Supporting a Diversely Connected World via Community Cellular Networking

Matthew W. Johnson

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Kurtis Heimerl, Chair

Richard Anderson

Shaddi Hasan

Program Authorized to Offer Degree:

Paul G. Allen School of Computer Science & Engineering

University of Washington

Abstract

Supporting a Diversely Connected World via Community Cellular Networking

Matthew W. Johnson

Chair of the Supervisory Committee:
Assistant Professor Kurtis Heimerl
Paul G. Allen School of Computer Science and Engineering

Community cellular networks (CCNs), small-scale cellular networks owned and operated by members of the communities they serve, have the potential to provide sustainable wide-area coverage across a diverse range of contexts on the margins of the Internet. This dissertation first characterizes some of the operational challenges observed in an existing CCN, and then develops new designs for future community cellular systems to address the identified challenges.

To ground the research, I gather a unique dataset characterizing the operation of a remote, expensive, data-only community LTE network. Through tight integration with the operator's infrastructure, the dataset has visibility to drill down to individual user behaviors in terms of both network traffic and prepaid credit purchases and transfers. I find that despite the network's constrained capacity, use is still highly unbalanced, with a handful of heavy consumers providing an outsized portion of network revenue. A large portion of the userbase only has intermittent access in this model. 45% of users are offline more days than online, and the median user consumes only 77 MB per day online and 36 MB per day on average, limiting consumption by frequently "topping up" in small amounts. The network is also only somewhat reliable, with downtime caused by software failures, network failures, and the generally harsh conditions on the rural edge.

To address the reliability challenge, I observe that multiple hotspot providers were already operating in the area, and propose a new authentication architecture to allow fallback to an overlapping provider even when the user's home network is offline. This system, dAuth, is fully backwards compatible with standard off-the-shelf devices, making it feasible for immediate deployment. I design, prototype, and evaluate dAuth against existing private edge-core and cloud-core architectures, and find it comparable to the status quo while providing robust operation in the presence of failures and good scaling properties as the number of small operators increases.

To address the challenge of providing sustainable access while avoiding intermittent availability for low-income users, I explore the design of community-based congestion-management policies and network management mechanisms through the lens of network capacity as a common pool resource. I present qualitative insights from a series of workshops and interviews exploring designs for community-scale networks with resource sharing. Participants expressed conflicting desires for preserving individual privacy while collecting longitudinal data to track the network's impact on the community, prioritization of high-value applications, equal link allocation between users, and human-mediated congestion management in lieu of automated enforcement.

Overall my work provides new directions for the design of next-generation community cellular networks that can be more reliable and operate with a wider range of governance structures to sustainably serve a wider range of communities.

Table of Contents

List of Figures

List of Tables

Table Number Page

Glossary

5G: The 5th generation of ETSI/3GPP cellular network standards, providing more flexibility in the Radio Access Network than LTE and moving to a service-based core network architecture.

5GC: The core network of a 5G cellular network, providing user management and traffic control. The 5GC is made of several components, including the AMF, SMF, UPF, AUSF, SEAF, and many other functions.

ACCESS AND MOBILITY FUNCTION (AMF): The 5GC component responsible for managing the users' mobility as they move between different RAN elements in the 5G network.

AUTHENTICATION SERVER FUNCTION (AUSF): The 5GC component responsible for anchoring user authentication in the home network. It communicates with the serving network's SEAF during roaming.

CITIZEN'S BROADBAND RADIO SERVICE (CBRS): A three-tiered spectrum access regime currently operational in the USA covering spectrum from 3.55GHz to 3.7GHz. Access is mediated between existing incumbent users, priority access users, and general access users by an automated spectrum access system. Now marketed under the trade name "OnGo."

COMMON POOL RESOURCE (CPR): A resource capable of being drawn sustainably up to some rate, but vulnerable to over-use and collapse if the rate is exceeded.

COMMUNITY CELLULAR NETWORK (CCN): A cellular network owned and/or operated by the users it serves.

CORE NETWORK: The part of the cellular network architecture responsible for coordinating multiple RAN elements and managing the UE's long term state as it moves through the network. Traditionally core network elements were physically centralized in the "core" of the network at central offices and regional switching centers.

ENODEB: An individual radio basestation in the LTE RAN, possibly connected to multiple antennas and providing multiple cells of service. While logically fulfilling the same function, can range in size from a small cell implemented as a single physical box to a macro cell comprising multiple racks of equipment.

EVOLVED PACKET CORE (EPC): The core network of an LTE network, providing user management and traffic control. The EPC is made of several components, including the MME, HSS, SGW, PGW, and other network functions.

GEOSYNCHRONOUS ORBIT (GEO): The orbit at exactly 35,786Km, with the property that the satellite's orbital period exactly matches to rotational period of the earth. This allows for a stationary ground antenna but at the expense of high latency due to the speed-of-light propagation delay across the large distances involved.

GNODEB: An individual radio basestation in the 5G RAN, possibly connected to multiple antennas and providing multiple cells of service. While logically fulfilling the same function, can range in size from a small cell implemented as a single physical box to a macro cell comprising multiple racks of equipment.

HOME SUBSCRIBER SERVER (HSS): The EPC component responsible for holding long-term state for each user and handling user authentication. It serves as the source of truth for the user's last known location when new data for the user is received in the EPC.

INTERNET SERVICE PROVIDER (ISP): An entity or organization responsible for providing downstream Internet service to end users or other organizations.

LONG TERM EVOLUTION (LTE): The 4th generation of ETSI/3GPP cellular network standards, defining the radio protocol between a user's mobile device and the cellular network and the signalling traffic for managing communication on the network. LTE is a notable departure from previous standards since all communication is handled as packet-switched data.

LOW EARTH ORBIT (LEO): The set of orbits ranging from 100Km to 2,000Km above the surface of the earth. New communications satellite constellations are being built in these orbits, which are relatively close to the earth and allow for communication with latency similar to existing terrestrial infrastructure.

MOBILE NETWORK OPERATOR (MNO): The organizational entity responsible for operating a cellular network. Traditionally a large and heavily regulated organization.

MOBILITY MANAGEMENT ENTITY (MME): The EPC component responsible for managing the user's connection state as they move between different RAN elements in an area of a cellular network.

PACKET DATA NETWORK GATEWAY (PGW): The EPC component responsible for receiving and de-encapsulating the user's dataplane traffic from the SGW. The PGW is the user's mobility anchor in LTE and is responsible for managing the user's IP address.

RADIO ACCESS NETWORK (RAN): The part of the cellular architecture responsible for connecting to users to the network. The RAN is traditionally physically distributed throughout the environment to provide ubiquitous coverage and support mobility.

SECURITY ANCHOR FUNCTION (SEAF): The 5GC component responsible for anchoring user authentication in the serving network.

SERVING GATEWAY (SGW): The EPC component responsible for tunneling the user's dataplane traffic from the RAN into the EPC and towards the user's current PGW.

SESSION MANAGEMENT FUNCTION (SMF): The 5GC component responsible for managing and updating the user's data sessions in conjunction with the AMF as they move between different RAN elements in the 5G core network.

USER EQUIPMENT (UE): The end-user device in the cellular architecture. Often it takes the form of a mobile phone, but could also be a standalone WiFi hotspot or embedded radio.

USER PLANE FUNCTION (UPF): The 5GC component responsible for forwarding, measuring, and policing user dataplane traffic. The UPF is the mobility anchor for a particular slice of the user's traffic in the 5G architecture.

Dedication

To my family, housemates, and Erin for supporting me through this journey, challenging me when it was necessary, and inspiring me to look beyond myself.

Chapter 1

Introduction

At the close of 2021, the ITU estimates that 62.5% of people globally use the Internet, and 95% are covered by a wired or wireless mobile broadband network (3G, 4G, or 5G) [26]. Yet as has been well-documented, the exact nature and definition of use can vary widely, and challenges remain in providing meaningful capabilities to all users of the network [48]. Furthermore, while 95% coverage is a great achievement, the expansion of Internet connectivity has slowed in recent years as relatively easier to serve markets, with dense populations and supporting infrastructure like roads and grid power, saturate [184]. This leaves hundreds of millions of people still uncovered, mostly in rural, remote, and small communities which are difficult to serve with traditional centralized communications infrastructure [79, 74, 149].

My lab and I have explored community networks as one possible solution to facets of these infrastructural challenges. Community networks are networks deployed, operated, and maintained

by and for the community members they serve [150]. Researchers have proposed that community-based networks could help to bridge the coverage gap by providing access to new areas by changing the economic balance of providing connectivity [148, 101], and the connectivity gap by allowing locally engaged organizations to reach sub-populations excluded by traditional networking systems and structures [21]. Community network operators can take advantage of local knowledge, social connections, and existing community resources to provide connectivity with substantially lower capital and operational costs than traditional operators [18, 120].

My work focuses on community *cellular* networks, a type of community network built with cellular networking technologies (2G GSM/EDGE, 3G UMTS, 4G LTE, and now 5G). Community Cellular Networks (CCNs) blend the organizational affordances of community networks with the technological affordances of wide-area cellular networks. Cellular networks offer several compelling advantages over other community networking technologies, particularly in the case of wide-area access at rural scales (see Section 2.1 [The Case for Cellular]).

Notably and in contrast to existing work on GSM-based CCNs though, my work specifically dives into the challenges and opportunities provided by *IP-based* CCNs built with LTE and 5G technologies. GSM provides a different authentication architecture and GSM CCNs classically only support voice and SMS services, leading to very different operational and management affordances than the secure, IP-based, and Internet-first networks provided by LTE and 5G networks. While bringing new challenges, my works finds that these challenges can be largely overcome, and that these networks can provide high-speed data connectivity across wide areas that makes them a promising candidate for future community cellular network deployments.

1.1 This Thesis

Next-generation LTE and 5G community cellular networks can provide a set of affordances and scalability to enable economically sustainable and reliable connectivity in communities otherwise

poorly served by existing Internet access technologies.

1.2 Summary of Work

The overarching goal of my work is to begin to understand and address the challenges of managing and sharing resources in data-based community cellular networks. I hope that this research can help to realize the vision of CCNs extending new backhaul connectivity to provide truly universal access. In it I address three primary research questions:

a) How are rural community cellular networks used for Internet access today?
b) How can cellular technologies be extended in a backwards-compatible way to facilitate redundancy and reliability across individually unreliable operators?
c) What design affordances are necessary to facilitate value-sensitive management of shared network resources in a small community setting?

Taken together, I hope that the answers to these questions will not only be of academic interest, but also help facilitate the creation of new community-scoped network deployment models to allow users currently outside the margins of the Internet to gain affordable, acceptable, and reliable connectivity.

1.2.1 Mixed Methods

In order to better understand this nuanced problem space, my work leverages a relatively wide variety of methods. These vary from customized Internet measurement and data science, to systems development and artifact performance evaluation, to value sensitive design and participatory design. My overall goal has remained consistently focused on building a holistic understanding of the role IP-based community cellular networks can play in the broader connectivity ecosystem and building tools to support their use in this domain. Leveraging multiple different methodologies has let me approach the problem space from more angles than would be possible with a single methodological toolset.

1.2.2 How are rural community cellular networks used for Internet access today?

To explore this first question, I undertook a longitudinal measurement study of a production rural community cellular network. The network uses LTE small cells and provides service to handsets across a 2km diameter area. The gathered study dataset spans a full year of operations, containing aggregated and anonymized Internet traffic logs as well as anonymized transaction data (the network operates with a prepaid credit model) to draw conclusions about both traffic and financial patterns of the network. My and my collaborators' findings that most users connect intermittently, while financial sustainability hinges on a small subset of heavy users, motivate my following projects to explore alternative resource sharing approaches. Additionally, with permission from the operator, we anonymized, packaged, and published this unique dataset for future analysis by the wider networking research community.

Methods

This project leverages methods drawn from Internet measurement and data visualization/data science. I built a custom data gathering pipeline appropriate for use on an extremely constrained rural edge network, and integrated this pipeline with our partners' infrastructure to collect longitudinal aggregated measurements in the network. The gathered dataset is non-trivial in size, with tens of millions of rows and over 30GB of data in memory, requiring care in its analysis. Once the data was collected and anonymized, I leveraged a variety of data analysis and visualization techniques to understand the gathered dataset and explore the traffic patterns and economics of the network.

Core Contributions

This part of my work contributes:

- A methodology for gathering and analyzing traffic data from remote edge networks with

constrained uplinks

- A unique dataset gathered, anonymized, and publicly released from a remote constrained network with both financial transactions and information about the traffic generated in the network
- Analysis discovering a wide range in the frequencies that individual devices used the network and had credit stored for on-demand use
- Analysis discovering the skewed and uneven role video played in this type of network
- Analysis showing the financial support of the network came from relatively few users, and that a large population of users used the network consistently but in small amounts

More details are provided in Chapter 3 [CCN Traffic Measurement: Whale Watching in Inland Indonesia] and Chapter 4 [CCN Traffic Measurement: COVID-19].

1.2.3 How can existing cellular technologies be extended to facilitate redundancy and reliability across a diverse set of individually unreliable operators?

To explore this second question, I designed the dAuth system, which adds redundancy to cellular authentication by allowing a network to delegate limited authentication capability to a set of semi-trusted backup networks. The system uses threshold key shares and a manipulation of how SIM cards manage authentication vectors to remain backwards compatible with off-the-shelf radio equipment and user devices. With my lab mates, I developed a fully functional prototype of dAuth and integrated it with the open source Open5GS cellular core network. We confirmed dAuth is backwards compatible through a test deployment with off-the-shelf CBRS radio hardware in partnership with the Seattle Community Network. We then evaluated our prototype against the current state of the art, a hosted cloud core, with a simulated large-scale RAN to test its scaling properties. We found that at low levels of load our system does add approximately one RTT of latency, but that at medium and high load its superior scaling properties outperform a cloud core while providing additional redundancy and resiliency to system failures.

Methods

In order to uncover the challenges of realizing my vision of multiple independent network operators collaborating to co-operate, I designed dAuth and co-developed a working implementation of the design which operates with off-the-shelf cellular hardware. This required developing a technical understanding of the constraints imposed by the cellular standards, developing an operational system following the standardized protocol and meeting protocol time constraints, and then methodically evaluating the performance tradeoffs of the design against related solutions in a realistic environment.

Core Contributions

This part of my work contributes:

- A novel application of threshold key sharing to the cellular authentication process
- An existence proof that such a system can be realized while maintaining backwards compatibility with off-the-shelf hardware and without compromising the link security of the LTE and 5G network architectures

More details are provided in Chapter 5 [dAuth: Distributed Authentication for CCN Federations].

1.2.4 What design affordances are necessary to facilitate value-sensitive management of shared network resources in a small community setting?

To explore this final question, I led a series of workshops and informal interviews with participants in a small town in Oaxaca, Mexico in collaboration with the NGO Rhizomatica. The community has a history of communal resource management, and strong sets of values which make traditional pricing-based management inappropriate for a community organized network. After analyzing the workshop transcripts and notes, my collaborators and I were able to extract concrete design elements absent from existing tools available to manage and operate the network. Among others,

we found needs to enable community-wide monitoring of total use without revealing individual use, integrate with "professional" interfaces like Diameter and Gx for resource management, prioritize high-value applications, ensure equal link sharing between users, and support mediated dispute resolution outside the network.

Methods

In order to explore the design constraints and feasibility of implementing community-based congestion policies and mechanisms in a CCN context, this part of my work relies on a qualitative research methodology operationalizing elements of Participatory Design and Value Sensitive Design. With Rhizomatica and the Santa Inés community telecom operator, I facilitated three public workshops to both educate the community about congestion in the LTE network and gather ideas for how to manage it, held two formal meetings with town leadership, and conducted two opportunistic interviews. I adopted mainly empirical and technical VSD, using knowledge of cellular network affordances, combined with participatory methods, to reveal users' preferences and concerns and explore a wide space of possible congestion management strategies. I generated transcripts from the interviews and workshops, and then coded and analyzed the transcripts and my field notes to identify themes in the responses across participants.

Core Contributions

This part of my work contributes:

- The qualitative insight that current network management tools are not sufficient to directly meet the needs of the study participants in managing congestion in the LTE network, but that a value compatible tool could likely be developed for this use case
- The design parameters for such a value-compatible tool
- The qualitative insight that participants valued the ability to prioritize different types of traffic, coupled with the technical insight that this is difficult, if not impossible, to accomplish

reliably in today's Internet ecosystem without resorting to intrusive and value-incompatible deep packet inspection and flow classification tools

More details are provided in Chapter 6 [Network Capacity as Common Pool Resource].

1.3 Sites and Context

Through my PhD I have undertaken a set of interrelated user studies, systems developments, and subsequent evaluations. This work leverages established partnerships with three operational community cellular networks in Bokondini, Indonesia, Oaxaca, Mexico, and the Puget Sound, USA. Collaborating with deployed networks has been an important component of my work, grounding it in the realities of real-world network operation. I have been fortunate to be able to work in the field with each network, and continue to maintain relationships with each organization. None of this work would have been possible without the support and inspiration provided by these networks.

Each site is unique, and taken together they span a diverse set of different network architectures, organizational structures, and user demographics that have helped me identify both common challenges across sites and unique strengths and weaknesses within each site. Each chapter of this document referencing a field deployment contains a "context" section with more detailed information about each partner organization, the research sites, and their influence on the research (See sections 3.3, 5.2, and 6.3).

Network Architectures

All of the networks I have collaborated with are relatively constrained, with only a few low-powered "small cells" for their radio access networks and simple control planes co-located with the radio access at the edge. Figure 1.1 shows a set of hardware typical of these types of networks. There are many different types of CCN deployment though, as diverse and customized as the

Figure 1.1: Edge CCN hardware. An eNodeB small cell (left) providing coverage over the town center in Santa Inés, and the standalone edge EPC computer (right) serving as the network control and data plane connected to it from the main office. For more information about this network, see Section 6.3

communities they serve, and I present a high-level overview of CCNs and some of the different CCN deployment architectures in Chapter 2 [An Overview of Community Cellular Networks (CCN)s].

Positionality

My own positionality as a researcher undoubtedly colors my understanding and lead me to choose these particular research topics of interest. These biases influence the results of my work reported here. I am from a relatively affluent urban area, got my first cellphone in the 7th grade when SMS was expensive, generally avoid social media, and cannot remember a time when I was not able to access and afford some form of Internet connectivity. I have a set of expectations around Internet

performance, an inherent interest and learned understanding of the underlying technology, and notions of the utility of connectivity that are my own.

Chapter 2

An Overview of Community Cellular Networks (CCN)s

My work specifically focuses on Community *Cellular* Networks, a type of community network built with cellular networking technologies (2G GSM/EDGE, 3G UMTS, 4G LTE, and now 5G). CCNs blend the organizational affordances of community networks with the technological affordances of wide-area cellular networks.

2.1 The Case for Cellular

Several factors differentiate CCNs relative to other types of community networks (based on WiFi or wired connections) that make them a promising area of study. Firstly, while device ownership is far from universal, cellphones are pervasive and are often present even in areas without existing connectivity [10, 168, 81]. Building a community network compatible with cellular standards allows the network to leverage these existing user devices, supply chains with economies of scale, and repair ecosystems [91].

Secondly, and particularly in the context of rural access, cellular networks offer the advantage of mainstream device availability in bands with wide-area propagation, and a waveform designed for long links with non-negligible propagation delay. In contrast to WiFi-based community networks, community cellular networks can provide blanket coverage over an entire settlement from a small number of basestations (often one). This allows the maintenance burden to be centralized [74, 90], provides continuous coverage for mobile users, and crucially reaches community members with limited mobility (or agency) who may not have been able, comfortable, or allowed to visit a central WiFi hotspot or cyber cafe [25].

Equipment in Diverse Frequency Bands: LTE and 5G support over forty different bands encompassing both licensed and unlicensed frequencies. Cellular small-cell basestations and clients are available at reasonable prices in bands with better propagation and higher allowed power than the ISM and U-NII bands used by WiFi. The range of band options already available allows operators to use frequencies appropriate for rural access without being confined to the limits of the unlicensed bands or needing expensive custom hardware.

LTE and 5G Waveforms Support Longer Range: Cellular technologies like LTE and 5G-New Radio (5G-NR) can outperform WiFi over the more tenuous links common in rugged areas. They are explicitly asymmetric, optimizing for an advantaged basestation and a low-power handset. Modern cellular's SC-FDMA uplink modulation allows higher efficiency transmission from mobile devices than WiFi's OFDM, and hybrid ARQ increases throughput under weak signal conditions. LTE and 5G-NR's schedulers also handle longer links by explicitly compensating for propagation delay. These characteristics map well to rural deployments, where a single basestation can be deployed on existing structures with reliable power (like barns or grain silos) to cover a large area with a single point of maintenance. Technologies designed for local area networks in urban areas have insufficient range for rural areas [185]; "wide area" technologies operate at scales more

appropriate to farms, ranches, and fields.

2.1.1 Spectrum Access

Despite cellular's advantages and appropriateness for extending access, deploying cellular equipment previously required access to limited frequencies licensed to national incumbents that strongly limited availability. Particularly since 2019, cellular spectrum has become more accessible to small organizations and community networks.

On one front, 3GPP Release 16 standardized NR-U, the operation of 5G New Radio in unlicensed bands, either with a licensed carrier anchor or standalone [61]. The Citizen's Broadband Radio Service (CBRS) also allows small operators a straightforward path to access licensed spectrum with good propagation characteristics. Lun et al. argue that the CBRS band's wide bandwidth and amenability to massive MIMO antennas of reasonable size make it an choice ideal for rural access from existing towers [110]. CBRS in the United States uses a novel three-tiered licensing scheme for the mid-band spectrum between 3.55GHz and 3.7GHz. Access to the spectrum is controlled by a "Spectrum Access System" automated database, where navy radar receives priority, followed by priority access license holders, and ultimately general authorized access users [11]. CBRS grants are leased out on a 10 minute basis and are accessible by API to certified devices. Commercial turnkey basestations using this spectrum are now readily available for self-hosted cellular networks.

Looking forward, researchers and activists have long been arguing to expand secondary use licensing regimes for more frequencies in rural areas too [142, 31]. Strict licensing is critical for urban areas, where spectrum is highly occupied, yet most spectrum in remote areas lacking network coverage is, by definition, available. Sankaranarayanan et al. observed that licensed cellular spectrum is often under-utilized, and proposed a standalone mac protocol for secondary users to communicate in slots unused by a primary GSM network [156]. Hasan et al. later proposed

a different solution, where secondary GSM networks detect empty channels via the neighbor cell sensing primitives built in to the GSM protocol, that has the advantage of being backwards compatible with standard UEs [76]. Baig et al. explored the specific question of how to deploy LTE in TV white space for rural Internet coverage [17], and developed a practical system, CellFi, to allow LTE deployments in secondary use spectrum with a coordinating spectrum database for channel allocation and intra-channel resource sharing. Hasan et al. in collaboration with a large MNO built tools to allow MNOs to share unused spectrum resources with community networks [74].

While spectrum access and spectrum policy continues to be an important and contentious issue for community wireless networks, options exist today to allow widespread deployment and experimentation with community cellular networking hardware. In my work I assume that users already have access to spectrum, and focus on the operational constraints encountered running their networks.

2.2 Cellular Architecture

Every community cellular network is unique, but all must implement the functionality required by the general cellular network architecture to remain compatible with standards-compliant user devices. The cellular architecture is broken into several parts, with a separation of concerns between components. The "User Equipment" (UE) is the user's device connected to the network, often a mobile phone or hotspot. The "Radio Access Network" (RAN) is physically distributed through the environment and is responsible for the actual radio communication with the user's device. The RAN is connected to a "Core", a logically (and often physically) centralized component that coordinates the RAN's operation. The connection between the core and the RAN is the "backhaul," and in modern cellular can be implemented with any IP-capable technology. The radio-emitting components (UE + RAN) in a deployed network must be certified for safe and legal

operation, limiting the amount of experimentation that can be done outside a lab setting. My work uses existing off-the-shelf UEs and RAN basestations, but explores how they can be uniquely adapted for the community network use case.

2.2.1 Deployment Architectures

There are a variety of different ways to assemble these components, which have ramifications on the way community cellular networks are physically built and perform in different deployment scenarios. In the traditional cellular architecture, the user's traffic is encapsulated over the backhaul and through the core network to allow the core to enforce quality of service guarantees on the traffic, provide precise traffic accounting, and allow legal intercept. The core is implemented with specialized physical hardware. Figure 2.1 shows the architecture of a traditional MNO deployment. This architecture allows masking mobility, makes it easier to coordinate between many basestations, and gives the operator strong control over their network.

For smaller organizations though, the overhead of establishing their own physical sites to host the core network pushes them towards alternative architectures. Many use a "cloud core" architecture (Figure 2.2), where the core network functions run virtualized in the cloud, either managed as a turnkey vendor solution or self-managed by the organization themselves [101, 122, 127, 8]. Due to the economies of scale with cloud networks, many large tier-1 operators are moving towards this type of architecture as well [160, 15, 189]. With a cloud core, the user plane data can be split off either at the edge for scalability and low latency, or tunneled through the core to mask mobility depending on the deployment's requirements.

A cloud core provides an easy on-ramp to operating a CCN, but in areas with a constrained uplink or where keeping the data local is important, an "edge core" architecture is often used instead (Figure 2.3). Pulling the core functions out to the edge keeps all of the cellular control traffic local, decreasing the overhead on the constrained backhaul link and improving control

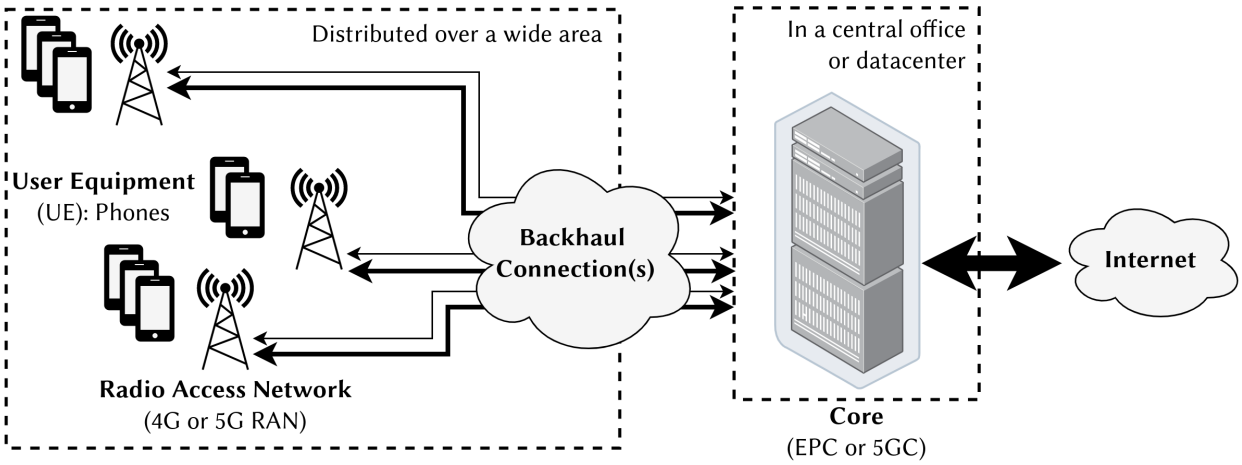


Figure 2.1: Traditional MNO architecture: Many RAN elements are connected via backhaul, often owned by the operator, to a small number of region-scale core networks which coordinate the RAN's operation. Both control (thin line) and user dataplane traffic (bold line) passes through the core for quality of service management and accounting.

plane responsiveness and performance [167, 83]. Keeping the core local also allows the edge network to continue to provide local service even if the backhaul link goes down [74]. All of the community cellular networks I have collaborated with have used this architecture, either for performance or data locality reasons.

2.3 The Impact of New Backhaul Technologies

Today small and extremely remote sites are typically served by tightly constrained backhaul links (the connection between the site and the wider Internet), utilizing geosynchronous (GEO) very small aperture terminal satellite (VSAT) or point-to-point microwave infrastructure [84]. Supporting disconnected and asymmetric operation where the local "radio access network" (RAN) has much more capacity than the backhaul has been a consistent thread through my and my collaborators' prior work [167, 96, 74, 79, 95]. Many initiatives have attempted to provide affordable high quality backhaul to remote areas with a wide range of approaches, and until recently all have

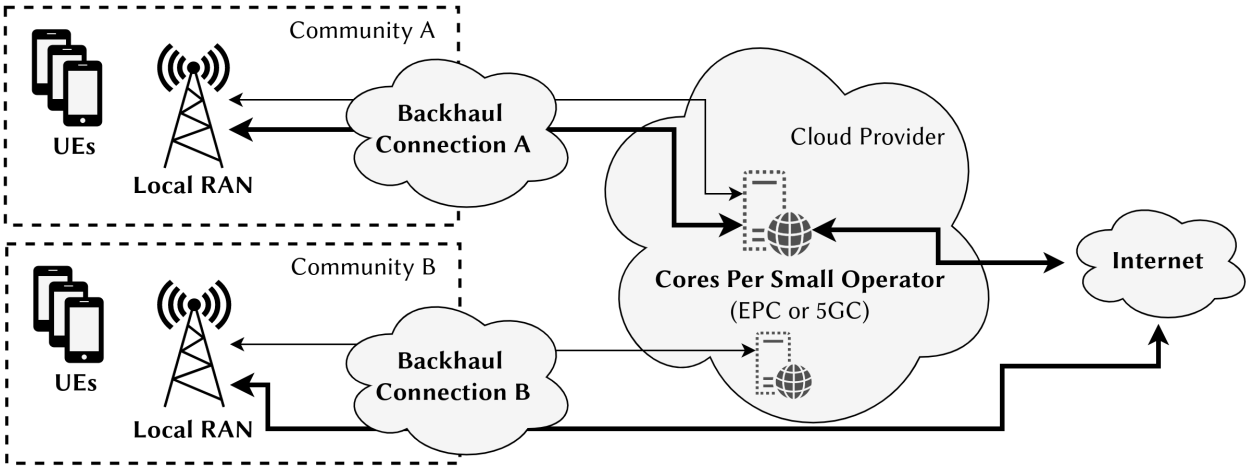


Figure 2.2: Cloud core architecture: To lower the burden of operating a network, the core network functions can be moved to the cloud, and either self-managed by the community network or provided as a turnkey service. Control traffic (thin line) always goes to the core, while user dataplane traffic (bold line) can either be forwarded through the core (Community A) or broken out locally (Community B).

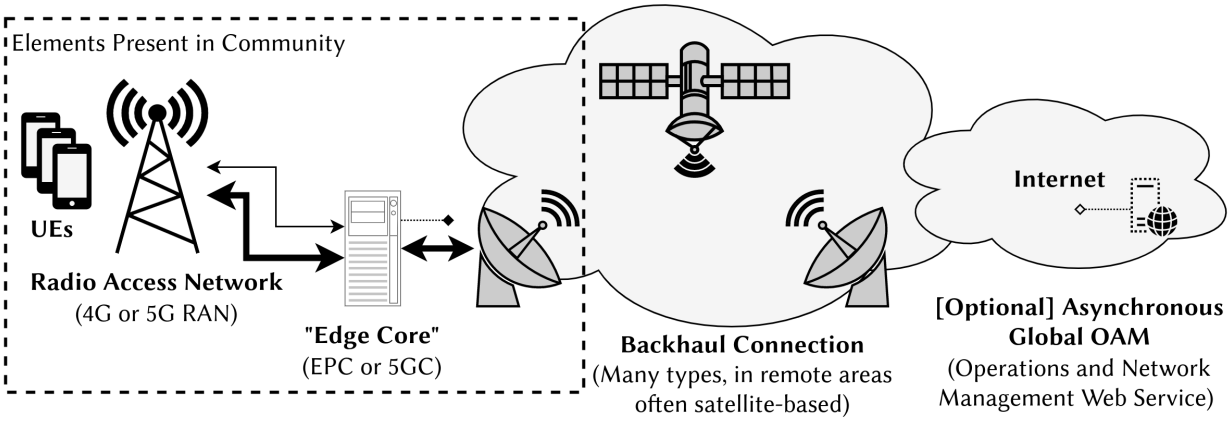


Figure 2.3: Edge core architecture: When the backhaul link is constrained, moving many of the core network functions to the edge of the network near the RAN improves performance. In this architecture all control traffic stays local. The user's dataplane traffic is plain IP at the point it traverses the backhaul connection, and can be routed directly to the Internet. Some systems add an asynchronous management layer to consolidate configuration and status reporting of multiple edge nodes in one place, but importantly the edge nodes can continue to operate fully if this connection is not available. The community networks I have collaborated with predominantly used this architecture.

failed to see meaningful sustainable deployment.

Yet after decades of high-profile failure[1], new networks incorporating large constellations of low earth orbit (LEO)[2] satellites are poised to begin truly offering high-bandwidth and low-latency *global* connectivity. In contrast to traditional GEO satellites orbiting at an altitude of 35,786km, LEO communications satellites orbit at only 500~2000km. These lower orbits require hundreds or thousands of satellites for continuous connectivity, but allow reduced round trip latencies (bounded by the speed of light, ~20-40ms) and higher capacity (due to shorter links and improved spatial re-use, ~100-1000mbps). Bhattacherjee et al. find that "this new breed of satellite networks could potentially compete with today's ISPs in many settings, and in fact offer lower latencies than present fiber infrastructure over long distances" [23].

This begs the question: have these networks finally "solved" global connectivity? I argue that while undeniably amazing feats of engineering, a panacea they are not.

High-bandwidth access to LEO satellite networks currently requires investment in a specialized ground terminal with a large phased antenna array (~$500USD) which must be independently powered, secured, environmentally protected, and placed outside in an area with a wide line of sight to the sky (LEO satellites transit the field of view quickly). Also, due to the weight, power, and thermal constraints of satellite operations, and the spectrum used for these networks, it is unlikely that this upcoming generation will ever be able to serve end-user devices, especially indoors, with a broadband-quality connection directly from orbit.

---

[1]Iridium: bankrupt 1999 [175], CDMA-450: failed to see wide adoption [147], Facebook's Aquila Drone: shutdown 2018 [115], OneWeb: bankrupt 2020 [49], Alphabet's (Formerly Google) Loon: shutdown 2021 [194], among others

[2]New medium earth orbit (MEO) constellations are also coming online [128]. MEO is defined as the range of orbits between LEO and GEO (from 2000km to 35,786km). Throughout this document I focus on LEO networks since they are more revolutionary, while the new MEO networks represent a more gradual evolution from existing services. In practice this work is applicable across backhaul technologies.

Community Cellular Networks Can Enable Wider Reach

Acknowledging the limitations of these new satellite networks alone, I see an opportunity for a deployment model leveraging the strengths of community cellular networks to complement the revolutionary backhaul capacity of LEO satellite networks. In this model members of a community share a smaller number of terminals, amortizing the cost and maintenance burden, while using medium range terrestrial wireless to distribute connectivity at the town/village scale. As a first order approximation, fewer than 10% of accounts in the dataset detailed in section 3.5.1 have data purchases sufficient to offset the monthly cost of their own next-generation satellite subscription. Many users consistently used the network at extremely low levels, purchasing just enough data to exchange text-based messages and/or accessing the network inconsistently. Without some way to share connectivity, these users can not be sustainably served by the deployment models used for these new satellite networks today.

Sharing a terminal introduces new challenges around resource coordination and management though, and the propagation of low-band cellular spectrum means the chance that multiple community cellular access points could have overlapping coverage is high. Rather than assume the access points must be independent, I envision them exchanging information to coordinate their operation regardless of who may deploy and maintain each one. Taking advantage of the small size and limited scope of community networks, it becomes feasible to build systems that can allow more optimized operations through human mediated coordination.

Coverage Beyond Convenience

The physical coverage of the network has implications beyond just allowing for more convenient access for existing users. In her ethnographic work characterizing operations in a range of community networks, Nicola Bidwell found that the networks based on wide-area cellular technologies allowed for more flexible access for a different demographic subset of the population than those

based on physically-limited WiFi hotspot networks [25]. Specifically women users with work duties that prevented them from spending time near the hostspot or with limited mobility due to social expectations were able to access community network services much more easily with mobile devices they could use at home or on the move in town while tending to other tasks.

Yet the current crop of next-generation satellite networks are all focused on providing access either only to large businesses (eliminating small community organizations) or to end users via WiFi attached to the terminal. Providing wide-area coverage via small community organizations, instead of just single-point access near the terminal, is about more than just improved economics and connection pooling. It is truly a matter of equity in reaching all types of people who may not have the freedom of mobility to use a low-range shared access point.

Chapter 3

| CCN Traffic Measurement |

Whale Watching in Inland Indonesia: Analyzing a Small, Remote, Internet-Based Community Cellular Network

Executive Summary

The goal of this project was to gain a grounded understanding of how an Internet-based community cellular network is used in the context of today's Internet device and service ecosystem.

Hypothesis

I hypothesized that due to the constraints of the backhaul link in our partner's network, that the network would be saturated most of the time, that many users would have similar levels of consumption capped at saturation, and that users would not consume high-bandwidth services

due to poor performance.

Summary of Key Findings

I found the network is profitable, but that average-based metrics fail to capture on-the-ground reality with a small population size. I observed individual use is highly intermittent; users elect to purchase many small amounts of data rather than one lump sum, and many are offline for days at a time. Use was also unequal. Heavy "whale" users contributed disproportionately to the network's financial sustainability, and light and heavy users differed both in the way they use the network as well as by how much. Video dominated the backhaul link, but video consumption was driven primarily by a small subset of users. Large platforms (Google & Facebook) sourced the majority of traffic and reached all users, but I found local trends also had strong effects on the network.

Next Steps & Connections

The findings from this project have helped inspire and inform my follow-on work. Finding that a small number of heavy consumers were effectively enabling on-demand access for the majority (by keeping the network profitable) led me to consider the potential impacts of next generation LEO satellite systems on the connectivity ecosystem. These systems currently market target toward individuals, and have the potential to capture the anchor users without a means to easily share connectivity with the rest of the community outside WiFi hotspot range.

3.1 Introduction

The GSMA published in 2017 that traditional mobile networks have largely expanded their footprint to *all* areas where service is profitable, that new deployments are slowing, and that new paradigms are needed if Internet access is to reach those who remain unconnected [184]. Community Networks, networks owned and operated by the community members that they serve, are a

promising paradigm that changes the economics of sustainable deployment [79, 51, 19, 141]. Community *Cellular* Networks, Community Networks based on mobile access technologies like GSM, UTMS, LTE, (and now 5G-NR), offer advantages relative to technologies like WiFi mesh networks in performance or TV-Whitespace in device availability, but at the cost of more complex licensing and deployment [96, 74, 167, 118]. Community Cellular Networks are particularly well suited to remote rural use cases, where spectrum is readily available (if secondary use is allowed) and wide-area coverage is desirable. Due to their remote nature, these networks often operate with satellite-based backhaul links, limiting performance in terms of both latency (500ms RTT, although this is changing– see 7.3) and throughput (1-10Mbps aggregate for all users). While these conditions are constrained, they are the reality of "access" to the demographics currently on the frontier of the Internet.

