristics compared to other higher frequency bands in terms of both range and non-line-of-sight (NLoS) propagation in presence of foliage and obstructions [21, 32, 49, 50, 56]. For example, it is possible to get 4 times greater range with the TVWS spectrum than with the 2.4GHz unlicensed spectrum used by Wi-Fi [21, 50]. This suggests lower infrastructure costs as fewer number of relays are sufficient to enable backhaul connectivity over long distances.

Unsurprisingly, the promise of TVWS spectrum for backhaul connectivity in rural and developing regions has been recognized in the literature [50, 52, 56, 57, 80]. However, existing TVWS systems (discussed further in §6) – both commercial solutions and research prototypes – fail to fully realize this promise as the throughput they can achieve is limited to a few tens of Mbps — insufficient for even a modest community of network users.

We present WhiteHaul, the first TVWS based system that can deliver an order-of-magnitude higher throughput (over 500Mbps) than the state-of-the-art by addressing several significant challenges and constraints pertinent to the backhaul use case, as outlined below and elaborated in §2. First, individual TVWS channels are narrow (6/8MHz depending on the regulatory regime). Second, available channels may not be contiguous depending on the presence of primary users (e.g., TV transmitters). So it is imperative to aggregate multiple possibly non-contiguous TVWS channels to realize high-speed TVWS backhaul connectivity. Third, it is desirable to use a single antenna at each backhaul link endpoint (even with multiple radio interfaces) because of the larger size of TVWS antennas, as also previously articulated in [78] – having multiple antennas requires separation in the order of meters and hence higher towers, steeply increasing the cost and deployment complexity [56, 72]. Fourth, TVWS spectrum exhibits a high degree of diversity in terms of chunk sizes, transmit power and interference levels, especially for long-distance backhauling, and these need to be taken into account in any design. Lastly, backhaul traffic exhibits high degree of asymmetry and temporal fluctuations.

In this paper, we make the following contributions pertinent to the design, implementation and evaluation of the WhiteHaul system:

- Considering Scotland as a representative country, we present a case study with an extensive analysis of salient aspects of TVWS spectrum characteristics from a backhaul use case perspective; and we analyze real-world rural backhaul traffic characteristics (§2).
- Informed by the above analysis, we design and implement the WhiteHaul system (§3 and §4) that features several innovations in both hardware and software. Chief among them are: (i) a TVWS conversion substrate to efficiently handle multiple non-contiguous chunks of TVWS spectrum with a single antenna using multiple low cost COTS 802.11n/ac cards; (ii) a novel way of leveraging MPTCP as an abstraction of a high-speed link-level tunnel, and efficiently aggregating multiple TVWS spectrum chunks (sub-links) via a novel uncoupled, cross-layer congestion control mechanism that is agile to underlying variations in available bandwidth of sub-links.
- We extensively evaluate WhiteHaul using our prototype implementation and simulations driven by real-world backhaul traffic traces to quantify the aggregate backhaul link throughput it can achieve in various network settings, e.g., number and width of TVWS spectrum chunks, and link conditions (§5). In particular, we show that WhiteHaul can aggregate almost the whole of TV band with 3 interfaces and achieve nearly 600Mbps TCP throughput. The WhiteHaul MPTCP congestion control algorithm is also shown to provide an order of magnitude improvement over the state of the art algorithms in the presence of typical loss rates experienced by TVWS backhaul.

From a wider perspective, NLoS propagation capabilities of TVWS, and the resulting cost advantages due to shorter towers or fewer relays, make our work complementary to alternative backhaul solutions based on licensed/unlicensed microwave and long-distance Wi-Fi (e.g., [70]) that require line-of-sight. Also, the use of TVWS for backhauling in general and WhiteHaul in particular compares favorably, especially in terms of cost, ease of deployment and robustness, with respect to alternative technologies being explored for addressing connectivity challenges in rural and developing regions (e.g., drones [4], free space optical communications [5], Google Project Loon [54], millimeter waves [61]). Our work is also complementary to other recent work that focuses on leveraging spectrum white spaces for access networks [18, 40] as well as works that focus on inter-working between commercial and community cellular networks [39, 79].

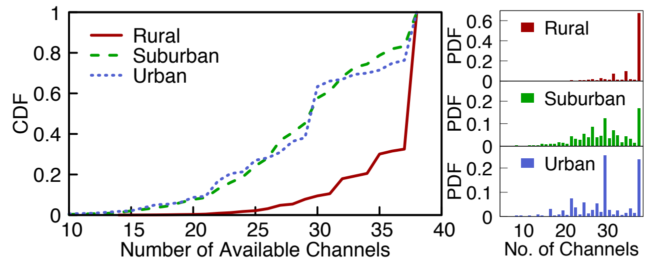


Figure 1: TVWS spectrum availability distribution in different types of areas across Scotland.

2 TVWS BACKHAUL LINKS: CHARACTERISTICS AND CHALLENGES

As a way to substantiate the challenges associated with using TVWS spectrum for backhaul links in underserved regions, here we consider Scotland as a representative country and present a case study examining the nature of TVWS spectrum in the 470-790MHz TV band as per the ETSI Harmonised Standard for White Space Devices [28]. To this end, we represent Scotland as a set of pixels, each corresponding roughly to a square of size 6km^2. For the center location of each such pixel, we query a commercial geolocation database to obtain the available TV channels at that location along with allowed power levels. The TVWS spectrum availability results in Fig. 1 for different area types, corresponding to area classification in Scotland based on population density, confirm the ample availability in rural areas (e.g., 38 8MHz wide TV channels available in 70% of rural locations), broadly in agreement with prior studies (e.g., [37, 85]). Our focus in the rest of this section is to shed light on several salient issues that need to be accounted for when designing TVWS backhaul links.

Spectrum fragmentation. We now focus on the inhabited rural areas to understand their TVWS spectrum characteristics. We first look into the extent to which the available TVWS spectrum is fragmented. For this, we define a *spectrum chunk* as a set of contiguous TVWS channels available at a given location. Fig. 2(a) quantifies the extent to which TVWS is fragmented in rural areas. For example, in 60% of the locations, the spectrum is fragmented into at least 6 chunks and in only 20% of the locations is the spectrum available in 4 or fewer chunks. Such a high degree of spectrum fragmentation is due to the presence of multiple lower power TV relays/amplifiers needed to extend the coverage of the primary DTT transmitters. This creates gaps in available TVWS spectrum as access to channels amplified by such relays are restricted by the geolocation database.

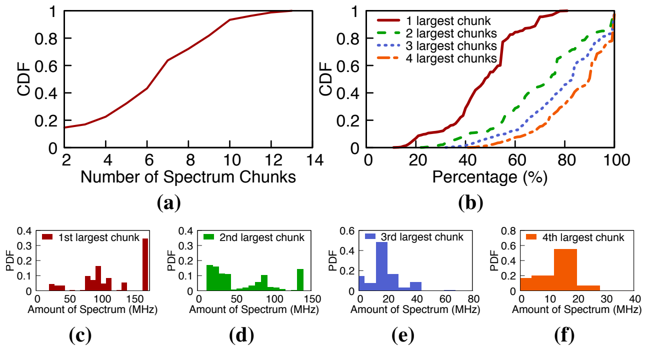


Figure 2: (a) shows the CDF of number of spectrum chunks across different locations. (b) the percentage of largest 4 chunks of the spectrum to the total amount of available spectrum. Distribution of the size of the 4 largest spectrum chunks is shown in (c), (d), (e) and (f).

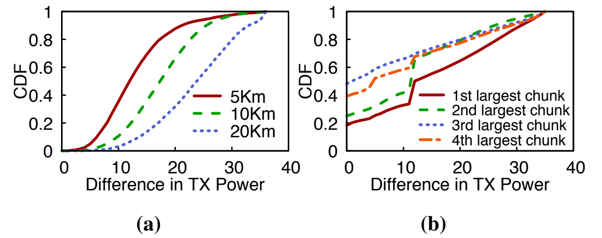


Figure 3: Power asymmetry effect: (a) at different link distances (5-20Km); (b) between channels within each spectrum chunks.

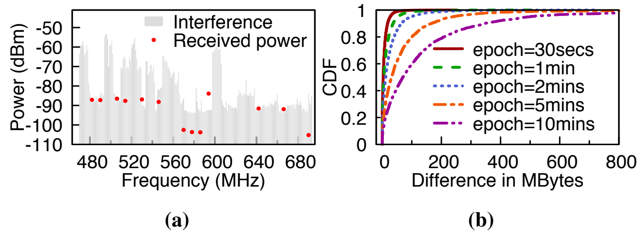


Figure 4: (a) Spectrum occupancy (reflecting interference levels) as measured by a spectrum analyzer at a receiver. The red points are the estimated received power from a remote TVWS transmitter on available channels at the latter; (b) Downstream traffic volume variation (in MBytes) between consecutive epochs for different epoch lengths.

Spectrum chunk size diversity. Fig. 2(b) shows how the spectrum is distributed across chunks through CDFs of the percentage of total available spectrum that is covered by the one, two, three and four largest chunks. We observe that only about half the available spectrum is covered by the largest chunk (red solid curve) in half the locations, and that even top four chunks combined cannot cover all the available spectrum in 80% of the locations. Distribution of chunk sizes within each of the four largest sized chunks is shown in Fig. 2(c)-(f). These results demonstrate *the significant diversity in chunk sizes*, which adds further complexity to the problem of aggregating TVWS spectrum in that traffic should be distributed across chunks accounting for this diversity.

Power asymmetry. With the TVWS spectrum, the geolocation database needs to be queried to determine the available set of channels and power levels at a location, and it is possible that these could be substantially different between link endpoints that are some distance apart. We examine the potential power asymmetry resulting from this issue in Fig. 3. Fig. 3(a) shows the allowed power level differences for different point-to-point (PtP) link distances for each channel in the common set of channels available at both ends of the link. We observe that the power differences can be quite significant for longer (20Km) links with a median around 25dB. This effect is also seen for shorter (5Km) links where only in 10% of the cases is the transmit power difference less than 6dB. Note that this power asymmetry effect is unique to TVWS based backhaul setting and not present in other approaches used in the literature for low-cost backhaul such as long-distance Wi-Fi [26, 31, 68, 70, 75, 82].

Power differences can also occur between different channels available at a link endpoint. This effect is quantified in Fig.3(b), where maximum power differences between channels within each of the four largest chunks is shown. We see that channels within a chunk can have quite different power levels; for example, the median power difference within 1st and 2nd chunks is around 12dB. As such, the set of chunks and their size need to be carefully chosen to minimize these above highlighted power asymmetry effects. Moreover, time allocation for each end of a backhaul link needs to account for the effective capacity of the link and traffic demand from each direction, the former being affected by transmit power used for each of the chunks. We factor these observations in our design (§4.2).