In this work, we explore a modern, extremely remote, data-only, Community LTE Network in Bokondini, Indonesia. The network is run by a nonprofit affiliated with a school situated in the town center, and covers most of the town. We were able to tightly integrate with the operator's systems to gather a dataset integrating both technical and business information, allowing us to analyze user traffic as well as how users purchase and spend prepaid balance in the network. Our data collection spans over a year of the network's operation, covering 53 weeks from midnight March 10, 2019 to midnight March 15, 2020 (local time UTC+9). We ask whether a prepaid, satellite-based, data-only, mobile network can be profitable without subsidies, what applications will be used in such a network, how frequently will users interact with the network, and how will users manage their prepaid credit?

We find the network is profitable, but that average-based metrics fail to capture on-the-ground reality with a small population size. We observe individual use is highly intermittent; users elect to purchase many small amounts of data rather than one lump sum, and many are offline for days at a time. Use is also unequal. Heavy "whale" users contribute disproportionately to the

network's financial sustainability, and light and heavy users differ both in the way they use the network as well as by how much. Video dominates the backhaul link, but video consumption is driven primarily by a small subset of users. We also note widespread use of over-the-top(OTT) messaging and calling apps like WhatsApp, Facebook Messenger, and QQ, and note that UDP streaming protocols consume a large fraction of the uplink bandwidth. Large platforms (Google & Facebook) source the majority of traffic and reach all users, but we find local trends can also have strong effects on the network.

We analyze how the network can remain profitable despite a relatively small number of users and an Average Revenue Per User (ARPU) of less than $1USD per day, and hope to offer a detailed characterization of the unique and understudied properties of an extremely remote network at the outer reaches of the mobile Web.

3.2 Related Work

3.2.1 Community Networking

This research builds on a long history of work on community networking. Community networks, networks owned and operated by users in some sort of collaborative way, have long been viewed as a promising mechanism for increasing access among rural and disadvantaged populations [59]. Community networks can operate using a variety of technologies including standard telephony [7], 802.11 WiFi [19], or cellular protocols [79] and examples exist in both rural [74] and urban environments [141], and developing [79, 74, 99] and developed countries [141, 151, 19, 51].

With the variety of community networks, there are similarly a variety of engagements with them in the networking literature. For example, the team behind Guifi.net, operating the preeminent community network with over 35000 nodes in Catalonia [19], has explored a wide swath of issues including topologies [187], cloud services [162], and sustainability [18]. Works related to other networks have explored appropriate network architectures [74, 118], licensing [151],

and many other issues in community networking. A notable consistent thread is the importance of human factors in the operation of the network. This is echoed in Jang et al. which explored leveraging local actors to conduct maintenance and repair [90] and Moreno et al. discussing the importance of user engagement for sustainability of community networks [150].

Most similar to this paper, there is a body of measurement studies in community networks. Heimerl et al. [79] presented measurements of inbound versus outbound traffic in an Indonesian 2G community network, finding outbound traffic was much more significant. They also explored the sustainability of the network, finding that it was economically viable. Follow-on research explored phone adoption in the same community [81]. Cerdà-Alabern et al. explored the financials of the Guifi network [34]. Lertsinsrubtavee et al. [108] recorded Web usage in a Thai community network, finding that usage behavior was similar to that of commercial networks in significant ways, such as a focus on social networks. Unfortunately, they also found that user apps performed similarly (e.g. downloading lots up updates) despite the limited bandwidth available. This is supported by Johnson et al. [94], where a rural Zambian WiFi network saw similar behavior. Our work expands this literature to explore 1) the specifics of service utilization in data-only LTE networks, 2) service use *combined with* service utilization, 3) addition of techniques for mitigating the operational difficulties of data collection behind limited backhaul for studies like these, and 4) a more modern (2020) look into the operation of these networks.

3.2.2 Rural and Developing Networks

Outside of community networks (comprised of both developing and developed regions) [94, 108, 34, 79], there is also a body of network measurement literature in rural and developing regions. Often utilizing traces from telecoms, ISPs, or regional IXPs, these works explore the unique circumstances of rural networks. These include studies on broadband performance and adoption in Nepal [100], Pakistan [16], or South Africa [41]; cellular performance in India [170]

and Pakistan [14]; censorship in Pakistan [6]; mobile phone properties in a Pakistani cellular network [9], and web latency in Ghana [198]. These have scaled up to continent-wide analyses, such as IPv6 adoption [109], interdomain routing [54], and inter-country latency [58] in Africa. World-wide studies exist as well, such as Schinkler et al.'s work on the performance of Facebook's edge caching [158].

The massive diversity of these works, inclusive of multiple continents, scales, technologies, populations, and venues, demonstrates the value of breadth in network measurement research. While each individual research agenda is not (and does not claim to be) "representative" of the Internet in whole, together they provide a holistic, diverse perspective on Internet use throughout the world. Our work contributes to this whole with the perspective of an extremely remote, data-only cellular network. In addition, we provide new insights by analyzing network traces together with records of user spending and the network's finances.

3.2.3 Small-scale Network Measurements

Lastly, a body of research focuses on small-scale networks, servicing households or small groups of people but not in an explicitly community-oriented fashion. One example is Maitland et al.'s exploration of Internet use in a refugee camp [116], where the network is run by the UN Refugee Agency. They found a wide range of experiences with the Internet in the camp, contributing to a set of barriers to access. Another set of works focus on tribal Internet, with Vigil et al. [190] explicitly focusing on failures in the use of apps like YouTube in the context of a TVWS deployment in US tribal lands. Even some of the community networking work cited above (notably Hasan et al. [74] and Martinez-Fernandez et al. [118]) involved partnerships with outside organizations, such as telecoms for spectrum, limiting the extent of the community participation. These works similarly broaden the range of measurements possible and inform elements of the novelty of our work.

We also note a large body of home and consumer Internet measurement [63, 179, 66]. While relevant in that community networks often leverage consumer Internet connectivity, their needs and expectations differ greatly from networks serving customers directly.

3.3 Context

3.3.1 Deployment Location

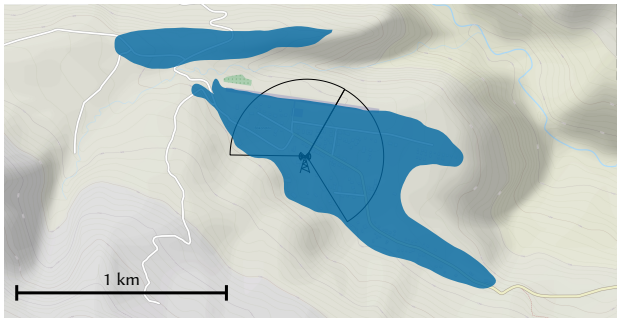
Bokondini, Indonesia, is a community of ~2,000 located a two hour drive (on a rough muddy road) from Wamena in the highlands of Papua. Bokondini, a central location for missionary activities for decades, still has limited infrastructure. There is a small airstrip, but no community-wide water or reliable power. The local regency government moved to Bokondini two years ago, and the town has been growing quickly. A national carrier has provided 2G coverage via a satellite-based small cell through a universal access program for about 4 years, but there was no Internet access until recently. During the final weeks of this study the carrier began offering LTE service nearby, still served by satellite. Mobile phones are common, with a critical number of LTE-capable devices already present [168].

3.3.2 Network

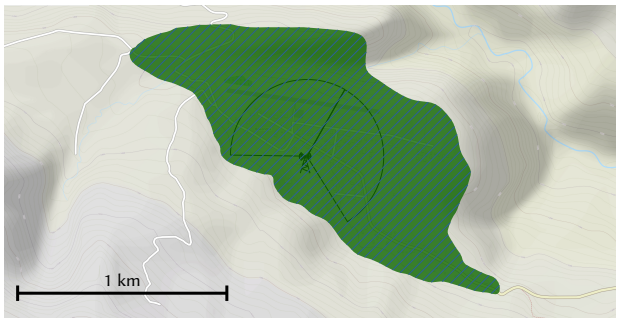
The network in Bokondini is an instance of a *Community Cellular Network*, owned and operated since 2013 by a missionary group whose primary function is running an elementary school in the area. They manage the day-to-day operations, including maintenance, credit sales to resellers, power management, and repairs. They operate a 5000KW microhydro and solar installation which powers the network as well as the school's lighting and computers. Unfortunately, the microgrid does not have enough reserve power to operate 24 hours a day, so the network is shut down manually between 11pm-5am (extended to 12-4:30am part way through the trial) to conserve power.

(a) The network's hardware deployed on a radio mast on top of the school gym.

(b) The cell site location and the orientation of the two antenna sectors in Bokondini. The town sits on a small plateau with a steep hill to the southwest and a ravine to the north and east. Map data from OpenStreetMap [132]



(c) The shaded area indicates a rough estimate of the most populated areas in Bokondini at the time of the deployment. Map data from OpenStreetMap [132]



(d) Interpolated coverage map from walk testing. The hashed green area indicates measured satisfactory coverage (-110dBm RSRP "two bars" or greater), capable of saturating the satellite backhaul link. Map data from OpenStreetMap [132]

Figure 3.1: Images and maps of the network deployment in Bokondini. The network uses two 33dBm small cells with 18dBi antennas arranged in two sectors to provide coverage over most of the area.

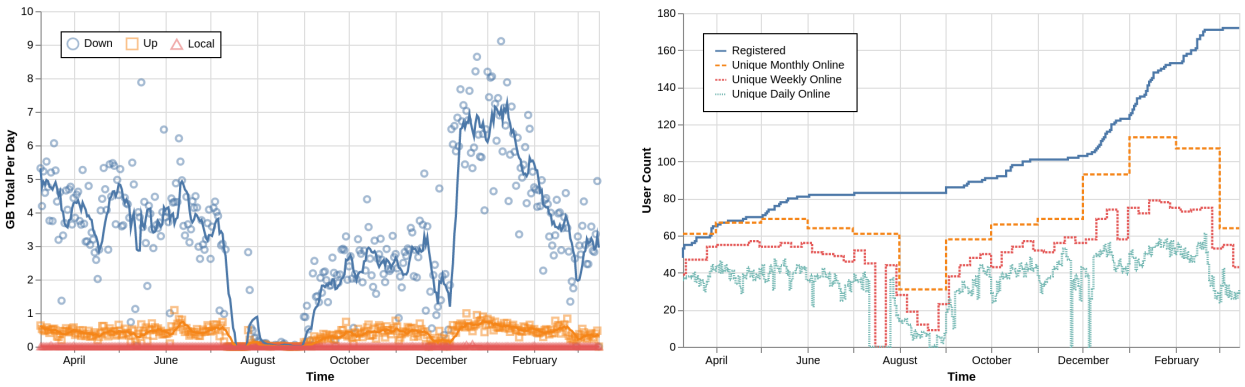


Figure 3.2: Network Activity vs. Time. The amount of traffic per day(Left) is highly variable and enveloped by the operational events detailed in table 3.2. Internet traffic eclipses local traffic by two orders of magnitude. The monthly online user count holds steady at around 60 users except for an extended outage in the summer and a peak in early 2020 as many new users join the network. Which users are online on particular days is variable and intermittent.

My collaborators and I have nearly a decade of experience working with them to explore different approaches to rural connectivity. The current iteration is a rural-optimized LTE network I helped design and install in 2018. Demand has fluctuated, serving between 40 and 80 users when operational. Its topology is relatively simple, with Radio Access (RAN) provided by a commodity eNodeB installed on the top of the school's gymnasium, connecting existing user handsets. An x86 mini-computer hosts an Open Source EPC to terminate all LTE signaling from the eNodeB, connected to a generic IP router with NAT, and ultimately a consumer-grade VSAT providing Internet backhaul. The site originally connected to another community via wireless relay to share a 3Mbps/1Mbps 8:1 (downlink/uplink & contention ratio) VSAT, but the relay was destroyed by lightning and a dedicated 1Mbps/256K 8:1 connection was temporarily established. This dedicated connection was later upgraded to 3Mbps/1Mbps 4:1.

Despite being an LTE network, no voice or SMS services are provided. The network operator encourages users to employ "over the top" (OTT) services like WhatsApp, Facebook Messenger, or Viber which are already extremely popular. Most users have dual-sim phones, and register

themselves via SMS on the national carrier 2G network. Providing only generic IP data greatly simplifies the network architecture and reduces associated costs and liabilities from interconnect and identity (phone number) management.

Credit Model

The operator uses a prepaid model, where users pay cash to a reseller in 1:1 exchange for "credit" on their account denominated in the local currency. Users later use a locally hosted web application to convert their credit into "data," denominated in bytes, allowing Internet access until the corresponding amount of data has been transferred and the balance falls to zero. The country's main commercial carriers also use prepaid models, and they are well-understood by local users. For distribution, the operator first generates credit via an admin interface and sells it to three different resellers in the community at a wholesale rate. The resellers own small stores, selling basic goods like rice, oil, and candy, and are open for most of the day. They pass credit on to end-users with a small margin using a locally hosted mobile web-interface developed by the researchers for the network.

Users can purchase arbitrary amounts of credit from resellers, and transfer it between users accounts. After loading their credit, there are three data packages available: 10MB, 100MB, and 1GB. Data pricing is flat, at Rp250,000(~$10USD)/GB. Local services are zero-rated, and all external traffic is billed equally.

Utilization

The Bokondini community network had ~50 users active at any point across the study period, where active means traffic to or from the user was measured in the network or the user made a credit transaction. Figure 3.2 shows the amount of traffic each day during the study period, the number of users active active at different aggregation levels, and the cumulative number of SIM

Table 3.1: Dataset summary statistics

	Log Count	%	GB	%
Internet Flows	56,001,999	74.5	1,324.9	98.9
Intranet Flows	19,179,804	25.5	15.1	1.1
Internet DNS Mapped	53,755,278	96.0	1214.6	91.7
Internet Org. Assigned	47,077,375	84.1	1250.7	94.4
Internet Categorized	46,826,941	83.6	1219.9	92.1
Transactions	40,450	100	-	-

cards registered with the network. We expected some attrition as SIMs are lost and replaced or users leave town. The combination of school break and several community members streaming a religious conference led to particularly high usage in late December and early January. During the last month of the study, a national operator began offering LTE service nearby, at a lower price point than the community network. This is correlated with the drop in traffic and active users observed in late February and early March. Despite the drop in activity, the network remained profitable.

3.4 Dataset & Methodology

We leverage two collected datasets: 1) network credit transaction logs, and 2) fused logs of IP and DNS metadata. We tightly integrated our instrumentation with the operator's existing systems to avoid overhead and gain access to ground-truth per-device information, but this required accepting the availability limitations of these existing systems detailed in section 3.3.2.

3.4.1 Data Collection

Credit Transactions

The credit transaction log records (1) when an administrator adds credit to the system, (2) when credit is transferred between users, and (3) when a user ultimately purchases data with their credit. Each entry contains a timestamp, user ID, transaction type (see section 3.3.2), amount, and for transfers, a destination user ID. It is implemented as an asynchronous Javascript module extending the network's existing admin application. This analysis contains 40,450 transaction logs (see table 3.1).

Data Flows

The data flow log records an entry for each flow in the community cellular network. Flows are defined by their "five-tuple", consisting of the IP source and destination address, the transport layer protocol, and transport layer port numbers if applicable. The data was collected IPv4/IPv6 agnostically, and we use the term "IP" here to refer to either IPv4 or IPv6. The system aggregated individual packets from each flow into 20 minute intervals, and at the end of the interval recorded the five-tuple, the interval start and stop timestamps, and the total number of bytes transferred in uplink and downlink. A shorter interval was not used due to anonymization concerns (see section 3.4.3) and limited local storage. By integrating with the network's policy control and billing system, our instrumentation could map each flow to a SIM card and user account. A post-processing step replaced the flow's local address with a scrambled user ID, coded with the same key as the credit transaction log to associate flows and transactions with the same entity. The raw data contains 75,181,803 flow logs.

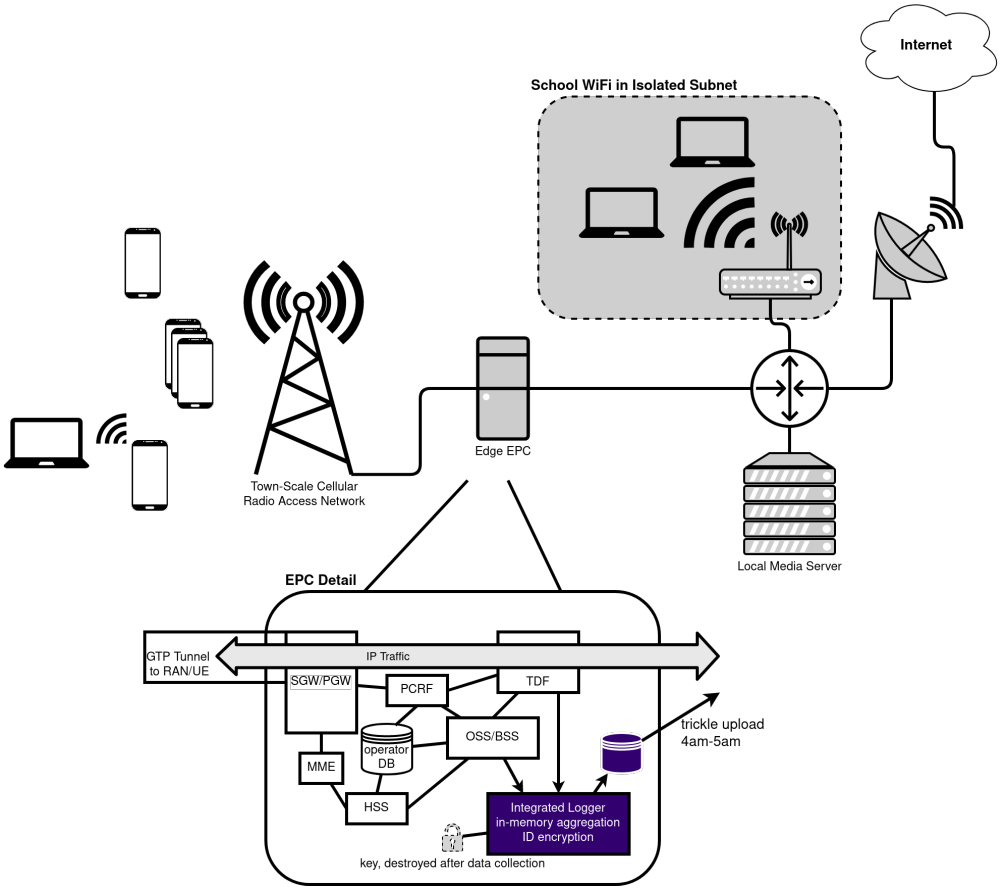


Figure 3.3: Data Collection Instrumentation Points. Instrumentation for flow collection occurs in the edge EPC deployed in Bokondini. Flows are aggregated and anonymized locally before being uploaded for privacy and to save on constrained upload bandwidth.

Intermediate DNS Responses

DNS establishes a likely mapping between observed destination IP addresses and the domain name a user is contacting. To augment the flow data with domain information, we collected an intermediate dataset of response timestamps, requesting scrambled user ID, domain requested, dns server response code, list of IP addresses returned from the DNS server, and list of address TTLs.

In post-processing we reconstructed the likely DNS state at each user's client over time, modeling each user as a single device with a shared DNS state, and iterating through the combined raw DNS and flow logs by time for each user. Each DNS response updates the client state to map the response IP addresses to the requested domain name. Since multiple DNS entries may point to the same IP address, the mapping can be ambiguous. We record the set of possibly ambiguous names and select the most recent for annotation. For each flow encountered, the current client DNS state is consulted. If the IP address has a known mapping, it is assigned from the DNS state. If the IP is not in the client state, we attempt a reverse DNS lookup for the IP address. Only if the reverse DNS lookup fails, we mark the host as unknown. For each flow we record the name, type of mapping (Client DNS or reverse DNS), and the count of possible ambiguities. While simplistic, we find this model mostly sufficient for this dataset, with 50,676,332 of flows covering 90.5 % of bytes coming from observed client DNS responses, and 41.1% of flows and 31.7% of bytes having an unambiguous mapping.

3.4.2 Data Processing

Once collected and anonymized, the data was uploaded to a central server for analysis. We periodically uploaded subsets of the data during off-peak hours to minimize the impact on user traffic. Before analysis, we removed users who join the network less than one week before the end of the data collection period (N=0) or who were active for less than one day (N=3). After filtering the final dataset contains 168 users and 72,278,238 flows.

Table 3.2: Notable events impacting network operation.

July 2017	Initial site visit and surveys
October 21, 2018	Initial launch of network with pilot users
October 31, 2018	Scaling of network by adding 10 new users
Feb 18, 2019	Transition to open network
March 10, 2019	**Beginning of dataset**
July 12-26, 2019	Extended outage due to relay lightning strike
July 26, 2019	Reconnected directly to school's VSAT
July 26-Sept 1, 2019	No credit sold while working with school
November 22, 2019	Operation extended to 4:30am to Midnight
December 1, 2019	VSAT Upgrade to 3/1 Mbps at 4:1 contention
February 20, 2020	National carrier begins operating 4G nearby
March 15, 2020	**End of dataset**

Classifying Domain Names

Through manual inspection we hand-built rules to classify domain names and tag them with an "organization" and "category." We built the classifiers by iteratively grouping flows by domain, sorting by the total bytes transferred, and looking for patterns in the top domains. As passive observers in the network, our methodology provides no ability to accurately determine the ground truth contents of encrypted flows. We consulted the domains themselves, whois data, publicly available documentation in the case of APIs, and sites hosted at the domain to determine the parent organization responsible for each domain and its contents. We classified all domains and IP addresses with 200MB total transferred in the network or more, assigning an organization and category to over 83 % of flows covering over 92 % of the total traffic.

We have made a best-faith effort to categorize traffic as thoroughly and faithfully as possible, preferring more detailed categories like "video" or "messaging" over generic ones like "social media" or "software and updates" where possible. We fully acknowledge the limitations of this approach, discussing them in more detail in section 3.4.4. Ultimately we mapped 98 organizations and 20 categories. We provide these artisanal classification rules and processing scripts for independent

scrutiny and reuse, see section 3.4.3.

Detecting Peer-to-Peer Flows

We found peer-to-peer flows facilitated by ICE (Interactive Connectivity Establishment) account for a notable fraction of traffic, particularly in the uplink. We separate these flows into their own category by reconstructing the ICE state at each client and in the network NAT/firewall. ICE works by having each peer open a connection from the client to a common server, establishing an open port in each peer's NAT at the NAT's public IP address, which is visible to the server. The server then shares the each peer's public port and address with their counterpart, which the peers attempt to re-use to establish direct connections if their NATs/firewalls allow it. If a direct connection cannot be established, the server falls back to relaying connection packets on behalf of the peers.

We reconstruct the ICE state by iterating through the flow logs for each user, and tagging any flows to the well-known STUN/TURN listening ports (UDP:3478 or TCP:5349) as likely ICE. After an ICE flow is detected, if a new flow starts within 1 minute to an unknown IP address but with the same client ephemeral port, we reclassify the flow as "Peer to Peer" instead of "Unknown." Due to port reuse, there is a low but nonzero probability of false-positive detection, so we do not reclassify flows already classified by domain.

3.4.3 Ethics

This work utilizes anonymized per-user flow metadata and network transaction history to better understand the dynamics and economic sustainability of Community Cellular Networks. In consultation with our institution's IRB and per locally applicable Indonesian data protection law, we determined that this work did not need explicit end-user consent since users were aware that this information was accessible to the network operator, there is low risk of harm, and a consent

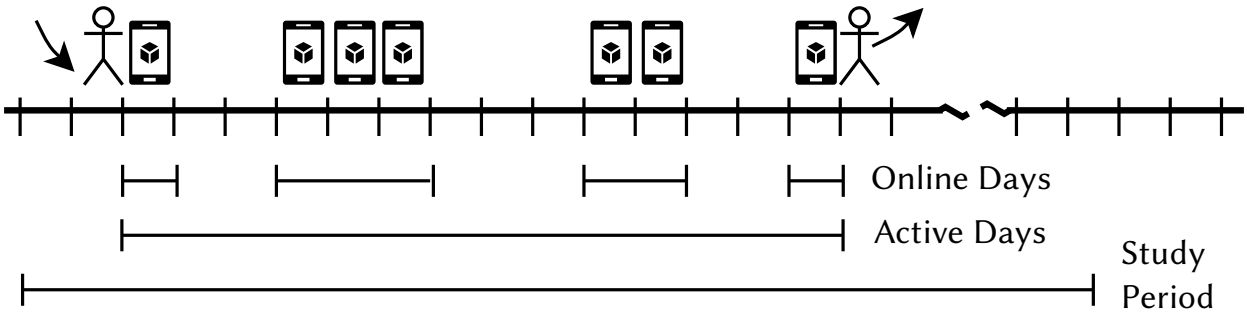


Figure 3.4: Visualization of Online vs. Active Days. For each user, we count an "Online" day when the user transferred at least one byte that day. We define the number of "Active" days as the span between their first and last interaction with the network.

process would require us to collect identifiable user information which we would not gain access to otherwise. Insights from this analysis have been shared with the network operator to improve the quality of service in the community.

Anonymization and Data Access

During data collection the operator scrambled all network IDs and transaction IDs with a key, which remained in the community and was destroyed after data collection. The key scheme allows correlations between the flow and transaction logs over the study period only. In generating flow metadata, only L3 addresses, the L4 protocol number, L4 ports, payload size, and unencrypted DNS responses were collected. Packets were binned into 20 minute flow chunks and aggregated before storage, preventing fine-grained timing analysis. All traffic to organizations with N<5 unique users was grouped and references to the original domains were dropped from downstream analysis. The dataset is available open-access at https://doi.org/10.6084/m9.figshare.13116740.

3.4.4 Limitations

Flow aggregation

Using aggregated logs instead of per-packet traces limits the analysis resolution and obfuscates protocol-level performance. In particular, we cannot comment on the efficiency of individual flows at the transport level, leverage deeper packet inspection to verify which higher-level protocols are in use, or use ML-based timing analysis to predict the flow contents.

Domain and IP-Based Classification

The process of content categorization is subjective by its nature, but essential for high-level analysis, since modern CDNs and distributed edge infrastructures mean that large numbers of unique domains map to the same organization and service. Our dataset's limitation to aggregated flows only adds uncertainty to this process. As an example, we categorize general infrastructure from traditional social media providers like Facebook or Twitter as general "social media," except if the subdomain explicitly indicates it hosts video (`video.*.fbcdn.net`), messaging traffic (`mqtt-edge-*.facebook.com`), or advertising (`lithium.facebook.com`). While some organizations have an infrastructure more amenable to categorization, which uses different domains for different types of content, others do not. For example, all TikTok content appears to come from one set of converged media servers, even though the platform supports both video and messaging. We classified converged services by their predominant category ("video" in the case of TikTok), or a mixed category if there is no dominant content type. In the cases of Google and Facebook, it is difficult to distinguish traffic from different user-facing applications but that are part of the same corporate conglomerate. For example, it appears that WhatsApp and Instagram use media CDNs hosted at `fbcn.net` subdomains, and YouTube pulls content from `video.google.com`. All users connected to IP addresses in Google-owned blocks that had no publicly queryable DNS information. We attempted to break applications into their own classes when possible, but were

not able to in all cases.

Generalizability

We partnered with the network in Bokondini due to its extreme remoteness and its uniqueness as a standalone satellite-backed LTE network with an approachable operator. All places are unique, but Bokondini is geographically and culturally similar to others in the remote highlands of Indonesian Papua. We don't claim to show generalizability due to the limitations of our dataset, but see no reason why our findings should be site specific.

3.5 Results

In this section we detail the major results of our analysis. Contrary to our expectations given the network's low throughput (3/1Mbps D/U shared across all users), we find highly unequal consumption between users, that many users consume with intermittent access rather than frugal access, and that video and major platforms still play a large role in the network. Additional non-essential information and supplementary plots can be found at https://github.com/uw-ictd/ccn-traffic-analysis-2020.

3.5.1 Inequality & Sustainability

A Wide Range of Spending

Network use was highly skewed, with some users spending 5.5x the average amount, and 8.3x the median. Figure 3.6 shows the distribution of total network revenues from each user. While over 20% of users spent less than $10USD and 50% less than $100USD, three spent over $1000USD equivalent, generating an outsized portion of total revenue. Figure 3.5 normalizes spending by the amount of time users are active. The three heavy spenders are visible as outliers in the top right of the chart, spending a large amount per day on average and connecting consistently.

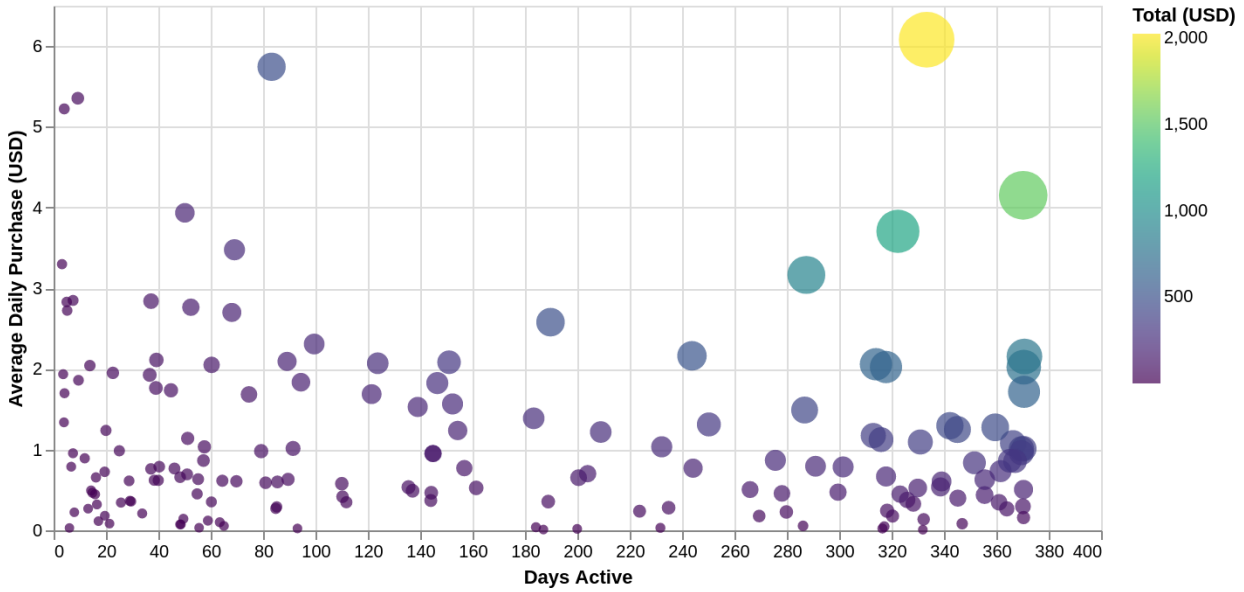


Figure 3.5: Daily Purchase vs. Days Active. Total purchased amount is encoded in the size and color of each point. The x-axis indicates how many days a user was active in the network, and the y-axis indicates the average amount of credit spent by the user when they were active. Most users make small purchases independently of how long they have been active, and thus contribute little to overall revenue (as indicated by their small dots) despite many being long-term customers.

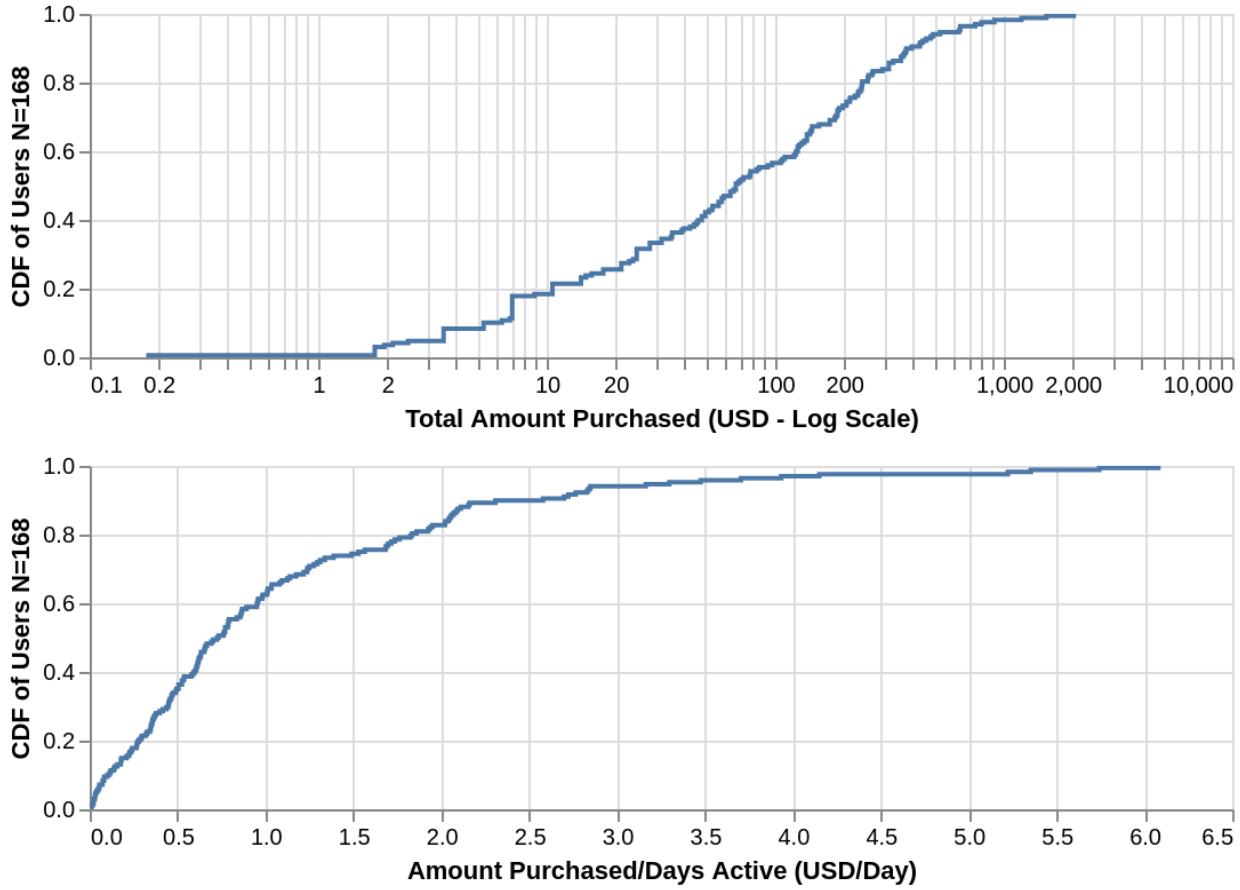


Figure 3.6: Distribution of spending across users. Consumption was highly unequal, with over half of users spending less than $100 USD while some users spent over $1000 USD.

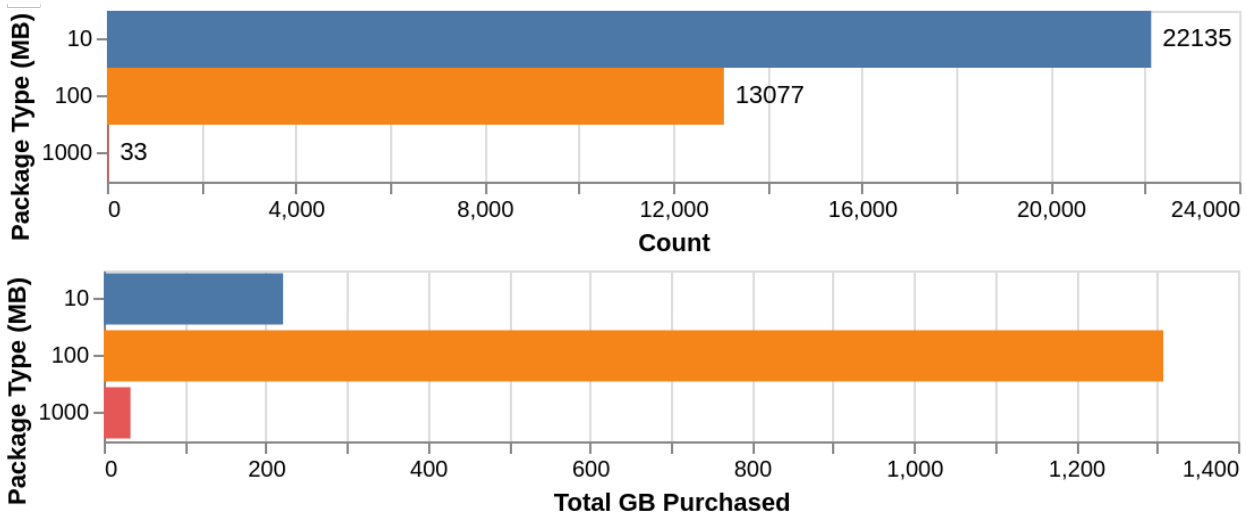


Figure 3.7: Relative use of data packages. The small (10MB) was most popular, but most bytes were purchased with the medium (100MB). The large (1GB) package was used rarely.

The heaviest users average over 300MB per day online, while the median consumes only 76.6MB. This disparity surprised us, since we hypothesized the network was bandwidth constrained, would not meet demand, and would have many users at the ceiling of available capacity. This has implications for the planning and operation of remote networks, further discussed in section 3.6.3.

Frequent Data Topups, But Sporadic Credit

The network billing system offers 3 data packages: 10MB, 100MB, and 1GB to purchase with account credit at a uniform price per MB (see 3.3.2). Examining the transaction logs, we found the 10MB package is the most popular, while most "bytes" are purchased in the 100MB sized package. The 1GB package is rarely used (see figure 3.7). The network's flat pricing schedule does not incentivize purchasing large packages, and users will often quickly purchase several small packages to synthesize the exact amount they want. We grouped the chains of purchases which occur within one minute from one to the next, and plot them in figure 3.8. Synthesized packages are still often small, with the bulk of packages coming in at 200MB or below. The network's

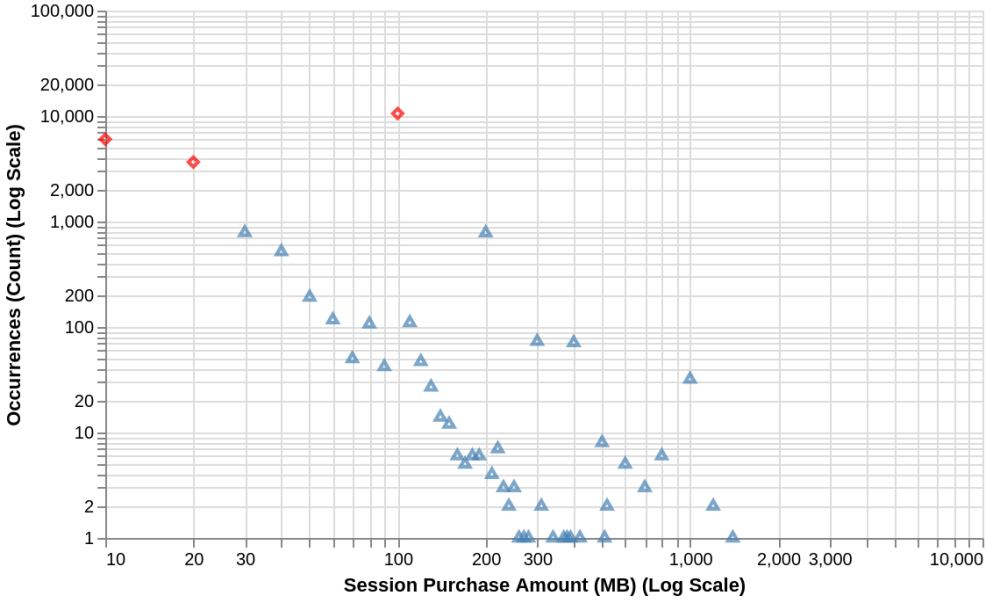


Figure 3.8: Amounts of data purchased in a single "chain". Amounts which occurred more than 1000 times are marked with red diamonds. The predefined 10MB and 100MB amounts were most popular, but users frequently synthesized non-standard packages to better suit their needs.

interface is designed to minimize friction for this workflow, requiring only two taps to make a data purchase.

Users tended to make frequent small data purchases through the web portal multiple times a day, even after accounting for purchase chaining. Figure 3.9 shows a CDF of the time between data purchase chains across users. 95% of users make a purchase every 10 hours or less on average, and over 90% of users have a 90th percentile inter purchase time of 10 hours or less as well. Frequent purchasing could help manage overall consumption and provide a sense of control over spending and background processes.

We hypothesized there would be a clear distinction observable between users who maintain a store of credit in the network for on-demand data purchasing and users who do not, but we found the reality to be much more ambiguous. 23.8% of users maintain a positive credit balance more than 95% of the time, but they tend to be new to the network. Of users active for more than 30

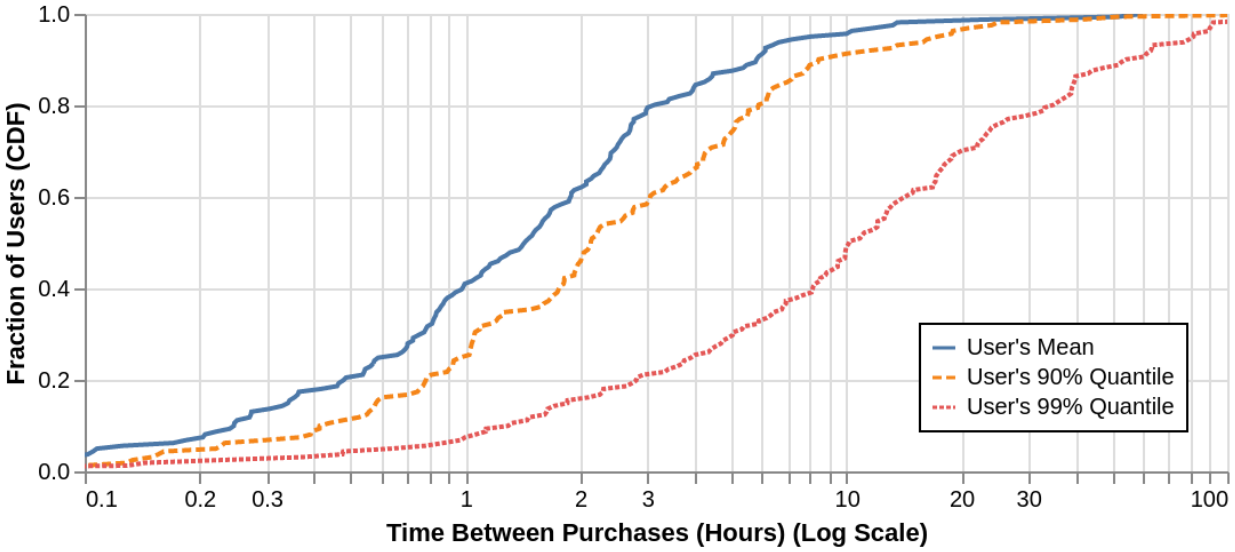


Figure 3.9: Time between user data purchase "chains". Most users commonly purchase data multiple times a day.

days, only 19.3 % have a positive balance more than 95% of the time. Figure 3.10 plots the fraction of time that each user had no credit while active versus the number of days they were active. We observe a wide variety of ratios at all time scales, indicating that the amount of time a user has zero balance may be more random and situational rather than a strategic choice.

Having a zero balance means that the user does *not* have data available on demand (in case of emergency or otherwise), and would need to visit a physical reseller to get access. In other contexts researchers have noted that users may prefer to not carry a digital balance to avoid pressure from friends and family to loan them data [197], or to keep reserves in more flexible cash with different affordances for bargaining and negotiation [105].

Intermittent Use

Network use, even among the heaviest users, is highly intermittent. Comparing the number of days a user is online to the number of days they are active (see figure 3.4), we find that the median

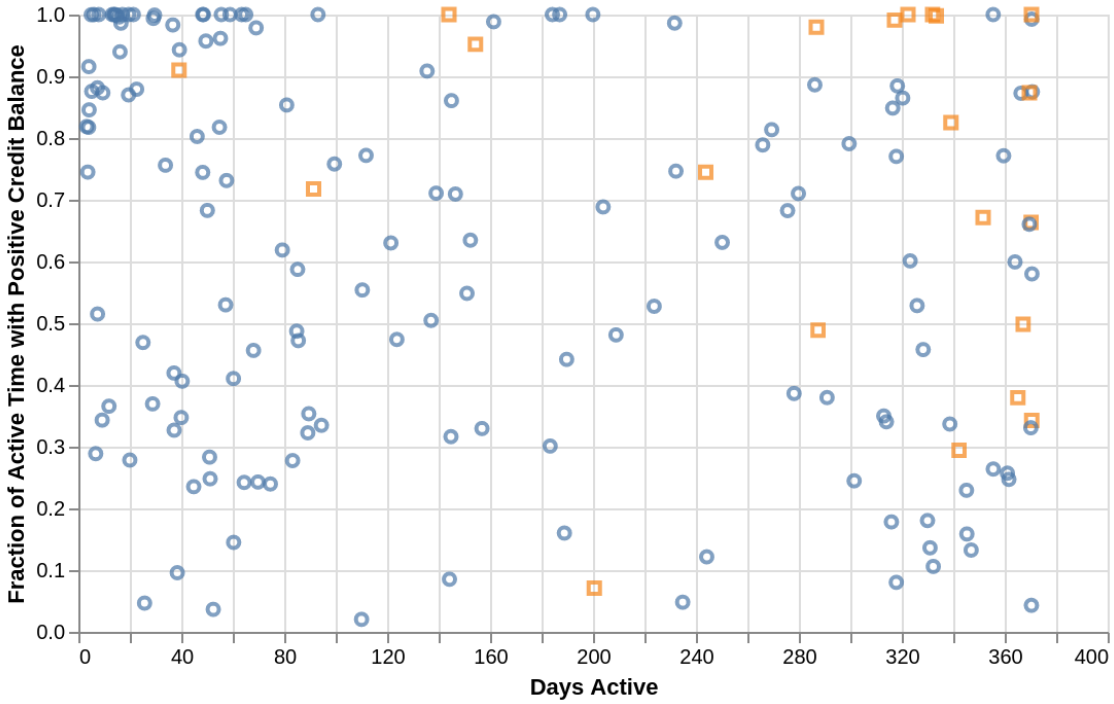


Figure 3.10: A cluster of new users tends to have nonzero balance most of the time, but a wide variety of ratios can be seen across both long-time and relatively new users. Orange square users were already members at the beginning of logging and have a manually adjusted start balance.

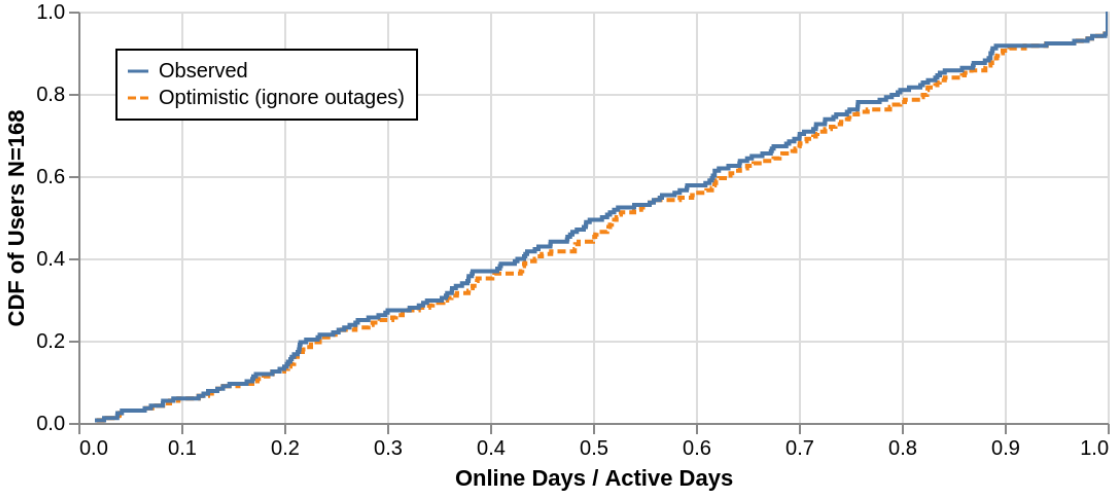


Figure 3.11: CDF of online ratios. The fraction of days users connect is highly variable. ~8% connect most days, but the remaining ~90% are distributed across a range of online ratios, even conservatively ignoring days with unplanned outages.

user is offline 53% of the days in their active time window. Only 7.7% of users access the network more than 95% of the days they are active. Figure 3.11 shows a CDF of the user Online Day/Active Day ratio, showing a roughly uniform distribution of the fraction of time online after accounting for the small number of users always online.

This itermittency impacts the operator's network planning and business operations, and is reflected in top-level statistics about the network. Figure 3.2 demonstrates this, where the count of unique daily, weekly, and monthly users differs substantially.

This intermittency combined with the varied amounts of time users spend without data-on-demand reflects on the core use-cases of the network. For the majority of users, their connection is *not* an always-available lifeline, but rather more sporadic and asynchronous. This may be partly due to the pace of life in general in the community, where residents are used to tasks taking days or weeks due to infrastructural limitations, or the availability of the national carrier's 2G network for small urgent messages, making the LTE network less essential.

The Network Is Financially Sustainable

Despite intermittent use and a relatively small number of users, the Community Cellular network in Bokondini is financially sustainable *without an external subsidy*. Re-use of local infrastructure and local installation labor kept the install capital expenses below Rp150M (~$10,000USD). Regular maintenance, mostly related to the power system, averages Rp1.3M (~$95USD)/month, and the satellite subscription costs $300USD/month (~Rp4.3M). Repairing the backhaul after the lightning strike cost two months of lost revenue and $1000USD (~Rp14M) in repair costs, but was covered by existing backup funds. Even with the downturn in use in February, revenues exceed costs, and are being invested in expansion to surrounding communities.

Figure 3.12 visualizes the network's cumulative revenue over the study period (ignoring early revenue from the pilot). Calling attention to the role of anchor users, we plot the revenue curves

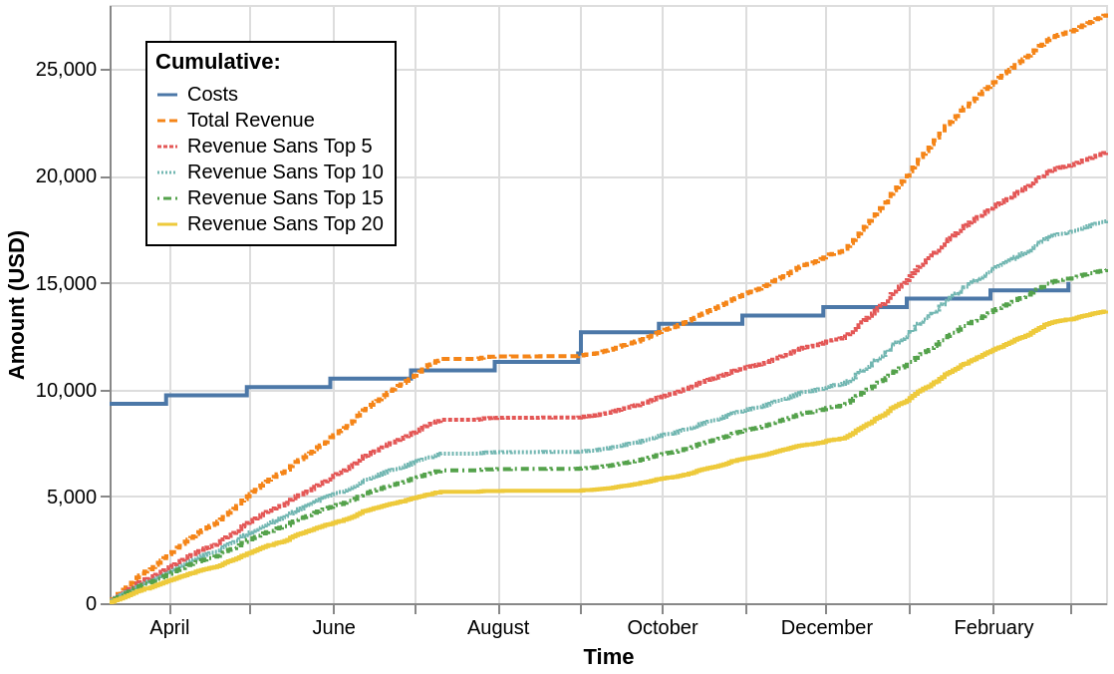


Figure 3.12: Revenue and Costs vs Time. The Bokondini network is financially sustainable, with top users contributing a large share of revenue. The costs line includes the upfront capital expense to deploy the network, regular operational expenses of monthly maintenance and backhaul, and the incidental operational expense to repair the network after lightning damage. User support is handled informally at the network's small scale and not accounted.

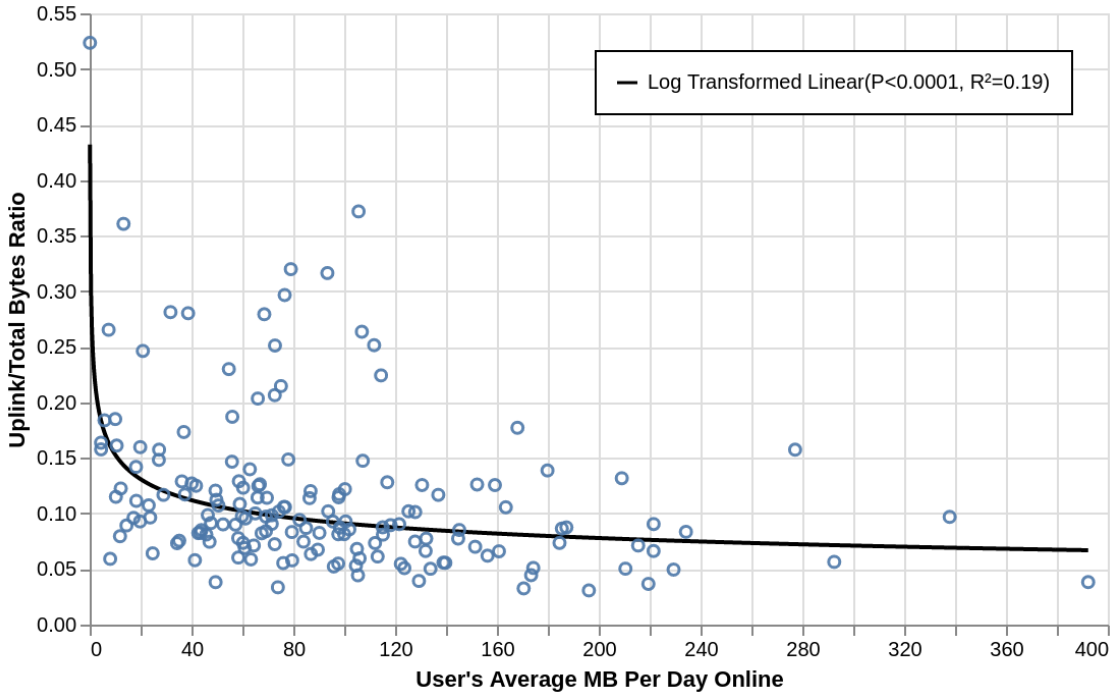


Figure 3.13: Uplink/Downlink Ratio vs. Consumption. Many light users have relatively more uplink traffic than heavy users, placing a different workload on the network.

excluding the top 5, 10, 15, and 20 overall users. The commonly quoted "ARPU (Average Revenue Per User)" metric does not capture the diversity of the underlying user population, and the impact that losing even a handful of these anchor users would have on the network. Without the top 15 users the network would still be sustainable after a year of operation, but a typical rural small cell site, with capex costs on the order of $40,000 USD [172], would not be. We discuss sustainability further in section 3.6.3.

3.5.2 Whales Engage Different Parts of the Web

We expected users would be limited similarly by the tight constraints of the Bokondini network on the modern web, but we find structural differences in the traffic of heavy and light users. Light users tend to have more balanced uplink/downlink ratios than heavy users, less video traffic, and

also to abstain from games, content uploading, and dedicated antivirus. This indicates that rather than just using the network less, lightweight users are actually using the network differently, consuming a different mix of content, likely encountering different performance constraints, and placing a different burden on the network. Figure 3.13 plots the uplink/downlink ratio of each user vs. their average consumption. There is a weak but significant overall correlation ($P < 0.0001$, $R^2 = 0.19$).

Video

Examining the categories of flows attributable to heavy and light users, we find that video traffic makes up a disproportionate fraction of content from heavy users, while other categories stay relatively constant. All users have at least some traffic in the video category, but figure 3.14 shows the explosive growth of the video category between light and heavy users, concentrated in the top 10% of users overall. General social media use increases for the top 50% of users, but does not see the dramatic increase video does. We note that video from mainstream applications like Youtube, Facebook, and TikTok significantly outweighed adult video sites (~9:1).

In the network as a whole, video (both adult and non-adult) only consumes 37% of the download bandwidth, compared to the global mobile Internet market where video makes up 65% of mobile download traffic [46]. Examining only the 10 heaviest users, video still only makes up 49% of download bytes. All users in Bokondini consume less video than the average global user, and the median user consumes significantly less (~1/3) as a share of her total consumption. The under-representation of video overall compared to global trends could indicate that prices are too high to support carefree video streaming and the media rich Web, or the network may not have sufficient capacity to meet demand.

Hotspotting: NBD

While anecdotally most users connect to the network via a smartphone, we observe traffic to domains commonly associated with computers, such as `update.microsoft.com` (Users=5), `download.adobe.com` (Users=7), and `cdn.mozilla.net` (Users=16). Although we expected PC users to consume significant traffic, and these users all fell into the top 50%, they were not the heaviest users in the network. Any educational outreach and/or tool development to manage network traffic will need to focus on mobile media consumption for the greatest impact.

3.5.3 Platforms, Reach & Utilization

Breaking traffic down by organization, we see that some organizations interact with almost all users, while others communicate with only a small subset. Unsurprisingly Facebook and Google receive traffic from all users, and WhatsApp from almost all, but TikTok, QQ Messenger, Twitter, ShareIt Games, W Share, and UShareIt are also popular, even though they account for a smaller share of overall traffic. Compressed web content, consisting of AMP pages or sites served through the UC browser, was also relatively common. All users interacted with local services to purchase more Internet data from their credit balance.

Large Platform Dominance

Traffic to and from Facebook owned properties (Facebook, WhatsApp, Instagram) made up 39.9 % of bytes, exceeding the Asia-Pacific regional average of 35 % [46]. Google and Google-affiliated services account for 31.5 %. Taken together, these two platforms alone account for 71.4 % of traffic in the network, supporting observations that large platforms (Facebook in particular) have wide reach in remote contexts.

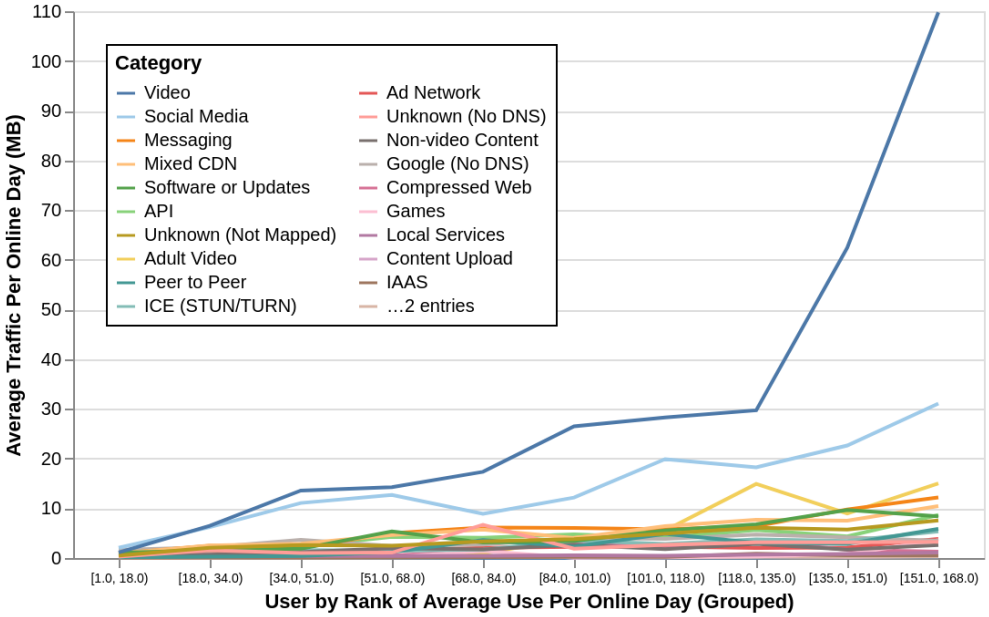


Figure 3.14: Bytes per Category vs. Decile. Video and Social media grows disproportionately between light and heavy users. The top 20% consumes a large share of total video traffic.

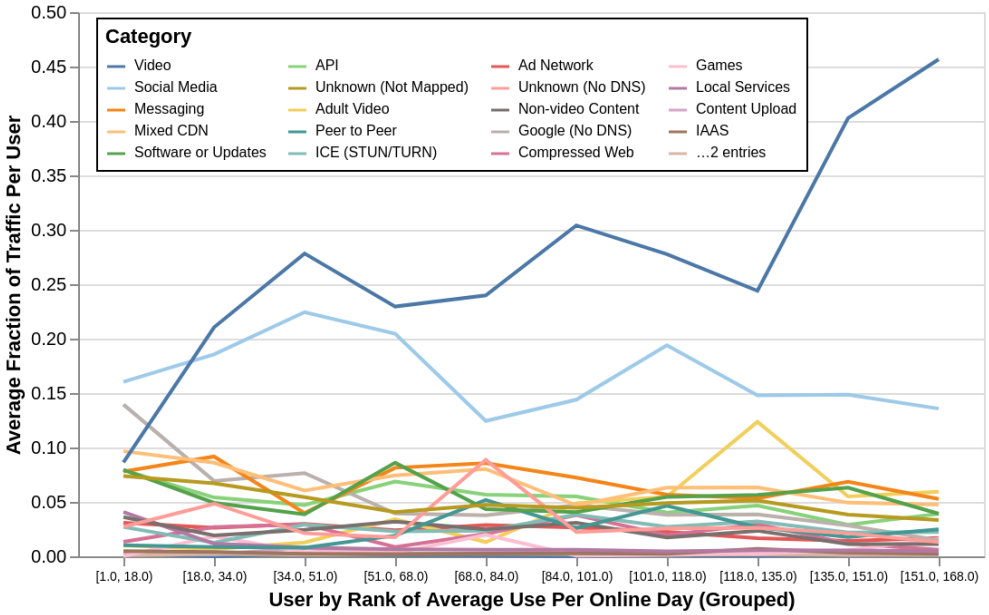


Figure 3.15: Share per Category vs. Decile. Video use as a share of total traffic increases sharply for the top 20% while most other categories remain flat.

Messaging Universality

While video and content consumption account for the most bytes on the network, we found that a significant amount of network resources go to messaging traffic and realtime UDP flows, particularly in the uplink. 90 % of users have some UDP uplink traffic facilitated by ICE (see section 3.4.2), and ICE or peer to peer bytes make up 6.1 % overall and 26 % of the uplink. Widespread messaging and communication via the Bokondini community network is surprising given how intermittently most users are connected and the availability of competing 2G voice and SMS services. Intermittent messaging use offers an example of how technology can be adapted to the constraints of remote edge networks in distinct ways from how it is used in well-connected areas. We discuss OTT applications and peer-to-peer communication further in section 3.6.2.

Local Trends

Due to the small number of users, local phenomena can cause large operational impacts to the network. In December and January we observe 22 users start streaming sessions from a month-long conference, consuming ~20 % of the network's resources for most of the month. The conference site was the 8th heaviest destination for the year overall, even though it was only visited for a little over a month by a small set of users. Ways to cache or "re-stream" content on the local side of the satellite link could greatly mitigate the impact of similar local phenomena.

Resource Utilization

The total data transferred per day had high variance, with a mean of 3.82GB and standard deviation of 2.15GB. There are some notable outlier days, where the network saw more than twice as much traffic as usual. Within each day, visualised in figure 3.16, there tends to be a slight bump in utilization around 12 noon (lunch time) and a stronger increase in use in the early evening from 6-10, peaking at around 9.

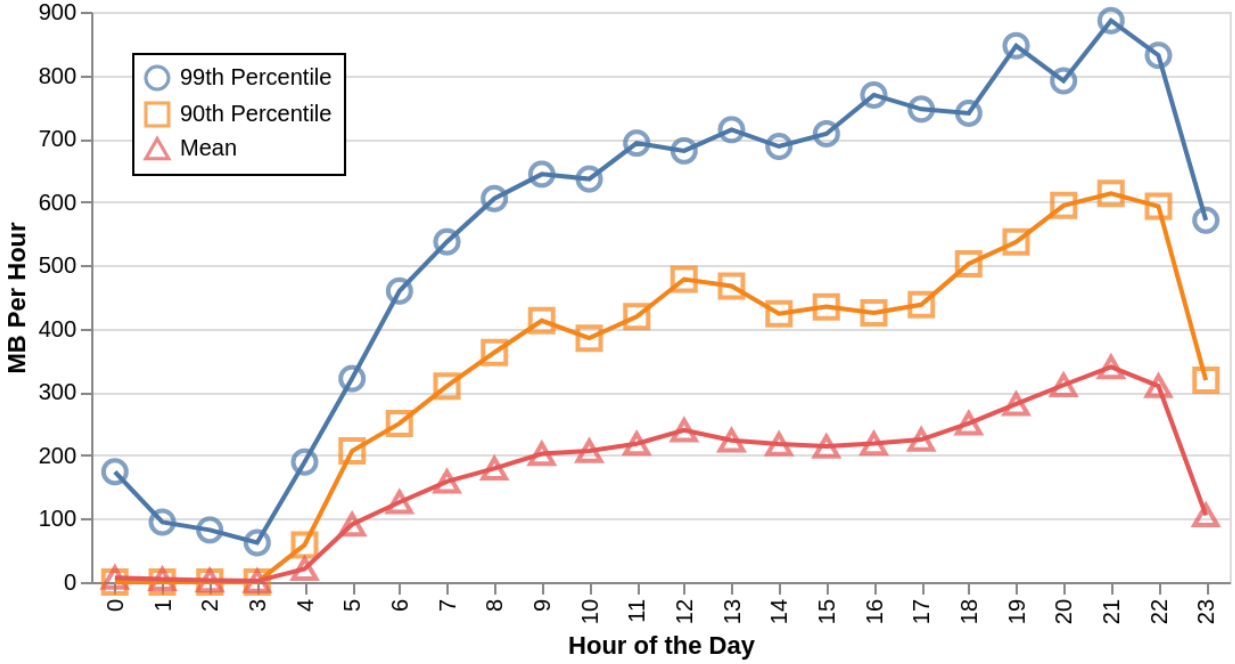


Figure 3.16: Bytes per time of day. The network is usually powered down at night, but it is a manual process and is skipped some days if there is power to spare.

Although the operator maintains a symmetric infrastructure internally, their satellite backhaul is asymmetric with a 3/1 mbps down/up ratio. Observed traffic actually exceeds this ratio, with an overall down/up ratio of 8.5:1 (but which varies per user, see Figure 3.13). This indicates downlink is the likely bottleneck for the system workload, and the uplink could be better utilized or reduced.

Local Services and Local Only Traffic

The network has two zero-rated local services: one is a portal where members can view their balance, purchase more data with credit, or transfer credit to another user, and the other is a local media server stocked with material from the school hosting the network. All users interact with the portal, but it is a basic web application and does not contribute much to the total traffic. The media server is heavier, serving video and other rich content, but sees much less regular demand.

While the network allows peer-to-peer communication within the community, there is only

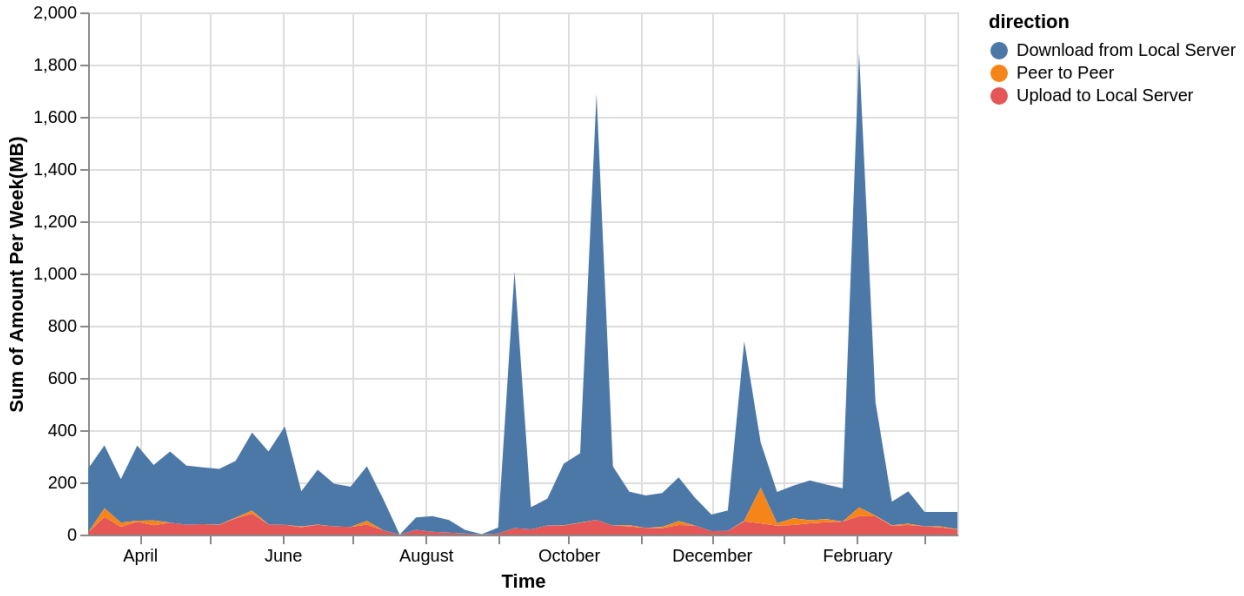


Figure 3.17: Local Traffic. The amount of local traffic is bursty, but relatively small in magnitude compared to Internet traffic, even though local traffic is zero-rated (free).

very sparse peer-to-peer traffic. Almost all local interactions are between a network user and one of the two provided services. This may be due to a lack of knowledge that peer-to-peer is available, or assumptions built into the wider ecosystem that most LTE networks do not allow local connectivity. The widespread use of sharing apps indicates that there is demand for sharing, but almost all sharing activity in our dataset is mediated by Internet services rather than relying on direct discovery in the local network. This is supported by prior work [79] that found that community cellular networks are used disproportionately for external communications.

3.6 Discussion

3.6.1 Practical Challenges Measuring Small Nets

This study presents a longitudinal deep-dive into the operations of a small Community Cellular Network providing sustainable data services in an area only recently reached by other operators.

Conducting this study required significant multi-year effort to coordinate and integrate new systems, which does not scale down proportionately with network size. By web standards our dataset is miniscule, and a similar level of integration effort with a larger network would have yielded a much more statistically powerful result, just due to the scale of the targeted network. Yet small operators play a large role in rural access [75], and face unique constraints and challenges. Tools and techniques, both technical and organizational, that lower the burden of conducting research and developing technologies for these environments will likely play an important role in expanding Web access over the next decade.

3.6.2 Network Design for Content

Video and Social Media in Constrained Networks

Despite operating at the extreme edge, we saw video and social media weigh heavily in the composition of the network's traffic. Video is delay tolerant and resource intensive, but currently peaks at the same time as real-time communication flows. Prior studies have found that media often circulates locally in close-knit communities, and is delay tolerant by nature [190]. Traffic shaping tools to identify heavy flows and prioritize realtime traffic on the satellite backhaul downlink could improve the experience of a large number of light users at the expense of only a minor delay in a video download. Incentives to demand-shift video consumption (by marking videos to download in the early morning), or encourage local peer-to-peer sharing could also be explored.

Policy around video opens interesting ethical questions about the values embedded in the network's operations. Is video more or less important than other types of traffic, and is it acceptable that video consumes such a large fraction of the link?

Messaging Applications & Peer To Peer

A key differentiator between the network profiled in this research and prior work is that it is a **data only** LTE network. The Bokondini network operates neutrally, charging a price per byte independently of the type of traffic. This forefronts the primacy of flexible over-the-top messaging services, such as WhatsApp, Messenger, or Viber, over in-network protocols like Voice-over-LTE (VoLTE). While an all-data approach eschews traditional network-based quality of service differentiation, our experience in Bokondini shows IP-based designs to be a massive success in the context of rural networks with extremely limited backhaul, even for voice.

Anecdotally the network operator reports users think the call quality of WhatsApp over the community network substantially exceeds the quality of calls via the national operator's existing 2G network and new 4G network, yet it is not immediately clear why this is the case. Telecom standards by their nature lack flexibility to adopt new technologies, and can lag behind state-of-the-art techniques than be quickly pushed out to OTT applications. OTT services, unlike centralized telecom voice services, also naturally support peer-to-peer communication, and transparently establish low-latency connections within the community when possible. WebRTC extends this capability to browser-based Web applications as well [192]. While local call routing is possible in LTE, it is inconsistently implemented since it adds signalling complexity and limits the telecom's ability to correctly track calls. Peer-to-peer may be a nice-to-have feature in well-connected areas, but it can be the difference between usability and frustration on the extreme edge.

In future work, we hope to rigorously measure the impact of OTT app design on service quality, digging into the anectdotal evidence that OTT apps are both less expensive and more performant in extreme-edge conditions despite the lack of network integration.

3.6.3 Sustainability & Whales

The sustainability of rural networking solutions is a hot topic in many policy circles. A host of models exist, leveraging measurements such as the network ARPU (average revenue per user) to show where networks are and are not viable. Our analysis provides a more thorough look into user behavior, specifically the *distribution* of subscribers in an area. Figure 3.5, in particular, shows that these networks have a wide range of types of subscribers, perhaps more akin to the "whales" present in the freemium games literature [171]. In our case, three users dominate revenue generation, pushing the network ARPU (Rp190,308/mo) significantly above the country average reported by Telkomsel (Rp47,000/mo in April 2019 [33]). Coarse metrics like ARPU, and even average installation costs, obscure the reality that each community is a unique location with unique citizens. While the network is still sustainable without these three individuals, we believe it is important to understand that "there is no average user" [30].

In terms of building models for future deployments, a survey of the community may not find these individuals or may find just them. This variance, inherent in operating where the overall number of residents is low when compared to dense urban situations, makes predicting the sustainability of these networks more difficult. Because of this, we argue that bottom-up decision making, likely via *local* entrepreneurs interested in providing connectivity to their communities, is a more efficient and sustainable way to allocate resources on the edge.

3.7 Future Work

Edge Measurement

The practical difficulty in gathering measurements from the network was surprising. Part of this was due to the fact that we were not colocated with the network, but with nearly 30GB of data to analyze, being present in Bokondini would have stripped us of the ability to investigate the

data at scale using cloud resources. Our group is currently running or assisting in the operation of networks in the Philippines, Mexico, Hawa'ii, and the Arctic, so we expect this issue will continue to manifest.

We hope to develop a network-side application which could perform an efficient, potentially streaming, first-pass analysis and transfer only compressed results to a longitudinal telemetry service. This approach could draw from advances in federated learning [27], to keep sensitive data in the community while improving the quality of high-level analysis. This application could take into account network behavior and patterns, optimizing compute and network utilization against demand in the network. Optimizing the size and structure of the produced model given the expectation of future bandwidth availability will be a focus of our research.

Edge Caching

One result from this work is the importance of reducing backhaul reliance through caching and local loops. This is not a new topic in community networks; Guifi.net has done extensive work in local services [161] (though due more to political desires than backhaul limitations) and others have explored novel caching schemes in the developing world (notably Raza et al. [145]). Siskin [178], from Google, is a similar initiative to enable peer-to-peer connectivity in disconnected environments.

While these efforts bring novel ideas, there is no satisfactory answer for edge caching yet in an HTTPS world [57]. We do want to call attention to WPack [107], one particular effort that we find intriguing. WPack is a proposed standard for downloading and signing web content explicitly providing support for redistribution and caching. WPack seems well-suited to the remote community network environment and could re-enable network caching, while also supporting secure service utilization, if included in browsers. We are tracking WPack development closely and hope to eventually use it to implement a media cache for remote networks.

COVID-19 Analysis

While this chapter limits its analysis to only the time period before the wider onset of the global COVID-19 pandemic, we were able to collect data over the course of the lockdowns and following months. It is difficult to draw meaningful conclusions from traffic data without a clear picture of how the situation evolved on the ground, but in the following chapter I provide some high-level results observable from the network-level data over this time period.

3.8 Conclusion

Through integration with a local operator's infrastructure, I gathered a unique dataset to characterize and report a year of finances and utilization in a, remote, data-only Community LTE Network. With visibility of single users, I found use highly unbalanced and the network supported by only a handful of relatively heavy consumers ("whales"). 45% of users were offline more days than online, and the median user consumed only 36 MB per day on average, making frequent purchases in small amounts. This data shows that Internet-only Community Cellular Networks can be profitable despite most users spending less than $1 USD/day, and provided a characterization of some of the unique properties of the network that make it distinct from other large-scale Internet measurement datasets.

Chapter 4

| CCN Traffic Measurement |

COVID-19

Executive Summary

This line of inquiry reports on an opportunistic observational measurement of the impact of COVID-19 restrictions on network operations in a small, remote community cellular network.