Interference. While consulting the geolocation database is mandatory to access the TVWS spectrum, it cannot capture *receiver-side interference* characteristics that could affect the quality of TVWS transmission from a distant transmitter. To illustrate this point, we suppose a TVWS receiver at the rooftop of our office building and consider a transmitter on top of another building 3km away. Based on available TVWS channels and allowed power levels at the transmitter location by querying the geolocation database, then accounting for path loss (calculated using the SPLAT! RF planning tool [48]) and antenna gains, we obtain expected signal power at the receiver on channels available at transmitter side as shown in Fig. 4a (the red dots). At the receiver, we use a spectrum analyzer to estimate the level of interference on different channels across the whole TV band, overlaid in the same figure. We see that the channels available at the transmitter can be very different from the receiver in terms of received signal power and interference levels. This highlights the importance of measuring and considering receiver-side perspective in choosing spectrum chunks, which we do in our design (§3, §4.2.1).

Traffic characteristics. Effective backhaul network design requires a good understanding of the characteristics of traffic it is expected to carry. To this end, we collected trace of traffic from Tegola [10], a long-running rural community wireless access network, as seen by its leased fiber backhaul link (with 200Mbps symmetric upstream and downstream bandwidth limit). This network serves at least 250 households and local businesses. Fig. 5 shows a week long backhaul traffic profile in the downstream (towards access network) and upstream directions. We observe a high degree of asymmetry with average daily downstream traffic much higher, around 63GBytes, compared to average upstream traffic each day (only about 4GBytes). We also note that, although there is some apparent diurnal pattern, traffic in both directions fluctuates substantially over time. To understand the degree of fluctuation at different time granularities

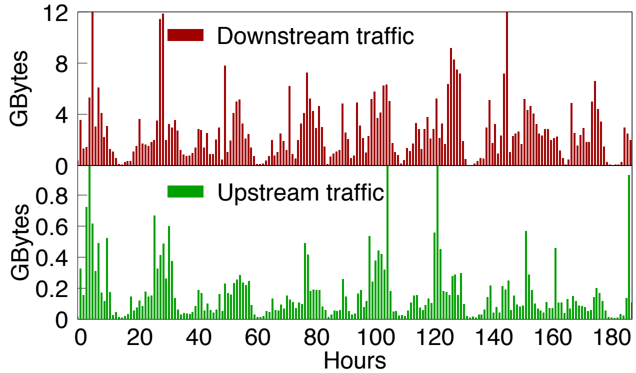


Figure 5: Weekly backhaul traffic volumes for Tegola network.

(epochs), Fig. 4b shows the CDF of variation in downstream traffic from one epoch to the next for different epoch lengths. These results indicate considerable degree of traffic variation for longer epochs but marginal variation with the smallest epoch size we could obtain for this trace, which is based on data collected every 30s.

3 WHITEHAUL OVERVIEW

This section gives an overview of our proposed WhiteHaul system for TVWS based backhaul. Its design addresses the various pertinent challenges and constraints highlighted in the previous section. First, the TVWS channels are narrow and the available spectrum is fragmented. This requires aggregation of contiguous and non-contiguous spectrum chunks across the UHF TV band spanning a few hundred MHz. However, wide band spectrum aggregation is complex and expensive as it involves dealing with voltage-controlled oscillator (VCO) pulling effects and careful RF filter design, which are evident from the RF front end design challenges for LTE/5G carrier aggregation [60, 69]. Given the somewhat relaxed form factor requirement in our setting and keeping cost in mind, we take an alternative approach that combines multiple low cost COTS Wi-Fi cards with a custom designed frequency conversion substrate for down/up conversion to TV band frequencies. Second, given the large size of TVWS antennas, we constrain each backhaul link endpoint to use a single antenna to avoid higher towers, which steeply increase the cost and deployment complexity. Allowing multiple antennas requires sufficient separation between them in the order of meters to counter side/back lobe interference effects that can hurt link throughput by 2.5x [47], and this in turn increases the tower height. Higher towers can also result in reduced transmit power limit due to TVWS regulatory constraints on antenna height [44]. Third, given the potential asymmetry in power, interference and traffic over TVWS based backhaul links as demonstrated in our characterization study, our design allows for flexible and dynamic time allocation for communication in each direction of a WhiteHaul link. Fourth, having the hardware capability to communicate over multiple different TVWS spectrum chunks is alone insufficient to realize high-speed TVWS based backhaul. What is needed is a glue to combine the capacity across the multiple spectrum chunks and accounting for their diversity, thereby creating a link level tunnel abstraction. Among the different multipath tunneling approaches, we find MPTCP to be the best fit for our use case. Even so, the existing MPTCP congestion control schemes (e.g., LIA[71] and OLIA[53]) are slow in responding to the change of underlying link capacity of

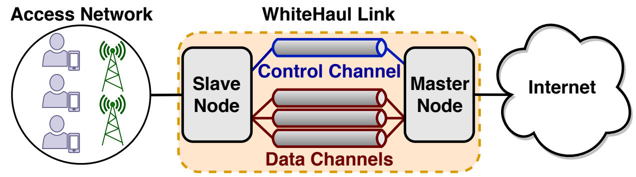


Figure 6: High-level schematic of WhiteHaul system in an end-to-end application scenario.

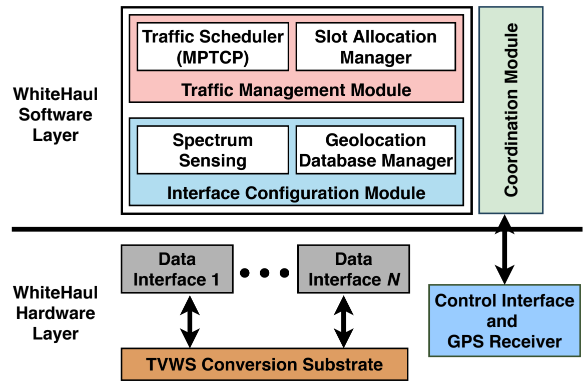


Figure 7: WhiteHaul node architecture.

each subflow/sub-link. This motivates us to develop a new cross-layer MPTCP congestion control algorithm for fast adaptation, thereby achieve high aggregated throughput across sub-links.

WhiteHaul is intended to realize point-to-point (PtP) TVWS based backhaul links, meaning a WhiteHaul node forms the endpoint of such a link. Fig. 6 shows a basic application scenario where a WhiteHaul link connects an access network (serving end-user devices) and the Internet. We do not make any assumptions on the nature of the access network, which could take several forms including a community cellular network (e.g., [80]). While it is fairly straightforward to go from the case of one WhiteHaul link to a path or mesh network with multiple WhiteHaul links with suitable spectrum usage coordination amongst them, we focus on the single link case in this paper. Also, as shown in Fig. 6, we adopt a `Master-Slave` model in that one end of a WhiteHaul link acts as the `Master Node` and the other as the `Slave Node` with the former responsible for link configuration decisions (e.g., spectrum to use for interfaces on both sides). The `Master` and `Slave` coordinate over an *out-of-band control channel* while carrying user traffic over *multiple data channels that each operate on a separate TVWS spectrum chunk*. There are as many data channels as the number of interfaces at WhiteHaul end nodes making up the link. For transporting data packets over the data channels bidirectionally we create a MPTCP tunnel by leveraging a SOCKS5 proxy, a commonly used approach in several MPTCP studies [23, 64, 77]. We install the client and server side of the SOCKS5 on the `Slave` and `Master` nodes respectively. For reasons elaborated later in section 4.2, WhiteHaul links operate in time-division duplexing (TDD) mode, meaning the `Master` and `Slave` take turns in time, possibly of different durations in each direction, to communicate over their respective data interfaces.

Fig. 7 shows the schematic of the WhiteHaul node architecture, that consists of two layers: the *Hardware Layer* and the *Software Layer*. The hardware layer is composed of the physical wireless interfaces used for both control and data communication as well as the TVWS conversion substrate for the data interfaces. For the data interfaces, we use COTS 802.11n/ac Wi-Fi cards operating in 5GHz band. The conversion substrate is responsible for frequency up/down conversion between available TVWS spectrum chunks and 5GHz Wi-Fi channels. For the control interface, we use LoRa [74], a low-power wide area network technology, that costs a few dollars a piece, operates in unlicensed sub-GHz bands, and provides data rates of tens of Kbps over long distances up to 40Km.

The WhiteHaul software layer orchestrates the underlying interfaces to maximize the overall system performance. It is made up of three modules: (i) the *Coordination Module* facilitates communication between the `Master` and `Slave` nodes via the underlying LoRa control interface; (ii) the *Interface Configuration Module* configures the TVWS spectrum chunks and transmit power of data interfaces as decided by the `Master` node. These configurations are based on the TVWS spectrum availability information obtained from the geolocation database and local low-cost spectrum sensing from both ends of the WhiteHaul link; (iii) the *Traffic Management Module* performs two functions. One, by the *Slot Allocation Manager*, is to adapt the time allocation between `Master-Slave` (*forward*) and `Slave-Master` (*reverse*) directions, every *epoch*, depending on the effective capacity and traffic demand of forward and reverse links. The *Traffic Scheduler* is responsible for the other function to efficiently schedule the traffic among the underlying data interfaces using a modified variant of MPTCP, aided by advisory signals from the *Slot Allocation Manager*.

4 SYSTEM DESIGN & IMPLEMENTATION

4.1 Hardware Layer

4.1.1 TVWS Conversion Substrate. As previously stated, this part of the hardware layer converts between 5GHz and TVWS spectrum. Fig. 8 shows the schematic of its design for the case of two 802.11n/ac data interfaces. As shown, data interfaces as well as the substrate, specifically the local oscillator (LO), are dynamically controlled by the *Interface Configuration Module* to set parameters such as oscillator frequencies, channel bandwidths and power levels while ensuring compliance with TVWS spectrum regulations. In general, we target a flexible and modular design for the substrate to allow its realization with replaceable/configurable components as per link requirements and cost considerations (e.g., trade off between noise and linearity). Our implementation consists of a desktop (running Ubuntu Linux 14.04) connected to a set of Mikrotik RB922UAGS-5HPacD Router-Boards with 802.11n/ac cards, acting as data interfaces, via Gigabit Ethernet. The desktop also hosts a USRP B210 [12] per data interface for use as a LO, as described below.