Hypothesis

I hypothesized that the Bokondini network would see an increase in demand and utilization, similar to those seen in larger upper-income country datasets, as pandemic restrictions were put in place on travel and work in the community.

Summary of Key Findings

My collaborators and I found no significant difference in total utilization of the network before and after lockdown restrictions were put in place. Any variation was dominated by seasonal variation already visible in the network.

Next Steps & Connections

This work further supports the conclusion from chapter 3 that behavior in this extremely remote community cellular network is distinct along many dimensions from more general Internet networks, and that takeaways drawn from general Internet trends do not always apply in these more remote and constrained contexts. Many questions remain due to the limited nature of this study, and I wonder if there could have been ways for the network to provide more utility to its users during the pandemic crisis that would have measurably increased utilization.

4.1 Introduction

The importance of robust and performant Internet connectivity has become much clearer with the onset of the global COVID-19 Pandemic. Work, education, and social interaction moved online for many with the onset of lockdowns and travel restrictions for social distancing to slow the spread of the virus. According to the International Labour Organization, COVID-19 resulted in a loss of 8.23% world-wide labour income to lost working hours, and global unemployment increased by 33 million in 2020 [85]. Their study shows an even more significant impact on lower-middle income countries, which experienced greater losses in working hours than upper income countries, well above the global average by as much as 11.3%.

Many Internet organizations in upper income countries measured and released data showing significant changes in network demand and performance with the onset of the pandemic [44, 146, 112, 56, 28]. Given the unique characteristics of the network in Bokondini and our ongoing

measurement partnership with the community cellular network there, my lab mates and I wondered if we would see similar clear trends in the utilization and transfer data in the network. To our surprise, we found little to no measurable change in network utilization in Bokondini with the onset of pandemic travel restrictions in the area. I characterize our measurements and results in this short chapter, which can be seen as a supplement to Chapter 3.

4.2 Related Work

4.2.1 Community Network Measurement

Existing work has measured many technical aspects of community network operations, exploring the implications of the community network organizational models on the design and performance of the network. These analysis include the types of traffic and overall network resource utilization in particular networks [108, 79, 94, 74, 118], the types of devices present in the networks [81, 168], and the finances and sustainability of networks, large [34] and relatively small [95]. Our work builds from this existing measurement literature to explore the impact of a specific shock, the global COVID-19 pandemic and lockdowns, on the observed traffic in a small community network.

4.2.2 Measurements in Non-Community Small and Rural Networks

Outside of community networks (comprised of both developing and developed regions), a body of network measurement literature focuses specifically on rural and developing regions. Often utilizing traces from telecoms, ISPs, or regional IXPs, these works explore the unique circumstances of rural networks. These include studies on broadband performance and adoption in Nepal [100], Pakistan [16], or South Africa [41]; cellular performance in India [170] and Pakistan [14]; censorship in Pakistan [6]; mobile phone properties in a Pakistani cellular network [9], and web latency in Ghana [198]. These have scaled up to continent-wide analyses, such as IPv6 adoption [109], interdomain routing [54], and inter-country latency [58] in Africa. World-wide studies exist as

well, such as Schinkler et al.'s work on the performance of Facebook's edge caching [158].

A body of research also focuses on explicitly small-scale networks, servicing households or small groups of people but not in an explicitly community-oriented fashion. One example is Maitland et al.'s exploration of Internet use in a refugee camp [116], where the network is run by the UN Refugee Agency. They found a wide range of experiences with the Internet in the camp, contributing to a set of barriers to access. Another set of works focus on tribal Internet, with Vigil et al. [190] explicitly focusing on failures in the use of apps like YouTube in the context of a TVWS deployment in US tribal lands.

Unlike these existing studies focusing on typical networks operations, our work explores the specific implications of the global response to COVID-19 and ensuing lockdowns on a rural network. Our analysis focuses on comparing the characteristics of the network before and after COVID-19 responses begin, rather than attempting a general characterization or broad snapshot.

4.2.3 COVID-19 Lockdown Network Measurement

Following the COVID-19 response, researchers and industry practitioners have reported initial findings on how COVID-19 impacted the operations and traffic of Internet networks. Feldmann et al. characterized the impact of lockdowns in Europe and the United States by examining anonymized flow traces provided by partnering Internet service providers (ISPs), Internet exchange points, and academic institution networks [56]. They find an increase in overall traffic following regional lockdowns, a change in the daily timing of network traffic towards a more "weekend-like" pattern spanning all days of the week, increased use of both network-based entertainment and productivity applications, and a shift in traffic towards consumer ISPs and away from places of work like universities. Lutu et al. provide complementary insights from the vantage point of a major mobile network operator in the UK, finding a substantial *decrease* in mobile data traffic as demand shifted to home broadband networks coupled with a large *increase* in cellular voice

as users rely more on digital communication technologies amid lockdown restrictions. Candela, Luconi, and Vecchio use the global-scale RIPE Atlas platform [176] to observe an increase in latency and packet loss corresponding to increased strain on ISP networks. Böttger, Ibrahim, and Vallis report the observations of network demand changes at a global scale through instrumentation of Facebook's global edge infrastructure, confirming the global increase in traffic demand with lockdown, particularly for livestreaming and messaging services [28]. Many non-academic reports and posts from major Internet infrastructure groups corroborate these findings [146, 88, 44].

In contrast to these studies and reports, our work zooms in to focus on the impact of the local COVID-19 response on a small community network in a remote area at the edge of the Internet. As noted in Chapter 3, these rural edge networks have different, constraints, use-cases, and behaviors that can be lost in the noise of global-scale analysis. This work provides a characterization of the impact of COVID-19 from this missing remote perspective and finds results contrasting those reported by existing work.

4.3 Context

The data collected for this study also comes from the same community LTE network in Bokondini, Papua, Indonesia as discussed in Chapter 3. Bokondini is a small mountainous village with ~2,000 residents, located on the highlands of Papua province. Members of my lab have engaged with a local wireless Internet service provider and school for the last decade to explore unique solutions to rural access, with an original deployment of a 2G voice and SMS community cellular network in the early 2010s [79], followed by an upgraded LTE-based Internet-capable network in 2018 [167]. The network operates on extremely constrained resources, with energy from a local renewable micro-grid, and backhaul Internet connectivity provided via a shared 3Mbps satellite link. A more thorough overview of the context in Bokondini is provided in Chapter 3, Section 3.3.

4.3.1 Covid Response

Due to limited healthcare capacity, the local government took aggressive action to quarantine area communities beginning on March 25, 2020. This began with closing schools, then all roads and markets the following week. The local school hosting the cellular network shifted to a homeschool-based system in the last week of March. The network ecosystem in the community also changed drastically; the operators of two competing hotspot installations left town without providing infrastructure (e.g., network credits) for continued operation, and person to person contact within Bokondini was extremely limited. In practice restrictions slowly eased, but I do not have clear specific dates for when different services and travel became available.

4.4 Dataset & Methodology

In this section, we list the details of the data collection process and the information collected. Multiple anonymization steps were taken both prior and after data collection to protect user privacy.

4.4.1 Data Collection and Preparation

Data Collection

The data reported is collected from a small community cellular network in rural Indonesia, using the same process detailed in section 3.4, but extended to include data during the height of pandemic restrictions.

The network measurement system aggregated the individual Internet packets in each flow into 20-minute binned intervals, associating each flow with a user account and SIM card. The measurement framework inspects unencrypted DNS requests to associate particular flow addresses with a corresponding destination domain. Additionally, the network's credit billing system was

Table 4.1: Notable events impacting network operation across COVID-19.

March 10, 2019	**Beginning of dataset**
July 12-26, 2019	Extended outage due to relay lightning strike
July 26, 2019	Reconnected directly to school's VSAT
July 26-Sept 1, 2019	No credit sold while working with school
November 22, 2019	Operation extended to 4:30am to Midnight
December 1, 2019	VSAT Upgrade to 3/1 Mbps at 4:1 contention
February 20, 2020	National carrier begins operating 4G nearby
March 25, 2020	Schools closed due to COVID
April 1, 2020	Road to Capital closed due to COVID
May, 2020	Travel possible within town but external travel still limited
May 24, 2020	Free credit exploit first used
June 9, 2020	Credit exploit in wide use and detected by operator
June 10, 2020	Credit exploit loophole closed
June 13, 2020	Exploit credits traced and removed from exploiting users
February 7, 2021	**End of dataset**

instrumented to record when users purchase or transfer network credits between accounts. All user identifiers in both the flow and transaction logs are anonymized by the operator within the community before information was collected by the research team. See sections 4.4.1 and 3.4.3 for details. Data collection was conducted in accordance with Indonesian data protection law, and the data collection protocol was reviewed and approved by the researchers' Institutional Review Board.

Flows

Each of the flow logs records an entry for each flow in the network and consists of the flow "five-tuple"; the IP (either IPv4 or IPv6) source and destination address, the transport layer protocol, transport layer port numbers if available, timestamp of the flow, and number of bytes transferred in the uplink and downlink. Before processing, the flow source address is replaced with its anonymized user identifier and the associated domain, if available, is appended to the flow record.

The raw data contains a total of 205,220,004 flows.

Anonymization

All network IDs and transactions IDs were anonymized by the operator with a random key, which was destroyed after the data collection process was completed. A key was required to consistently associate the network flows with the transactions logs during data collection. To avoid fine-grain timing analysis, packets were binned into 20-minute flow blocks and aggregated before being stored for further processing. Additionally, rare organizations, IP addresses, and domains with fewer than 5 unique users in the dataset were grouped into an "other" category and replaced before the analysis to additionally protect user privacy.

4.4.2 Data Processing

Flow Domain Classification

To analyze the impact on specific types of traffic, we categorized flows from domains possibly related to telehealth and the pandemic. All domains which contained any (case-insensitive) entry from a pre-compiled list of health keywords (see Appendix B.1) were hand-inspected to confirm if they contained health information. The initial candidate set of domains contained 169 domains from 38 primary domains, which after inspection was narrowed to 53 domains from 12 primary domains.

4.4.3 Limitations

Free Credits

During the study, we detected an implementation issue in the network's billing system. The defect allowed a subset of users to double their number of network credits when they transferred credits back to themselves. Exploits of the flaw began on May 24, and were in widespread use by

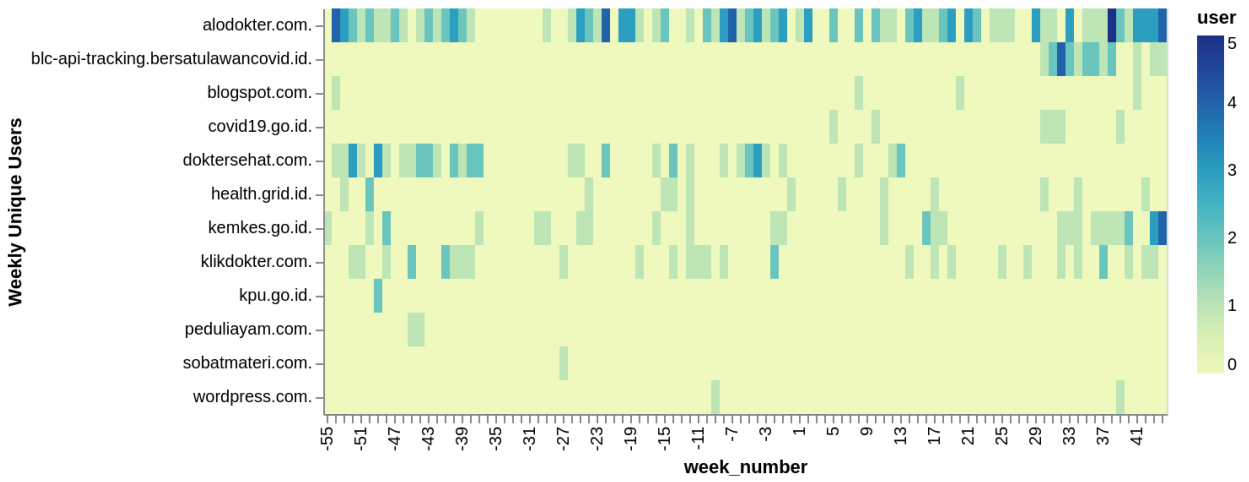


Figure 4.1: A heatmap of the number of users connecting to particular health-related domains versus time. Week 0 indicates the time lockdown restrictions were first instituted in the community.

June. The operator detected the issue on June 9th and the defect was corrected by early June 10th. Even though the tainted credits were traced and removed from all users June 13, 2020, the data purchased from the exploited credits continued to be consumed for several following weeks and continued to impact the overall traffic volume through July.

4.5 Results

In this section, I describe the results from my brief analysis. Additional information and analysis source code can be found in our GitHub repository at https://github.com/uw-ictd/ccn-covid-analysis. My lab mate Firn also led a different line of inquiry on the user transfers collected as part of this dataset. The results are omitted from this dissertation, but can be found in *The Low Impact of COVID-19 on Rural Community Network Traffic* [182].

4.5.1 The use of telemedicine and digital health resources

We found that telemedicine resources were already been in-use prior to the coronavirus pandemic and lockdowns. The weekly unique users to various well-known telehealths in Indonesia, such as alodokter.com, increased only slightly, but received visits regularly through the observation period. Some health domains even saw a decrease in the number of users per week after the lockdown restrictions were put in place (such as doktersehat.com). Figure 4.1 shows a heatmap of the count of distinct users per week to each relevant domain prefix. Unlike in well-developed regions with plentiful in-person healthcare infrastructure, residents of Bokondini were already using telemedicine services for basic healthcare in lieu of traveling to a nearby city to see a doctor.

A cluster of users of the app bersatulawancovid ("United Against Covid") was detected, although usage decreased over time. I was unable to detect any usage of the official government contact tracing app, PeduliLindungi, but this may be due to the app's implementation. I did not detect many users visiting the coronavirus digital resources that were provided by the Indonesian government, such as covid19.go.id. A cluster of users did begin visiting the national department of health website near the end of the study, correlated with the beginning of vaccine availability.

4.5.2 Minimal change in traffic volume

While other large-scale datasets show clear dramatic differences in the Internet traffic demand corresponding to the institution of regional lockdown orders, there is no clear change present in our dataset. Seasonal variation in demand around the winter holidays dominates any difference correlated with the institution of lockdown orders. Figures 4.2 and 4.3 show the network's overall traffic vs. time.

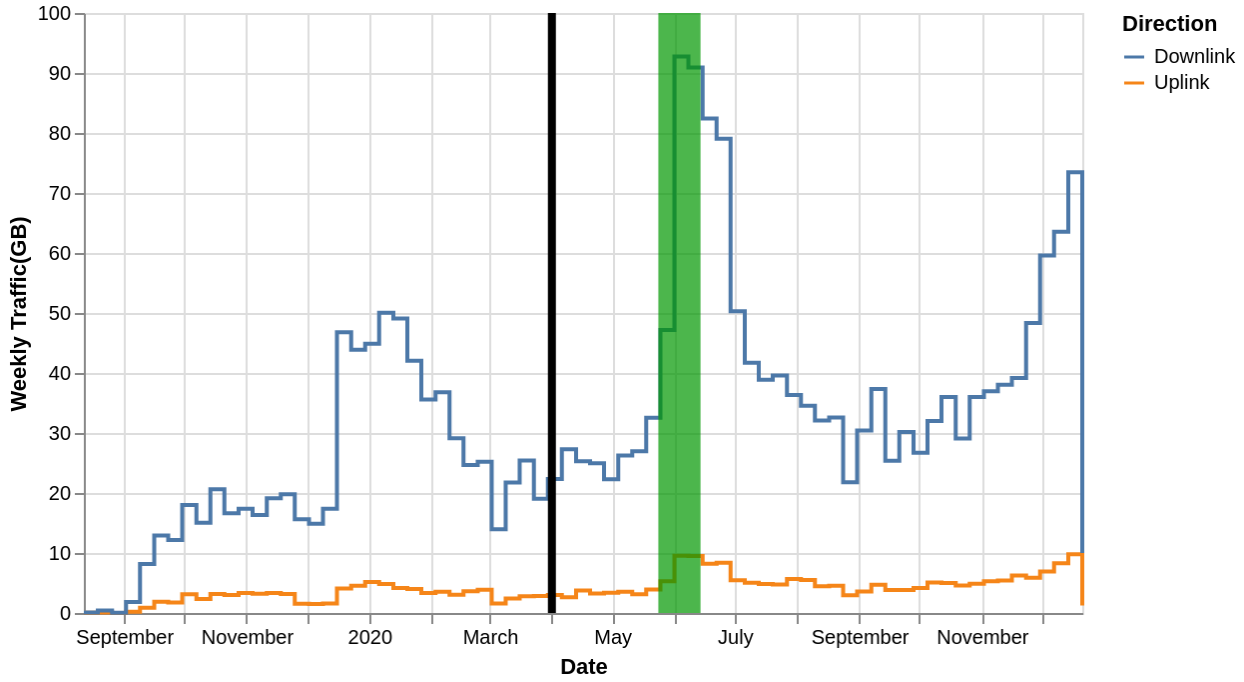


Figure 4.2: The total weekly traffic in the network versus time. The black bar on April 1, 2020 indicates the beginning of regional lockdown measures. The green area from May 24, 2020 to June 14, 2020 shows when a bug in the network's billing system was exploited to get free credits for a subset of users. These credits were removed on June 13, but data purchased with the free credits continued to be consumed over the next several weeks. Seasonal variation and the impact of the free credits dominate the impact of the lockdowns on the observed traffic volume.

4.6 Conclusion

As countries transition out of COVID-19, the unexpected crisis will likely leave lasting economic and social scars across the world. In this chapter I leveraged my existing relationship with the rural community cellular network in Bokondini, Indonesia to report on the pandemic's impact to our measured network metrics. I found national digital health resources from the Indonesian Government were not popularly visited until the beginning of vaccine availability, but that some well-known telemedicine services showed regular usage throughout the pandemic.

Contrary to other datasets, in this network the overall traffic volume did not significantly

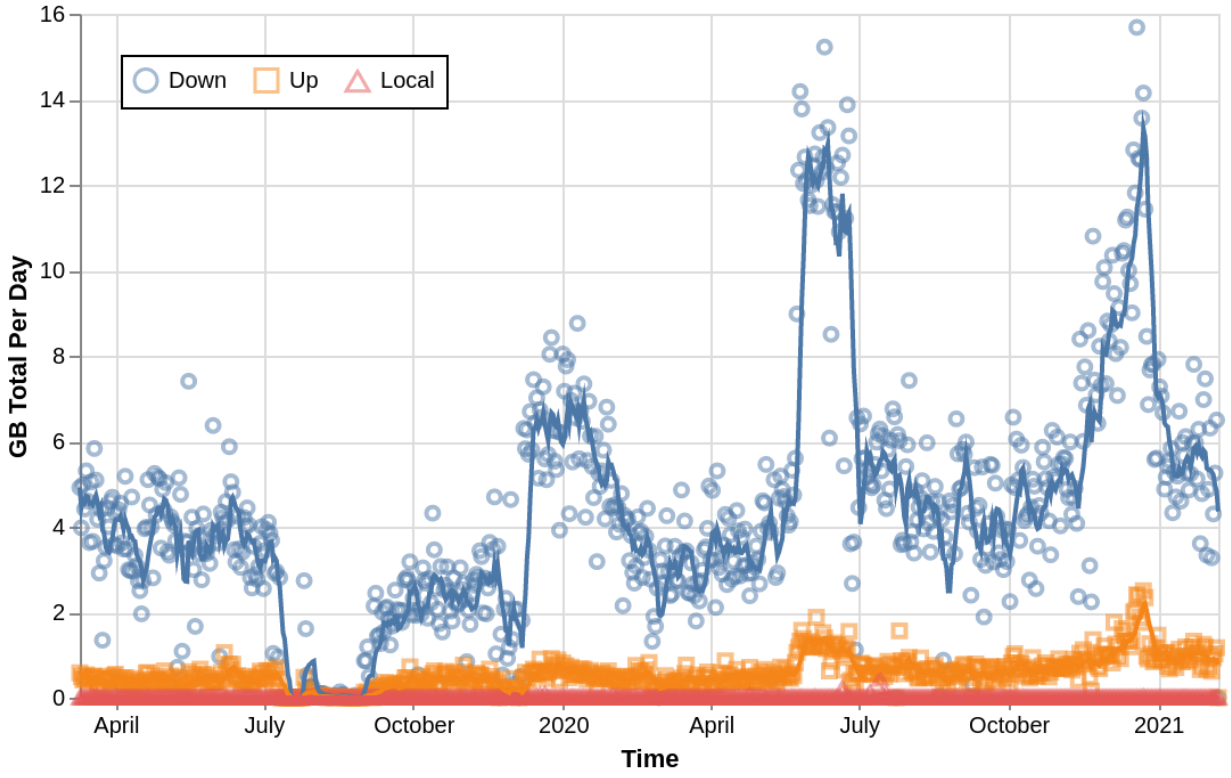


Figure 4.3: The total bytes transferred per day versus time. Despite the availability of free and zero-rated educational resources on the school's local media server, traffic to the local server was dominated by general Internet traffic even while schools were closed.

change before and after lockdown, and any shift was completely overshadowed by seasonal variation. In contrast, a dramatic shift was clearly visible when credits were accidentally free in the network, providing a natural, if unintended, control case. The lack of utilization change in the network with the institution of pandemic restrictions implies it likely was not providing enough additional utility to justify more use at its current pricing structure. Future work could explore why this was not the case, and what about the context in Bokondini made the network less essential than as seen in other measured contexts where utilization increased.

Chapter 5

dAuth: Distributed Authentication for CCN Federations

Executive Summary

This study explored the feasibility and constraints of making cellular authentication and the cellular roaming process more redundant for the community cellular network setting. With my lab mates I developed a system, dAuth, using secret sharing and a division of concerns between sensitive data stored with backup networks and non-sensitive public directory data to securely scale authentication across multiple redundant nodes in different organizations via federation. The system allows a collection of preconfigured backup networks to authenticate users on behalf of their home network while the home network is unavailable.

Hypothesis

I hypothesized that it should be possible to break up the cellular authentication process using multiparty computation in a way that allows for a more robust authentication process in the presence of individual network failure, while still maintaining its security properties with a subset of malicious nodes.

Summary of Key Findings

My collaborators and I designed the dAuth protocol and built a prototype of the system for testing. We evaluated its backwards compatibility through testing with an off-the-shelf LTE RAN and UEs, and performance using a simulated 5G RAN. We found that it performed comparably to a standalone cloud-based 5G core at low load, and outperformed a centralized core at high load due to its innate load-sharing properties.

Next Steps & Connections

The existence proof of dAuth shows us that cellular technologies can be extended in backwards-compatible ways to facilitate redundancy and reliability across individually unreliable operators. This lowers the threshold to running cellular infrastructure for smaller anchor institutions and lays the foundation of building a larger reliable network from many individually operated nodes.

5.1 Introduction

Researchers have proposed that community-based networks could help to bridge the internet usage gap, both by providing access to new areas by changing the economic balance of providing connectivity [148], and by allowing locally engaged organizations to reach sub-populations excluded by traditional networking systems and structures [21]. In particular, community-based cellular networks offer affordances well-suited to bridging the access gap, including wide-area

coverage [80], a design focused on concurrently serving many users [167], a robust market for user handsets (UEs), and the ability to provide indoor coverage to reach people who may not have the ability to come to traditional access locations like cyber-cafes or public hotspots [25]. Private edge deployments of cellular access network infrastructure, marketed as "neutral host" architectures when the hardware extends access for existing large-scale carriers, or "private LTE"/"private 5G" solutions when the provided connection is local, appear at first glance a perfect fit for the community cellular network use case. Yet while new licensing regimes (CBRS) are beginning to see success in allowing general community-driven deployment of infrastructure in cellular spectrum, challenges remain in enabling a broad swath of small local organizations to independently deploy and operate their own cellular networks.

Unfortunately the design of the cellular device/network authentication process limits the utility of the private edge approach for community networking, requiring each edge network to either register every individual or rely on the roaming process to authenticate with the user's home network as new users appear. Furthermore, assumptions embedded in the design of the cellular roaming architecture limit the total number of networks and require the home network to be highly available to authenticate roaming users on demand. The cellular network architecture divides functions between a radio access network, a logically distributed set of access points, and a "core network," a logically centralized set of operational functions for managing long-term user state and routing user traffic. The architecture requires that the core network have high-availability and reachability (discussed in Section 5.2.2), which drives small networks to use managed core networks provisioned by a traditional telecom service provider (Telco) or cloud provider in practice. Yet this outsourcing approach removes the agency of the local networks, requires users to put their trust in an external entity, and results in coupling that harms resiliency in the face of an outage at the central service provider. In one well-publicized example, a physical infrastructure outage (due to an explosion) at a single site caused a multi-day regional outage of telecommunications

services across the southeastern USA [183].

In this work, we propose a design based agreements between many operators that takes advantage of the decorrelated operations of the local actors to provide a robust and resilient service to end-users even in the face of individual operator unreliability. We envision a diverse collection of local institutions, grounded in physical-world relationships, serving as the anchor for their local users and use the number of registered nonprofits registered in the USA (1.5M) as a benchmark for the number of operators we would expect in such a network[177]. Unlike the traditional cellular architecture, our design allows the network to more gracefully tolerate transient unavailability of its individual components, making it more feasible that less professionalized organizations operate the networks themselves. This architecture would allow local organizations to outsource operations to an external provider if desired *but not require it by design.*

Specifically, we define a community-based federated trust model that allows organic scaling of a wide-area network deployment with only incremental trust between partner organizations, and then build an authentication and authorization scheme using this model for granting access on a serving network (even when the user's home network may be offline). Unlike traditional roaming, we introduce an additional layer of abstraction which removes the need for each network to have a pre-existing relationship with the serving network and allows our system to scale to larger number of organizations. Our design, dAuth, takes great pains to remain compatible with off-the-shelf devices, without requiring changes to device firmware or the over-the-air radio interface, to take advantage of the existing install base of cellular devices and robust manufacturing and repair ecosystems around these devices originally designed for nation-scale carrier networks. While our goal could possibly be achieved more efficiently with a completely new protocol, we believe that "Greenfield" approaches requiring low-level changes to the UE's interface are not realistic for deployments of such a network over the next decade. The initial 5G standard was frozen in 2018 and is only now seeing wider deployment.

We develop a prototype of dAuth to evaluate its performance, and demonstrate its compatibility and technical feasibility with a testbed deployment in partnership with the Seattle Community Network, a regional federation of community cellular networks in an urban area in the USA. We evaluate our system's performance against the status quo of a centrally hosted "cloud core" and a non-roaming "edge core" private network.

In addition to allowing for organic decentralized growth of a mobile access network based on federation, we believe that the benefits of our proposed architecture go beyond this use case and offer opportunities for the design of more resilient cellular networks in general.

5.2 Background & Context

Our work is situated at the intersection of work on community-based networks and cellular networks. In this section we provide background context for both of these domains.

5.2.1 Community Networks

Community networks are physical instantiations of networks built, owned, and operated by the community the serve. There are a huge diversity of different types of community networks as diverse as the varied communities they come from.

This project was motivated by work with the Seattle Community Network (SCN), a network founded to explore the use of community cellular in urban contexts and connect marginalized populations who have not been able to or do not want to receive service from traditional providers. SCN has a strong organizational mandate to empower local users and seeks to de-mystify cellular infrastructure through co-ownership.

SCN currently has 6 deployed sites (with 2 more in deployment) with a variety of partners including two public libraries, the local school district, two cultural community centers, and a hackerspace. The sites are diverse and have a variety of different backhaul connections from

Org	Backhaul	Availability
Hackerspace	P2P to University Campus Fiber	99.02%
School 1	Fiber ISP A	98.998%
Community Center 1	Fiber ISP A	95.815%
Library 1	Fiber ISP A	91.821%
School 2	P2P to University Hospital Fiber	89.562%
Community Center 2	Fiber ISP B	87.171%

Table 5.1: Deployed sites in the Seattle Community Network, sorted by their observed uptime. The sites use a variety of backhaul technologies and providers. Uptime in practice has been determined more by site equipment than any particular ISP. No sites have achieved "3 nines" of availability (corresponding to less than 8 hours and 47 minutes of downtime in a year).

various Internet Service Providers. Table 5.1 lists details for each site. Importantly, all sites are maintained and operated by volunteers or available staff at each organization, and see a range of uptimes which are individually not sufficient for a reliable service.

5.2.2 Cellular Networks

This work focuses on cellular networking technologies, which are heavily standardized and have a specific set of roles and functions divided between the "Radio Access Network" (RAN), the actual radio basestations providing access across the network, and the "Core Network" (Core), the set of more centralized systems connecting together the RAN elements. The cellular network architecture evolved in the context of wide-area telecommunications services, and core network functions are responsible for city or region-scale sets of resources operated by a professional "mobile network operator" (operator/MNO) of region, city, or even national scale.

Cellular networks have standardized interfaces for "roaming", the ability to use a device from one operator's network via another operator's network [1]. In cellular roaming, a connection from the "visited" or "serving" network is made back to the "home" network to establish the user's

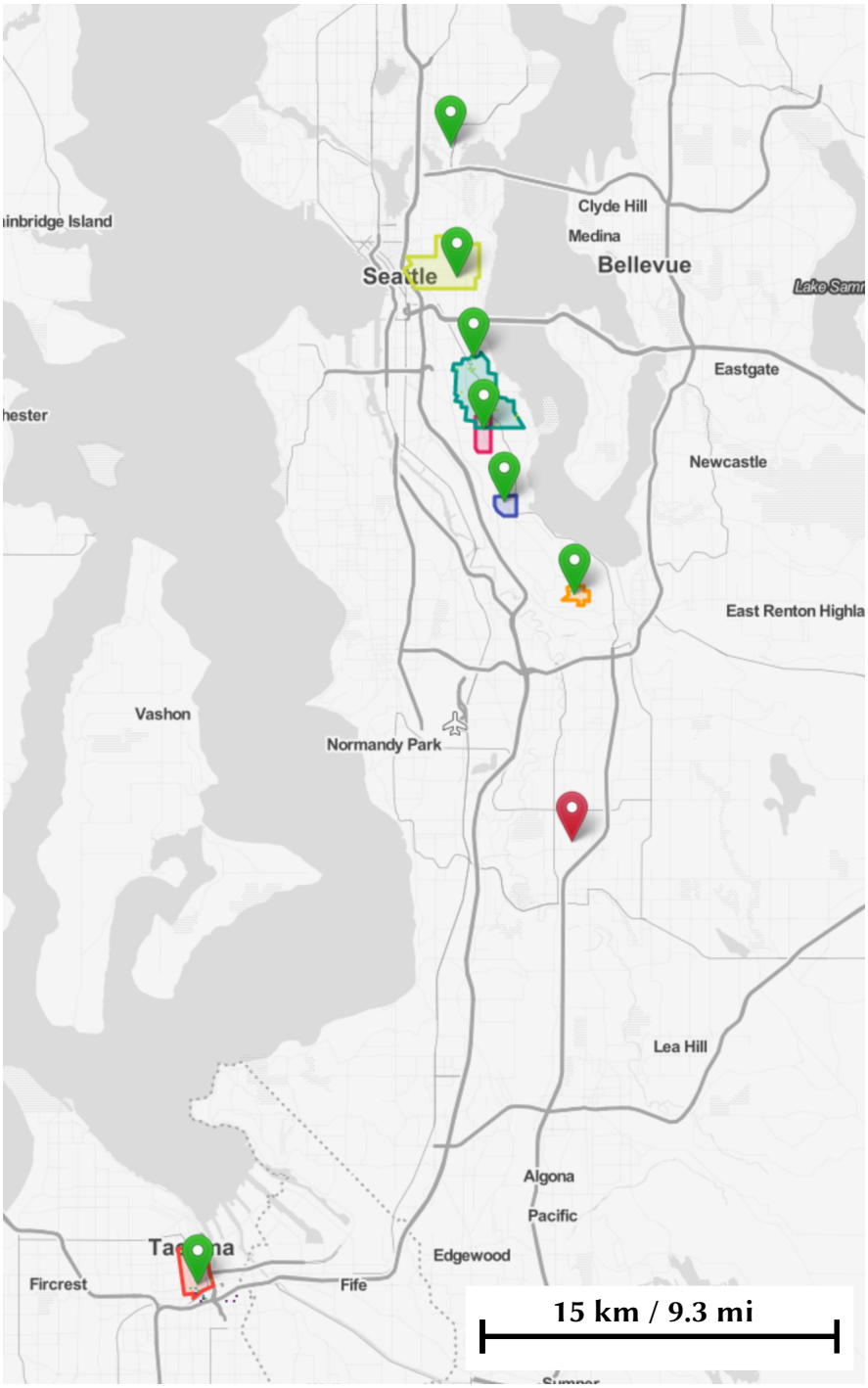


Figure 5.1: A map of the sites in the Seattle Community Network. The sites span a wide area, and the network already provides some overlapping coverage in the dense mid-part of the network.

identity. The home network completes the Authentication and Key Agreement (AKA) with the UE based on the secret key held by the operator and stored in the user's SIM card. At the conclusion of the AKA procedure the serving network is provided a key to communicate with the UE, and the user can be confident that their home network has approved of the operations of the visited network they are communicating with.

There are many variations of roaming architectures between operators, and in some cases the user's traffic is tunneled all the way back to their home network [117].

5.2.3 LTE & 5G Networks

Starting with 4G-LTE (3GPP Release 8+) and continuing with 5G (3GPP Release 15+) modern cellular networks have moved to an ip-centric design where packet-switched services are used to provide both plain data and QoS-enhanced voice/sms/video/emergency services to end users over a common substrate.

Partly due to the ease of IP interconnect and partly due to the availability of new hardware in lightly-licensed regimes like CBRS or general secondary use [12], researchers have explored smaller and more standalone deployments of cellular networks for community connectivity over a town-scale area from a single tower [167, 87, 101]. These networks have either operated independently in a standalone mode, with no roaming capability, or relied on a third-party service to hold user keys and serve as a broker for authentication [111].

The importance of network auth

While most higher-level user traffic is end-to-end encrypted on modern mobile devices, they are still vulnerable if connected to a compromised access network. In addition to blackholing or redirecting the encrypted user traffic, a fully authenticated mobile network has much more control over the user's device than in more distributed standards like WiFi.

Along one dimension, the RAN has tight control over the UE radio, and can cause it to tune to different frequencies and transmit at different power levels, potentially turning the UE into an unwitting jammer or quickly draining its battery. Along another, in LTE (which a 5G network can downgrade a connection to for most phones) the network also has the ability to craft and intercept SMS messages from arbitrary numbers, creating spam or more malicious phishing messages from sensitive protected numbers the user trusts. Additionally, the cellular standards allow an authenticated RAN to query sensitive information from the device like permanent hardware identifiers in the UE baseband or the user's current precise location. Even at the application layer, the cellular network provides time and coarse location data which can be spoofed and consumed by a phone's operating system and/or user applications.

The importance of a non-malicious serving network in particular for off-the-shelf cellular technologies drives dAuth to not immediately dismiss link authentication like some other prior works.

5.3 Design Elements

In this section we explain the constraints driving dAuth's design and how it fits into the wider design space of cellular core networks. As discussed in 5.2.1, the primary purpose of dAuth is to allow a wider variety of smaller anchor community organizations to control and operate mobile access networks in a municipality-scale federation.

5.3.1 Key design goals

Our goals break down along these three major axes, which we detail below.

Tolerate Failures: Tolerate temporary failure of a subset of nodes without losing liveness or safety of the overall system.

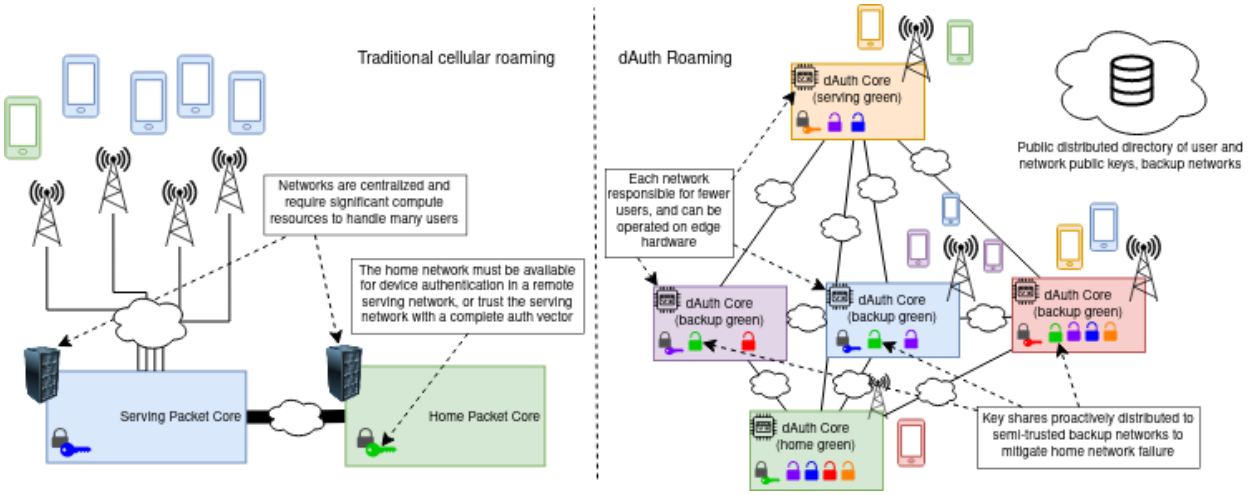


Figure 5.2: The high-level design of dAuth versus traditional cellular roaming. dAuth factors some non-sensitive information into a distributed public directory service, and allows home networks to proactively distribute shares of key material to backup networks to improve resiliency in a semi-trusted environment.

Tolerate Malicious Nodes: Tolerate the presence of malicious nodes outside the user's home network without compromising the user's security.

Compatibility: The system must operate with existing off-the-shelf hardware built for standardized 4G and 5G networks and without requiring a dAuth-specific upgrade from the manufacturer or new certifications.

Tolerate Failures

To lower the threshold to include organizations in the federation with a wide-range of networks operations experience and technical sophistication, the overall federation must be resilient to small hiccups in operation of any particular organization's network. This greatly reduces the operational burden on individual anchor institutions, and allows for operational shortcuts like planned maintenance windows and system reboots with minimal disruption to end users. Specifically dAuth should allow a user to gain access through an alternative serving network in the

federation even when the user's home network core is temporarily unavailable. We parameterize our prototype around an expectation of home network outages lasting on the order of 1 day.

Tolerate Malicious Nodes

Driven by our desire to increase the number of operating organizations and lower the trust barrier to entry, our system needs to tolerate the presence of a subset of malicious nodes in the overall federation. It is even possible that some previously trusted nodes could become malicious over time due to compromised security. While we assume a specific user can always trust their own home network, dAuth should provide a way to prevent a user from attaching to untrusted serving networks and tolerate the compromise of a subset of backup networks. From the serving network's perspective, dAuth should also provide a strong guarantee that the authenticating user is indeed a member of their claimed home network for compliance and accounting.

Compatibility

Deploying and exploring the non-technical challenges of the SCN vision without losing the economies of scale of the existing cellular equipment ecosystem requires a design that can be realized with unmodified off-the-shelf hardware built for standards-compliant 4G and 5G networks. Long-term real-world deployment also requires radio hardware with certification from the relevant standards bodies for radio emissions and devices rugged and robust enough to support everyday use by non-experts, which is difficult to achieve with most research prototypes.

In 4G-LTE there was only a single authentication scheme (4G-AKA) required for user devices, and while 5G does specify additional optional support for the IEEE Extensible Authentication Protocol (EAP), the only authentication all devices are required to support is 5G-AKA [3]. 5G AKA is very similar to 4G AKA, with an additional round of confirmation that the user is present in the serving network before exchanging the session key. In order to support the broadest number of

devices possible, we have designed dAuth to look exactly like 4G/5G AKA from the basestation and handset's perspectives. We do assume that the core network functions (in software) can be modified at will and the user's SIM card is under the operator's control.

5.3.2 Refactoring Auth For Trust at Scale

In order to allow dAuth to scale, the system design distinguishes between non-sensitive public information, like the user's public key and home networking mapping, from sensitive personal information like when and where the user has authenticated and which traffic they send. dAuth factors out the non-sensitive global information to allow using a traditional large distributed directory like a verifiable key directory, DNS, or a distributed ledger. This directory can scale up publicly with very low overhead as demonstrated by DNS and other public ledgers.

dAuth leaves the sensitive information about authenticating the particular user and making authorization decisions for that user onto a particular serving network to a much smaller set of trusted entities. These entities communicate directly, limiting the exposure of the user's more sensitive telemetry to only the nodes participating in the current authentication. While authentication information is split across multiple backup networks, the fan-out of information from each home network is limited to a constant factor (at most 31 in the current design). This constant factor, independent of the total number of networks, allows the overall system to scale naturally as the number of participating networks increases.

Community-based Federated Trust Model

Decentralized and federated systems allow users to limit the scope of who they trust with their information. With dAuth, we envision a model where users can establish a real-world relationship with an organization they trust. This trusted organization is the user's *home* network, and serves to anchor their identity in the wider dAuth ecosystem. The dAuth home network fulfills the role of a

traditional operator home network, holding the user's AKA private key, generating authentication vectors for the user, and serving as the anchor the user's identity when they roam onto a different serving network.

Our desire to enable resiliency in the face of home network failure (to allow a wider scope of organizations to serve as home networks) is at odds though with the interactive authentication between the home HSS/AUSF and the serving network in traditional 3GPP roaming. To navigate this tension, we define a third intermediate level of trust in between the home network and the untrusted serving network, the *backup* network. In our trust model, the set of backup networks can cooperate with a dAuth serving network to complete the interactive portion of 5G-AKA while the home network is down, but never have access to the user's root key or the derived key used by the serving network to provide access within its security context. Additionally, unlike a traditional 3GPP home network, dAuth backup network grants are revokable, allowing the user or the user's home network to remove trust from a backup network through a revocation procedure.

5.3.3 System Entities

UE The user's off-the-shelf device.

SIM The user's SIM card, issued by the user's home network. It can be customized but must maintain compatibility with the UE baseband interface. The SIM holds K_i^u, the user's milenage key, $\langle SQN \rangle^u$, the user's vector of used SQN values, SQN_{max}^u, and the user's current identity $GUTI^u$

Home Network The user's home network, which in dAuth is run by a small community network at a single community organization. Runs the dAuth service, and optionally fulfills the role of a serving network for this user, or a backup network or serving network for other users. It holds the signing key Sk_h for its published Pk_h, and key material for all of its users

$(\mathrm{K}_{i,u}, \mathrm{SQN}[v \forall v \in \mathrm{SIM}]) \forall u \in HNet$

Serving network An off-the-shelf Radio Access Network with a customized software-defined core network and dAuth service, providing coverage to the user when away from the home network. The SNet holds a private signing key Sk_s corresponding to its published public key Pk_s.

Backup Network(s) A set of networks semi-trusted by the home network to collectively hold single-use key material for the user. The backup network holds a signing key Sk_b corresponding to published Pk_b, a set of assigned auth vectors for each backed up user $(\mathrm{AUTH}, \mathrm{H}_1(\mathrm{HRes}^*), \mathrm{RAND})_{i,n}$, and the key shares corresponding to auth vectors held by other backups $\mathrm{Share}(\mathrm{Res}, \mathrm{K}_{\mathrm{seaf/asme}})_{\mathrm{H}_1(\mathrm{HRes}^*),i,n}$

Directory Visible to all participants, and can be based on existing verifiable public key directory schemes [119, 36], or a hierarchy like DNSSEC [13]. All information in the directory is public information with no controlling organization. The directory contains Pk_n and an address for each Network. It also contains a mapping from user to home network for each user, and a set of backup networks for each home network. Entries in the directory are signed by their respective parties, and are assumed to change rarely.

5.3.4 Specific Protocol Components

Sequence Number Handling

In 4G and 5G AKA, there is a sequence number associated with each authentication attempt to prevent replay attacks against the symmetric ciphers used in the network. The SIM is responsible for storing a record of the used sequence numbers and ensuring that no sequence numbers are ever repeated. The 3GPP recommends an implementation in TS 33.102:Annex C [2] where the SIM

tracks the greatest sequence number seen across 32 independent slices. dAuth takes advantage of this behavior to allocate independent authentication vectors to each backup network that do not have to be synchronized across the backup networks at authentication time, lowering coordination overhead between the networks. It also takes advantage of the ability for new sequence number to supersede older issued sequence numbers within a slice to allow the home network additional control to revoke published auth tuples from backups at a later time. See 5.4.3 for details on the revocation procedure, and C.1 for more general information on SIM card behavior.

User Identifiers & Privacy

While other work has explored modifications to cellular networks to anonymize users and enhance privacy, dAuth is focused on operational networks where some form of user identity is necessary for compliance and to prevent abuse of the network. dAuth leverages the existing cellular identity architecture, but does come with the privacy advantage that by spreading access across many small operators in independent organizations, any one serving operator will only be able to identify the user while they are connected to their particular network.

In 5G there are two identifiers which can be used to authenticate to a network: the GUTI, a globally unique temporary identifier assigned by the last network to serve the user, and the SUPI, a long-term permanent identifier corresponding to the user's SIM. In the event that a GUTI cannot be converted to a permanent identifier because the current serving network cannot communicate with the previous one, 5G supports requesting the user's permanent identifier directly. In order to protect the user's identity from passive over-the-air sniffing though, the SUPI is encrypted into the SUCI with asymmetric cryptography and a key corresponding to the home network. dAuth currently requires that this identity key be shared with the backup networks, although we hope to explore ways to also protect it with threshold cryptography or a multi-party computation in future work.

Key Shares

The dAuth protocol leverages threshold-protected key shares to provide resistance a subset of compromised backup networks. The key shares are constructed with Shamir secret sharing [169], originally described by Adi Shamir in 1979. The shares are constructed together by the home network, and then split up among the different backup networks as part of key dissemination.

Shamir secret sharing provides the guarantee that if M or more of N total shares are combined, the original secret (in dAuth K_{seaf}) can be losslessly recovered. If fewer than M shares are available, no information is leaked about the underlying secret. We leave the threshold M configurable on a per-network basis, since there is a tradeoff in the robustness of the level of protection provided by a high threshold against the performance and availability of the network when some backup networks are slow or unreliably online.

Shamir secret sharing also does not inherently provide a way to validate a received share is valid, and is subject to tampering if a node contributes a malformed share. There are extensions to Shamir sharing that do provide validation at the expense of extra overhead [55, 138], but we use simple Shamir sharing in dAuth because the shares are always part of larger messages already signed by the home network.

5.4 Operation

In this section we detail the design of a basic scheme which allows mutual authentication, and then extend this scheme to allow the user to continue to authenticate when the home network is down for an extended multi-hour period.

5.4.1 Basic Authentication Scheme

When the home network is online, dAuth functions similarly to standard 5G roaming but with special handling of the authentication sequence number to not interfere with existing generated

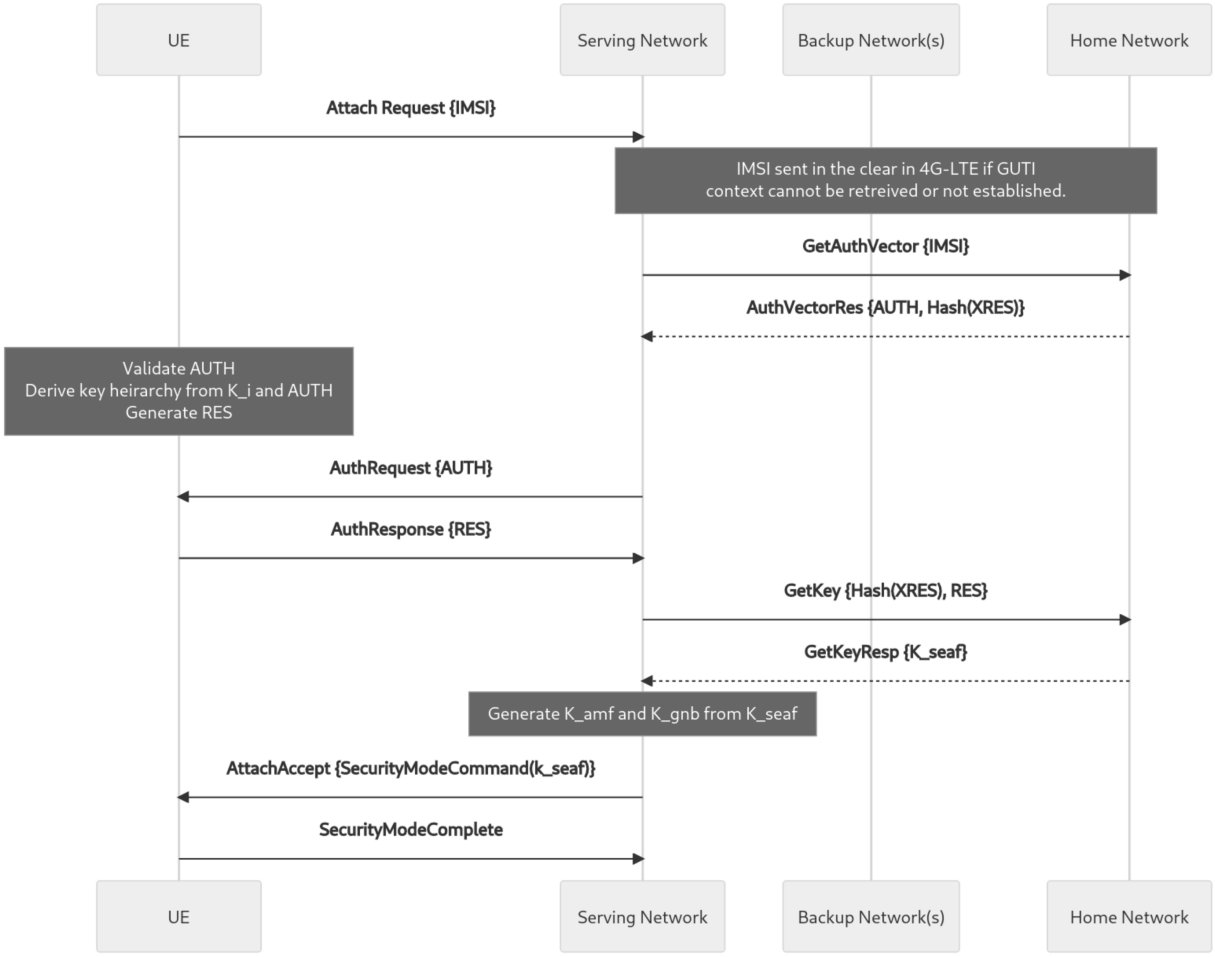


Figure 5.3: The basic authentication flow used between a serving network and the home network when the home network is available.

backup authentication vectors (see 5.3.4). The authentication process starts once the UE synchronizes its radio with the serving network and sends an attach request message. The attach request message contains the user's ID in the form of an IMSI/SUPI, SUCI, or GUTI depending on the UE's previous connection. If the ID is an IMSI/SUPI the serving network can query the public directory service directly to look up the user's home network. (Recall that the public directory does not change often, so it can be implemented as a widely distributed system with relatively low latency like DNS.) If the ID is a SUCI, the home network ID is directly embedded in the message. If the ID is a GUTI, the serving network receives a pointer to the prior serving network it can contact for the user's identity and home network. If this contact fails, the serving network can request that the UE provide a long-lived identifier and receive an IMSI/SUPI or SUCI.

Regardless of which ID lookup path is taken, once the user's home network identity is established, the serving network opens a direct connection to the home network and requests to begin authentication. At this time the home network and serving network validate each other's identity via Internet standards (TLS and PKI). The home network still has no way to know yet if their user is present at the serving network, but if it trusts the serving network to provide access to the user, it generates a one-time-use authentication challenge and a hashed version of the user's expected response from the user's secret key K_i and a reserved slice of the sequence number space. It then sends the auth challenge and hashed expected response to the serving network, which stores the hashed expected response and forwards the challenge along to the UE.

If the UE is indeed the user, they can validate the challenge using K_i to confidently determine it came from their home network, and send back the correct response. Upon seeing the UE's response, the serving network can validate it is a preimage for the hash it received from the home network, and have confidence the UE does belong to the home network. It then forwards the response back to the home network, who can also validate it and gain confidence that its user is indeed present at the serving network. Only at this point does the home network generate a

session key for the user's connection and send it to the serving network, allowing the serving network to establish a validated control connection to the UE to manage its radio and serve the user's traffic.

5.4.2 dAuth Backup Authentication Scheme

When the home network is not online, dAuth nodes can seemlessly fall back to the backup scheme. It has three phases: key material dissemination (while the home network is online), backup authentication (while the home network is offline), and reporting (when the home network is online again).

Key Material Dissemination

In this phase the home network generates a set of AKA authentication vectors, a hash of the expected UE response, and corresponding key shares for each user to be backed up. The vector sequence numbers are chosen carefully so each backup network's sequence numbers are in independent dimensions of the sequence number space. Once generated, the auth vector, hashed expected response, and random salt from the hash are serialized into a binary bundle, signed by the home network, and sent to a specific backup network. The corresponding key shares are also added to bundles with the hashed expected response (used as in index) and random salt value, and each key share bundle is also signed by the home network. The key share bundles are split up, with each being sent to a different backup network.

The backup network must generate N^2 key shares and N auth vectors to provide a single auth vector to all backup networks. In practice N is limited by the number of practical backup networks and the number of sequence number slices in commonly available SIM cards (32, with one reserved for the home network), so the N^2 scaling is not an issue in practice since the maximum number of backups is limited to a relatively small constant factor.

Additionally, if 5G ID encryption is used by the home network, the home network shares the ID decryption key with the backup networks.

Backup Authentication

Backup authentication occurs between the set of backup networks and a serving network when the home network is unavailable. After determining the user's home network via the same procedure as basic auth, the serving network looks up the backup networks elected by the home network from the directory. It then opens a secure connection to the closest backup network and sends the user id (IMSI/SUPI or SUCI) to request an auth vector. If the ID is a SUCI the backup decrypts it with the id key shared by the home network.

After obtaining or receiving a plaintext ID from the serving network and ensuring the serving network is of sufficient reputation, the backup network looks up the next corresponding auth vector in the series assigned to it for that user, and returns the home-network-signed auth vector bundle to the serving network. The serving network validates the home network's signature, and then forwards the authentication vector to the UE. Upon reception the UE validates the vector as in the basic scheme and returns its response to the serving network. The serving network then can validate the response against its received hash as in the basic scheme.

After validating the response, the serving network creates and signs a bundle of the $H_1(HRes^*)$ and salt it originally received with the auth vector, and the received response from the UE that is a preimage for the hash. This bundle serves as the proof that the UE was indeed present at the serving network, and is forwarded by the serving network to *all* backup networks in the next stage to request the key share for the corresponding auth.

Upon receiving the key request from the serving network, the backup networks validate the serving network signature and the hash preimage. If valid, each looks up the key share it has stored at the $H_1(HRes^*)$ index. The backups store the received bundle from the serving network

in persistent storage to report to the home network later once it is back online. Each backup then forwards its key share bundle to the serving network, which after receiving and validating enough bundles to meet the network's configured threshold assembles the session key and completes the authentication process with the UE.

Event Reporting

Once the home network is online, the backup networks report authentication events to the home network and request new auth material to replace used vectors and shares. This triggers the home network to update any backup networks that were not part of the authentication event to replace their now obsolete key shares. The home network is also able to validate if there is any inconsistency between backup network reports, and potentially stop trusting particular backup networks or serving networks involved with inconsistent authentication attempts.

5.4.3 Revoking a Backup Network

dAuth allows the home network to revoke a backup network's ability to authenticate a particular user in the future in the event a backup network is ever compromised or otherwise becomes untrustworthy. The auth vectors given to the backup network are revoked by initiating a special authentication within the revoked network's sequence number slice at a value greater than the greatest sequence number ever disseminated to the now-revoked network. If the user is currently attached to a serving network, the home network contacts the serving network directly and requests the network perform a network-initiated reauthentication immediately. The serving network returns the UE's authentication response to prove it completed the handshake. The home network can then notify the remaining backups to delete the now obsolete key shares.

If the user is not currently attached, the home network sends a "flood vector" request to all remaining backup networks, providing the vector and key shares to be used for the next auth for

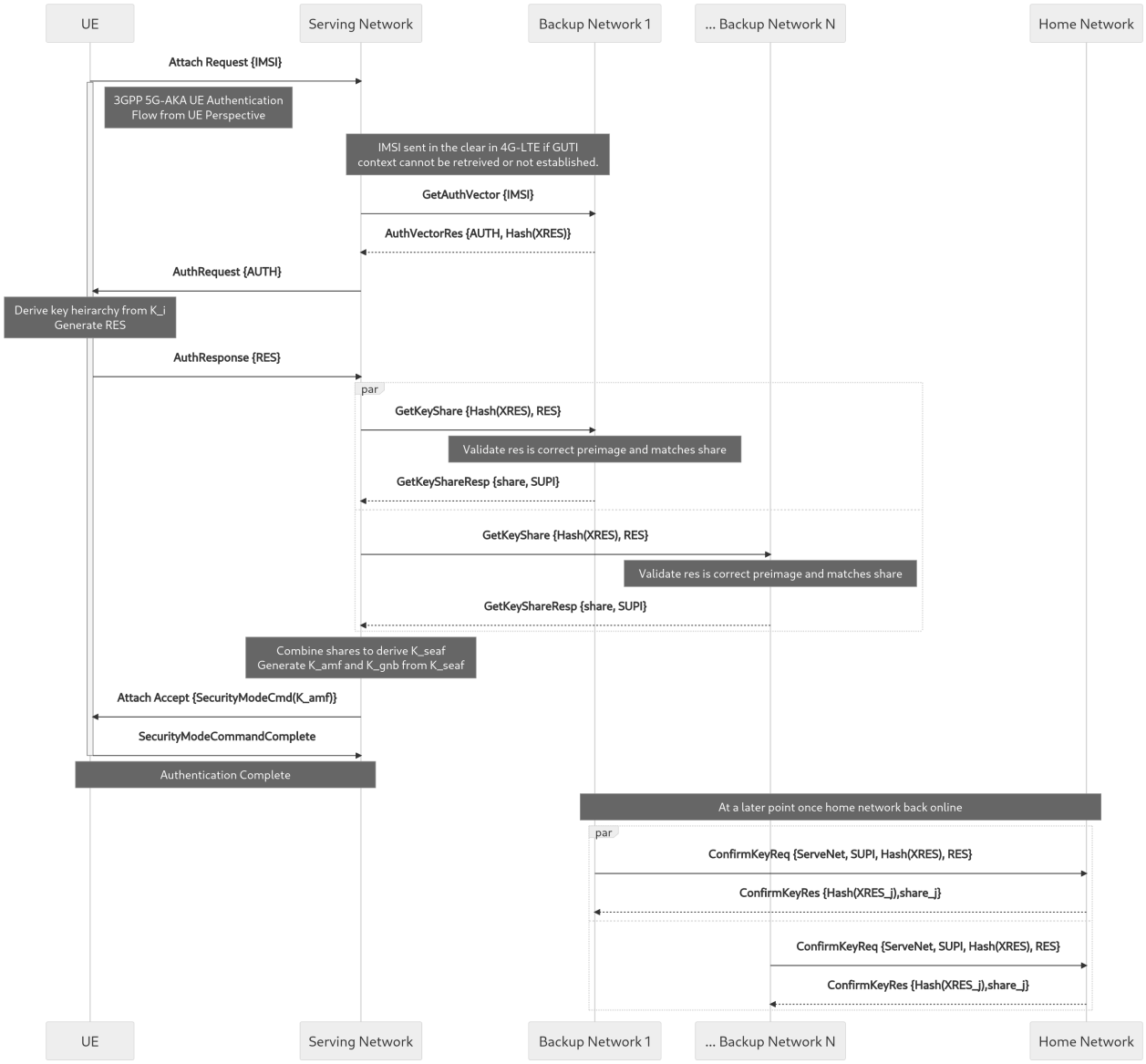


Figure 5.4: The dAuth authentication flow when the home network is offline. dAuth allows a (sub)set of backup networks to authenticate a user on behalf of the home network when the home network is unavailable. As long as one of the participating backup networks follows the protocol, the home network will receive confirmation of where the user was authenticated and can detect malicious or suspicious activity. Additionally, a serving network of insufficient reputation will not be able to establish radio control over the UE as long as one of the threshold backup networks faithfully enforces the user's trust preferences.

the user instead of the backup's usual series of vectors. This alone does not guarantee that the user will not initiate an authentication with the now-revoked backup. At the same time, the home network instructs all backup networks to invalidate and delete their key shares corresponding to the vectors given to the deleted backup network. Even if the now-revoked network were to be selected by a serving network and provide its auth tuple, as long as N − threshold of the user's N backup networks have received the revocation notice, the untrusted backup will be unable to complete UE registration since it will not be able to assemble the session key.

5.5 Implementation

We developed an open-source prototype of the dAuth system as a proof of concept of its feasibility and to evaluate its performance characteristics. Our prototype consists of three main components:

- A dAuth daemon running on each edge-core, responsible for tracking state relevant to each particular network and interfacing with the local core network, the directory service, or other dAuth daemons.

- A modified version of the Open5GS core network stack which interfaces with the dAuth daemon when new users are authenticating to the network.

- A directory service for initial prototype testing.

5.5.1 dAuth Service

The dAuth daemon is implemented as an asynchronous gRPC server written in Rust with the ToNIC gRPC framework. The server provides three endpoints, a `Local Auth` endpoint to interface with the edge-core, a `Backup Network` endpoint to provide key shares and accept authentication proofs when acting as a backup network, and a `Home Network` endpoint to distribute key shares

to backup networks and accept asynchronously reported authentication proofs. The `Backup Network` and `Home Network` endpoints communicate with other instances of the dAuth service running in other networks. It uses SQLite to store persistent state (user keys as and sequence number state as the home network, and any delegated auth information as a backup network). Each backup network stores any authentication events pending to report to the home network, and periodically polls the home network availability when any events are pending.

While only a proof of concept and not heavily tuned or optimized, our implementation does include 3 notable optimizations: the ability to cache and re-use gRPC connections across RPCs, local in-memory caching of directory information, and the ability to race concurrent requests to multiple backup networks when getting auth vectors and assembling the final threshold key. We found these optimizations significantly improved the performance of the network, particularly for repeat requests as would be expected for a network operating in a local area with a set of consistent users.

5.5.2 Modified Open5GS

Our prototype is integrated with the Open5GS (the core network currently used by SCN), although it should be portable to other core network implementations. Open5GS is written in C/C++, and we use Google's open-source C++ gRPC client implementation to communicate with the dAuth service. We modified the AUSF, the function responsible for authentication information in Open5GS, to query the dAuth daemon via the `Local Auth` endpoint for new authentications for 5G connections, and the MME to query the `Local Auth` endpoint for 4G connections. Our implementation allows the AUSF and MME to concurrently answer other queries while waiting for a response from dAuth. We also modified the PCF, UDM, and PCRF functions to provide a default network slicing and QoS configuration to dAuth UEs not otherwise present in the Open5GS database.

5.5.3 Directory Service

The last component of our prototype is a simple directory service to emulate the public directory. We also implement the directory service with Rust and the ToNIC gRPC server. The prototype directory stores its information in SQLite, and contains functions to query and update network keys public keys, addresses, and the mapping from users to the networks and backup networks.

5.6 Evaluation

To evaluate dAuth's feasibility, we tested its performance across 4 scenarios informed by our experience operating community cellular networks. We explore how dAuth performance scales as the security parameters of the system change and load increases in each of the four scenarios, comparing dAuth to both a local "edge-core"[1] which does not allow roaming, and a cloud-based centralized core emulating a traditional network deployment. We used Open5GS v2.4.7 for all core network instantiations (both on the edge and in the cloud) to avoid confounds from performance optimization of a particular core network stack.

5.6.1 Test Network

Our test network consists of 12 nodes with a variety of processors, memory, and disk configurations matching the diversity of machines seen in production community networks. Two machines are actual nodes in the Seattle Community Network open to us running experiments from them. Four nodes are cloud machines at four different major cloud providers, three nodes are low-power edge computers deployed on residential cable Internet connections, and the last three nodes are in a university lab with a high-quality Internet backbone. Table C.3 provides details of each machine.

Due to the presence of NAT on some connections and to mitigate risks of deploying our prototype software on the SCN network nodes, we used Tailscale [139] to establish a mesh VPN

[1] the irony of this phrase is not lost upon the authors

between all endpoints for testing, and run dAuth over the Tailscale connection. We characterized the overhead of Tailscale versus a direct IP connection, and found it to be comparable with a slight fixed latency penalty of ~3ms RTT and a throughput penalty of <10%.

5.6.2 Physical Testing

To validate dAuth's ability to interface with off-the-shelf equipment, we integrated it with production hardware loaned from the Seattle Community Network. The physical testbed RAN was a Baicells Nova 233 CBRS eNodeB with 20MHz of bandwidth and a CBRS-compatible TDD duplex mode, deployed in a university lab. We conducted two types of physical tests, a compatibility test, with an off-the-shelf phone and sim, and performance tests, with an instrumented software-defined UE where we could measure precise authentication timing.

Physical Compatibility Testing

For the compatibility test UE, we used an unmodified Google Pixel 4 running Android 12 with an off-the-shelf sim conforming to TS 33.102:Annex C [2]. We tested all three configurations of dAuth, including local authentication to the home network, roaming authentication to the home network, and backup authentication when the home network was offline. In all cases the UE was able to connect with no issues. Figure 5.5 shows the off-the-shelf UE successfully authenticated via backup networks.

We also validated that the UE was unable to connect if fewer than the key share threshold of the backup networks approved of the authentication request from the serving network, and that authentication attempts were promptly reported to the home network when it returned online.

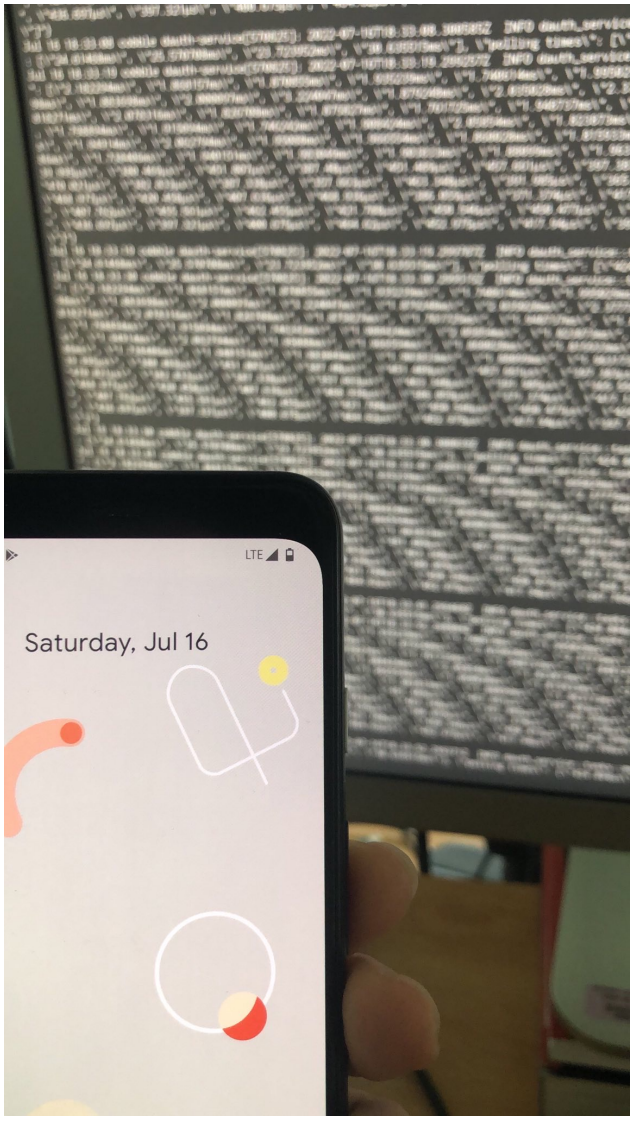


Figure 5.5: dAuth Off-The-Shelf Operation: This figure shows a successful authentication with an off-the-shelf UE to a dAuth federation using backup networks, where the phone's home network is offline. In the background of the image is the log output from the dAuth daemon processing the received key shares. A video demonstration is available at https://youtu.be/obpqBO6oAV8

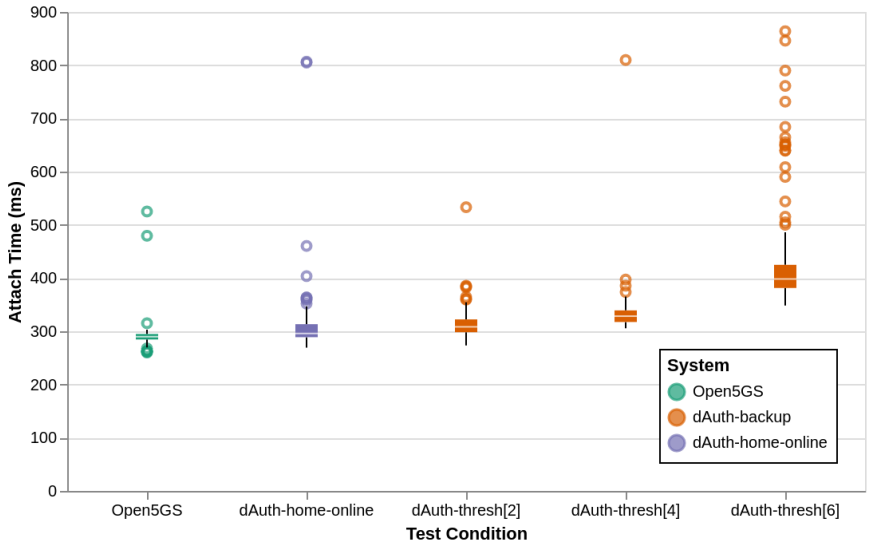
Physical Performance Testing

To test the system's performance, we turned to the srsRAN UE [174], a software-defined UE implementation which can operate with common software-defined radios like the Ettus USRP. We started with the latest released version 22_04, and modified it to capture fine-grained timing measurements during the attach and network authentication process. We used a USRP B210 as the SDR controlled by srsRAN. Our USRP did not have a functioning GPSDO and suffered from clock instability, so 39 of 1986 samples (1.96%) were excluded from further analysis due to synchronization loss during the authentication process. Since this SDR-based system is not approved for the CBRS bands, tests had to take place in an rf-isolated environment.

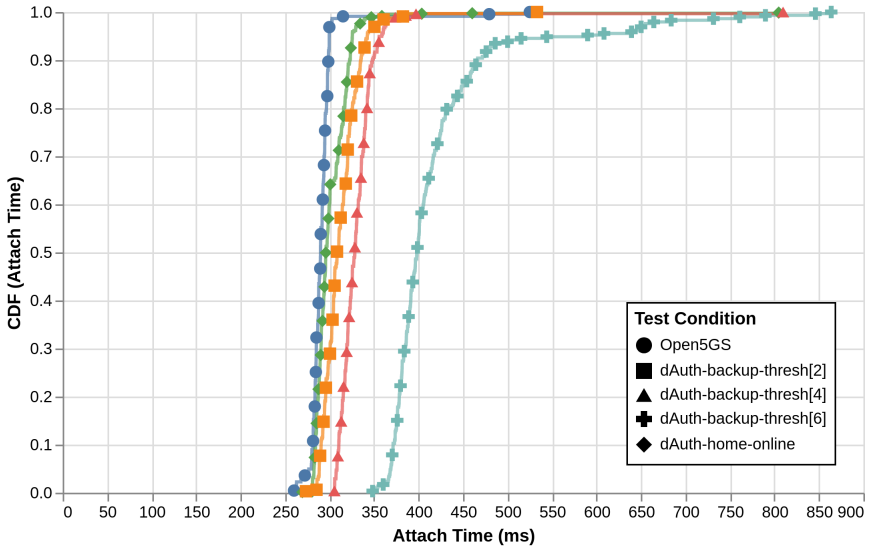
During each test, the srsUE continuously attached and detached from the test core network and recorded each attach duration. The LTE and 5G protocols include optimizations for re-attachment, so to accurately test the performance a UE new to the dAuth network would see, we additionally modified srsUE to discard its connection state information and start attaching from scratch in every iteration. For each test we collected at least 250 attach samples, and repeated this process with an edge deployment of Open5GS, dAuth when the home network was online, and dAuth when using 6 backup networks in SCN at various key share thresholds (excluding the four cloud-based test networks and two UERANSIM host nodes). Figure 5.6 shows the results form this test. Overall dAuth performed well, adding less than 50ms of additional latency for the backup authentication process when the authentication threshold was low. At the highest threshold, the authentication process was limited by the least performant node, a low-powered atom-based device with a relatively high latency backhaul connection.

5.6.3 Simulation Performance Testing

While srsUE allows us to make precise authentication timing measurements, we are unable to generate more than one authentication at a time. To test the scaling performance of dAuth,



(a) BoxPlots of Attach Times Across dAuth Test Conditions: dAuth adds a small but acceptable amount of additional latency when resorting to backup networks. All cases include rare outliers when packets must be retransmitted and/or inter-function connections established.



(b) CDF of Attach Times Across dAuth Test Conditions. dAuth is competitive with unmodified Open5GS when the home network is online and when the backup threshold is low. As the backup threshold increases, the serving network must wait for responses from straggling backup nodes before it can proceed.

Figure 5.6: Multiple visualizations of the authentication performance of dAuth vs. Open5GS with an off-the-shelf Baicells eNodeB and srsUE when a single UE continuously attaches.

we used the UERANSIM open-source 5G-RAN emulator (v3.2.4) to programatically generate a configurable number of authentication events in each test [67]. UERANSIM has both a gNB and UE component, and simulates connection overhead and the full 5G connection state machine to appear to the core network as a fully-functional 5G RAN. We modified the UE component to record a high-precision timestamp when it starts the connection process and when the connection is complete. We then programatically launch new UEs at a regular interval for each load level to simulate new users entering and authenticating to the network, possibly overlapping, and analyze the recorded timestamps to determine the connection latency experienced by each UE.

Connection Scenarios

We identified four scenarios of interest for testing representing different styles of community network endpoint. We provide results for each scenario.

1. An “edge core” deployment on an embedded computer at a site with high-quality Internet access. This corresponds to the majority of sites in SCN. For dAuth tests in this scenario we configure the RAN emulator to be attached to the serving network directly, and communicate with other dAuth instances via the backhaul connection. For Open5GS tests in this scenario we run a local instance of Open5GS on the edge with no roaming support, emulating a private 5G network.

2. An “edge core” deployment on an embedded computer at a site with residential-quality Internet (asymmetric, higher latency, typically cable). dAuth and Open5GS are configured the same as in scenario 1.

3. A managed “cloud core” deployment at a site with high-quality Internet, representing an approachable turnkey deployment solution for less technical organizations. For dAuth tests in this scenario we configure the RAN emulator to attach to a serving network hosted in

the closest datacenter of the 4 providers in the test network. For Open5GS tests we run Open5GS in the cloud, emulating a cloud provider or major operator's hosted core network.

4. A managed "cloud core" deployment at a site with residential-quality Internet. dAuth and Open5GS are configured the same as scenario 3.

General Performance

Overall our dAuth prototype performed well relative to the baseline of Open5GS, and shows that the additional capabilities for inter-site roaming and backup authentication provided by the dAuth architecture do not significantly impair performance.

In our first test we compare the latency of dAuth when a home network is available to a standalone Open5GS core. Results are shown in Figure 5.7. The additional overhead of dAuth relative to the standalone core is noticeable but acceptable at low load levels. At higher load levels dAuth actually outperforms the standalone Open5GS core and is able to maintain relatively consistent performance due to how load is distributed across machines in the network.

In our second general performance test we explored how dAuth behaves when operating in the backup network mode. The ability to authenticate when the home network is offline is a new capability relative to existing cores and comes with additional communication and cryptographic overhead, so we expected performance to degrade. We do see that overall the backup mode is slower than both standalone Open5GS and dAuth in the home mode at low load levels. As load increases each serving network has to communicate with multiple backup networks for each authentication, so there is less gain from load sharing across the network and performance degrades similarly to the centralized core. Figure 5.8 shows the detailed backup network test results.

We do observe that dAuth in backup mode still marginally outperforms Open5GS at higher load, and anticipate this is due to there being more nodes in the network than UEs were configured

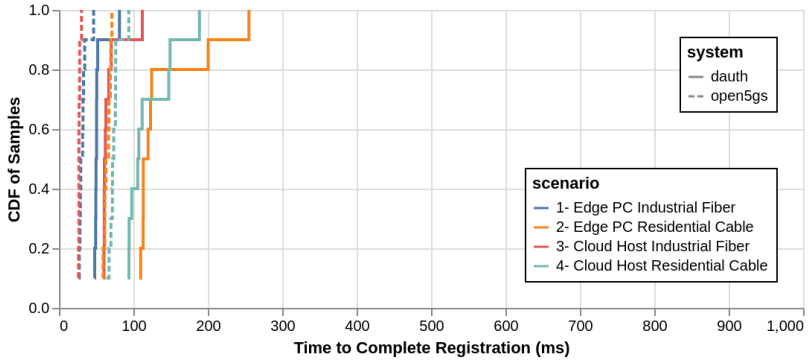
to back up to, still allowing for some level of load sharing. This leads us to believe that as the number of participating nodes grows larger than the number of backups, the load will reach a steady state per node while the total network capacity can continue to scale.