SDR based Local Oscillator. The LO component plays the key role of translating RF/IF signal up/down to a different frequency band by multiplying it with the sinusoidal signal it generates. Previous works employing the frequency conversion concept (e.g., [16, 63]) have relied on Voltage Controlled Oscillators (VCOs) to generate the LO signal; specifically, these VCOs take a control voltage as input to determine the frequency of output LO signal. As VCOs operate under high non-linearity and produce unwanted emissions, they cause frequency fluctuations (harmonics and phase noise) of the output signal that blur the output IF signal when used in down-conversion

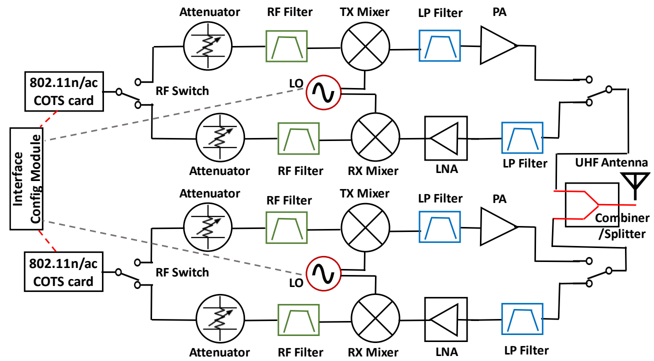
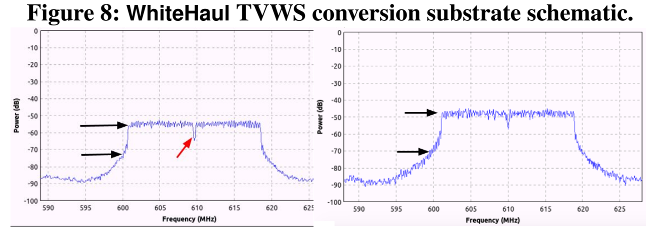


Figure 8: WhiteHaul TVWS conversion substrate schematic.



(a) (b)

Figure 9: Comparison of down-converted signal using VCO in (a) with that of our SDR-based LO approach in (b).

and degrade its Signal-to-Noise Ratio (SNR). Our experiments validate this effect and show, for instance, that a VCO oscillator (with $-69dBc/Hz$ phase noise at 1kHz offset frequency) used to down-convert from 5GHz to UHF band can degrade the SNR value of the output IF signal by 5dB (Fig. 9a). Having two of these VCOs, one in the transmit chain and another in receive chain, can reduce the overall SNR level of the system by 10dB.

So to avoid such degradation, we take a different approach and use a SDR board (USRP B210 in our prototype) to generate a sinusoidal signal without any distortion or phase noise. The SDR-generated LO signal can then be fed to the mixer (see Fig. 8) to generate the up/down converted signal. Our approach not only results in a higher-SNR signal compared to VCO oscillators by 4dB (Fig. 9b) but also provides high flexibility in (re-)configuring the center frequency of the generated signal by the Interface Configuration Module. On the other hand, lack of such flexibility with the VCO due to the low granularity of the tuner voltage (steps of 0.25V which maps to steps of 20MHz in the generated LO signal) leads to misalignment in center frequency (as indicated by the red arrow in Fig. 9a).

Separate Transmit and Receive Paths. Our design is also distinct from prior work in that it uses two separate transmit and receive paths, which allows for fine-grained configuration for RF components to optimize signal quality of each path. To realize this separation, we make use of two fast Single Pole Double Throw (SPDT) RF switches with 35ns switching time, one interfacing with the 802.11 card and the other before combiner/splitter and UHF antenna. These switches have a wide bandwidth range from 500 to 6000 MHz, allowing operation in both 5 GHz and UHF bands.

The transmit path has a configurable RF attenuator followed by a high-pass filter to cut out spurious emissions from the 802.11 interface. Next is a highly linear down-conversion mixer that supports wide range of frequencies from 3700 to 7000 MHz and is driven by the SDR based LO as described above. The resulting IF signal goes through a low pass filter to remove mixer related non-linearities. Then a low noise power amplifier (PA), capable of 27dBm output power and with a low noise figure of 1.2dB, is used. This is a voltage-controlled amplifier which can be adjusted (by changing the voltage level) to stay in the allowed transmit power. The amplified signal so generated is then fed to combiner/splitter through another RF switch.

On the receive path, we have a low-pass filter to eliminate unwanted signals before going through a low noise amplifier (LNA), which in our prototype has a high gain of 22.5dB, ultra low noise figure of 0.5dB and wide operational bandwidth range from 50 to 3000MHz. The up-conversion mixer translates the received UHF signal into the 5GHz band with help of LO, as above. We only use one LO for a pair of transmit and receive paths to make sure that the center frequency of the up-converted signal is the same as that of the RF signal down-converted in the transmit path. Following the mixer, we have a high pass filter to remove the unwanted signal from the mixer. Finally, a configurable attenuator to avoid saturating the receive chain on the 802.11 interface.

Combiner/Splitter. To satisfy our design constraint of using a single antenna, while using multiple interfaces towards a high capacity backhaul link, we have a RF power combiner/splitter in the design that can combine different transmit paths (from different 802.11 interfaces) into one output that is fed to the antenna or split the received UHF signal into multiple receive paths. In our prototype, we use a combiner/splitter with high isolation (25 dB typical) to prevent leakage between paths. Moreover, it can handle high transmit power up to 10W (aggregated), and has 5 input ports (for combining up to 5 different 802.11 cards) with a total bandwidth of 320MHz that can span the whole TV band.

4.1.2 Coordination and Synchronization. The control interface (see Fig. 7) enables the coordination between the `Master` and `Slave` ends of a WhiteHaul link, which is useful for two purposes. First, for the `Slave` node to notify the `Master` about the spectrum sensing information at its location as well as its traffic demand every epoch (30 seconds in our implementation, see Fig.10; the red slots represent the control channel transmission). Second, for the `Master` to notify the `Slave` about the set of TVWS spectrum chunks to use for its interfaces as well as the time slot duration in the reverse direction. In our implementation, we realize this control communication channel using low-cost Pycom LoRa gateways [7] that operate on very narrow channels (7.8 - 500KHz) in 868MHz spectrum band and are capable of achieving few tens of Kbps data rate over long distances, up to 40Km. Along with a control interface, each WhiteHaul node is also equipped with a GPS receiver to facilitate localization and time synchronization, latter for TDD operation.

4.2 Software Layer

4.2.1 Interface Configuration. This module is responsible for the configuration of data interfaces at WhiteHaul link endpoints with spectrum chunks and power levels via coordination over the control interface. As a basic step, the `Master` periodically checks with the geolocation database (as required by the regulator) about TVWS spectrum availability and allowed power levels at both `Master` and `Slave` node locations. In addition, we equip each WhiteHaul node with a low cost spectrum analyzer (RF Explorer [9] in our implementation) to estimate the interference in available TVWS channels.

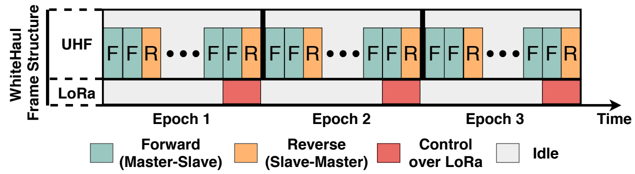


Figure 10: Example illustrating WhiteHaul slot allocation.

This is motivated by the observation made in §2 about interference and recent work (e.g., [42]) that observes that aggregate interference from multiple nearby transmitters can impact the quality of available channels. So, using the spectrum analyzer, each WhiteHaul node sweeps the whole TV band periodically to obtain the signal level on all the available TVWS channels as an estimate of interference on those channels. Note that it takes 5.6 seconds to sweep the whole TV band, which is sufficient to obtain an up-to-date measurement of interference on individual channels (and thereby the link quality) – link quality varies less frequently in our setting compared to Wi-Fi spectrum due to fewer and relatively static number of interferers. The `Slave` node then syncs all its sensing information with the `Master` node. Here the spectrum analyzer helps to quantify the interference level (at both endpoints); however, this is insufficient to determine the allowed TX power on different TVWS channels. To this end, WhiteHaul endpoints consult the geolocation database.

Using the sensing information, the `Master` then estimates the SINR for each of the TVWS channels commonly available between the endpoints in both forward and reverse directions. This involves estimating the received signal power on the available TVWS channels common to both endpoints. For this, we use the SPLAT RF planning tool[1] [48] to estimate the path loss in each direction along with allowed TX power and antenna gains. Based on the above, the link SINR for each TVWS channel is estimated as the lower of the two values for forward and reverse directions. With this channel-level SINR for the link in hand, the question is to decide on the TVWS spectrum chunks for the node interfaces. For this, we start with identifying the potential spectrum chunks considering contiguous set of available channels on both ends. Then, keeping in mind the number of data interfaces available, a subset of chunks (with center frequencies and chunk sizes that are greater or equal to 802.11ac channel widths (20/40/80 MHz)) are selected from all possibilities with the aim of maximizing the minimum channel-level SINR for each chunk. This may result in picking smaller sized chunks (with higher SINR yielding TVWS channels) instead of larger chunks with lower average SINR. Note that the power level chosen for a chunk is limited by the lowest power allowed among the constituent TVWS channels. So by maximizing the minimum SINR in a chunk and across both link directions, we address the power asymmetry effect highlighted in §2. We let the modulation and coding scheme (MCS) be automatically adapted by the 802.11 interfaces via the default built-in mechanism. Note that transmission over UHF spectrum via the conversion substrate is oblivious to MCS used by endpoint interfaces.

4.2.2 Slot Allocation. As noted earlier, we design WhiteHaul links to operate in TDD mode. This not only provides more flexibility and control for traffic scheduling across interfaces but also is essential

[1] SPLAT parameters (e.g., terrain profile, climate conditions, earth conductivity) are fixed, and input to the interface configuration module at the time of deployment.

given our choice to use to low-cost COTS 802.11 cards based on CSMA. Concerning the latter, the root of the issue is that in 802.11 each node senses the spectrum for a certain time period given by DIFS (34μs) before attempting to transmit. When the link distance is longer than 10.2Km, one end of the link cannot hear the other within this period, which then leads to collisions from potential simultaneous transmissions by both ends. Indeed, this observation has motivated the shift to time division multiplexing (TDM) based MAC protocols in the literature for long-distance Wi-Fi (e.g., [70]).

While our TVWS conversion substrate with separate transmit and receive paths allows for fast switching between the transmit and receive modes (35ns), the key issue with TDD is deciding on the time allocation for each direction (forward/reverse). The long-distance Wi-Fi literature (e.g., [26, 68, 70]) has taken the simplest approach to this issue by using static equal time share in both directions but we observe this is not an efficient approach for backhaul and TVWS settings. First, as shown earlier in §2, backhaul traffic can be highly asymmetric, with more traffic in the downstream (access network) direction, and also varies over time. Second, with TVWS spectrum, power asymmetry and interference effects may result in different effective capacities between forward and reverse directions. The above suggests an adaptive time split to counter these effects.