5.6.4 Impact of Security Parameters

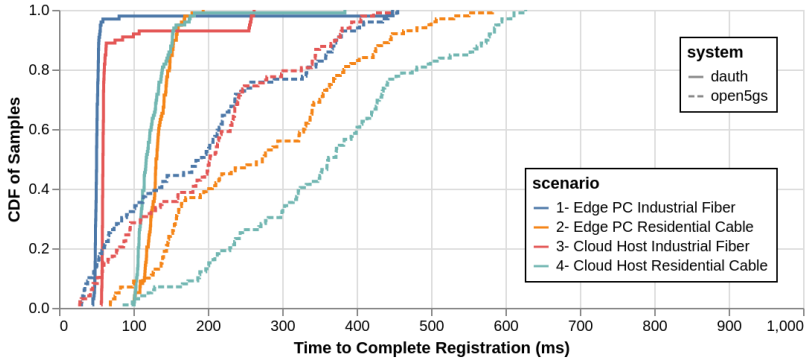
The level of security given by dAuth is directly related to the threshold of collaborating backups required to reconstruct the key share. The system is more resistant to collusion as the threshold increases, but at high thresholds serving networks must wait for many responses before proceeding with authentication. Similarly, the total number of backup networks also impacts performance and is interrelated with the threshold. A wide gap between the number of backups and the threshold gives the most flexibility to the serving network, but adding additional backup networks requires trusting them to not collude and adds additional message overhead (since messages will be concurrently sent to all available backups during auth).

In the single UE case, Figure 5.6 shows how the latency increases as the threshold is increased given a fixed set of available backup networks. In the many UE case though, the threshold does not have a consistent impact on latency, likely indicating the bottleneck at higher load is elsewhere in the system. Figure 5.9 shows the performance at varying load levels and thresholds with a fixed total number of backup networks.

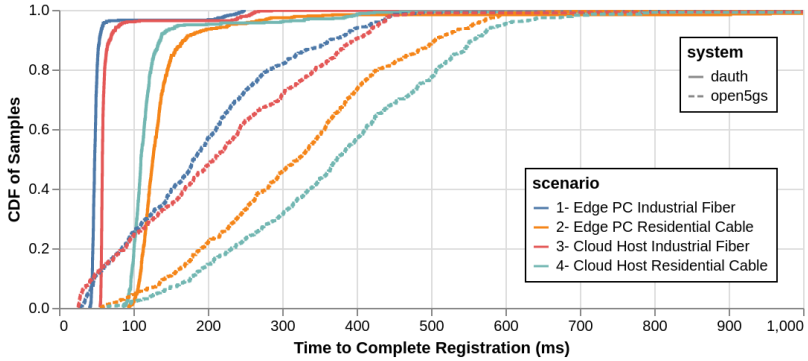
While the threshold alone does not significantly change the system's performance under the simulated load, increasing the total number of backups does have a measurable impact as load increases. In particular, the system saturates and tail latency degrades for a given threshold level as the number of backups decreases. The backups are queried in parallel for both initial authentication vectors and key material, but only the key material requires a threshold response, so having more backup networks allows more opportunities for the auth vector to be received quickly and the auth handshake to proceed. While the RAN and UE are processing the auth



(a) 20 registrations per minute

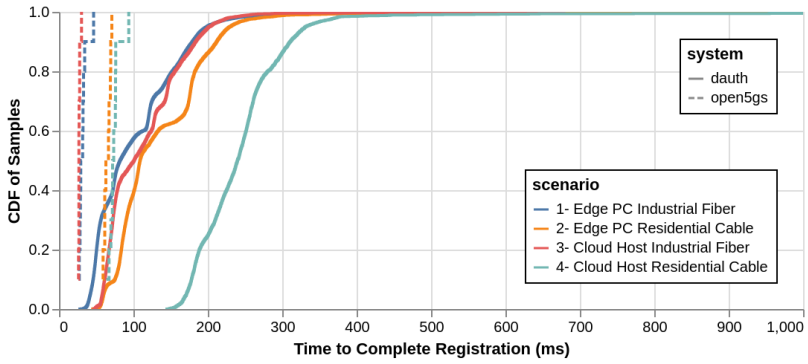


(b) 200 registrations per minute

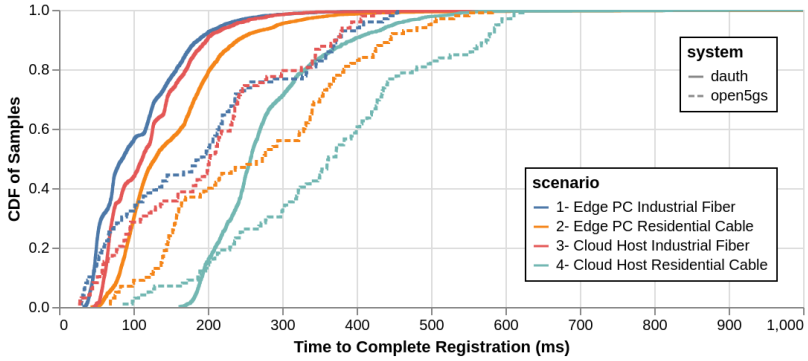


(c) 1000 registrations per minute

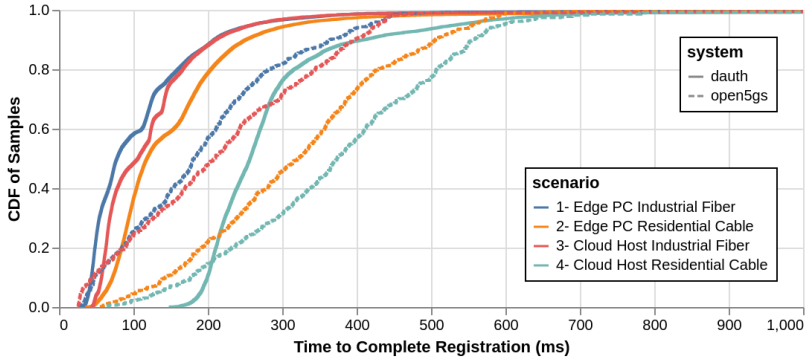
Figure 5.7: Attach latency of dAuth to a nearby home network vs Open5GS in a (~5ms rtt) data-center region. At low load the inter-core communication for dAuth roaming gives the advantage to Open5GS, but as load increases dAuth spreads processing across multiple machines.



(a) 20 registrations per minute



(b) 200 registrations per minute



(c) 1000 registrations per minute

Figure 5.8: Attach latency of dAuth using backup networks vs Open5GS hosted in a nearby (~5ms rtt) datacenter region. At medium and high load the scaling behavior of dAuth outperforms the centralized cloud core on less capable edge hardware, particularly over a less capable residential-class Internet connection. For multiple samples at each load dAuth is reconfigured with 8 random backups and a key threshold of 4.

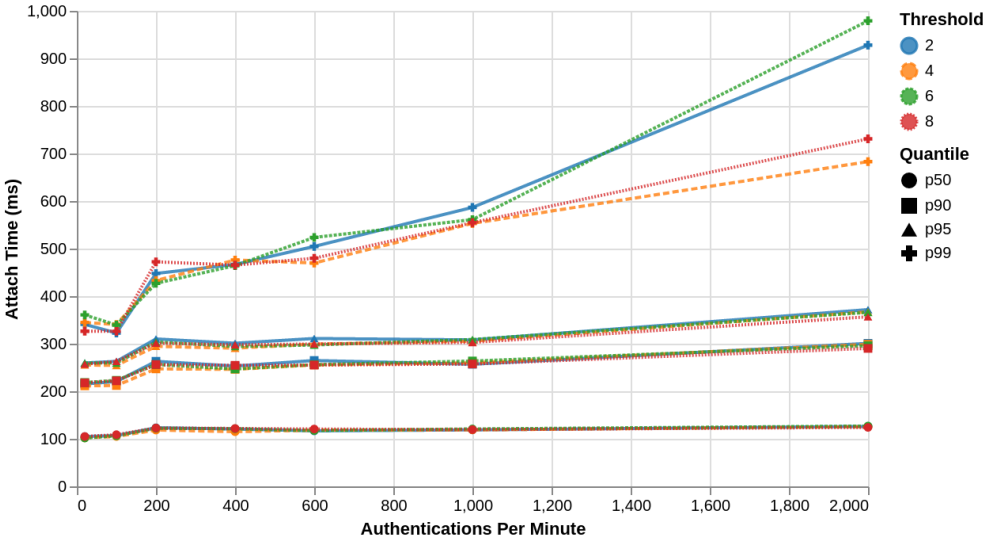


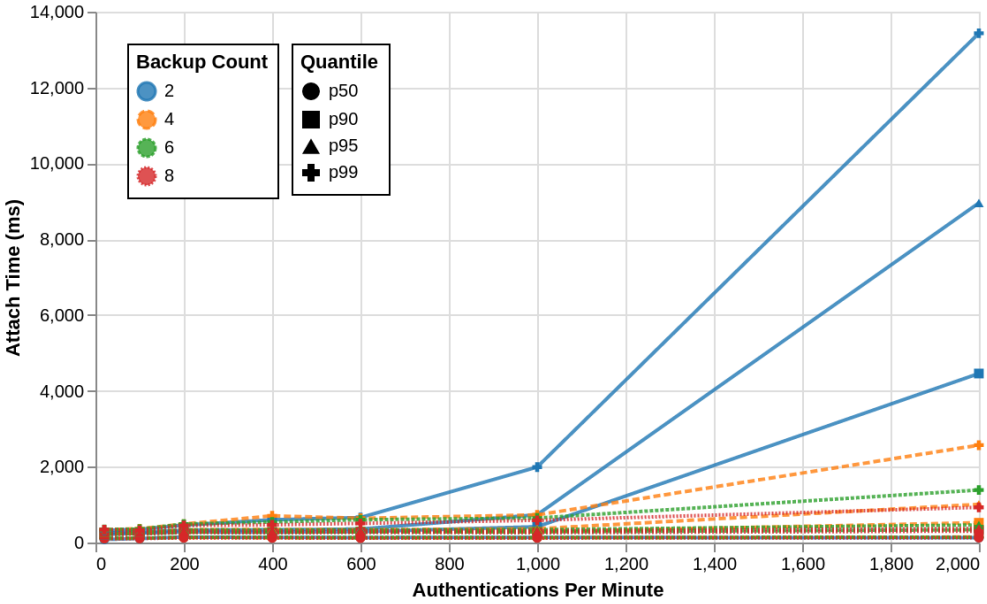
Figure 5.9: dAuth Loaded System Performance vs. Threshold. This figure plots the authentication latency at different load levels for different key share thresholds. In all cases the backup network size is fixed at 8 nodes. Under load there is no consistent relationship between the threshold and relative performance, neither in median performance nor tail performance.

vector and preparing the RES, gRPC connections are still being established to the lagging backup networks, and all backup networks can proactively read key shares into memory, making the key share query following receipt of the RES from the RAN much faster in the normal case.

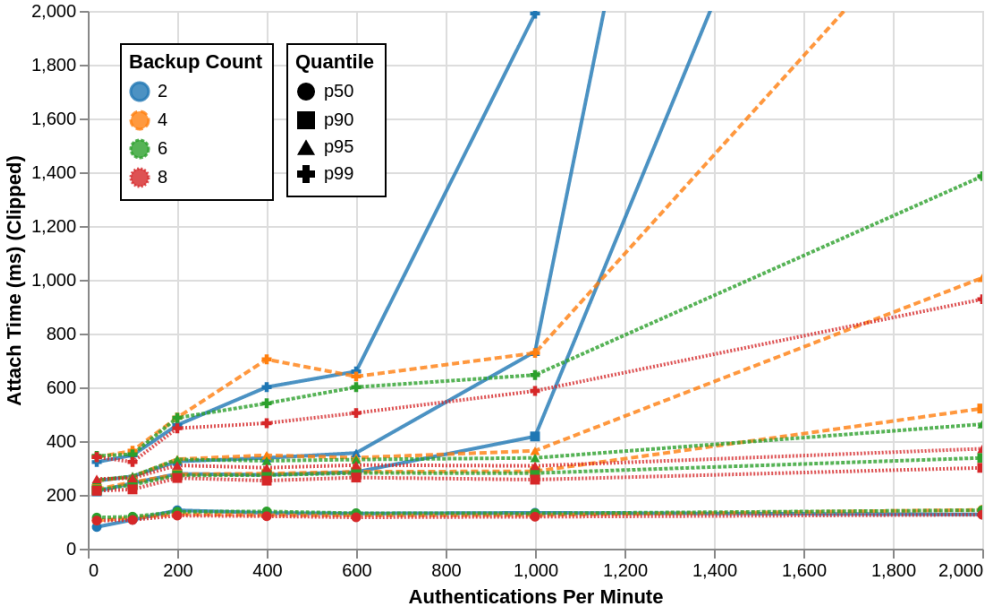
5.7 Discussion

5.7.1 Why not eliminate the home network?

The main benefit of dAuth is that the home network no longer needs to be online to allow a UE to authenticate to a serving network. This opens up an interesting possibility of actually eliminating the home network entirely, and implementing the home network functionality on the user's device itself. In this mode the secret key would only exist on the user's device, and would be used to generate auth tuples and key shares then proactively distributed across the backup networks. After the UE is initially bootstrapped to provide its own keys to a set of backup networks, the UE



(a) Performance vs. Backup Count, Full Graph



(b) Performance vs. Backup Count, Clipped for Scale

Figure 5.10: dAuth Loaded System Performance vs. Backup Count. This figure plots the authentication latency at different load levels for different numbers of configured backup networks. In all cases the key share threshold is fixed at 2 nodes. Tail latency degrades as the number of backup networks decreases and there are fewer nodes available for natural load balancing.

itself has all of the data necessary to act as a home network with only one user. This would raise challenges in the provisioning enough network IDs for every user, but could be surmounted by aggregating multiple UEs together into a virtual pseudonetwork, while still keeping the secret keys only ever present on the UE.

5.7.2 Future Work

This work raises implications for how to design fallback to allow connectivity when networks inevitably fail, even within the status quo of a few large networks, and we're excited to see this research continue. In particular we hope to dive into the nuances of how to better handle the SUCI key, and do a more thorough and formal analysis of how reputation could be managed and correct behavior incentivize in a production dAuth network. We also think there is interesting work to be done in applying the same ideas from dAuth to explore how to perform handover between small networks in our envisioned federation.

Another interesting direction could be extending this work to explore how to handover, not just roam, between networks operated by different organizations. Allowing for performant and secure inter-organizational handover likely requires additional changes to the cellular standards that were out of scope for dAuth, but solving it would make a large-scale dAuth system much more performant and suitable for more rapid mobility scenarios.

5.8 Related Work

5.8.1 Community Networks

dAuth is motivated by the wider body of work building and characterizing community networks, networks owned and operated by users in some sort of collaborative way. Community networks have long been viewed as a promising mechanism for increasing access among rural and disadvantaged populations [140, 149]. They come in as many diverse forms as there are communities

to host them, ranging from large urban areas [141, 98] to small rural sites [74], in the global north [141, 18] and global south [108, 83]. Examples of operational community networks include Guifi.net [20], Digital Tribal Village [155], TakNet [108], and many others.

Community networks span a wide spectrum of organizational structures [120], with some growing organically as user-to-user meshes [45, 50], while others adopt more centralized but still community-oriented structures [167, 19]. Researchers from the Universitat Politècnica de Catalunya, Barcelona, Spain in particular have thoroughly characterized many aspects of the Guifi community network from both technical [186] and operational [19] perspectives. In dAuth, we take advantage of the natural decorrelation that can occur in more decentralized community networks with multiple owning and operating organizations with different supporting infrastructure, and motivate our assumptions by this existing body of work.

Community Network Resiliency and Repair

There is a long history of research across the networking and HCI communities in the design [187, 131, 94], operation [151, 68, 19], and resiliency [50, 90] of community networks. In their early work in the space, Surana et al. find that real-world rural community networks face many operational challenges, primarily from backhaul reliability, power, and system configuration [180]. I note that the characteristics of many of these failures are isolated to single sites, further motivating dAuth's approach to building system resiliency from individually less reliable networks. Bidwell characterizes the social repair ecosystems around community networks extensively [25, 24], Dye et al. documents the person-to-person repair practices in Havana's StreetNet [50], and Garrison et al. find that repair and operation of a community network often times is used as a pretense for maintaining social connections and bonds between operators of different network nodes. The dAuth system design leverages these qualitative findings to build a practical system based on the existing trust relationships between different small operators that often exist in the community

networking context.

Community Cellular Networks

Community networks can grow using a variety of technologies from 802.11 WiFi [37], to legacy non-Internet cellular protocols [79], to analog telephones bridged over wireless [7], to 5G [101, 113]. dAuth focuses on Community Cellular Networks due to their unique blend of affordances and the rapidly evolving cellular device and spectrum access ecosystem that makes them a timely area for further research (see Section 2.1 [The Case for Cellular]). Most existing CCN research has focused on the challenges of building and deploying CCNs in a single network context [101, 166, 79], focusing on rural access and coverage. dAuth builds from this work's development of community-appropriate cellular technologies, but targets a different operational point where there are multiple coexisting operators in an area rather than a single operator. This brings additional challenges and a different trust model in the protocol design. Such an operating point is likely more realistic in urban and suburban settings in the short term, but could feasibly extend to small rural towns in the future.

5.8.2 Shared Cellular

Others are also exploring different models for shared networks in the cellular context. FreedomFi is a startup in the United States building hardware and software to enable individuals to deploy basestations and receive credit on the Helium blockchain for providing service [5]. Through partnerships with existing network operators, they allow for both private users and capacity offload from national networks to individually-deployed basestations. While this model decentralizes the RAN, it still relies on centralized authentication and billing infrastructure from the partner MNO, leaving the same single points of failure as existing traditional cellular networks.

Hasan et al. developed Community Cellular Manager, a system which allows an anchor

MNO to leverage federation with community cellular network more easily than permitted by existing 3GPP protocols [74]. Its architecture accounts for backhaul unreliability by pushing core network functions to the edge like dAuth, but assumes a setting where either nodes are trusted and authentication information can be shared directly with the edge nodes, or they are not and the authentication information must remain with the anchor operator. Some of the original authors of CCM went on to work on Magma, a 4G/5G-capable distributed core network with a similar high-level architecture [181]. Neither CCM nor Magma directly address the challenge of safely distributing security information across edge nodes when edge nodes cannot be independently trusted. dAuth addresses this problem directly, and the techniques developed for dAuth could be applied in the FreedomFi or CCM/Magma federated network architectures.

5.8.3 Cellular Core Reliability and Performance

Beyond community networks, general reliability in cellular networks has received considerable attention from the networking community. In Skycore, Moradi et al. sought to build a reliable UAV-based network, and arrived at an architecture where each UAV runs its own core network stack at the edge, similarly to CCM, Magma, CoLTE, and dAuth [123]. Additionally, to address the power and capability constraints of lightweight UAV hardware, they pioneered pre-computing and pre-distributing the authentication and security information tuples, not just the key material, ahead of time to lower the processing overhead on each UAV. dAuth builds on this work in a very different operational domain, and adds additional guarantees to safeguard the pre-computed authentication information in the event that a subset of backup nodes are compromised.

Much attention has also focused on possibilities for refactoring and redesigning cellular core networks for increased performance, both in terms of scalability and latency/throughput. Moradi et al. decomposed the monolithic cellular core into smaller pieces with better support for reconfiguration to improve flexibility, resiliency, and performance [124, 122], Qazi et al. proposed

re-architecting the protocols between core network components for better state locality [144], and Mohammadkhan et al. completely redesign the division of concerns between core network components, again to improve performance [121]. In the context of large-scale networks on virtualized platforms, Katsarakis et al. developed rVNF to make it possible to implement reliable stateful cellular core network functions more efficiently on cloud infrastructure through a cellular-optimized replicated data store [102], and Nguyen et al. re-allocated the responsibility for particular pieces of state between the core and the edge to allow implementing the core on performant but unreliable cloud infrastructure [127]. While these systems for the most part improve reliability and performance within an individual network, and are backwards compatible, none of them address the natural correlation inherent in the operations of single-operator systems and are still vulnerable to configuration or software rollout failures. By allowing for collaboration between multiple organizations, dAuth provides an additional layer of resiliency and decorrelation beyond single-operator approaches, which could theoretically be combined with these proposed intra-operator architectures in future work.

5.8.4 Alternative Authentication Schemes

Researchers have explored alternative authentication schemes for cellular networks, seeking to also bring more flexibility into the authentication process. Johnson et al. proposed a network architecture based on publicly releasing the symmetric keys for all user SIM cards, effectively removing any authentication protections and relying on over-the-top services to provide security via VPN connections or TLS [96]. Schmitt and Raghavan propose using the same SIM credentials for all users in the network to preserve anonymity at the link-layer, and use signed tokens at the application layer to pay for service while remaining relatively anonymous [159]. Both of these approaches are backwards-compatible with the existing device ecosystem like dAuth, but by relinquishing control over the connection between the UE and the basestation, leave the user

vulnerable to attacks abusing the basestation's trust in the device ecosystem to cause the UE to behave maliciously, potentially illegally interfering with other networks, draining the device's battery, or spamming the user with spurious alerts and messages. dAuth provides additional guarantees to the user that the serving network is trusted by their home organization while still remaining fully backwards compatible with off-the-shelf devices.

In Cellbricks, Luo et al. outline a vision similar to ours, were many small organizations are capable of joining together into a federated cellular network [111]. Unlike dAuth, their approach relies on a centralized data broker to be online to validate user credentials (leaving users vulnerable to an outage at a single organization), and requires modifying the baseband of the user device to use asymmetric cryptography. Jover and Lackey similarly propose a non-backwards-compatible architecture based on asymmetric cryptography with a coordinating blockchain in dHSS [97]. In contrast to these works, dAuth avoids asymmetric cryptography to maintain compatibility with existing 3GPP standards. This allows for immediate and incremental deployment in today's CCNs with unmodified off-the-shelf user and RAN hardware.

5.9 Conclusion

Our system, dAuth, enables real-world deployment of small cellular networks with standards-compliant Commercial Off The Shelf (COTS) user devices widely available today. We take advantage of the details of the battle-tested AKA authentication scheme to allow networks to proactively share authentication material to allow redundancy in the case of local failure and share the load of authentication across multiple nodes through natural sharding of user state.

Chapter 6

Network Capacity as Common Pool Resource: Community-Based Congestion Management in a Community Network

Executive Summary

In this work I conducted an initial study of the feasibility and appropriateness of applying established CPR governance principles (see Section 6.2.4 [Common Pool Resource Governance] and Table 6.1) to community network congestion management. My collaborators and I drew our methodology from elements of both Value Sensitive Design and Participatory Design, soliciting participants' feedback on a range of possible congestion management system designs while probing the value tradeoffs embedded in each.

Hypothesis

Drawing from my lab's collective experience building and working in community networks, my collaborators and I hypothesized that existing economic theories of Common Pool Resource (CPR) Governance, formalized by Elinor Ostrom and the Bloomington School of Economics [135, 43], could apply to congestion management in community cellular networks.

Summary of Key Findings

Among participants, my collaborators and I confirmed a preference to avoid pricing-based mechanisms when allocating network resources and find desires to non-neutrally prioritize person-to-person communications, support local human-mediated management tools, and balance respect for individual data privacy with informed, community-based network governance. Our initial results support that Ostrom's CPR governance principles are compatible with the values of local users and could help structure efforts to achieve the goals identified by the workshop participants for the network.

Next Steps & Connections

Building from this work I prototyped a user-facing system for managing bandwidth in community cellular networks according to the principles outlined above, and am continuing to work with Rhizomatica to support their deployment of values-compatible cellular networks in Oaxaca. This project also inspired me to re-think the sets of privacy tradeoffs and assumptions which become embedded in low-level protocol designs, connecting to my work examining how networks can collaborate to provide more reliable service.

6.1 Introduction

As communities build and begin operating networks, physical limitations such as device power, legal limitations such as spectrum availability, and financial limitations, such as equipment cost, lead to practical constraints of the materialized network. Network congestion, the state of one or more of the network's links being insufficient to support the traffic generated by end-user applications, is a common phenomenon that results in poor quality of service and poor experiences using the network. In Internet Protocol (IP) networks, congestion causes packet delay and loss, experienced as slow page loads, stutters and drops in audio calls, or in extreme cases, resource timeouts that make entire sections of the modern web unavailable to these users [64].

We draw a distinction between *congestion control*, the automatic scaling-back of traffic to coordinate multiple devices on a network at machine-appropriate timescales (nanoseconds to seconds), and *congestion management*, the process of allocating bandwidth between users and tasks at human-appropriate timescales (minutes to years). Commercial Internet Service Providers (ISPs) traditionally perform congestion management through pricing, where users pay a higher price for a greater amount, faster rate, or increased Quality of Service (QoS) [47, 163, 114, 164, 193, 78, 53]. These policy decisions are then enforced through changes to network parameters which impact machine-level congestion control and ultimately users' traffic.

"Data plans" with a fixed speed (i.e. 10Mbps) and/or quantity (i.e. 5GB) are straightforward and common, but lead to under-utilization and wasted capacity. Dynamic "smart data pricing" mechanisms are economically efficient [163], but are complex, can have adverse side-effects on real users who rarely act completely rationally [129], and break incentives when networks face divergent short and long-term objectives [152]. Managing congestion through pricing also introduces practical challenges to billing and payment collection [18]. Importantly, allocating resources via pricing in a *community* network can contradict the values inherent to that particular network [120], which may be modeled by its community as a public or common good inappropriate

for pricing.

Drawing from our group's collective experience building and working in community networks, we hypothesized that existing economic theories of Common Pool Resource (CPR) Governance, formalized by Elinor Ostrom and the Bloomington School of Economics [135, 43], could apply to congestion management when pricing is inappropriate. While Ostrom's work focuses on physical resources such as fisheries or watersheds which are vulnerable to tragic long-term collapse [135, 133, 136], persistent network congestion collapse can lead to frustration and an inability to accomplish productive or time-sensitive tasks. Network congestion emerges when the collective demand of local users exceeds the carrying capacity of the network, mirroring how shared physical resources collapse in the existing economics literature. The research team has observed multi-hour outages where although a connection was available, even relatively lightweight mobile-optimized applications timed out and failed due to excessive congestion despite congestion control.

In this work we conduct an initial study of the feasibility and appropriateness of applying established CPR governance principles to community network congestion management. Management of a network with the principles outlined by Ostrom requires rules for access and allotment of resources in the face of scarcity that carry embedded values. As primarily ICTD researchers, we are sensitive to the centrality of the Global North to the design of existing tools for network management. We draw our methodology from elements of both Value Sensitive Design and Participatory Design, and solicit the users' feedback on a range of possible designs while probing the value tradeoffs embedded in each.

Working with a local partner, we held a series of workshops and interviews in Santa Inés, a small community with a history of communal resource management in Oaxaca, Mexico near Asunción Nochixtlán. We identify concerns held by members of the community around privacy and information in a locally owned IP network, explore values around how their community would define fair sharing of network throughput, and gather opinions for how such sharing should be

structured. Community network management presents a unique challenge since operators often lack deep technical knowledge of how IP networks function [90], but are tasked with managing a network serving the entire town. At the same time, their responsibilities cross privacy and trust boundaries more typical of wide-area networks, serving a variety of users with whom they may have very different relationships.

We do not claim to establish which approach is best for this community or other communities in general, but identify directions for the design of future systems in this context which diverge from existing approaches. Among participants, we confirm a preference to avoid pricing-based mechanisms when allocating network resources and find desires to non-neutrally prioritize person-to-person communications, support local human-mediated management tools, and balance respect for individual data privacy with informed, community-based network governance. Our initial results support that Ostrom's CPR governance principles are compatible with the values of local users and could help structure efforts to achieve the goals identified by the workshop participants for the network.

6.2 Related Work

6.2.1 Community Network Operations

A significant volume of work from CSCW, HCI, ICTD, and other non-academic contexts has explored the challenges of building and maintaining community networks. Examples of community networks include Guifi.net [20], Digital Tribal Village [155], TakNet [108], and many others. Community networks span a wide spectrum of organizational structures [120], with some growing organically as user-to-user meshes [45, 50], while others adopt more centralized but still community-oriented structures [167, 20].

Previous literature has explored long-term maintenance and upkeep of these networks, with Surana et al. describing early technical challenges in network maintenance [180], Bidwell ex-

amining the role of women in sustaining community networks in the Global South [25], Dye et al. documenting the care and inter-personal coordination of maintaining Havana's organic StreetNet [50], and Jang et al. exploring crowdsourced local repair [90]. In contrast to work focused on construction or repair, we consider the day-to-day resource management in an operational network, but which is constrained even in its fully operational state. Dye et al. do briefly discuss bandwidth management and conflict remediation practices implemented within the StreetNet organizational hierarchy, but in a very different context than Santa Inés. Our work additionally builds on theirs to explore a wider design space of possible management practices.

Prior work has also characterized how resource management impacts sustainable long-term operation of community networks. In African community networks, Rey-Moreno et al. [148] found that backhaul (the connection between the community network and the public Internet) bandwidth was a significant part of operational costs. This is the case in Santa Inés as well. Rey-Moreno et al. note the importance of local services to remove some pressure on the backhaul, but did not explore the details of how the limited bandwidth and network congestion was managed. Baig et al. [18] recounted the challenges scaling and managing resources within Guifi.net, the largest and most successful community network in the world. The Guifi.net model relies on a complex set of accounting and cost sharing agreements between operators and maintainers, and uses a traditional monthly billing model for end-users. In contrast, the network in Santa Inés operates at a much smaller scale and faces fundamentally different challenges since there is only one operator (the community's telecom coop) that operates much closer to its users. Additionally the dominating financial constraint in Santa Inés is the operational cost of a long-distance Internet connection rather than capital cost of new infrastructure.

6.2.2 Existing Approaches to Wide-Area Congestion Management

Pricing

Much scholarship has been dedicated to the question of how to *price* Internet and mobile phone services in the face of limited capacity and users' quality of service expectations. A first wave of Internet pricing research in the '90s and early 2000s focused around the challenges of using pricing to control demand and maintain service quality in rapidly growing Internet and mobile networks. A variety of pricing mechanisms based on auctions [114], congestion marking [62], priority bands [130], flows [137, 92], and hybrid approaches [195] were proposed, but Internet providers predominantly continued to use basic pricing schemes [129]. DaSilva [47] and Falkner et al. [53] survey this era of research.

A second era of pricing research began in the early 2010's and continues today. Sen et al. and Chiang et al. argue that "smart data pricing" is still relevant for congestion management in modern networks [164, 163, 42]. Ha et al., Joe-Wong et al., and Sen et al. argue for time dependent pricing in both fixed and mobile networks [93, 69, 165]. Ha et al. additionally developed a user-facing application, TUBE [71], to help end users manage the complexity of optimizing use under these time varying schemes.

While pricing has the advantage of good scaling properties, pricing may not always be appropriate in all contexts, particularly in close-knit communities like Santa Inés. In all of the research on economic means of network congestion management, relatively little exploration has been done towards non-pricing-based mechanisms. MacKie-Mason and Varian note as an aside: "There are many ways to deal with congestion externalities. One way is to establish social norms that penalize inappropriate behavior. Such norms can work well in small groups where there is repeated interaction, but they often do not scale well to a system with millions of users" [114]. In this work we explore whether non-pricing congestion control mechanisms and policies can

apply to medium-sized groups of hundreds, not millions, and allow small community networks a different way to manage available capacity.

Embedded Automatic Protocols

"Congestion Control" is a well-studied technical domain in computer networking essential to the more general task of congestion management. In the Internet architecture, end-to-end transport protocols like TCP [35] or QUIC [106], using congestion control algorithms like CUBIC [70] or BBR [32], sense network congestion through packet loss or delay and then scale back the amount of data transmitted by each client until the congestion is resolved. These protocols are fundamental to the stability of Internet and largely operate automatically outside the knowledge of end users. Yet on extremely constrained links, like those found in remote networks like in Santa Inés, automatic congestion control protocols can only divide the limited resource so much. In a network serving tens to hundreds of users with only a few megabits per second of Internet throughput, this can result in per-flow allocations of 100Kbps or less, insufficient for the modern Internet. Higher level decisions about what traffic should be allowed when, and by whom, are required. We see low-level automatic techniques as compliments to, rather than replacements for, higher-level logic to manage network demand as explored in this work.

6.2.3 Existing Approaches to Home and Personal Connection Management

Insights from home and personal network management can also apply if we model a community network as an entity managed by a group of non-experts with a shared commercial backhaul connection and many connected devices. The CSCW and HCI communities have explored some of the approaches taken by individual users to manage their home connections. Grinter et al. [65] make a case for CSCW researchers to pay attention to how home networks are used, the maintenance and operations work that they generate, and their impact on the overall home

environment. Yang et al. [196], Chetty et al. [38, 39], and Mortier et al. [125] develop and deploy several different home network management tools, delving into the social implications of revealing network status and giving network control to users in a shared space. These tools could be deployed as mobile applications or as views in a central management interface, but would not be appropriate for a context like the Santa Inés network (violating privacy values discussed in section 6.5.1). Chetty and Mortier both note social tensions and conflict caused by the information made visible and the control capabilities provided by their technologies. We anticipate that these issues would be exacerbated in the community network setting, where users have looser social connections and less frequent direct interaction, and designed our line of inquiry to explore these tensions.

Researchers have also examined how individual users manage their personal connections and developed a range of applications to assist in managing them under different network regimes. Chetty et al. developed uCap to help users plan for data caps in the home [40], Sambasivan et al. [153] developed SmartBrowse for managing mobile data consumption, and Ha et al.'s TUBE [71] allows users to automatically optimize consumption with time-varying prices. Im et al.'s AMUSE [86] helps users plan for when low-cost connections will likely be available, time shifting non-essential traffic. Similar tools could be deployed in the community networking context to help make users more aware of their consumption, smooth demand, and provide visibility to more effectively conserve network resources. While these applications give power and control to end-users, they all assume the network itself is fixed and that users interact with it independently. Community networks offer an additional opportunity, unconsidered in existing tools, for users to collectively optimize network policies that could lead to higher performance than individual-level optimizations alone.

6.2.4 Theoretical Framings

Throughout our work, we operationalize Ostrom's theory of Common Pool Resources (CPRs) [135] as well as Participatory and Value Sensitive Design (VSD) methodologies [60].

Common Pool Resource Governance

Within economics, significant study and theoretical work has been dedicated to the management and sustainability of common pool resources (CPRs), resources which can be appropriated (used) communally without private ownership, but can be overused without management and coordination. Elinor Ostrom's Nobel Prize-winning work [135] characterized successful and unsuccessful approaches to the management of common resources and popularized a framework for understanding how communities can effectively manage CPRs over time through collective action. Prior to her work, it was widely believed that the only ways to sustainably manage CPRs were external government regulation or privatization, but she finds stable counterexamples from a wide variety of real-world institutions, and devises game-theoretic models for how these counterexamples operate [134].

We conceptualize community network bandwidth as a Communal Pool Resource in Santa Inés, where network users are appropriators and network congestion corresponds to states of overuse. Prior works from Bernbom [22] and Hess [82] explore modeling the macro-scale Internet as a CPR using Ostrom's principles, but neither explore design implications nor consider how Ostrom's principles could be scaled appropriately to the community network context.

Ostrom outlines common principles shared by successful institutions (summarized in table 6.1), which we explore applying to community network congestion with the members of Santa Inés.

Table 6.1: A brief summary overview of Ostrom's design principles of CPR governance, detailed in Chapter 3 of Governing the Commons [135]. We find that congestion in the LTE network in Santa Inés could be managed in a way compatible with community values and all 8 principles, and discuss this further in sections 6.5.4 and 6.6.4.

CPR Design Principle	Summary Description
1. Clearly defined boundaries	The boundaries of the resource and those allowed to use it are defined and enforceable.
2. Congruence between appropriation and provision rules and local conditions	The way the resource is actually used is reflected in any rules applied, and those rules are responsive to changes in local conditions "on the ground".
3. Collective-choice arrangements	Most users impacted by operational rules can participate in rule modification.
4. Monitoring	Monitors are able to track use, and are accountable to local users or are local users themselves.
5. Graduated sanctions	Punishments for breaking rules are contextual and flexible given the seriousness of the offense.
6. Conflict-resolution mechanisms	Inevitably arising conflicts can be resolved quickly and at low cost.
7. Minimal recognition of rights to organize	The rights of users to organize and self-regulate are not challenged by external authorities.
8. Nested enterprises	If the CPR is part of a larger system, activities are organized in multiple layers.

Participatory & Value Sensitive Design

Value Sensitive Design (VSD) methodologies as outlined by Friedman, et al. are an attempt at supporting ethical technology design by investigating the relationship between human values and designs that may align or conflict with those values through their functions or affordances [60]. VSD proposes three major lines of investigation: conceptual, focusing on forming theories about stakeholders and their values; empirical, focusing on observing and clarifying stakeholders' competing values, practices, and motivations around use; and technical, focusing on properties and mechanisms of technology and the values that they support or hinder. We conduct participatory

Figure 6.1: From left to right: the trial LTE hardware, the view of the centro in Santa Inés, and the town hall and community telecom office.

design workshops [103] with the help of our partner Rhizomatica to elicit Santa Inés network stakeholders' values and desires for network management, which will inform future technical designs for their network.

6.3 Context

This research is part of a long-term collaboration between the researchers and Rhizomatica/TIC, a nonprofit organization specializing in rural connectivity and community media, to develop new forms of Internet-capable community networks. Here I outline the context to understand the community's current Internet access, unmet technical needs, and relationship with Rhizomatica/TIC. Rhizomatica has a multi-year relationship with Santa Inés, and they jointly run a sustainable GSM (2G) (voice + messaging only) community cellular network. Rhizomatica introduced the research team as a possible partner in exploring management options for a new Internet-capable LTE (4G) network, and Santa Inés let the researchers stay for two months of field work and helped organize public workshops in exchange for assistance soliciting and organizing ideas for managing the LTE network.

6.3.1 Rhizomatica/TIC

Rhizomatica/TIC has operated in the Oaxaca region for the past 12 years, offering technical and legal support for the development of Community Cellular and Radio Networks. Rhizomatica/TIC's mission is "to increase access to and participation in telecommunications by supporting communities to build and maintain self-governed and owned communication infrastructure," through "regulatory activism and reform, critical engagement with technology and the development of decentralized telecommunications infrastructure, and direct community involvement and participation." They explicitly encourage local, value-sensitive governance of telecommunications, with implicit political biases towards local autonomy, sovereignty, and community solidarity. These biases come into play through Rhizomatica/TIC's role in workshop facilitation, subtly impacting themes and results. However, we perceive that Santa Inés's long-standing relationship with Rhizomatica/TIC has also led to general alignment on values and goals.

Santa Inés

Santa Inés can be broadly considered *rural*, which shapes residents' access to the Internet and other technical infrastructure. I follow Hardy, Wyche, and Veinot's recommendation [73] to characterize the research context's rurality along descriptive and sociocultural dimensions.

Santa Inés is a primarily agricultural area with around 1000 residents located 30km from Asunción Nochixtlán, the nearest larger town with a petrol station, supermarket, and high-speed connectivity. The surrounding terrain is rugged, with sharp hills and mountains on all sides. Regular communal taxis go between Santa Inés and Nochixtlán, 25 minutes each way (plus wait time to collect a full car).

There is a strong sense of shared identity in Santa Inés, reinforced by multi-generational familial, economic, and geographic ties. A rotating local government is elected every 3 years to manage day-to-day operations, but all large decisions are put to a vote before "la asamblea," a

monthly meeting of all heads of household. Multiple participants mentioned a sense of duty to the community's welfare, for example to prioritize local over distant sales of corn in times of scarcity.

Wireless ISPs offer Internet connectivity for ~$50USD/Mbps/month, but few purchase a home connection since the links are expensive and unreliable. A shop and small cyber cafe in the town center sell hourly WiFi access ($0.50USD/hour), and there is a free connection and computer lab at the town hall. The free connection is time-limited per person per week since it is provided by the national government with an expensive satellite link. Santa Inés residents are not new to the Internet, but must travel to the town center or Nochixtlán for the services they need.

6.3.2 Existing Cellular Networks

2G Network Providing Congestion Management Experience

Santa Inés owns and operates a nonprofit GSM (2G) cellular network serving ~400 users, with technical assistance and training from Rhizomatica/TIC. The access point (or *basestation*) is located on a tower on a tall hill, covering the centro and many of the surrounding ranchos. The network provides calls and texts, both locally and with long-distance interconnect to the global phone network. Users pay a small monthly fee to become network "members," which grants unlimited local calling and texting, as well as the ability to receive calls from outside the community. Outbound long-distance calls are charged per minute depending on the destination, but rates are set as low as possible while covering costs. The network is a nonprofit, and the rates are generally considered reasonable by the residents.

The Santa Inés network only supports 15 simultaneous calls, and is commonly saturated since local calls are unlimited and free. Rather than putting a price on local calls, the community addresses congestion by limiting local calls to 5 minutes when the network is busy. This: a) breaks up long calls to allow new users to connect, and b) reminds local callers talking for a long time that they should consider continuing the conversation face-to-face. We revisit this alignment of

policy and values in the context of IP networks in section 6.5.2:Embodiment of Local Values.

LTE (4G) Network Trial and Adapting to Internet Congestion

Santa Inés community telecom and Rhizomatica/TIC are running a trial of a new LTE network to eventually provide high-speed Internet over a similar coverage area as their GSM network. The LTE network currently runs at low power due to limitations in the trial equipment, covering only the centro. It initially served ~15 users, and had expanded to ~40 by the end of this study. Rhizomatica/TIC has not decided how they would recommend Santa Inés manage network congestion in the LTE network and users had not yet considered the problem.

At the start of this research the trial network was operating as designed, but was already suffering from Internet congestion issues though the community had not identified them. The LTE and GSM networks shared their backhaul connection, and excess traffic from the LTE network was congesting this link, leading to packet loss which manifested as stuttering and dropout of long-distance GSM voice calls. Without management tools to make the problem visible or networking expertise to intuit what was happening, the president of the cooperative consulted the backhaul provider, who "fixed" the problem by disconnecting the LTE network. The LTE network was later reconnected by Rhizomatica/TIC and the research team, with a coarse rate limit in place to avoid future interference for the duration of the trial.

6.3.3 A History of Communal Operations

Santa Inés has a history of communal resource management and established values around appropriate and inappropriate allocation of shared resources. For example, water resources are collected in a communally managed reservoir and a rotating committee is charged with allocating water fairly between different families based on the size of the household and each family's needs. Distribution is based around perceived fairness, rather than a cost per liter. Similarly, communal

taxis charge a low flat price and pride themselves on delivering essential transportation to the community.

Funding to build the 2G cellular network's tower and purchase radio equipment was provided by the Santa Inés government, and the network occupies communal land and office space. Residents view the network as a shared resource which should be operated altruistically to the benefit of all. These values inform the network's management and pricing structures: the fixed monthly fee equally distributes the costs of maintenance and operations across all users, users pay at cost for long-distance calls, and local resource utilization is managed through the 5 minute local call limit applied equally to all when the network is busy. While there is a desire to apply the same high-level principles in the new LTE network, it is unclear how best to rectify the nature of modern Internet congestion with these existing values.

6.4 Methodology

Operationalizing elements of Participatory Design and Value Sensitive Design (see 6.2.4: Theoretical Framings), we sought viewpoints from multiple stakeholders, encouraged participants to think about community-wide interests and policy preferences, and explicitly elicited values associated with the network. Ostrom's principles informed our line of inquiry and helped structure our results. We adopted mainly empirical and technical VSD, using knowledge of cellular network affordances, combined with participatory methods, to reveal users' preferences and concerns and explore a wide space of possible congestion management strategies. In our analysis we synthesize our observations to develop a conceptual understanding of the values expressed and distill concepts to high-level themes such as *privacy* of usage data and *equality* regardless of financial means.

With Rhizomatica and the Santa Inés community telecom operator, we facilitated three public workshops to both educate the community about congestion in the LTE network and gather ideas for how to manage it. We also held two formal meetings with town leadership and conducted two

opportunistic interviews. All interactions were in Spanish; the research team leads speak proficient Spanish and Rhizomatica staff are a mix of fluent and native speakers. Over the course of this research, one researcher lived in Santa Inés for a month and facilitated workshops 1 and 2. Two other researchers joined for two weeks at the end of the study for all meetings, interviews, and workshop 3. Even though it was not a research outcome, the researchers present in the community made themselves available to the network operator for direct assistance with the trial network, troubleshooting several technical issues and helping implement a rate limit in coordination with Rhizomatica at the operator's request.

6.4.1 Workshops

The workshops extended from the residents' existing understanding of management strategies used for different resources in Santa Inés, network management experience with their existing 2G cellular network, and exposure to commercial prepaid and postpaid plans used when outside Santa Inés. Two facilitators from Rhizomatica were present for all workshops in addition to the researchers. The workshops took place in an outdoor gathering space near the town hall, and were planned for an hour of content but tended to run long due to questions and discussion. The Santa Inés network's leadership suggested that the most people would be able to attend if the workshops were held in the early evenings on weekdays, and they scheduled each workshop. Participants were not monetarily compensated, but refreshments were provided to attendees. The workshops were approved by our University IRB.

Rhizomatica indicated that some participants would likely be illiterate, so the workshops were designed to not require individual literacy. All written artifacts pictured were discussed verbally and written by identified scribes. Consent to participate was gathered verbally as well, and written demographic surveys were not conducted. Audio recordings and the researchers' field notes are the primary outcomes of each workshop.

Table 6.2: The count of participants in each workshop, their gender presentation, and the number of participants who were returning to participate in a second workshop.

	Total	Female	Male	Returning
Workshop 1	18	7	11	N/A
Workshop 2	6	4	2	2/6
Workshop 3	15	3	12	4/15

Workshop 1

Workshop 1 covered background of how the LTE network operates, how congestion impacts Internet networks, and generated initial use cases and ideas for congestion management. Participants were recruited from the general population of Santa Inés via a broadcast SMS message sent to all members of the 2G network two days beforehand. The operator also sent a broadcast reminder message and made a loudspeaker announcement in the town square half an hour before the workshop. 18 people participated (7 women & 11 men) and the workshop lasted for 1 hour and 40 minutes.

The workshop contained two parts, an overview of the network and congestion concepts introduced in a brief lecture, followed by an interactive question and answer session and large group discussion. Network congestion was explained (and discussed later) with both road traffic and carrying weight analogies to give an approachable way to relate to the abstract concept of congestion. Facilitators did not attempt to teach the details of how low-level protocols operate, but did offer feedback throughout the discussion that some strategies may be easier or harder to implement due to the types of information available in the network. Participants already understood that networks have a capacity limit from their experiences with the 2G network, and intuitively grasped the relative burdens different types of traffic place from experience with prepaid plans on commercial networks in larger cities.

The open discussion began by answering participants' questions about congestion and the differences between the new LTE network and the existing GSM network. After answering these questions, the facilitators steered the discussion towards the more open-ended topics of *Which applications and use cases are most important in Santa Inés?* and *What are some ideas for how the community could/should manage the network?* Facilitators emphasized that the network would belong to the people of Santa Inés at the end of the technical trial and that they had the agency to decide how they would like to operate it.

Workshop 2

Workshop 2 followed later in the same week as workshop 1, with the goal of generating more ideas for congestion management and making a rough rank order of the appeal of different policies. Participants were recruited through an announcement at the end of workshop 1 and a targeted SMS to residents using the trial LTE network or who had explicitly expressed interest in the LTE network to the telecom administrator. Workshop 2 had only 6 participants (4 women & 2 men), including 2 return participants from workshop 1, and lasted for 1 hour and 20 minutes. Workshop 2 was accidentally scheduled in conflict with a church service and following party which we believe limited participation in combination with the more focused recruiting.

Workshop 2 was structured as a group brainstorming session to explore possible network management approaches. The workshop began with a review of concepts from workshop 1 given by one of the returning participants with help from the facilitators. Afterwards the participants reviewed the ideas from workshop 1 and additional example proposals from the researchers, and then brainstormed new ideas together. For each idea, facilitators elicited the participants' thoughts about *Do you think this approach is a good or bad idea? Why?* and *How would this policy help or hinder using the network?* At the end of the workshop the participants came to a consensus on the top 3 ideas they would most like to see implemented.

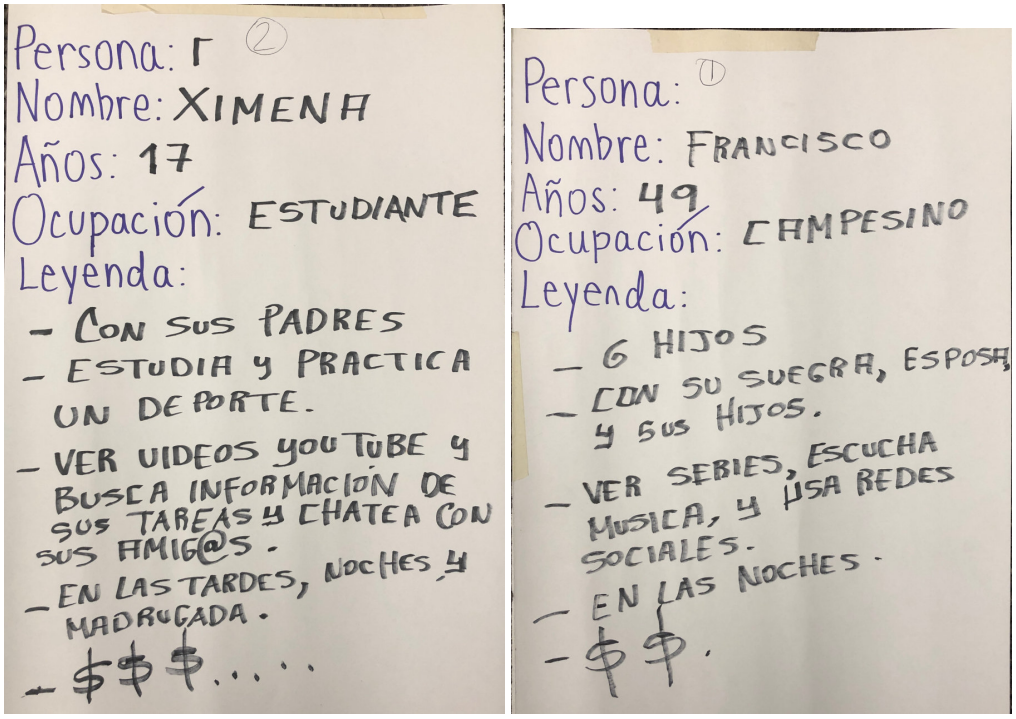


Figure 6.2: Two of the five personas created for workshop 3. Personas contained basic demographic information, as well as a freeform "legend" section where participants could describe important parts of the persona's life, including how the persona interacts with the network.

Workshop 3

Workshop 3 followed two weeks later, and was designed to gather feedback and opinions on the proposals from a wide variety of participants. Participant recruiting was done via a broadcast SMS message one day in advance, with a personalized followup reminder message targeting users of the LTE trial network the afternoon before the workshop. Workshop 3 had 15 participants, 3 women and 12 men, and lasted 1 hour and 35 minutes. Three participants had been present during workshop 1, and one returned from workshop 2.

Workshop 3 began with the generation of 5 user personas [143], which were then used to evaluate sets of theoretical network policies. Participants divided into 5 groups, where each group

represented a persona in addition to themselves. The researchers generated 5 sets of policies concretely implementing the ideas gathered in workshops 1 and 2, and presented these policies for evaluation. Unlike in workshops 1 and 2, the facilitators did not attempt to explicitly explain congestion, but provided a short practical example justifying each policy and then asked the participants *Is your profile in favor of this policy? Are you in favor of this policy? Why?* and *Does your profile thing this policy is fair and effective? Do you think this policy is fair and effective? Why or why not?* This sparked followup questions and debate, which the facilitation team encouraged.

6.4.2 Other Interactions

In addition to the main workshops, the field researchers arranged two meetings with Santa Inés' government authorities and the telecom coop leadership, and conducted two informal interviews. Only notes were taken during these conversations and after each interaction the research team met to debrief and record detailed notes. The researchers invited participants from workshops 1 and 2 to contact them through the telecom with any questions or if they wanted to share additional opinions, but none reached out.

Leadership Meetings

The first meeting took place between workshop 2 and workshop 3, and was attended by the mayor, two other government representatives, the cooperative president and treasurer, and the researchers. All non-researchers were male, and the meeting lasted for about half an hour. The researchers reviewed the policies gathered from the first two workshops to gauge their acceptability to these key stakeholders, and sought input from the leaders about what information they would want from the network to evaluate its utility to the community.

The second meeting took place in the afternoon before workshop 3 to include representatives from Rhizomatica who had come to facilitate the workshop. The mayor, two other government

representatives, the cooperative president and treasurer, the cooperative network administrator, two residents, and two representatives from Rhizomatica were present. All non-researcher participants except the administrator were male, and the meeting lasted for approximately one hour. This meeting focused on setting boundaries for metadata collection, planning for future Internet health workshops hosted by Rhizomatica, and finalizing commitments between all parties for how results would be shared and next steps decided since the current study was coming to a close.

Impromptu Interviews

The field researchers interviewed one resident, a middle-aged woman and mother of three, who had befriended the researchers but was unable to attend the workshops. The interview focused on how she and her family would like to use the network, her opinions on the proposed policies from workshops 1 and 2, and her experience as a user of the current GSM network. The interview lasted approximately 40 minutes.

The field researchers also interviewed the telecom network administrator, a middle-aged woman. The interview focused on her experiences running the GSM network, and the processes and tools she uses for network management. The research team presented some of the metadata that could be gathered and used for administration of the LTE network, and sought her feedback on what information she felt would be most useful. The interview lasted approximately one hour.

6.4.3 Analysis

Transcripts were generated from the audio recordings of workshops 1 and 2, but the audio from workshop 3 was unfortunately unintelligible due to a microphone misconfiguration. My co-author and I conducted thematic analysis on the transcripts and field notes, ultimately generating 103 codes grouped into relevant high-level themes. We then examined the themes according to how they inform implementing Ostrom's CPR governance principles in this context. All participant

responses in the paper are translated to English by the research team with names anonymized. Some responses from workshop 3 are paraphrased from field notes instead of quoted, indicated by *italic text* with no quotation marks. We shared a draft of this paper with Rhizomatica, who agreed it is a faithful representation of the field outcomes. We also shared a translated summary with the community via Rhizomatica, but have not received direct feedback.

6.4.4 Limitations

This study only includes one site, with a somewhat small sample size (33 unique participants) skewed towards men. Additionally, only a fraction of participants had direct experience with the LTE network at the time the study was conducted, and their views could change with experience and once possibilities are no longer hypothetical. The workshops were scheduled and advertised by the Santa Inés telecom authorities, and power dynamics and politics unobserved by the researchers and Rhizomatica could have impacted who chose to attend. There is also likely some bias introduced by Rhizomatica's participation in the workshop execution, since their existing relationship with Santa Inés would influence which topics participants chose to broach and elaborate.

Transferablity & Reproducibility

The confluence of experience with communal infrastructure management (see 6.3.3) and ongoing engagement in community-based telecommunications (see 6.3.2) distinguishes Santa Inés as a research partner. Directly reproducing this study in another context will likely require additional capacity building prior to conducting workshops, but we believe that most insights from this work are not tightly coupled to Santa Inés and could be applied to designs for other community-area networks. Nevertheless, many of the insights in this study are particular to mid-sized "community-area" networks like the one in Santa Inés, which fall technologically and organizationally between wide-area networks (commonly operated by a third party provider) and local-area networks (com-

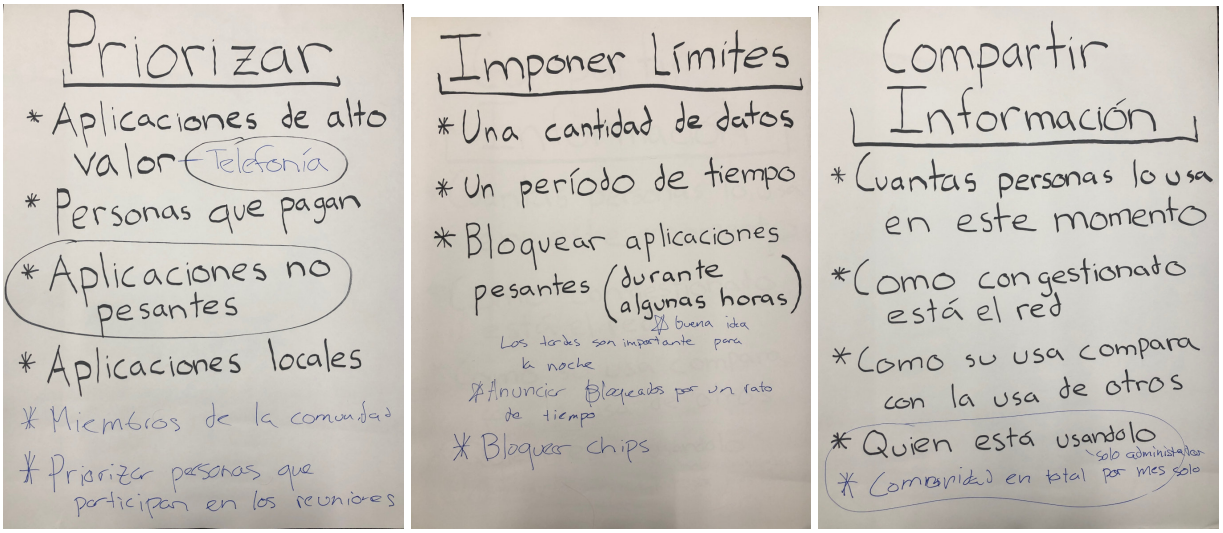


Figure 6.3: Policy ideas generated and evaluated by participants in workshop 2. The top three policies are circled.

monly operated privately). Ostrom's principles may not be an appropriate model for understanding the dynamics of these distinct network structures, which have different sets of stakeholders and power dynamics.

6.5 Findings

In this section we highlight common themes to inform the design of future community-based congestion management tools in Santa Inés and the values embedded within them.

6.5.1 Tradeoff between Individual Privacy and Collective Awareness

As in any network or platform, community networks' generated metadata is both a powerful tool for understanding the network and a liability to individual privacy. Participants understood this tradeoff and were willing to explore compromises between these two extremes.

Value of individual privacy

Many participants felt it was important that their individual usage not be exposed or recorded. There was some confusion about how much information could be seen by the network operator, but even when it was clear that the operator could not read individual messages due to encryption, these participants preferred that their individual app usage history not be recorded [1]. One participant offered (paraphrased) *I am a taxi driver, and some of my competitors live in this town. I don't just want everyone to be able to see how I am using my phone, what sites I visit, what messages I send. Maybe someone could use this information to steal my business secrets or my customers!*

Participants were willing to allow an administrator to see the aggregated amount of data they use as long as it did not show specific sites and apps. This would allow identifying network hogs, but still provide some privacy. This arrangement is analogous to the policy used in Santa Inés's 2G billing system, where the administrator sees the total number of SMS and calls per user, but not the full call data record. In the LTE network people did not want any per-person statistics to be stored for longer than a month or two. The length of the telecom membership cycle is one month, so a short holding policy would allow investigation of any issues without preserving data unnecessarily.

Importance of collective awareness

The community in Santa Inés has an interest in understanding the long-term benefits or detriments of the network, since they are ultimately responsible for its operation. Yet the need for informative longitudinal data runs counter to the value placed on individual privacy discussed above. In traditional ISP networks, the disjoint nature of the user and the operator hides this tension, but it is readily apparent in the context of the Santa Inés network. We discuss this further in section 6.6.2.

[1] We believe it is worth noting that when asked directly, participants express a desire for individual privacy, but that most participants are using mainstream Internet applications which conduct substantial background tracking and hidden data collection [199].

Participants were open to gathering aggregate statistics about usage of specific sites for the purposes of making decisions about the network, as long as they were for the whole community and not individuals. Participants suggested that this data could even be useful in developing workshops customized for Santa Inés to teach the tradeoffs of applications that see wide use.

6.5.2 Embodiment of Local Values

Since the network is owned and operated by the community, local values inform a locally appropriate definition of network fairness.

Allocate resources for each user

Participants discussed models for fairly allocating bandwidth in all three workshops. There was general agreement that all else being equal, instantaneous bandwidth should be divided per-user, not per-application as occurs naturally in IP networks. A user with a heavy application should not take up more of the link than other users with light-weight ones. In one representative exchange between a facilitator and a participant: "*How would you want to share the connection? If there are 200 people who want to use the half megabyte, should it be all equal? / Yes, everyone equal. / Per person, or per application? / It should be per person... yes, per person.*" While dividing bandwidth per user is an intuitive high-level concept, implementing it requires explicit traffic shaping in the network, and the community and Rhizomatica were previously unaware of these requirements.

Tradeoff between fairness and waste

Participants were more divided on the best way to define a fair distribution of data on longer timescales. Some advocated for daily or weekly data caps as is done in many commercial networks, since this is easily measurable, familiar, and treats everyone's data equally. Other users argued this would not be fair, since it would penalize those who want to use the network heavily when it

is otherwise lightly loaded, leading to wasted resources. One participant in particular strongly advocated against data caps. Paraphrasing her argument: *it's not fair to have a daily quota if I am only in town on the weekend when there are not so many people. I hardly use the network at all during the week, and would lose that data. Even with a weekly quota, it's not fair. If the road is empty, why should someone have to pay extra to use it just because they already used their quota? It's a waste.* Others suggested that after exceeding the cap users should not be disconnected, but rather just deprioritized. Even this was controversial though, since responsible heavy use in the morning when the network is lightly loaded could make it harder to accomplish everyday tasks later in the day. In general, there was no clear consensus on whether it is fair to have someone's use (or lack of use) prior to the current moment impact the instantaneous allocation of bandwidth.

Opposition to paid priority

Some participants were in favor of allowing people to pay for extra capacity on the constrained backhaul link as long as this capacity did not come from the fixed capacity shared by the entire community. The distinction between whether a user is paying for *extra* backhaul or whether they are *cutting in line* on the shared communally purchased resource is subtle but extremely important. This point was debated extensively during workshop 3 due to a misunderstanding. Once participants realized that the facilitator was trying to propose that "paid priority" simply meant that the payer's traffic went first on the existing link, all participants immediately united to oppose the policy. One explained (paraphrased): *Just because someone has money, that doesn't mean that they can get special treatment or are more important than anyone else. This proposal goes against our values as a community.* An arrangement where additional funds went towards additional resources without taking away from the common allocation would be acceptable to some, but would require a new arrangement with the backhaul provider (who currently only offers a fixed rate service).

Concern for existing WiFi providers

Several participants expressed concerns about how a community data network would impact existing local businesses selling WiFi or computer access. The workshop attendees did not come to a conclusion about the best course of action, but did want to consider the impacts of any network on their fellow neighbors' businesses as the details of the network are finalized. One participant noted, *"It seems to me that the people who sell Internet, like Luis and Jose, are not going to stay in business. It won't work because people are going to prefer to pay monthly [with the community network], but currently have to pay per package (one hour session)."*

Some options discussed were explicitly restricting the community IP network to critical traffic to allow businesses to serve users seeking entertainment, or limiting the public network throughput to only a basic level. Limits on the public network would leave the businesses to offer higher speeds for a better browsing experience to those who need it. Another suggestion offered by Rhizomatica would be to involve these businesses in the operations and maintenance of the community network, since they already have some IP networking experience and could offer customer support. Participants did not come to a conclusion since details about the costs and capabilities of the community network after the trial have not been finalized.

Capacity building and preparedness

At several points during the workshops and in the meetings with leadership, the community expressed a desire for additional workshops and resources to better understand the network and how to be safe online. User education can be an important tool for long-term network health, and integration of information from the network with education could open up opportunities to better achieve the community's goals than could be done with network management alone. One resident elaborated, *"You know, the best thing you could do in the workshops, for example, would be to explain each app, if there are consequences to the application. For example, Youtube, gives this benefit, brings*

this cost, and Facebook, everything... we want to know really what each application brings to us."

Participants suggested that locally accessible videos of any workshops or other educational materials would be useful. One community member proposed requiring all users complete some form of training before receiving full access to the network. This training could give users both an understanding of how their use impacts others in the community, and also resources to protect themselves and their families on the Internet.

6.5.3 Use Cases and Application Prioritization

The goals of the Santa Inés network's users shape the design of an appropriate network management approach. Participants identified calling or chatting with friends and family, listening to music, watching videos, using social networks, and searching for information as common tasks. Video streaming services like Youtube were mentioned as both a source of entertainment and educational content. Wikipedia was also commonly cited as an example of a useful information service, and one participant explicitly mentioned searching for concepts to help her son with his homework.

Importance vs. priority

Entertainment and media were clearly important use cases in the community; one participant noted (paraphrased): *a network that didn't support videos would be ineffective.* Even so, participants were enthusiastic about de-prioritizing bandwidth-hogging media applications if doing so would improve performance for high-value, if rarer, traffic such as emergency calls. While opposed to paid priority (see 6.5.2), participants were open to applying task-based prioritization equally across all users. Participants who discussed both prioritization and application blocking tended to prefer prioritization as a more flexible approach to traffic balancing. As a concrete example, in workshop 2 one participant proposed blocking video streaming during busy hours. Others responded that

only de-prioritizing those applications would be better, since if the network were full they would be blocked, but if there was extra capacity it could be used and not go to waste.

Supporting person-to-person communication

Above all other use cases, participants consistently stressed that the network must provide reliable telecommunications for the community– that voice calls and messages are the most important function of the network. Calls and messaging allow community members to maintain social connections with far away family members as well as accomplish practical tasks like calling a taxi or ordering supplies from town. One participant put it plainly: *"What is the point of giving data to the whole community if we can't make a long distance call? Understand, most people just want to communicate with their family and friends."* When asked to select their top three desired features in the network, participants in workshop 2 agreed that prioritization of telephony was the single most important feature, and there was unanimous agreement in workshop 3 that telephony prioritization would be a good general policy.

Small applications

Most participants thought it would be a good idea to give preferential treatment to lightweight traffic that places a low burden on the network, especially after discussing the relative traffic magnitudes of different Internet activities. Participants grasped that prioritizing lightweight traffic makes it easier to interactively search and browse text information, while only slightly delaying heavy tasks like media streaming or file downloading. One facilitator was explaining, *"You can decide which applications are limited and which get more data..."* and a participant interrupted *"Ah, so for example Youtube could be last? ... This would be a good idea."*

Supporting education

Several participants raised the point that students often receive homework requiring the Internet for research or to watch an educational video, and that supporting education is one of the primary goals of the network. According to one, *"And here we need the [general] Internet connection above all, for the students."* A parent recounted that she and her school-aged son live in an outlying rancho, and they sometimes have trouble completing his homework if it requires the Internet. Either she has to wait in the town center after school while he completes the assignment, or if no computers are available, attempt to return later. Returning is a heavy cost in terms of both time and resources (in the form of gasoline with the family car or a taxi fare), and sometimes she tells him that he just won't be able to finish his homework that day.

Implementability

Not all proposed ideas are easily or sustainably implementable, and coming up with feasible approaches that can approximate or functionally replace other solutions is an important part of real-world system development. While control of the network is powerful, it is also limited by features of the modern Internet (https + encryption) that restrict the information available at the network layer. After learning about these restrictions, participants identified that it would be difficult to prioritize an important task (such as doing doing homework) when all the network can see is which addresses and sites are being accessed. One of the participants responded to a proposed deprioritization of Youtube: *"But then, let's say that we have done this and are watching Youtube, how slow is it? Can you? Now you can't. If one might want to look for information, then, you have to wait."*

6.5.4 Design for CPR Governance

In this section we synthesize our findings to identify design opportunities for supporting network congestion management compatible with the values and ideas expressed during the workshops while satisfying Ostrom's Principles for CPR Governance (see Table 6.3). Successful governance will enable the citizens of Santa Inés to control their own network, applying the resources of the network where they see them as most beneficial.

Santa Inés already has established practices and technology for maintaining boundaries (1), time-tested collective choice arrangements (3), and a recognized right to organize (7). The network is independent and small enough that nested enterprises (8) are not yet required. To support the remaining principles in a value-compatible manner, we see a need for new technical systems in the community network context enabling the creation and testing of network traffic policies, ethical data collection and monitoring, flexible sanction application, and efficient conflict resolution.

Congruence Between Appropriation and Provision Rules and Local Conditions (2)

Ostrom's second principle addresses the ability to craft and adapt rules appropriate for changing local conditions. Participants proposed a wide range of theoretical policies, from priority access for attending training, to application bans, to quiet hours with limited speeds, among many others. However, with today's tools, it would be difficult for non-expert users to craft, experiment, and deploy arbitrary management policies. There is an opportunity to build functionality into the management system to support users in crafting new policies and testing their effects in a controlled setting. There is also an opportunity to support ongoing efforts in training, education, and capacity building by making network operation and status more transparent.

Data about collective use of the network will be needed to inform rule changes over time, but its collection should balance respect for individual privacy. Community leaders expressed a desire for operational data, but models for collecting useful longitudinal data for non-expert users, while

maintaining individual privacy, agency, and data sovereignty, remain to be developed. Specifically, any deployed analysis tools must avoid storing detailed individual metadata traces, and will need to be flexible to support current use cases without precluding future policy evolution.

Supporting iterative, responsive decision-making aligns well with Ostrom's principles, but presents challenges when developing software. Current design best practices streamline information flow to preserve attention, but this assumes the designers can know what information is relevant, and often comes at the expense of adaptability to future unforeseen use cases. When discussing plans for the network, local leaders were hesitant to commit to a plan, preferring a flexible ad-hoc approach. The cooperative president remarked: *"We don't get even a megabyte with what we have right now, in the 4G [network], but it's the Internet... I don't know who it matters to, but if you all want to add a megabyte, it will be faster, I'll tell them [the provider] to add more Internet."* It will be important to build tools with flexibility in mind that communicate ground truth data in a way that could be applicable to a wide variety of policies and questions years down the road.

Monitoring (4)

Monitoring resource appropriation is critical for preventing or penalizing excessive use. Participants desired a monitoring framework that could identify persistently heavy users while maintaining appropriate privacy safeguards, and supported establishing a position similar to the network admin in the 2G network. The admin would have privileged access to the monitor, and would rotate with the changing government duties in the town. Certain aspects of monitoring in the Santa Inés community network could be simpler than in a natural CPR (such as a fishery or forest), due to the ability to directly measure network traffic. Like the statistical data collection tool discussed in the previous section, a network monitoring tool will need to thoughtfully to preserve end-user values towards privacy and autonomy while still providing enough data to enforce flexible community rules.

Graduated Sanctions (5) & Dispute Resolution (6)

Rather than attempting to encode strict rules in the network software, participants suggested deferring decisions when possible to a human administrator, as is done in the existing 2G network. The human in the loop can reach out via SMS to inform users if they are hogging or abusing the network, and apply rate limits or disable access only if they are non-responsive and continue bad behavior. An admin tasked with monitoring conditions during peak hours can identify problems, gather information both in-band and out-of-band, and implement reasonable, timely solutions taking into account all available context. Extending their toolkit for monitoring LTE network utilization and placing rate limits or bans on bad actors would be relatively straightforward, and allow more flexibility in enforcement than hard-coded rules.

There is also a design opportunity to facilitate communication and dispute resolution between the admin and users. On the management side, any tools should provide a means for the admin to save some kind of record of the behavior in question to consult when mediating a dispute. On the user side, network management tools should provide status to end users, allowing them to see if they are being sanctioned and why. A record should be kept of sanctions applied, so end users can appeal to higher authorities if they are being treated unfairly. While the admin is available in a physical office in the town center, sanctions should not prevent the user from using their phone to contact the admin or make an emergency call, since they may not be able to easily travel to the town center to otherwise discuss the problem.

6.6 Discussion

6.6.1 Net Neutrality versus Network Ownership

Many of the congestion control/avoidance strategies discovered and discussed in the workshop run afoul of the principle of net neutrality: that network providers should not prioritize or block

traffic based on the destination, application, or service [52]. For example, workshop contributors clearly indicated a preference for voice and messaging over other traffic. We believe these results speak strongly to the limitations of net neutrality policies in backhaul-limited and congested rural access networks. In these networks, there are both critically important services and limited capacity to handle all traffic. This combination requires network providers to have prioritization schemes (or congestion management) to resolve this conflict.

We aim to move away from simplistic notions of net neutrality, primarily a policy designed to keep powerful ISPs from engaging in business practices that harm customers, to *network ownership*, which is when users can decide network policies based on their needs. In this case, assuming ongoing community participation and ownership, traffic prioritization can be done in service of improved network performance and user experience. With network ownership, we believe that non-net neutral policies can be explored in more depth, especially mechanisms to ensure they remain in control of, and support, their users.

6.6.2 Privacy in Community Networks

Privacy is an important counterpoint to the above discussions on mediated access and community control. A strange advantage of the traditional ISP model is that service is provided by a large corporation which has (or should have) little interest in your daily life. With tight bonds between members of small communities, other users may have an unhealthy personal interest in your use (or non-use) of the network.

This will be a key challenge going forward, as the tension between providing a rich dataset for the community to use for decision making is inherently at odds with anonymous Internet access. For example, a power user may be downloading too many movies via YouTube and need to be throttled. While inherently innocuous, this may cause the organizing committee to ask what videos are being seen and, even though the content in anonymized, they could continue to limit

Internet access until that information is received.

This issue has come up before in our partner's GSM (2G) community networks. In a separate network, the community asked for the ability to gather call logs from the access point to monitor youth and shared accounts. This was not implemented by the partner, who was concerned it would create an incentive and means to monitor romantic relationships, specifically romantic partners.

6.6.3 Differing Scales of Congestion Control

Network congestion can happen at a number of time scales. It can happen at the millisecond scale when two competing services require more network capacity than is present. It can happen in the span of hours as multiple users, with services that require reasonable amounts of data, all sum to more than the network can handle. It can also happen over the course of a week, as the backhaul link temporarily fluctuates in capacity due to network failures outside of the community.

Each scale may require different congestion control mechanisms with differing levels of automation. The rapid microsecond congestion control remains the domain of traditional congestion control protocols such as TCP. At longer human-appropriate time scales, it may make more sense to allow operators to selectively disable or enable classes of service or switch entirely from one automated prioritization regime to another. We believe building usable network administration tools that embrace the time-varying nature of rural edge networks, at human time scales, could be an interesting new direction for research.

While participants proposed many automated mechanisms, such as the prioritization of certain content, there was also a repeated suggestion to designate a person "in charge" who could serve as a point of contact and leverage outside context and human sensibility in management decisions. Mediated access has a long history of research in ICTD [154], but in this case maps closely to governance structures already present in the community. In Santa Inés, individuals are given

control over key community resources with the expectation that they are trained to maintain and support the infrastructure until their time has passed. Santa Inés uses a similar strategy to manage water resources, where trusted members of a water control committee have the authority to disperse allocations from a communal reservoir and responsibility to manage overall consumption to avoid a shortage. As such, developing congestion management mechanisms with a "human-in-the-loop," but that are compatible with individual Internet privacy and security expectations, is an exciting area of future research. We anticipate developing novel systems that both provide "knobs" for operator control as well as mechanisms for users to observe and interact with the operator's choices, such as network status and views into currently applied rules.

6.6.4 Technologies for Community-area Networks

Supporting infrastructure, both technical and non-technical, for community networks falls into an under-explored middle-ground between wide-area networks (e.g., regional ISPs), and local-area networks (e.g., home WiFi). In wide-area networks, powerful tools cater to professional network operators [157], providing advanced monitoring, filtering, and shaping capabilities. Yet they fall short in the community setting because they have a high usability threshold, gather and make visible too much data about individual users (see section 6.6.2), and assume a highly asymmetric power relationship between the operator and the end-user. They also lack affordances for transparent sanctions or dispute resolution like user-accessible sanction logs or a privacy-respecting means to store evidence of bad-behavior. In local-area networks, home-targeted tools have a lower usability threshold, but often lack facilities for flexible traffic shaping and don't integrate with commercial-grade cellular equipment (via DIAMETER and the Gx interface) to enforce radio resource limits or identity primitives like SIM-card authentication. Local-area tools also gather and make visible too much data about individuals, assuming all users are part of the same trust circle (whether it be a family or small business), and do not consider auditing and

dispute resolution in their designs.

Building an effective community-area network management tool simultaneously compatible with Ostrom's Principles and the values expressed by workshop participants in Santa Inés will require re-evaluating design assumptions around the relationship between the network operator and end-users. Such a tool would require a mix of affordances from today's wide-area and local-area tools, while adding capabilities to support dispute resolution and balance end-user privacy with collective awareness of network operation.

Collateral Damage from the Surveillance/Privacy Arms Race

During the workshops, participants commonly desired high-level task-based policies, like prioritizing realtime voice calls, education, and messaging applications while deprioritizing entertainment videos. In practice though, Transport Layer Security (used by HTTPS) masks task information from the network, severely restricting what can be made visible for the sake of congestion management. The tension between *privacy of content*, and *privacy of general task*, is an important tradeoff in the community-area network setting rarely considered in wider Internet privacy discourse.

Quality of Service (commonly referred to as QoS) tags theoretically allow applications to signal how packets should be handled by the network, and can differentiate realtime communication traffic from background bulk transfers, but are often not implemented and subject to abuse. In a sample of three end-user devices in the Santa Inés LTE network, none consistently tagged packets from popular communication applications WhatsApp and Signal. Heuristics or AI-based classifiers on flow characteristics like protocol, address, packet size, and frequency can be developed for some tasks, but would be subject to errors which could exclude traffic the community would ideally want to prioritize and visa-versa.

Ultimately, reconsidering how much information is revealed to the network and fixing QoS tagging would require long-term changes across the device and application stack, but would open

up new classes of enforceable management policies. Such changes do not preclude development of community management tools now, as long as such tools are flexible enough to make QoS, or some other type of task information, visible if it were to provide a meaningful signal.