In view of the above, we seek a dynamic time allocation for forward and reverse link directions, driven by traffic demand and effective sub-link capacities (latter obtained via the Interface Configuration module). Note that each sub-link here corresponds to each of the underlying data interfaces. We adapt the time allocation for each direction at a coarser time granularity of *epochs*. The epoch duration in our implementation is chosen to be 30s to match with traffic variability; effective capacities of sub-links vary even less frequently in our setting. However note that WhiteHaul design is agnostic to the particular choice for epoch duration. Within each epoch, to avoid undesirable TCP effects we switch the use of the link between forward and reverse directions at a finer time granularity of *slots* (20ms in our implementation as per [70]). Allocation of slots for each direction within an epoch is proportional to the relative time allocation at the epoch level between forward (`Master-Slave`) and reverse (`Slave-Master`) directions. Fig. 10 shows an example where two thirds (one third) of the slots within each epoch are allocated to the forward (reverse) direction. The decision on relative time allocation between forward and reverse directions is made at the `Master` based on locally available information and that obtained from the `Slave` over the control interface. It is worth noting that the time allocation decision is made after configuring the interfaces with the available spectrum chunks. The key idea behind our approach is to choose the fraction of time allocated to forward and reverse directions in an epoch so that the total backlogged traffic in both directions is minimized. More precisely, at the start of each epoch t, the `Master` solves the following optimization problem:

$$
\begin{aligned}
&\max_{t_F, t_R} \left(\frac{\sum_{i \in F} \theta_i(t) \times t_F + \sum_{j \in R} \theta_j(t) \times t_R}{V_F(t) + V_R(t)} \right) \text{ s. t.} \\
&\sum_{i \in F} \theta_i(t) \times t_F \leq V_F(t) \quad \forall i \in F \\
&\sum_{j \in R} \theta_j(t) \times t_R \leq V_R(t) \quad \forall j \in R \\
&t_F + t_R = 1 \qquad AND \qquad t_F, t_R > 0 \qquad\qquad (1)
\end{aligned}
$$

Where $\theta_i(t)$ is the effective sub-link capacity via interface i at the start of epoch t. t_F and t_R are the fractions of time within the upcoming epoch allocated for the forward and reverse directions respectively. $V_F(t)$ and $V_R(t)$ are, respectively, traffic volumes in forward and reverse directions for the upcoming epoch t. Each of these traffic volumes are the sum of forecasted traffic demand in the coming epoch t and backlogged traffic carrying over from the previous epoch. We consider a simple forecasting approach of assuming that the demand in the coming epoch will be same as in the previous epoch, which is justified from the results in §2 for small epoch durations (30s in our case).

4.2.3 Traffic Scheduling. The TVWS conversion substrate in the WhiteHaul hardware layer described earlier in this section provides the physical capability to aggregate TVWS spectrum across multiple spectrum chunks over different interfaces. But translating this capability to higher layer aggregated data rates requires a way to concurrently use multiple interfaces and distribute backhaul traffic among them. We seek to realize this aggregation in a transparent manner to end-user traffic so as to give the *abstraction of a high-speed link-level tunnel* through which user traffic is transported across. We find Multipath TCP (MPTCP) a natural fit for our use case to meet the above mentioned requirements. Moreover, it offers a reliable bit pipe like TCP while also handling packet reordering from striping user traffic across multiple interfaces. It also automatically adapts to any changes to underlying sub-link capacities or slot durations. In contrast, the other multipath tunnelling solutions (e.g., MLVPN [6] that is based on the TUN/TAP technique [11]) that operate over multiple UDP or TCP sessions simply do not meet our key requirements; e.g., quickly adapting to sub-link capacities, handling packet reordering, and re-transmitting lost packets within and across sub-links.

However the default MPTCP with a coupled congestion control algorithm (e.g., LIA [71], OLIA [53]) is not suitable for our purpose as its main focus is on shifting traffic from congested to less congested paths to preserve network fairness among competing TCP flows. We have a single logical flow that needs to efficiently utilize the available capacity to maximize aggregate link data rate. Even the uncoupled variant with the commonly used CUBIC [36] fails to quickly and accurately track the changes in time split between link directions, sub-link capacities and/or losses, as we demonstrate in §5.

Motivated by the above, we propose a new cross-layer and uncoupled congestion control algorithm that is tailored for MPTCP use in WhiteHaul. Information from the Slot Allocation Manager is used to dynamically adjust the congestion window (`cwnd`) size of each individual subflow (bound to a different interface) according to its current effective capacity. This allows WhiteHaul MPTCP to exploit the available capacity of each subflow even as it varies over time. The *Slot Allocation Manager* periodically queries the rate adaptation module in each of the underlying COTS Wi-Fi cards for the current data rate used. At the start of each epoch, the *Slot Allocation Manager* sends an advisory signal to the MPTCP congestion control module (through Linux filesystem, sysfs) indicating the estimated effective capacity of each subflow, determined as the product of corresponding sub-link's effective capacity (the average data rate of the sub-link over epoch duration) and the fraction of time allocated to the endpoint in question. MPTCP then calculates a target cwnd value for each subflow by multiplying the advertised effective capacity

Algorithm 1: WhiteHaul MPTCP Congestion Control

```
   input: Ω_i, is the target window for subflow_i
   input: λ_i, is the delay budget for subflow_i
   input: α_i, is the increase weight for cwnd_i
1  CongestionAvoidance(subflow_i)
2      if A new window of data begins then
3          aqd ← AvgQueueingDelay(subflow_i)
           /* Update α_i                                  */
4          if aqd ≤ λ_i then
5              if cwnd_i ≥ Ω_i then
6                  α_i ← 1
7              else
8                  α_i ← 100
9          else
10             α_i ← 1
           /* Periodic cwnd reduction                      */
11         if aqd > λ_i & cwnd_i > Ω_i then
12             cwnd_i ← Ω_i
13             ssthresh_i ← cwnd_i
       /* Response to new ACKs                             */
14     if cwnd_i ≥ ssthresh_i then
15         cwnd_i ← cwnd_i + α_i/cwnd_i
16     else
17         tcpInSlowStart()
   /* Response to 3 duplicate ACKs                         */
18 DecreaseCWND(subflow_i)
19     cwnd_i ← max(cwnd_i − (cwnd_i * 0.1), 2)
```

and minimum RTT estimate it holds[2]. The *Slot Allocation Manager* realizes a time synchronous mechanism (aided by GPS based clock synchronization across link endpoints over the control interface) that holds the packets or passes them to the underlying layer based on the current time slot schedule (forward or reverse time slot). Note that we do not change the underlying COTS Wi-Fi cards nor their rate adaptation module; however, we adapt the sending rate over each subflow through careful adjustment of the corresponding congestion window. WhiteHaul initially increases cwnd rapidly (100 packets per RTT in our implementation) until it reaches the target value. Thereafter, WhiteHaul keeps increasing cwnd slowly (one packet per RTT) while monitoring the queuing delay. When the queuing delay exceeds a certain delay budget, cwnd is reduced back to the target value. In this way, WhiteHaul prevents self-inflicted packet losses by controlling the queue occupancy of network interfaces. Additionally, even when a packet is dropped for any reason (e.g., low SINR) WhiteHaul is designed to quickly return back to its target, ensuring that the subflow capacity is fully used.

Algorithm 1 highlights the key part of our algorithm, which we implemented in the Linux Kernel. In short, when a new ACK is received on subflow (i), which begins a new window of data, the average queuing delay (aqd) is computed (lines 2-3). Thereafter, the weight factor (α_i), which dictates the increase rate of $\mathtt{cwnd}_i$, is updated (lines 4-10). The value of α_i becomes large when $\mathtt{cwnd}_i$ is lower than the target window (Ω_i) and also the average queuing delay is smaller than the queuing budget (λ_i), ensuring that during the congestion avoidance phase (lines 14-15) $\mathtt{cwnd}_i$ is increased rapidly until it reaches Ω_i. Otherwise, $\mathtt{cwnd}_i$ is increased slowly (similar to standard TCP). At the beginning of every window of data, WhiteHaul examines whether it is allowed to reset $\mathtt{cwnd}_i$ to Ω_i (lines 11-13), making sure that the queue occupancy of underlying wireless interface of subflow (i) stays below a certain threshold. Finally, when a loss is detected on a subflow, by receiving three duplicate ACKs, WhiteHaul reduces cwnd gently (lines 18-19).

5 EVALUATION

In this section, we present a wide ranging evaluation of WhiteHaul, assessing all aspects of its design. This includes: (i) benchmarking its spectrum aggregation and conversion efficiency relative to the ideal case using our prototype implementation; (ii) experimentally evaluating WhiteHaul MPTCP using our Linux kernel implementation to compare its congestion control algorithm with commonly used OLIA (coupled) and CUBIC (uncoupled) algorithms; (iii) additional measurement based and simulation studies (mostly driven by real-world traffic traces) to evaluate other aspects of WhiteHaul.

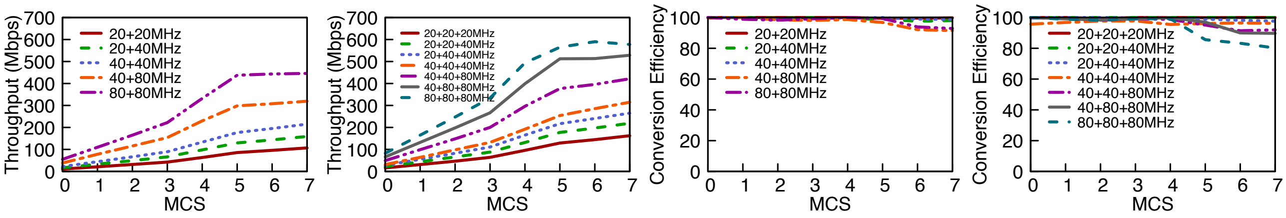
Compliance with regulations. Besides having to periodically query a geolocation database for available TVWS channels and their allowed power levels at the operating location, another key requirement for white space devices is to comply with the transmit spectrum mask related regulations. The latter is to ensure power leakage into adjacent channels under the prescribed limit so as not to cause interference to incumbents. To verify that our WhiteHaul prototype meets this requirement, we measured its out-of-band (OOB) emissions at different 802.11ac channel widths it can be configured to (20/40/80 MHz) using a spectrum analyzer (from Keysight) and find that OOB from our prototype is well within the limit specified by ETSI [28] for Class 1 TVWS devices (the relevant class for our outdoor and backhaul use case). For instance, with 80MHz channel, the OOB EIRP spectral density (P_{OOB}) with WhiteHaul is -103.8 dBm/Hz for the adjacent 100 KHz spectrum, which is as per Class 1 device requirements. We omit the detailed results with spectrum analyzer screenshots due to space restrictions.

Spectrum conversion and aggregation efficiency. We now study aggregate TCP throughput obtained with WhiteHaul in all possible two and three interface combinations of 802.11ac channel widths supported by our interface cards (20, 40 and 80MHz[3]). The flexibility with WhiteHaul allows us to down-convert any 5GHz channel to TV band by appropriately configuring frequency of the SDR based LO.

Our experimental setup consists of two Intel i7 machines (7567U processor at 3.5GHz, 8GB of RAM) that run the software layer (see Fig. 7) of two WhiteHaul nodes. Both machines have Ubuntu 16.04 with our modified MPTCP Linux kernel implementation (as described in §4.2.3). Each machine is connected to a WhiteHaul TVWS conversion substrate through a GbE interface. The two WhiteHaul nodes making up the endpoints of the link under test are placed in the same lab few meters apart. We use the unoccupied 5GHz Wi-Fi spectrum in our environment in the range of 5500-5750MHz. For traffic generation, we use *iperf* to generate TCP traffic between the endpoint

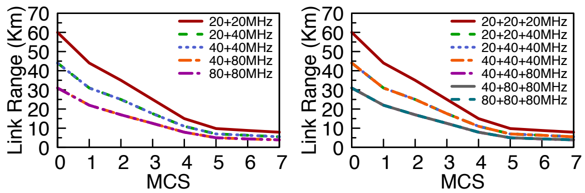
[2] The minimum RTT estimate is the smallest observed instantaneous RTT estimate obtained through ACK packets.

[3] WhiteHaul can use 80+80MHz non-contiguous channels; however, this mode is not widely supported in practice including the cards we used for our prototype.



(a) (b) (c) (d)

Figure 11: (a) and (b): WhiteHaul aggregate TCP throughput performance in all two and three interface scenarios. (c) and (d): WhiteHaul conversion and aggregation efficiency as percentage of best case in all two and three interface scenarios.



(a) (b)

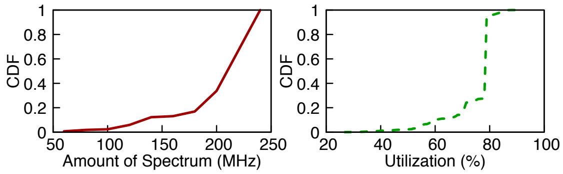
Figure 12: Relationship between link range and MCS for different channel widths combinations assuming free-space pathloss and receive sensitivity values with commodity 802.11ac cards.

machines. Note that this setup reflects the whole WhiteHaul system except for the control interface part, which is evaluated separately. We have experimented with two variants of this setup, one with two and the other with three 802.11ac interfaces at each endpoint.

Figs. 11a and 11b, respectively, show the throughput results with WhiteHaul for two and three interface scenarios. Here MPTCP throughput achieved using the above setup with WhiteHaul after down conversion and aggregation is reported. In effect, each interface ends up using a TVWS spectrum chunk of the specified width. We see that WhiteHaul can provide maximum throughput up to 446Mbps when using two 80MHz channels. And with three interfaces using three 80MHz channels packed together with little channel spacing (aggregation of 240MHz spectrum), it provides up to 590Mbps aggregated throughput; and up to 640Mbps with adequate channel separation (not shown here). To put these achievable throughput results into perspective, we show in Fig. 12 the relationship between estimated link range (in Kms) and MCS, based on receive sensitivity values for the 802.11ac cards in our prototype and assuming a free space pathloss (reflecting a best case link deployment scenario with LoS propagation), for different channel width combinations. With multiple channel widths are used together, the wider channels limit the range. This allows us to infer the achievable range for each of the various scenarios considered in Figs. 11a and 11b, thereby provide insight into rate-range tradeoffs. For example, considering a two interface scenario using MCS 7, the combination of 40+40MHz provides a throughput of 215Mbps up to a range of 5.5Km, whereas 40+80MHz combination reduces the maximum range to 3.9Km but with a higher achievable throughput of 320Mbps.

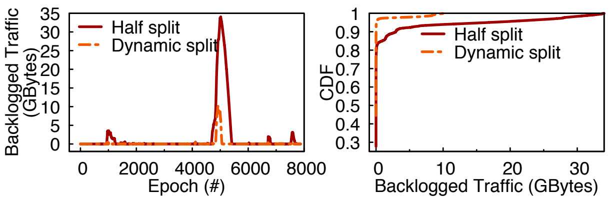
We benchmark these throughput results against the best case overall throughput, which is computed as the sum of maximum TCP throughputs achievable per interface for a given channel width and MCS in 5GHz band in isolation, i.e., without any frequency conversion or aggregation. The corresponding throughputs obtained with WhiteHaul as a percentage of the best case overall throughput, referred to as "Conversion Efficiency", are shown in Figs. 11c and 11d. This metric characterizes the effectiveness of frequency conversion and aggregation in WhiteHaul by comparing the throughput it achieves against the ideal case. We observe that in all combinations of 20 and 40MHz channels, WhiteHaul TVWS conversion substrate and aggregation technique achieves 99% efficiency across all MCS values. This is because of almost perfect frequency conversion without any signal distortion and effective aggregation with our MPTCP variant. When using even wider 80MHz channels with reduced inter-channel spacing, however, higher adjacent channel leakage lowers the efficiency. For low MCS values (up to MCS5), WhiteHaul is still able to achieve close to best performance with average efficiency of 98% (as shown in 80MHz scenarios of Figs. 11c and 11d). But the average efficiency drops down to 89% for higher MCS values (MCS6 and MCS7) and to the lowest level of 80% in the case of 3 immediately adjacent 80MHz channels. However, with adequate inter-channel separation, as would be the case in a practical deployment, leads to higher efficiency. Equally, more number of interfaces and smaller widths can provide high efficiencies (results not shown here).

Impact of using 802.11ac channel widths. As we rely on 802.11ac interfaces in WhiteHaul, the set of channel widths we can use are limited to those that come with 802.11ac (which in practice are 20/40/80MHz). We now examine the impact of restricting to these few channel widths on the use of available TVWS spectrum when aggregating individual narrower (6/8MHz) TVWS channels. We follow the methodology earlier used in §2. We first obtain the amount of available TVWS spectrum in rural locations of Scotland by querying a commercial geolocation database. Then, for each of those locations, we find how much of the available spectrum can be covered by any combination of 20/40/80MHz channel widths and supposing three 802.11ac interfaces at each WhiteHaul node. For example, suppose at a particular location, 14 TVWS channels (of size 8MHz each) are available for WhiteHaul system but they are fragmented into two chunks of 56MHz each. The total amount of available spectrum is 112MHz, which can be partially utilized by two 802.11ac radios configured with 40MHz each, and frequency converted to TVWS spectrum. Thus, the estimated utilization for that setup is (80/112) or 71.4%. In other words, the utilization represents the ratio between the actual used spectrum by the 802.11ac radios to the total amount of available TVWS spectrum, expressed as a percentage. Results show that with 3 interfaces: more than 200MHz of spectrum can be utilized in more than 70% of the locations (Fig. 13a); and this is equivalent to nearly 80% utilization (Fig. 13b). The utilization can be improved further with additional interfaces and/or using smaller



(a) (b)

Figure 13: The amount and percentage of utilized spectrum for the case with three 802.11ac interfaces.



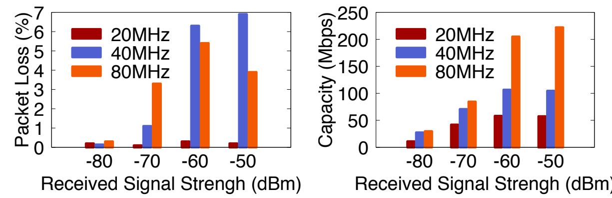
(a) (b)

Figure 14: Performance benefit with WhiteHaul dynamic time slot allocation compared to commonly used static, half-split approach: (a) the amount of backlogged traffic per epoch for a one week period; (b) CDFs of backlogged traffic.

channel widths. Overall, these results indicate that using 802.11ac interfaces and their widths only have a marginal negative impact in being able to fully exploit the available TVWS spectrum.

LoRa based control interface. Here we present our experimental assessment of using LoRa based control interfaces for coordination between WhiteHaul link endpoints. To this end, we experimentally study the data rates and reliability with LoRa under wide range of radio conditions in a lab environment. For reliability, we used packet reception rate (PRR), defined as percentage of successful packet reception, as the measure. Recall that the coordination between the `Master` and `Slave` over the LoRa control interface is driven from the `Master` end, i.e., the `Master` either consults the `Slave` to retrieve the spectrum analysis information, or instructs the `Slave` to change some configurations. This makes collisions unlikely. We find that LoRa based communication is reliable with PRRs between 99.8% to 100% across a diverse range of link qualities (RSSIs ranging from -115dBm to -70dBm). And it achieves a data rate of $2.8Kbps$ even with the very low RSSI of -115dBm. To put these results into perspective, note that WhiteHaul endpoints need to exchange control information every few tens of seconds to aid decision making at the `Master` every epoch (30s). And the size of the control commands from the `Master` to the `Slave` are quite small (few tens of bytes). The largest control message size in WhiteHaul is the spectrum sensing information from `Slave` to the `Master` with a payload size of 797 bytes but this can be exchanged at the timescale of tens of minutes. With a duty cycle of 0.8%, even this message can be delivered 2.3 seconds at the achievable data rate every 5 minutes, and other smaller messages in less than a second.

Slot allocation. Here we evaluate the dynamic time slot allocation between forward and reverse directions in WhiteHaul in comparison with static, equal time split baseline. We use the total backlogged traffic carried forward from epoch to epoch as the metric. This evaluation is based on Monte Carlo simulations (1000 iterations per data point) using MATLAB, each spanning about 8000 epochs (30s long as per analysis in §2) and assuming two interfaces per WhiteHaul



(a) (b)

Figure 15: (a) Packet loss rate and (b) effective capacity at different link qualities (RSS values) and channel widths.

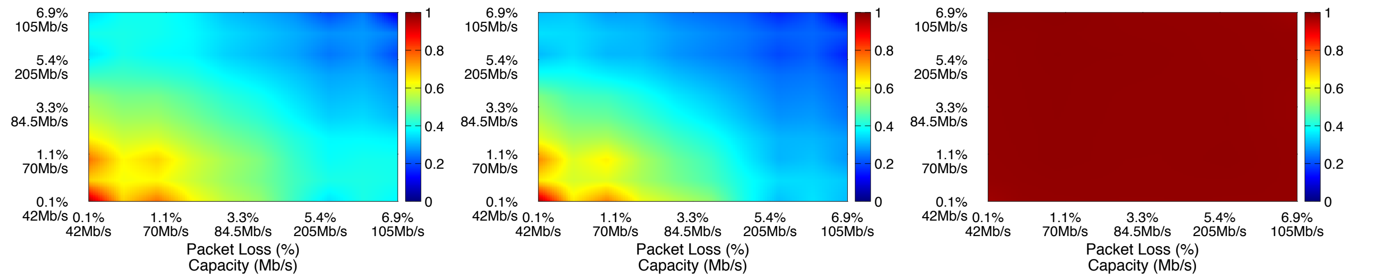
node. Effective sub-link capacities are randomly generated within the range of 5 to $60Mbps$ for this case but traffic volumes in forward and reverse directions are from the real-world network trace as described in §2. Fig. 14a indicates that the dynamic time allocation approach significantly reduces the maximum backlogged traffic across the one week period from up to 30GB with half-split approach to less than 7GB. This improvement is also evident from the CDF plot in Fig. 14b where there is no backlogged traffic in nearly 90% of the cases with dynamic split as opposed to under 60% with half-split.

Traffic scheduling. Here we experimentally assess the effectiveness of the WhiteHaul MPTCP aggregation in a wide range of capacity and packet loss scenarios, relative to two MPTCP alternatives using coupled congestion control with OLIA [53] and uncoupled congestion control with CUBIC [36]. We approach this in two steps. We first characterize packet loss and effective capacity in different link quality conditions and with different channel widths. This characterization then is used as the basis for emulation based evaluation of WhiteHaul and other MPTCP alternatives.

Packet Loss and Capacity Characterization. Using wider channel bandwidths for high capacity TVWS backhaul links can potentially impact transmission range and cause higher susceptibility to interference due to distribution of the same power over a wider bandwidth [25]. These both manifest as increased packet loss rates, which can have adverse effect on most TCP based alternatives. We therefore conduct a controlled study to understand the nature of packet losses and effective capacities in a TVWS-based long distance setting. For this, we attenuate the transmitted TVWS signal using a combination of step attenuator and transmit power adjustment. The resulting received signal strength (RSS) values range from $-80dBm$ to $-50dBm$, reflecting increasing link distances.

To measure the loss rate and effective capacity, we use *iPerf* with CBR UDP traffic streams and also experiment with three different channel widths (20, 40 and 80MHz). As shown in Fig. 15(a), on narrow channels, e.g., $20MHz$, the link has very low loss rate, almost negligible, and the maximum loss rate observed across different RSS values was 0.3%. At higher channel width, packet losses increase up to 6.9% (with 40MHz channel). Packet loss differences between 40MHz and 80MHz channels are an artifact of the mechanics of underlying rate adaptation mechanism. On the other hand, the effective capacity, measured in terms of UDP throughput, shown in Fig. 15(b) is expected – increasing with increasing RSS and channel width.

WhiteHaul MPTCP in Diverse Conditions. We use the observed packet loss and effective capacity results from the above study to evaluate the performance of WhiteHaul MPTCP in different conditions. We use our Linux kernel implementation for WhiteHaul and compare it with MPTCP using CUBIC [36] and OLIA [53]. The setup for



(a) CUBIC **(b) OLIA** **(c) WhiteHaul**

Figure 16: Aggregation efficiency comparison between WhiteHaul MPTCP, uncoupled MPTCP with CUBIC [36] and coupled MPTCP with OLIA [53], in different conditions shown as a heatmap – red is best.

this experiment consists of two Linux desktops, each with two Gigabit Ethernet interfaces. We use Linux Network Emulator (NetEm) and Traffic Control (tc) to emulate both losses and sub-link capacities from the above characterization study. It is important to note that the following performance results reflect the case of WhiteHaul MPTCP deployment in the field because the used Netem configurations are chosen based on real-world losses and capacities observed with actual WhiteHaul hardware in the above study. For WhiteHaul, we empirically set the queuing delay budget (λ) as $50ms$ and the target `cwnd` size is set as described in §4.2.3. We use *iPerf* tool send uni-directional TCP traffic and measure the aggregate throughput across the two interfaces.

Fig. 16 displays the results for the three alternatives as heatmaps of efficiency, defined as the ratio of achieved throughput (in presence of losses) for a given alternative – WhiteHaul, CUBIC or OLIA – to the maximum achieved throughput (without losses) with the same link capacities. We observe that CUBIC in general performs better than OLIA in presence of losses. This can be attributed to the fact that with CUBIC each sub-flow behaves independently with a separate congestion window, so the impact of packet losses of one sub-flow does not affect the congestion window of the other sub-flow. Overall though, WhiteHaul algorithm clearly achieves superior performance, by an order of magnitude, compared to both CUBIC and OLIA. In some scenarios like in the upper right corner, WhiteHaul, CUBIC and OLIA achieve aggregate throughputs of 403.8Mbps, 31.6Mbps and 22.3Mbps, respectively. This significant improvement with WhiteHaul is due to two reasons. First, WhiteHaul relies on a periodic advisory signal from the slot allocation manager to keep track of the optimal target congestion window size and uses it to ramp up the rate instantly to that level and then increases linearly while keep monitoring the queuing delay. If the queuing delay exceeds the delay budget, the congestion window shrinks back again to the target value, preventing self-inflicting losses due to RTOs. Second, the WhiteHaul algorithm is robust in presence of losses to maintain sending rate close to sub-link capacities whereas CUBIC and OLIA have drastic responses to lost packets.

MPTCP use in WhiteHaul is transparent to end-to-end traffic flows (between user devices in the access network and the Internet – see schematic in Fig. 6) as it seeks to create a link-level tunnel abstraction. We evaluated the effectiveness of WhiteHaul to this end through an experiment varying number of end-to-end TCP flows. In this experiment, the WhiteHaul link consists of three sub-links respectively

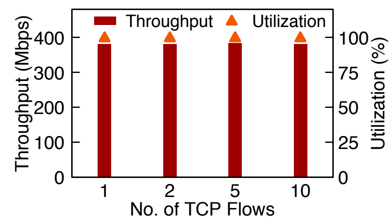


Figure 17: Effectiveness of WhiteHaul MPTCP link-level tunnel with varying number of end-to-end TCP flows.

using 20, 40 and 80MHz chunks and aggregate capacity of 385Mbps. Results (Fig. 17) confirm that WhiteHaul link throughput and utilization stay unaffected by the number of end-user TCP flows, indicating no undesirable interactions between them and WhiteHaul MPTCP.

Cost analysis. We close this section with a detailed cost analysis of WhiteHaul, which is primarily CAPEX, as the OPEX cost includes small nominal fee to access a commercial geolocation database. Table 1 provides a breakdown of WhiteHaul link CAPEX costs. We assume a LimeSDR based LO. The basic version consists of one interface per WhiteHaul node making up a link and costs about 1600 USD. This version can aggregate up to 10 contiguous TVWS channels (80MHz spectrum) and provide a throughput up to 240Mbps based on results from earlier in this section. Some items are per node and unaffected by the number of interfaces per node (combiners/splitters, antennas and compute platforms) while the rest scale with the number of interfaces. With 3 802.11ac interfaces per node, WhiteHaul link costs around 4500 USD and can deliver nearly 600Mbps. For a marginal additional cost, we can use dual polarized Yagi antennas to get the MIMO benefits and double the achievable throughput to over 1Gbps. Cost of our solution is comparable to existing commercial TVWS solutions (e.g., RuralConnect by Carlson Wireless, ACRS2 B1000 Adaptrum, GWS5002 by 6Harmonics, and XR by Redline) that cost around 4500-5000 USD but with an order-of-magnitude or higher throughput. The licensed commercial microwave solutions provide higher link capacity but also come with spectrum licensing cost which could be up to several tens of thousands of dollars per annum [66]. On the other hand, unlicensed commercial microwave solutions (e.g., Ubiquiti AirFiber-24-HD) costing around 3500 USD per link only operate in LoS conditions, and over relatively shorter distances. In NLoS conditions, both microwave and mmWave backhaul solutions suffer from poor performance, latter also sensitive to weather conditions. In contrast, WhiteHaul, thanks to the superior propagation characteristics in the UHF band, is better suited for

Table 1: WhiteHaul link CAPEX costs

Item		Cost (USD)
TVWS Conversion Substrate (per interface per node)	2 IEEE 802.11ac interfaces	200
	2 SDR-based LO	600
	RF Accessories (Amp, Attenuators, filters, etc.)	550
2 Combiner/Splitter		270
2 TVWS Yagi antennas		160
2 Compute Platforms (Raspberry Pis)		75
Link cost with one interface per node		1585
Link cost with two interfaces per node		3205
Link cost with three interfaces per node		4555

NLoS settings; this can lead to cost savings through reduction in required tower heights or number of relays.

6 RELATED WORK

TVWS Spectrum Aggregation. In IEEE 802.22 [22], the early TVWS standard targeting rural and remote areas, only a single (6/8MHz) TVWS channel is used. In the more recent 802.11af standard framework that is aligned with regulatory requirements to use geolocation databases, aggregation up to 4 contiguous or non-contiguous TVWS channels (i.e., up to 32MHz) is allowed [32]. The various commercial TVWS solutions that currently exist (some compliant with 802.11af) support aggregation of at most 2-3 contiguous TVWS channels [1–3, 8] with a total bandwidth less than 20MHz. In the research literature, Holland [42] analyzes the achievable capacity with aggregating available TVWS channels with results showing promising theoretical data rate improvements but does not discuss how such rates can be achieved by a practical system.

From the systems perspective, there exists a long line of research prototyping and experimenting with TVWS platforms that use commodity 802.11 cards via frequency down-conversion [16, 43, 45, 49, 56, 62, 63]. These systems are only capable of limited TVWS spectrum aggregation (up to 4 6MHz TVWS channels) and provide a maximum TCP (UDP) throughput of 23Mbps (35Mbps). In contrast, our WhiteHaul system, while leveraging the same frequency down-conversion concept, achieves 500Mbps+ TCP link throughput by enabling efficient aggregation of an order-of-magnitude more (even non-contiguous) TVWS spectrum. By focusing on the PtP backhaul use case, our work also complements the extensive TVWS systems research till date aimed at several other use cases (e.g., Wi-Fi like wireless LANs [17, 29, 89, 90], TVWS based point-to-multipoint (PtMP) for agriculture, sensor, IoT and other applications [76, 78, 86]).

Support for Spectrum Aggregation in Other Wireless Standards. The IEEE 802.11 wireless LAN (WLAN) standards have supported channel bonding (contiguous channel aggregation) since 2009 when 802.11n introduced 40MHz bonded channels as a feature to provide high user/network data rates. This continued with the subsequent and current 802.11ac standard [19], which additionally mandates support of 80MHz channels and operation exclusively in the 5GHz band. 802.11ac also has optional support for contiguous 160MHz or non-contiguous aggregation of two 80MHz wide channels. Although the focus of 802.11 is on the WLAN setting involving communication between an AP and associated clients, equivalent to a PtMP scenario, and is largely limited to aggregation of contiguous spectrum, we leverage the 802.11 channel aggregation feature as an element in the design of WhiteHaul towards a low-cost PtP backhaul solution. On the other hand, spectrum aggregation, or more precisely carrier aggregation (CA), has also been an integral part of 3GPP cellular network standards since 2011 (Release 10) when LTE-Advanced (LTE-A) was introduced [58, 81]. The essential idea behind the CA is to aggregate two or more component carriers (CCs) of different bandwidths (between 1.4MHz and 20MHz) up to 100MHz with five CCs of 20MHz. The target usage scenario for CA is different from ours on PtMP and the focus there is mainly on how to select the CCs and schedule resource blocks to users across CCs. CA also limits the amount of spectrum that can aggregated to 100MHz. Other physical layer aggregation approaches have been proposed in [46, 84] to utilize multiple spectrum fragments using one radio and incorporating frequency reshaping techniques. Their target use case is again different with focus on enabling coexistence between different technologies in narrow shared spectrum. Moreover, the use of a single radio limits the aggregated spectrum to 40MHz in these works.

Higher Layer Aggregation Approaches. Other aggregation techniques exist at higher layers, such as MPTCP at transport layer [33, 34], Multipath QUIC (MP-QUIC) [24, 87] at application/transport layer, and LTE-WLAN Aggregation (LWA) at LTE PDCP layer [14]. Different from the lower layer techniques, these approaches can be technology agnostic which allows more flexible aggregation not only between transceivers of the same technology but also across different technologies. In recent years, there have been several traffic scheduling algorithms proposed in this context targeting heterogeneous settings and considering different characteristics (bandwidth, latency, etc.) [30, 35, 38, 55, 59]. In contrast to the above works, our WhiteHaul system takes a cross-layer/hybrid aggregation approach combining the lower layer aggregation capability with a higher layer technique, the latter in our case realized using MPTCP variant with new tailored uncoupled congestion control algorithm.

7 CONCLUSIONS

We have presented WhiteHaul, a cost-effective and efficient system for high-speed backhaul over TVWS spectrum. WhiteHaul is the first TVWS system that is capable of aggregating the whole UHF band, utilizing both the contiguous and non-contiguous chunks of available spectrum. WhiteHaul is made up of innovations across both hardware and software. The hardware layer is represented by a conversion substrate that leverages SDR-based local oscillator to down/up-convert between available TVWS spectrum and 5GHz Wi-Fi channels, and to effectively combine non-contiguous chunks of TVWS spectrum. The WhiteHaul software layer orchestrates the underlying COTS Wi-Fi interfaces to maximize throughput by efficiently distributing traffic among them. It features a novel use of MPTCP along with a novel cross-layer congestion control algorithm to quickly and efficiently utilize available capacity across sub-links. Our extensive lab evaluations of WhiteHaul using a prototype implementation we developed show that WhiteHaul can aggregate almost the whole TV band and can achieve nearly 600Mbps TCP throughput using a single antenna, and its MPTCP component outperforms the state-of-the-art by an order of magnitude. Our future work will focus on real-world trials of WhiteHaul as well as examining its wider applicability beyond the TVWS spectrum.

ACKNOWLEDGMENT

We thank our shepherd Mariya Zheleva and the anonymous reviewers for their helpful suggestions on improving the paper. We also thank Frankie Garcia for access to the Keysight spectrum analyzer.

REFERENCES

[1] *6Harmonics*. http://www.6harmonics.com/.
[2] *Adaptrum*. http://www.adaptrum.com/.
[3] *Carlson Wireless Technologies*. https://www.carlsonwireless.com/.
[4] *Facebook Connectivity*. https://connectivity.fb.com/high-altitude-connectivity/.
[5] *FSOC*. https://x.company/projects/fsoc/.
[6] *MLVPN*. https://mlvpn.readthedocs.io/en/latest/what_is_mlvpn.html.
[7] *Pycom - Next Generation Internet of Things Platform*. https://pycom.io/.
[8] *Redline Communications*. https://rdlcom.com/tv-white-space/.
[9] *RF Explorer*. http://j3.rf-explorer.com/.
[10] *Tegola Rural Community Wireless Access Network*. http://www.tegola.org.uk/networks.html.
[11] *Universal TUN/TAP device driver*. https://www.kernel.org/doc/Documentation/networking/tuntap.txt.
[12] *USRP B210*. http://www.ettus.com/all-products/UB210-KIT/.
[13] Measuring the Information Society Report 2018, 2018. https://www.itu.int/en/ITU-D/Statistics/Pages/publications/misr2018.aspx/.
[14] 3GPP:REL.14. Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification. Retrieved July, 2018 from https://www.etsi.org/deliver/etsi_ts/136300_136399/136323/14.03.00_60/ts_136323v140300p.pdf.
[15] Anand, A., Pejovic, V., Belding, E. M., and Johnson, D. L. VillageCell: Cost effective cellular connectivity in rural areas. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (2012), ACM, pp. 180–189.
[16] Ashok, A., Subbiah, I., Varga, G., Schrey, M., Achtzehn, A., Petrova, M., and Heinen, S. Whitelan: Facilitate cost-efficient SDR research with COTS IEEE 802.11 b/g devices. In *ACM SIGCOMM Computer Communication Review* (2014), vol. 44, ACM, pp. 45–52.
[17] Bahl, P., Chandra, R., Moscibroda, T., Murty, R., and Welsh, M. White space networking with wi-fi like connectivity. *ACM SIGCOMM Computer Communication Review 39*, 4 (2009), 27–38.
[18] Baig, G., Alistarh, D., Karagiannis, T., Radunovic, B., Balkwill, M., and Qiu, L. Towards unlicensed cellular networks in TV white spaces. In *Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies* (2017), ACM, pp. 2–14.
[19] Bejarano, O., Knightly, E., and Park, M. IEEE 802.11ac: From Channelization to Multi-User MIMO. *IEEE Communications 51*, 10 (Oct 2013).
[20] Biggs, P., and Lozanova, Y. THE STATE OF BROADBAND 2017: BROADBAND CATALYZING SUSTAINABLE DEVELOPMENT, September 2017. https://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-BROADBAND.18-2017-PDF-E.pdf.
[21] Chandra, R., Moscibroda, T., Bahl, P., Murty, R., Nychis, G., and Wang, X. A campus-wide testbed over the TV white spaces. *ACM SIGMOBILE Mobile Computing and Communications Review 15*, 3 (2011), 2–9.
[22] Cordeiro, C., Challapali, K., Birru, D., and Shankar, S. 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.* (2005), IEEE, pp. 328–337.
[23] De Coninck, Q., Baerts, M., Hesmans, B., and Bonaventure, O. A first analysis of multipath TCP on smartphones. In *International Conference on Passive and Active Network Measurement* (2016), Springer, pp. 57–69.
[24] De Coninck, Q., and Bonaventure, O. Multipath QUIC: Design and Evaluation. In *Proceedings of the 13th International Conference on Emerging Networking EXperiments and Technologies* (2017), ACM, pp. 160–166.
[25] Deek, L., Garcia-Villegas, E., Belding, E., Lee, S.-J., and Almeroth, K. The impact of channel bonding on 802.11 n network management. In *Proceedings of the Seventh COnference on emerging Networking EXperiments and Technologies* (2011), ACM, p. 11.
[26] Djukic, P., and Mohapatra, P. Soft-TDMAC: A software TDMA-based MAC over commodity 802.11 hardware. In *INFOCOM 2009, IEEE* (2009), IEEE, pp. 1836–1844.
[27] Ericsson,. Ericsson Mobility Report June 2019, June 2019. https://www.ericsson.com/en/mobility-report/reports/june-2019.
[28] ETSI. White Space Devices (WSD) Wireless Access Systems operating in the 470 MHz to 790 MHz TV broadcast band, 2018. ETSI EN 301 598 V2.1.1.
[29] Feng, X., Zhang, J., and Zhang, Q. Database-assisted multi-AP network on TV white spaces: Architecture, spectrum allocation and AP discovery. In *2011 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)* (2011), IEEE, pp. 265–276.
[30] Ferlin, S., Alay, Ö., Mehani, O., and Boreli, R. BLEST: Blocking estimation-based MPTCP scheduler for heterogeneous networks. In *IFIP Networking Conference (IFIP Networking) and Workshops, 2016* (2016), IEEE, pp. 431–439.
[31] Flickenger, R., Okay, S., Pietrosemoli, E., Zennaro, M., and Fonda, C. Very long distance Wi-Fi networks. In *Proceedings of the second ACM SIGCOMM workshop on Networked systems for developing regions* (2008), pp. 1–6.
[32] Flores, A. B., Guerra, R. E., Knightly, E. W., Ecclesine, P., and Pandey, S. IEEE 802.11 af: A standard for TV white space spectrum sharing. *IEEE Communications Magazine 51*, 10 (2013), 92–100.
[33] Ford, A., Raiciu, C., Handley, M., Barre, S., Iyengar, J., et al. Architectural guidelines for Multipath TCP development. *IETF, Informational RFC 6182* (2011), 2070–1721.
[34] Ford, A., Raiciu, C., Handley, M., and Bonaventure, O. TCP extensions for multipath operation with multiple addresses. *Internet Engineering Task Force* (2013).
[35] Guo, Y. E., Nikravesh, A., Mao, Z. M., Qian, F., and Sen, S. Accelerating multipath transport through balanced subflow completion. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking* (2017), ACM, pp. 141–153.
[36] Ha, S., Rhee, I., and Xu, L. CUBIC: a new TCP-friendly high-speed TCP variant. *ACM SIGOPS operating systems review 42*, 5 (2008), 64–74.
[37] Harrison, K., Mishra, S. M., and Sahai, A. How much white-space capacity is there? In *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on* (2010), IEEE, pp. 1–10.
[38] Hartung, L., and Milind, M. Policy driven multi-band spectrum aggregation for ultra-broadband wireless networks. In *Dynamic Spectrum Access Networks (DySPAN), 2015 IEEE International Symposium on* (2015), IEEE, pp. 82–93.
[39] Hasan, S., Barela, M. C., Johnson, M., Brewer, E., and Heimerl, K. Scaling community cellular networks with CommunityCellularManager. In *16th USENIX Symposium on Networked Systems Design and Implementation (NSDI 19)* (2019), pp. 735–750.
[40] Hasan, S., Heimerl, K., Harrison, K., Ali, K., Roberts, S., Sahai, A., and Brewer, E. GSM Whitespaces: An opportunity for rural cellular service. In *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)* (2014), IEEE, pp. 271–282.
[41] Heimerl, K., Hasan, S., Ali, K., Brewer, E., and Parikh, T. Local, Sustainable, Small-Scale Cellular Networks. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1* (2013), ACM, pp. 2–12.
[42] Holland, O. Some are born with white space, some achieve white space, and some have white space thrust upon them. *IEEE Transactions on Cognitive Communications and Networking 2*, 2 (2016), 178–193.
[43] Holland, O., Aijaz, A., Ping, S., Wong, S., Mack, J., Lam, L., and de la Fuente, A. Aggregation in TV white space and assessment of an aggregation-capable IEEE 802.11 white space device. In *Communications (ICC), 2016 IEEE International Conference on* (2016), IEEE, pp. 1–5.
[44] Holland, O., Ping, S., Aijaz, A., Chareau, J.-M., Chawdhry, P., Gao, Y., Qin, Z., and Kokkinen, H. To white space or not to white space: That is the trial within the Ofcom TV white spaces pilot. In *2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)* (2015), IEEE, pp. 11–22.
[45] Holland, O., Raman, A., Sastry, N., Wong, S., Mack, J., and Lam, L. Assessment of a Platform for Non-Contiguous Aggregation of IEEE 802.11 Waveforms in TV White Space. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)* (2016), IEEE, pp. 1–5.
[46] Hong, S. S., Mehlman, J., and Katti, S. Picasso: flexible RF and spectrum slicing. *ACM SIGCOMM Computer Communication Review 42*, 4 (2012), 37–48.
[47] Ireland, T., Nyzio, A., Zink, M., and Kurose, J. The impact of directional antenna orientation, spacing, and channel separation on long-distance multi-hop 802.11 g networks: A measurement study. In *2007 5th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops* (2007), IEEE, pp. 1–6.
[48] John Magliacane. *SPLAT!: an RF Signal Propagation, Loss, And Terrain analysis tool*. http://www.qsl.net/kd2bd/splat.html.
[49] Johnson, D., Zlobinsky, N., Lysko, A., Lamola, M., Hadzic, S., Maliwatu, R., and Densmore, M. Head to Head Battle of TV White Space and WiFi for Connecting Developing Regions. In *e-Infrastructure and e-Services for Developing Countries: 8th International Conference, AFRICOMM 2016, Ouagadougou, Burkina Faso, December 6-7, 2016, Proceedings* (2017), vol. 208, Springer, p. 186.
[50] Johnson, D. L., and Mikeka, C. Bridging Africa's broadband divide. *IEEE Spectrum 53*, 9 (2016), 42–56.
[51] Kassem, M. M., Marina, M. K., and Radunovic, B. DIY model for mobile network deployment: A step towards 5G for all. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies* (2018), pp. 1–5.
[52] Khalil, M., Qadir, J., Onireti, O., Imran, M. A., and Younis, S. Feasibility, architecture and cost considerations of using TVWS for rural Internet access in 5G. In *2017 20th Conference on Innovations in Clouds, Internet and Networks*

(ICIN) (2017), IEEE, pp. 23–30.

[53] Khalili, R., Gast, N., Popovic, M., and Le Boudec, J.-Y. MPTCP is Not Pareto-optimal: Performance Issues and a Possible Solution. *IEEE/ACM Trans. Netw. 21*, 5 (2013), 1651–1665.

[54] Kirkpatrick, K. Bringing the Internet to the (Developing) World. *Commun. ACM 61*, 7 (June 2018), 20–21.

[55] Kuhn, N., Lochin, E., Mifdaoui, A., Sarwar, G., Mehani, O., and Boreli, R. DAPS: Intelligent delay-aware packet scheduling for multipath transport. In *Communications (ICC), 2014 IEEE International Conference on* (2014), IEEE, pp. 1222–1227.

[56] Kumar, A., Karandikar, A., Naik, G., Khaturia, M., Saha, S., Arora, M., and Singh, J. Toward enabling broadband for a billion plus population with TV white spaces. *IEEE Communications Magazine 54*, 7 (2016), 28–34.

[57] Kumbhkar, R., Islam, M. N., Mandayam, N. B., and Seskar, I. Rate optimal design of a wireless backhaul network using TV white space. In *2015 7th International Conference on Communication Systems and Networks (COMSNETS)* (2015), IEEE, pp. 1–7.

[58] Lee, H., Vahid, S., and Moessner, K. A Survey of Radio Resource Management for Spectrum Aggregation in LTE-Advanced. *IEEE Communications Surveys & Tutorials 16*, 2 (2014).

[59] Lim, Y.-s., Nahum, E. M., Towsley, D., and Gibbens, R. J. ECF: An MPTCP path scheduler to manage heterogeneous paths. In *Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies* (2017), ACM, pp. 147–159.

[60] Liu, B., Yi, X., Yang, K., Liang, Z., Feng, G., Choi, P., Boon, C. C., and Li, C. A Carrier Aggregation Transmitter Front End for 5-GHz WLAN 802.11 ax Application in 40-nm CMOS. *IEEE Transactions on Microwave Theory and Techniques* (2019).

[61] MacCartney, G. R., and Rappaport, T. S. Rural macrocell path loss models for millimeter wave wireless communications. *IEEE Journal on selected areas in communications 35*, 7 (2017), 1663–1677.

[62] Mack, J., and Cartmell, J. Field trial results for a Wi-Fi based spectrum sharing technology in TVWS. In *IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014* (2014), IEEE, pp. 1–6.

[63] Narlanka, S., Chandra, R., Bahl, P., and Ferrell, J. I. A hardware platform for utilizing TV bands with a Wi-Fi radio. In *Local & Metropolitan Area Networks, 2007. LANMAN 2007. 15th IEEE Workshop on* (2007), IEEE, pp. 49–53.

[64] Nikravesh, A., Guo, Y., Qian, F., Mao, Z. M., and Sen, S. An in-depth understanding of multipath TCP on mobile devices: Measurement and system design. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking* (2016), pp. 189–201.

[65] OECD. International Cables, Gateways, Backhaul and International Exchange Points. *OECD Digital Economy Papers, No. 232, OECD Publishing, Paris* (2014).

[66] Ofcom. Fees for Spectrum Access 28 GHz Licences. https://www.ofcom.org.uk/__data/assets/pdf_file/0028/80938/28_ghz_fee_statement.pdf, Dec 2015.

[67] Ofcom. Airwaves opened up to support wireless revolution, July 2019. https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/airwaves-opened-up-to-support-wireless-revolution.

[68] Panigrahi, D., and Raman, B. TDMA scheduling in long-distance WiFi networks. In *IEEE INFOCOM 2009* (2009), IEEE, pp. 2931–2935.

[69] Park, C. S., Sundström, L., Wallén, A., and Khayrallah, A. Carrier aggregation for LTE-advanced: Design challenges of terminals. *IEEE Communications Magazine 51*, 12 (2013), 76–84.

[70] Patra, R. K., Nedevschi, S., Surana, S., Sheth, A., Subramanian, L., and Brewer, E. A. WiLDNet: Design and Implementation of High Performance WiFi Based Long Distance Networks. In *NSDI* (2007), vol. 1, p. 1.

[71] Raiciu, C., Handley, M., and Wischik, D. Coupled congestion control for multipath transport protocols. Tech. rep., IETF RFC 6356, Oct, 2011.

[72] Raman, B., and Chebrolu, K. Experiences in using WiFi for rural internet in India. *IEEE Communications Magazine 45*, 1 (2007), 104–110.

[73] Raul Katz,. The Impact of Broadband on the Economy: Research to Date and Policy Issues, April 2012. https://www.itu.int/ITU-D/treg/broadband/ITU-BB-Reports_Impact-of-Broadband-on-the-Economy.pdf.

[74] Raza, U., Kulkarni, P., and Sooriyabandara, M. Low Power Wide Area Networks: An overview. *IEEE Communications Surveys & Tutorials 19*, 2 (2017), 855–873.

[75] Reigadas, J. S., Martinez-Fernandez, A., Ramos-Lopez, J., and Seoane-Pascual, J. Modeling and optimizing IEEE 802.11 DCF for long-distance links. *IEEE Transactions on Mobile Computing 9*, 6 (2010), 881–896.

[76] Roberts, S., Garnett, P., and Chandra, R. Connecting Africa using the TV white spaces: From research to real world deployments. In *The 21st IEEE International Workshop on Local and Metropolitan Area Networks* (2015), IEEE, pp. 1–6.

[77] Saha, S. K., Kannan, A., Lee, G., Ravichandran, N., Medhe, P. K., Merchant, N., and Koutsonikolas, D. Multipath TCP in smartphones: Impact on performance, energy, and CPU utilization. In *Proceedings of the 15th ACM International Symposium on Mobility Management and Wireless Access* (2017), pp. 23–31.

[78] Saifullah, A., Rahman, M., Ismail, D., Lu, C., Liu, J., and Chandra, R. Low-Power Wide-Area Network Over White Spaces. *IEEE/ACM Trans. Netw. 26*, 4 (Aug. 2018), 1893–1906.

[79] Schmitt, P., Iland, D., Zheleva, M., and Belding, E. HybridCell: Cellular connectivity on the fringes with demand-driven local cells. In *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications* (2016), IEEE, pp. 1–9.

[80] Sevilla, S., Johnson, M., Kosakanchit, P., Liang, J., and Heimerl, K. Experiences: Design, Implementation, and Deployment of CoLTE, a Community LTE Solution. In *The 25th Annual International Conference on Mobile Computing and Networking* (2019), ACM, p. 45.

[81] Shen, Z., Papasakellariou, A., Montojo, J., Gerstenberger, D., and Xu, F. Overview of 3GPP LTE-Advanced Carrier Aggregation for 4G Wireless Communications. *IEEE Communications 50*, 2 (Feb 2012).

[82] Sheth, A., Nedevschi, S., Patra, R., Surana, S., Brewer, E., and Subramanian, L. Packet loss characterization in WiFi-based long distance networks. In *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications* (2007), IEEE, pp. 312–320.

[83] Talbot, D. *The Unacceptable Persistence of the Digital Divide*. MIT Technology Review, Dec 2016. https://www.technologyreview.com/s/603083/the-unacceptable-persistence-of-the-digital-divide/.

[84] Tan, K., Shen, H., Zhang, J., and Zhang, Y. Enable flexible spectrum access with spectrum virtualization. In *2012 IEEE International Symposium on Dynamic Spectrum Access Networks* (2012), IEEE, pp. 47–58.

[85] Van De Beek, J., Riihijarvi, J., Achtzehn, A., and Mahonen, P. TV white space in Europe. *IEEE Transactions on Mobile Computing 11*, 2 (2012), 178–188.

[86] Vasisht, D., Kapetanovic, Z., Won, J., Jin, X., Chandra, R., Sinha, S., Kapoor, A., Sudarshan, M., and Stratman, S. Farmbeats: An IoT platform for data-driven agriculture. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)* (2017), pp. 515–529.

[87] Viernickel, T., Froemmgen, A., Rizk, A., Koldehofe, B., and Steinmetz, R. Multipath QUIC: A deployable multipath transport protocol. In *2018 IEEE International Conference on Communications (ICC)* (2018), IEEE, pp. 1–7.

[88] Yaacoub, E., and Alouini, M. Connecting the Remaining 4 Billions: A Survey on Rural Connectivity. *CoRR abs/1906.11541* (2019).

[89] Ying, X., Zhang, J., Yan, L., Zhang, G., Chen, M., and Chandra, R. Exploring indoor white spaces in metropolises. In *Proceedings of the 19th annual international conference on Mobile computing & networking* (2013), ACM, pp. 255–266.

[90] Zhang, J., Zhang, W., Chen, M., and Wang, Z. WINET: Indoor white space network design. In *2015 IEEE Conference on Computer Communications (INFOCOM)* (2015), IEEE, pp. 630–638.

[91] Zheleva, M., Paul, A., Johnson, D. L., and Belding, E. Kwiizya: Local Cellular Network Services in Remote Areas. In *Proceedings of ACM MobiSys* (2013), ACM, pp. 417–430.