6.7 Future Work

This study is part of a long term collaboration between the researchers, Rhizomatica, and Santa Inés. While we have identified several directions for future work, we plan to focus immediately on the co-design and development of appropriate tools to support an administrator in managing a small to medium sized community LTE network. In the long run I think there is much potential to explore additional social mechanisms to improve the collective management of community IP networks, inspired by Ostrom's principles for common resource management and social-based conservation efforts.

6.8 Conclusion

In this work we conducted initial design explorations of community-based systems for congestion management in a rural community LTE network. Through a series of participatory workshops, interviews, and discussions, we gathered a wide range of policy proposals and feedback on their feasibility and appropriateness to Santa Inés. Among participants, we found a desire to non-neutrally prioritize person-to-person communications, an aversion to pricing-based mechanisms for allocating network resources, support for human-in-the-loop local management tools, and an aspiration to balance respect for individual privacy with informed network governance. We see an opportunity to apply collective-action approaches, informed by Elinor Ostrom's Principles for CPR Governance, effectively in the community cellular network context for a more humanistic and flexible approach to community network governance than traditional pricing and automated congestion control.

Table 6.3: This table details how Ostrom's principles can be operationalized for management of the Santa Inés community LTE network as a CPR. Asterisk entries (*) mark principles with a technological component, and **bold** entries indicate areas where the researchers hope to offer support through the participatory design process. We identify design opportunities to support rule creation (2) and enforcement mechanisms (5) in the network, as well as tradeoffs in providing a local admin with data and visibility into the network to facilitate monitoring (4) and conflict resolution (6) while respecting community privacy values.

CPR Design Principle	**Santa Inés LTE Community Network as CPR**
1. Clearly defined boundaries*	Potential users can be clearly delineated via possession of a network-specific SIM card required for access, and also blocked based on SIM card identity.
2. Congruence between appropriation and provision rules and local conditions*	Through participatory workshops and local governance structures, the community could craft rules such as cost structures and usage policies based on their needs and values. The rules could be updated as conditions change.
3. Collective-choice arrangements	Santa Inés has established governance structures for the telephony network, and users can modify management rules through collective action via the local government.
4. Monitoring*	A network admin, employed by and a member of Santa Inés, could observe others' usage through a management portal in a manner consistent with local privacy values.
5. Graduated sanctions*	Rules for sanctions, such as temporary blocking or de-prioritization, can be decided by the community along with other policies. A network admin can apply a range of sanctions on a case-by-case basis.
6. Conflict-resolution mechanisms*	The network admin can be consulted for minor conflict resolution, or serious issues can be escalated through existing community governance structures.
7. Minimal recognition of rights to organize	Independently managed community cellular networks in rural Mexico have been granted special permission to operate by the national government.
8. Nested enterprises	The community network operates independently and is small enough to be approachable.

Chapter 7

Concluding thoughts

Throughout my research and this dissertation I have been able to work with a wide range of collaborators who have taught me how to bring an equally wide range of different methodologies to the problem space. My work has varied from traditional data science, to systems development, to human-centric design, but all focused on building a holistic understanding of the role IP-based community cellular networks can play in the broader connectivity ecosystem. Community networking and ICTD in general is such a vibrant field partly because of this broad scope for relevant work and the value the field places on a wide range of methodologies. I would not have wanted to work any other way.

In my experience, building sustainable and appropriate connectivity solutions is not something that can be done in a one-size-fits-all manner. The rich history of ICTD and HCI in diving into how technology is used, adapted, tolerated, or opposed by all the humans around a particular technology has helped guide my work and shaped the way that I understand the problem. Much remains to be

done to build a diverse set of robust and performant connectivity solutions for everyone, especially as the ecosystem to continues to evolve and develop over the next decade(s). In particular, there is a large gap in both product and spectrum availability for small vs. large operators. Many systems and processes in connectivity have ossified to the point that starting a new service is a bureaucratic and regulatory challenge that doesn't scale down to small organizations.

In writing this conclusion I spent a fair amount of time looking at my early writing and reflecting on how my understanding of the problem space has changed over time. Unlike the prior sections, this chapter does not attempt to make a rigorous argument, but rather documents my observations and reflections throughout the process, hopefully to the benefit of those reading this document and continuing this work.

7.1 Threads

7.1.1 Distinctly Targeting the "Smedium" Scale

A common thread through my work has been the intentional embrace of the limited scale of small community networks at the protocol and technology level. By explicitly giving up on large scale and a one-size-fits-all objective, my work allows challenging the assumptions that individual nodes cannot come together to make cooperative agreements, and that all issues of fairness must be resolved mechanically. At the same time, my work explicitly targets operation between "economic entities," who are not part of the same family, business, or organization, and where complete transparency is not appropriate. This middle scale manifests a blend of the challenges, opportunities, requirements, and responsibilities of small and large networks, and I think there is an opportunity for interesting research leading to both novel organizational structures and technologies that could make a real impact on how medium-sized networks are used, managed, and understood.

7.1.2 Humans and networks

In the follow-on to my work in Oaxaca, I was planning to focus on the co-design and development of appropriate tools to support an administrator in managing a small to medium sized community LTE network. Partly as a result of covid and partly due to the general difficulty of field work, I didn't get a chance to finish this research. In the long run, I think there is lots of room to explore additional social mechanisms to improve the collective management of community IP networks. Beyond just community networks, increasing densification of networks leading to tough questions about network ownership and responsibility in general. Tools that help community networks operate could also be leveraged to apply to "neutral host" operators and federations of independent access points too. There is a lot of quality work in other fields analyzing common resource management and social-based conservation efforts to build on and adapt to the computer networking domain.

I think it would also be interesting to extend this work to investigate network ownership and user education. "How can systems give more control over the Internet back to the communities they serve?" is a difficult and nuanced question. To begin with, should systems do such a thing, or is it important that the Internet strives for a more individual-centric notion of neutrality? As every community is a collection of individuals, how should the network's affordances help or hinder the community in establishing and enforcing local norms? These types of discussions are not insulated from the question of power, and implementers will need to consider who will be able to use these enabling technologies to enforce their view of what is appropriate in the network. Coming to a single answer is likely impossible, but it does seem like there could be promising directions here for maintaining a community's shared culture and values while still providing additional utility through Internet connectivity.

All in balance: One thing I did take away from my field work though is an appreciation for how precious attention is, and the importance of letting the network get out of the user's way when it can. Given the possibility for many different failure modes in edge networks, applications for measurement and network understanding for end users should be able to provide a more useful status than "bars" to aid users in fixing or adapting to the problem at hand, without being overwhelming during normal operation. Such tools might also help users grok the increasing complexity of networks in the 5G era outside rural access.

There is an important tradeoff between giving enough information that users feel like they understand the network and would feel comfortable fixing or modifying it, while being considerate to the fact that most people are not networks researchers, don't love networks, and just want to use the network to accomplish their other goals. Robust research into this tradeoff seems like it would be challenging to conduct, but could have significant impacts on the design of user-facing networked systems going forward.

7.2 Challenges

While I believe that CCNs have an important role to play in the portfolio of connectivity options needed to solve practical access challenges, they are not on their own a perfect, or even sufficient, standalone solution. Other important challenges outside my work remain that will need to be solved for CCNs to be effective.

Backhaul

Most immediately, CCNs only address the "radio access" part of the connectivity puzzle. The backhaul connection from the community back to the wider Internet is taken as a given in CCN research, and it is up to each community to acquire a sufficient and economically sustainable backhaul connection. While many communities are able to creatively piece together backhaul,

many others still are not, and CCNs are silent on addressing the technical challenges of establishing a backhaul connection in remote or challenging areas. Low-earth-orbit satellite networks and other aerial platforms may help to address this challenge from a technical perspective for remote areas, but other technically and operationally creative solutions could be explored. In urban areas where the challenges often stem from empowerment and economics rather than physical access, partnerships with municipal fiber or wireless point-to-point or mesh networks could yield interesting possibilities for cross-layer optimization and customization of the network's characteristics to provide quality cellular access.

Operations

Everyone I have met in ICTD maintains "field work is hard," and my experience has been no different. Supporting operational networks as a researcher is a big challenge on its own. Supporting operational networks while maintaining work-life balance and research progress is a monumental one. Much of my learning through grad school is not reflected directly in this thesis, since it's all operational. I now know how to manage python dependencies, build and deploy debian packages securely through the repository infrastructure, update keys, ship software updates across 256K satellite links, roll back software updates over 256K satellite links, and back up three different kinds of database. Leaders and long-term participants in the field must continue to push for new and creative ways to maintain projects and support institutional capacity and knowledge transfer for collaborating with partner networks in the long term, beyond the tenure of a single student.

Device ecosystem and costs

There is a substantial capability gap between the traditional telecommunications equipment vendors and the smaller vendors willing to sell to community network organizations. These vendors' equipment offers lower performance equipment in terms of peak bandwidth, power

efficiency, scheduler performance, and radio sensitivity than the equipment available from the major RAN vendors. While the traditional vendors offer high performance small-cell hardware that would be technologically appropriate for community networks, they do not have sales and support structures that allow small operators to purchase and operate this equipment. There is no way to buy the equipment at a listed price in small quantities, and no way to get basic support (like firmware upgrades) without an expensive negotiated support contract. Furthermore, while the hardware is capable, these current carrier systems are designed to only integrate with large-scale customized Operations and Management (O&M) software suites developed for region-scale networks with many nodes. While theoretically possible to scale them down, these systems are onerous to license and deploy, and are often very expensive.

Further complicating the issue is that while there are projects in progress to better standardize RAN components for modularity and lower costs through competition [173] and new deployment models [5], there is not yet a robust, modular, or open-source RAN ecosystem. This means there's no open source baseline to bound costs, and no foundation to allow smaller players to compete and push the market towards customer focused products. Compared to the more diverse WiFi device ecosystem, cellular is improving but still relatively unhealthy. Devices are full of attempts at vendor lock-in via incompatible software and feature licenses, there is no stable open source firmware, and minimal transparency in pricing and features makes it hard to even understand what is available as a small player.

7.3 Towards Future Networks

The late 2010s and early 2020s have seen substantial developments in the wider connectivity ecosystem. Several companies (OneWeb, Amazon, and SpaceX, among others) are actively working to send thousands of satellites into low-earth orbit (LEO) with the promise of high-quality, low-latency global connectivity. As noted in Kleanse et al. [104] "Constellations of hundreds to

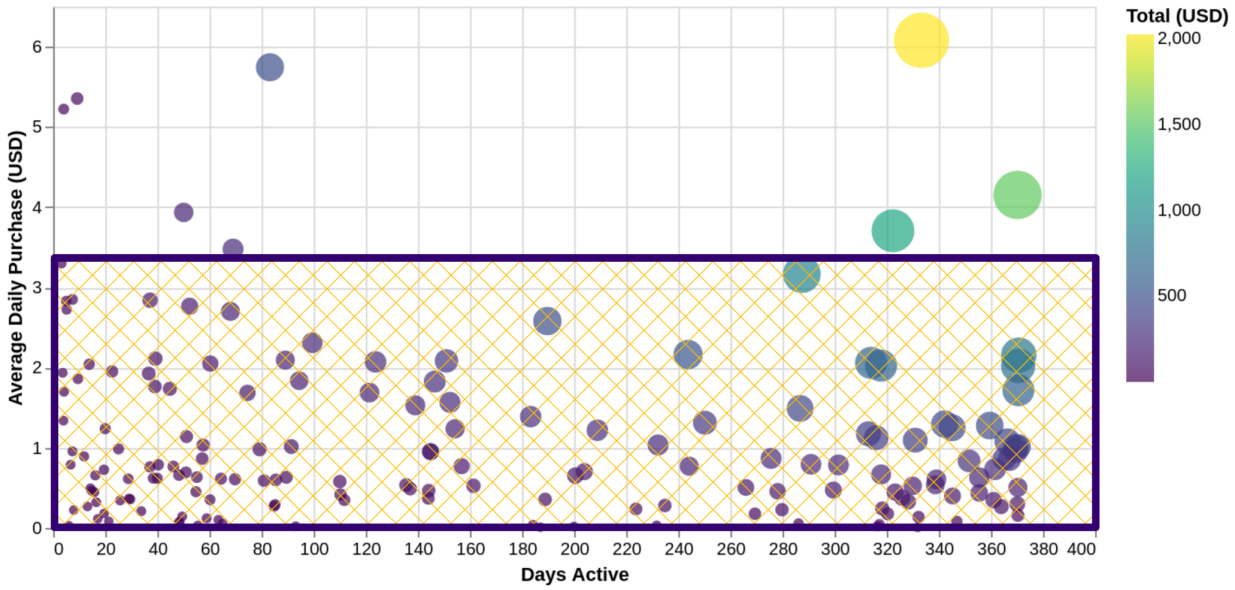


Figure 7.1: Bokondini Network Affordability. In this figure the total purchased amount is encoded in the size and color of each point. The x-axis indicates how many days a user was active in the Bokondini network, and the y-axis indicates the average amount of credit spent by the user per day when they were active. The yellow hashed area indicates accounts whose consumption would not directly justify a standalone satellite subscription. While consumer preferences and spending power naturally change over time, as a rough first-order estimate, very few of the accounts in Bokondini spent enough on network services to afford their own next-generation satellite terminal and subscription.

thousands of satellites promise to offer low-latency Internet to even the most remote areas." In other domains, the vision of ubiquitous machine to machine connectivity via the Internet of Things and low-power wireless is also coming to fruition [77, 72]. The development of the IoT ecosystem is bringing with it extremely efficient waveforms and low power [188], or even completely batteryless [89, 4], devices with new use cases for rural networks [185]. While the ICTD community has lately focused on creative applications of mature technologies, I think there is a timely opportunity to influence the rollout and applications of these new technologies in a way that could benefit from the more holistic perspective of ICTD research.

New Options in Low-band Spectrum

Wide area coverage for rural access benefits from the propagation and penetration characteristcs of (relatively low frequency for broadband applications) UHF and VHF spectrum. Due to its good propagation, this spectrum is highly sought after and relatively scarce. As a consequence of its physical properties, this spectrum is *also* the spectrum being targeted by Internet of Things protocols and the wider IoT device ecosystem. As discussed in Section 2.1 [The Case for Cellular], remote CCNs benefit from their ability to access this spectrum, but come with challenges in global deployment due to the need to negotiate secondary use spectrum access and address regulatory concerns to avoid interference with incumbent mobile network operators. In contrast the IoT device ecosystem, built from the ground up for end-user deployment, offers hardware in the (admittedly very constrained) UHF and VHF *unlicensed* ISM bands.

Combining community cellular in easier to license "mid-bands" for high-speed town-scale coverage with IoT protocols for coverage extension and continuity of access around the periphery of the area could allow for significantly more coverage at a capacity supporting messaging and photo sharing. In the ICTD and community networking space, researchers have already developed systems for mixed cellular/WiFi connectivity before the availability of high-speed

cellular connections [29]. Vigil-Hayes et al. recently built a system exploring a cross-layer model specifically for hybrid IoT and broadband connectivity using LoRaWAN for coverage extension of the mobile web beyond a wireless LAN [191]. There's a lot of opportunity to explore other models for how to expose this type of mixed connectivity to end users and effectively take advantage of the growing IoT device ecosystem for access extension.

New Options From Orbit

Given how much of my work, and community cellular networking work in general, focuses on the challenges of operation behind a backhaul-limited satellite link, these are exciting times. I am cautiously optimistic that the upcoming generation of low earth orbit satellite networks will finally deliver on the promise of service in remote areas with fiber-like latency and orders of magnitude more throughput than existing geostationary VSATs. Moving from a design regime of bandwidth scarcity to bandwidth abundance, on-par with metropolitan connections, invites new applications to remote areas while also bringing new challenges.

For one, formerly connectivity-limited communities and their local cultures will have to grapple with the availability of the media-rich modern Internet, no longer isolated by physical remoteness. This will be a problem *soon*, and is increasingly moving from a hypothetical to a real practical challenge that needs creative and powerful solutions, both technical and social, to address.

In addition to the possible cultural impacts, widely available high-performance satellite networks will directly and profoundly impact the sustainability and operations of community networks. While less expensive and more performant links are a huge advantage in building a remote network, it's also possible that the individual-focused nature of satellite connections could actually have the counter-intuitive effect of eroding access availability by poaching the highest value users from existing networks. Losing these users removes the implicit cross-subsidization of lighter users on existing shared infrastructure. Figure 7.1 is a re-render of the graph from Figure 3.5, with

a dividing area added to distinguish the accounts in Bokondini whose spending would support a personal satellite connection at current advertised prices, and those whose would not. The network would be much more challenging to operate sustainably without the contributions from the "whale" users.

Ubiquitous LEO satellite brings more uncertainty beyond impacts to culture and network economics too. Operating a community network takes work [50]; will an option to buy an individual high quality connection that solves their immediate problems erode the motivation of anchor users to invest time and skills in CCNs? Will LEO networks be able to continue to improve speeds and reliability to keep up with advances in urban networks, or will they soon feel as slow as VSAT feels today? Will rural site economics change such that traditional national operators networks cover the entire world? Will the satellite networks develop technology to directly reach end-user devices *without* a terminal after all [126]? The long-term costs, real-world performance, reliability, and market/regulatory impact of the LEO networks remain to be determined, but their nascent presence is already having an outsized impact on the state of rural and remote access.

A bright, and quickly approaching, future: Despite the looming challenges sparked by any substantial technological development, I think that these new tools are on their way to revolutionizing connectivity and bringing high-quality global access tantalizingly close to in reach. Now is the time to make sure they are built with affordances to enable the type of broader connectivity ecosystem we want to see in the world and leave for our children.

Acknowledgments

No work happens on its own, and this is especially true in the ICTD context when building systems and working with field partners for real-world deployments. My advisor, Kurtis Heimerl, has supported me throughout this journey and trusted me to go explore new ideas while introducing me to his network of potential collaborators, helping me build partnerships, and pushing me to drive quickly to working solutions. As my labmates and multi-project collaborators, Esther Jang, Nick Durand, Sudheesh Singanamalla and Spencer Sevilla have also provided an incredible amount of support and learning through my grad school journey, and have profoundly influenced the way I approach research.

The wider ICTD community and researchers exploring Community Networks both within UW and outside it have also been incredibly supportive and essential in inspiring, critiquing, and informing my work. In alphabetical order and surely missing important folks, Aditya, Anna, Claire, Emmanuel, Esther, Fahad, Firn, Frankie, Galen, Gaetano, Innocent, Lisa, Matt, Melissa, Naveena, Nick, Pat, Philip, Rachel, Richard, Sam, Samia, Shaddi, Spencer, Sudheesh, Trevor, and Waylon have all helped and supported me at different points throughout my degree.

Outside of the research context and the UW ICTD lab, I have also received essential support from my friends, Adrian, Alex, Amanda, Andrew, Ben, Bradley, Calder, Carissa, Daniel, Dave, Eric, Gabe, Geoff, Ivan, Jake, John, Josh, Joshua, Leena, Lindsey, Melody, Nathan, Nicasia, Pascal, Sasha, Spencer, and Zack; my housemates, Esther, Laura, Naveena, Philip, Soubhik, Stefania, Taylor, Tyler, and Veevee; everyone I've played sports with in IM and on the Chili's; long-term friend and mentor Scott; my family, Claire, Mark, and Pat; and my partner Erin.

Project Collaborators

I would like to specifically thank David Haag, Ibel Itlai, Marmulia Siahaan, Scotty Wisley, and the Ob Anggen School for their support in operating and measuring the Ungu network in Bokondini.

I also thank all of the interview and workshop participants in Santa Inés who graciously volunteered their thoughts and time to the community congestion management project, and the leaders of the Santa Inés cooperative telecom for hosting me and their flexibility in facilitating this research. This work could not have happened without Telecomunicaciones Indíginas Comunitarias and Rhizomatica, for their collaborative support and effort into the design and execution of the workshops and the network trial as well.

Claire, Flip, Maria, Ronel, Thessa, and the entire University of the Philippines team supported me in work that while not directly included in this dissertation, was formative in my understanding of the challenges of building and sustainably operating CCNs.

In Seattle, Esther Jang has built and led funding for the Seattle Community Network, which provided the motivation for the dAuth project and supported its evaluation through the lending of test nodes and radio equipment. The Seattle community network would also not be possible without the energy and support of Black Brilliance Research, the Filipino Community Center, Seattle Public Schools, the Skyway Library, Tacoma Community Network, Surge Tacoma, the Oromo Cultural Center, the University of Washington, and all of the volunteers who keep the

network running and healthy.

Funding

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1762114. I have also received support to teach, TA, and work with undergraduate researchers from the Paul G. Allen School, and was able to begin grad school with support from the Gaetano Borriello Endowed Fellowship for Change.

Prior Works

This document draws from the following works I previously published:

[95] Matthew Johnson, Jenny Liang, Michelle X. Lin, Sudheesh Singanamalla, and Kurtis Heimerl. "Whale Watching in Inland Indonesia: Analyzing aSmall, Remote, Internet-Based Community Cellular Network". In: *Proceedingsof the Web Conference 2021*. WWW '21. Ljubljana, Slovenia: ACM, Apr. 2021, p. 12.
DOI: 10.1145/3442381.3449996.

[96] Matthew Johnson, Spencer Sevilla, Esther Jang, and Kurtis Heimerl. "dLTE: Building a More WiFi-like Cellular Network: (Instead of the Other Way Around)". In: *Proceedings of the 17th ACM Workshop on Hot Topics in Networks* (Redmond, WA, USA). HotNets '18. New York, NY, USA: ACM, 2018, pp. 8–14.
DOI: 10.1145/3286062.3286064.

[167] Spencer Sevilla, Matthew Johnson, Pat Kosakanchit, Jenny Liang, and Kurtis Heimerl. "Experiences: Design, Implementation, and Deployment of CoLTE, a Community LTE Solution". In: *The 25th Annual International Conference on Mobile Computing and Networking*. MobiCom '19. Los Cabos, Mexico: Association for Computing Machinery, Oct. 11, 2019, pp. 1–16.
DOI: 10.1145/3300061.3345446.

[182] Nussara Tieanklin, Matthew Johnson, and Kurtis Heimerl. "Poster: The Low Impact of COVID-19 on Rural Community Network Traffic". In: *ACM SIGCAS Conference on Computing and Sustainable Societies*. COMPASS '21. New York, NY, USA: Association for Computing Machinery, June 28, 2021, pp. 417–422.
DOI: 10.1145/3460112.3472311.

Bibliography

[1] 3GPP. *TS-22-011:Service Accessibility (3GPP TS 22.011 Version 16.5.0 Release 16)*. Nov. 2020. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=566.

[2] 3GPP. *TS-33-102: Digital Cellular Telecommunications System (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); 3G Security; Security Architecture*. Aug. 2020. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2262.

[3] 3GPP. *TS-33-501: Security Architecture and Procedures for 5G System*. Aug. 2020. URL: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169.

[4] Ali Abedi, Farzan Dehbashi, Mohammad Hossein Mazaheri, Omid Abari, and Tim Brecht. "WiTAG: Seamless WiFi Backscatter Communication". In: *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*. SIGCOMM '20. New York, NY, USA: Association for Computing Machinery, July 30, 2020, pp. 240–252. DOI: 10.1145/3387514.3405866.

[5] *About FreedomFi*. FreedomFi. URL: https://freedomfi.com/about-freedomfi/ (visited on 08/07/2022).

[6] Giuseppe Aceto, Alessio Botta, Antonio Pescapé, M. Faheem Awan, Tahir Ahmad, and Saad Qaisar. "Analyzing Internet Censorship in Pakistan". In: *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow (RTSI)*. 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow (RTSI). Bologna, Italy: IEEE, Sept. 2016, pp. 1–6. DOI: 10.1109/RTSI.2016.7740626.

[7] Michael Adeyeye and Paul Gardner-Stephen. “The Village Telco Project: A Reliable and Practical Wireless Mesh Telephony Infrastructure”. In: *EURASIP Journal on Wireless Communications and Networking* 2011.1 (Aug. 25, 2011), p. 78.
DOI: 10.1186/1687-1499-2011-78.

[8] Mukhtiar Ahmad, Syed Usman Jafri, Azam Ikram, Wasiq Noor Ahmad Qasmi, Muhammad Ali Nawazish, Zartash Afzal Uzmi, and Zafar Ayyub Qazi. “A Low Latency and Consistent Cellular Control Plane”. In: *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*. SIGCOMM ’20. New York, NY, USA: Association for Computing Machinery, July 30, 2020, pp. 648–661.
DOI: 10.1145/3387514.3406218.

[9] Sohaib Ahmad, Abdul Lateef Haamid, Zafar Ayyub Qazi, Zhenyu Zhou, Theophilus Benson, and Ihsan Ayyub Qazi. “A View from the Other Side: Understanding Mobile Phone Characteristics in the Developing World”. In: *Proceedings of the 2016 Internet Measurement Conference*. IMC ’16. New York, NY, USA: Association for Computing Machinery, Nov. 14, 2016, pp. 319–325.
DOI: 10.1145/2987443.2987470.

[10] Abhinav Anand, Veljko Pejovic, Elizabeth M. Belding, and David L. Johnson. “VillageCell: Cost Effective Cellular Connectivity in Rural Areas”. In: *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*. ICTD ’12. New York, NY, USA: Association for Computing Machinery, Mar. 12, 2012, pp. 180–189.
DOI: 10.1145/2160673.2160698.

[11] Dani Anderson, K. A. Shruthi, David Crawford, and Robert W. Stewart. “Evolving Spectrum Sharing Methods, Standards and Trials”. In: *Spectrum Sharing*. John Wiley & Sons, Ltd, 2020, pp. 59–74.
DOI: 10.1002/9781119551539.ch4.

[12] APC News. *What’s New on the Spectrum? “Let’s Make Sure We Can Use It for What Is Needed”: A Conversation with Peter Bloom from Rhizomatica | Association for Progressive Communications*. Association for Progressive Communications. May 29, 2020.
URL: https://www.apc.org/en/news/whats-new-spectrum-lets-make-sure-we-can-use-it-what-needed-conversation-peter-bloom (visited on 06/29/2022).

[13] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. *DNS Security Introduction and Requirements*. RFC4033. RFC Editor, Mar. 2005, RFC4033.
DOI: 10.17487/rfc4033.

[14] Tayyab Arshad, Muhammad Faheem Awan, Tahir Ahmad, and Saad Qaisar. “Performance Evaluation of Mobile Broadband Cellular Networks in Pakistan”. In: *2016 IEEE 41st Conference on Local Computer Networks Workshops (LCN Workshops)*. 2016 IEEE 41st Conference

on Local Computer Networks Workshops (LCN Workshops). Dubai, United Arab Emirates: IEEE, Nov. 2016, pp. 104–111.
DOI: 10.1109/LCN.2016.035.

[15] *AT&T Moves 5G Mobile Network to Microsoft Cloud*. June 30, 2021.
URL: https://about.att.com/story/2021/att_microsoft_azure.html (visited on 07/10/2022).

[16] Muhammad Faheem Awan, Tahir Ahmad, Saad Qaisar, Nick Feamster, and Srikanth Sundaresan. "Measuring Broadband Access Network Performance in Pakistan: A Comparative Study". In: *2015 IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops)*. 2015 IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops). Clearwater Beach, FL, USA: IEEE, Oct. 2015, pp. 595–602.
DOI: 10.1109/LCNW.2015.7365903.

[17] Ghufran Baig, Dan Alistarh, Thomas Karagiannis, Bozidar Radunovic, Matthew Balkwill, and Lili Qiu. "Towards Unlicensed Cellular Networks in TV White Spaces". In: *Proceedings of the 13th International Conference on Emerging Networking EXperiments and Technologies*. CoNEXT '17. New York, NY, USA: Association for Computing Machinery, Nov. 28, 2017, pp. 2–14.
DOI: 10.1145/3143361.3143367.

[18] Roger Baig, Lluís Dalmau, Ramon Roca, Leandro Navarro, Felix Freitag, and Arjuna Sathiaseelan. "Making Community Networks Economically Sustainable, the Guifi.Net Experience". In: *Proceedings of the 2016 Workshop on Global Access to the Internet for All*. GAIA '16. New York, NY, USA: ACM, 2016, pp. 31–36.
DOI: 10.1145/2940157.2940163.

[19] Roger Baig, Ramon Roca, Felix Freitag, and Leandro Navarro. "Guifi.Net, a Crowdsourced Network Infrastructure Held in Common". In: *Computer Networks*. Crowdsourcing 90 (Oct. 29, 2015), pp. 150–165.
DOI: 10.1016/j.comnet.2015.07.009.

[20] Roger Baig, Ramon Roca, Leandro Navarro, and Felix Freitag. "Guifi.Net: A Network Infrastructure Commons". In: *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ICTD '15. Singapore, Singapore: Association for Computing Machinery, May 15, 2015, pp. 1–4.
DOI: 10.1145/2737856.2737900.

[21] Luca Belli, Bruno de Souza Ramos, Panayotis Antoniadis, Virginie Aubrée, Roger Baig Viñas, Aris Dadoukis, Paolo Dini, Mélanie Dulong de Rosnay, Nicolas Echániz, Kurtis Heimerl, Matthew Johnson, Pathirat Kosakanchit, Florencia López Pezé, Steven Mansour, Stavroula Maglavera, Jens Martignoni, John Mavridis, Sascha Meinrath, Leandro Navarro, Harris Niavis, Ramon Roca i Tió, Spencer Sevilla, and Félix Tréguer. *The Community Network Manual : How to Build the Internet Yourself*. FGV Direito Rio, Nov. 2018.
ISBN: 978-85-959702-9-8.

[22] Gerald Bernbom. “Analyzing the Internet as a Common Pool Resource: The Problem of Network Congestion”. In: Constituting the Commons: Crafting Sustainable Commons in the New Millennium, the Eighth Biennial Conference of the International Association for the Study of Common Property. IASCP’2000. Bloomington, Indiana, USA: International Association for the Study of Common Property, May 31–June 4, 2000, p. 28.
URL: http://hdl.handle.net/10535/1168 (visited on 05/31/2020).

[23] Debopam Bhattacherjee, Waqar Aqeel, Ilker Nadi Bozkurt, Anthony Aguirre, Balakrishnan Chandrasekaran, P. Brighten Godfrey, Gregory Laughlin, Bruce Maggs, and Ankit Singla. “Gearing up for the 21st Century Space Race”. In: *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*. HotNets’18. HotNets ’18. Redmond, WA, USA: Association for Computing Machinery, Nov. 15, 2018, pp. 113–119.
DOI: 10.1145/3286062.3286079.

[24] Nicola J. Bidwell. “Wireless in the Weather-world and Community Networks Made to Last”. In: *Proceedings of the 16th Participatory Design Conference 2020 - Participation(s) Otherwise - Volume 1*. PDC ’20. New York, NY, USA: Association for Computing Machinery, June 15, 2020, pp. 126–136.
DOI: 10.1145/3385010.3385014.

[25] Nicola J. Bidwell. “Women and the Sustainability of Rural Community Networks in the Global South”. In: *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development*. ICTD2020. New York, NY, USA: Association for Computing Machinery, June 17, 2020, pp. 1–13.
DOI: 10.1145/3392561.3394649.

[26] Doreen Bogdan-Martin. *Measuring Digital Development - Facts and Figures 2021*. ITU Statistics Report 2021. Geneva, Switzerland: United Nations International Telecommunications Union, Nov. 30, 2021, p. 31.
URL: https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx (visited on 06/06/2022).

[27] K. A. Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloé M. Kiddon, Jakub Konečný, Stefano Mazzocchi, Brendan McMahan, Timon Van Overveldt, David Petrou, Daniel Ramage, and Jason Roselander. “Towards Federated Learning at Scale: System Design”. In: *SysML 2019*. SysML. Stanford, CA: Google Research, Mar. 31, 2019, pp. 1–15.
URL: https://research.google/pubs/pub47976/ (visited on 02/12/2021).

[28] Timm Böttger, Ghida Ibrahim, and Ben Vallis. “How the Internet Reacted to Covid-19: A Perspective from Facebook’s Edge Network”. In: *Proceedings of the ACM Internet Measurement Conference*. IMC ’20. New York, NY, USA: Association for Computing Machinery, Oct. 27, 2020, pp. 34–41.
DOI: 10.1145/3419394.3423621.

[29] Waylon Brunette, Morgan Vigil, Fahad Pervaiz, Shahar Levari, Gaetano Borriello, and Richard Anderson. “Optimizing Mobile Application Communication for Challenged Network Environments”. In: *Proceedings of the 2015 Annual Symposium on Computing for Development*. DEV ’15. New York, NY, USA: Association for Computing Machinery, Dec. 1, 2015, pp. 167–175.
DOI: 10.1145/2830629.2830644.

[30] Neil F. Budde. “There’s No Such Thing as an "Average" User”. In: *Interactions* 11.2 (Mar. 1, 2004), p. 54.
DOI: 10.1145/971258.971275.

[31] Michael Calabrese. *Use It or Share It: A New Default Policy for Spectrum Management*. New America Foundation, Mar. 8, 2021.
URL: http://newamerica.org/oti/reports/use-it-or-share-it/ (visited on 07/10/2022).

[32] Neal Cardwell, Yuchung Cheng, C. Stephen Gunn, Soheil Hassas Yeganeh, and Van Jacobson. “BBR: Congestion-Based Congestion Control”. In: *Queue* 14.5 (Oct. 2016), 50:20–50:53.
DOI: 10.1145/3012426.3022184.

[33] CEIC Data. *Indonesia Telecommunication Statistics: Monthly Average Revenue per Unit*. Indonesia Telecommunication Statistics. 2019.
URL: https://www.ceicdata.com/en/indonesia/telecommunication-statistics-monthly-average-revenue-per-unit (visited on 10/20/2020).

[34] Llorenç Cerdà-Alabern, Roger Baig, and Leandro Navarro. “On the Guifi.Net Community Network Economics”. In: *Computer Networks* 168 (Feb. 26, 2020), p. 107067.
DOI: 10.1016/j.comnet.2019.107067.

[35] Vinton G. Cerf and Robert E. Kahn. “A Protocol for Packet Network Intercommunication”. In: *IEEE Transactions on Communications* 22.5 (May 1974), pp. 637–648.
DOI: 10.1109/TCOM.1974.1092259.

[36] Melissa Chase, Apoorvaa Deshpande, Esha Ghosh, and Harjasleen Malvai. “SEEMless: Secure End-to-End Encrypted Messaging with Less</> Trust”. In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. CCS ’19. New York, NY, USA: Association for Computing Machinery, Nov. 6, 2019, pp. 1639–1656.
DOI: 10.1145/3319535.3363202.

[37] Kameswari Chebrolu and Bhaskaran Raman. “FRACtel: A Fresh Perspective on (Rural) Mesh Networks”. In: *Proceedings of the 2007 Workshop on Networked Systems for Developing Regions*. NSDR ’07. New York, NY, USA: ACM, 2007, 8:1–8:6.
DOI: 10.1145/1326571.1326583.

[38] Marshini Chetty, Richard Banks, Richard Harper, Tim Regan, Abigail Sellen, Christos Gkantsidis, Thomas Karagiannis, and Peter Key. "Who's Hogging the Bandwidth: The Consequences of Revealing the Invisible in the Home". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA). CHI '10. New York, NY, USA: ACM, 2010, pp. 659–668.
DOI: 10.1145/1753326.1753423.

[39] Marshini Chetty, David Haslem, Andrew Baird, Ugochi Ofoha, Bethany Sumner, and Rebecca Grinter. "Why Is My Internet Slow?: Making Network Speeds Visible". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada). CHI '11. New York, NY, USA: ACM, 2011, pp. 1889–1898.
DOI: 10.1145/1978942.1979217.

[40] Marshini Chetty, Hyojoon Kim, Srikanth Sundaresan, Sam Burnett, Nick Feamster, and W. Keith Edwards. "uCap: An Internet Data Management Tool For The Home". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea). CHI '15. New York, NY, USA: ACM, 2015, pp. 3093–3102.
DOI: 10.1145/2702123.2702218.

[41] Marshini Chetty, Srikanth Sundaresan, Sachit Muckaden, Nick Feamster, and Enrico Calandro. "Measuring Broadband Performance in South Africa". In: *Proceedings of the 4th Annual Symposium on Computing for Development*. Symposium on Computing for Development. ACM DEV-4 '13. Cape Town, South Africa: Association for Computing Machinery, Dec. 6, 2013, pp. 1–10.
DOI: 10.1145/2537052.2537053.

[42] Mung Chiang, Rachid El-Azouzi, Lin Gao, Jianwei Huang, Carlee Joe-Wong, and Soumya Sen. "Guest Editorial: Smart Data Pricing for Next-Generation Networks". In: *IEEE Journal on Selected Areas in Communications* 38.4 (Apr. 2020), pp. 641–644.
DOI: 10.1109/JSAC.2020.2971899.

[43] Daniel H. Cole and Michael D. McGinnis. *Elinor Ostrom and the Bloomington School of Political Economy: Resource Governance*. Vol. 1–4. 4 vols. Lanham, MD, USA: Lexington Books, Sept. 4, 2015. 443 pp.
ISBN: 978-0-7391-9109-5.

[44] *Comcast Releases 2020 Network Performance Data, Highlighting COVID-19 Impact*. Mar. 2, 2021.
URL: https://corporate.comcast.com/press/releases/comcast-2020-network-performance-data (visited on 05/28/2021).

[45] Stefano Crabu and Paolo Magaudda. "Bottom-up Infrastructures: Aligning Politics and Technology in Building a Wireless Community Network". In: *Computer Supported Cooperative Work (CSCW)* 27.2 (Apr. 1, 2018), pp. 149–176.
DOI: 10.1007/s10606-017-9301-1.

[46] Cam Cullen. *The Mobile Internet Phenomena Report*. Plano, TX: Sandvine, Feb. 2020, p. 14.
URL: https://www.sandvine.com/download-report-mobile-internet-phenomena-report-2020-sandvine (visited on 10/19/2020).

[47] Luiz A. DaSilva. "Pricing for QoS-enabled Networks: A Survey". In: *IEEE Communications Surveys Tutorials* 3.2 (2000), pp. 2–8.
DOI: 10.1109/COMST.2000.5340797.

[48] Jonathan Donner. *After Access: Inclusion, Development, and a More Mobile Internet*. Information Society Series. Cambridge, Massachusetts ; London, England: The MIT Press, 2015.
ISBN: 978-0-262-33125-8.

[49] Katie Dowd. *OneWeb Files for Chapter 11 Restructuring to Execute Sale Process*. OneWeb. Mar. 27, 2020.
URL: https://www.oneweb.world/media-center/oneweb-files-for-chapter-11-restructuring-to-execute-sale-process (visited on 03/11/2021).

[50] Michaelanne Dye, David Nemer, Neha Kumar, and Amy S. Bruckman. "If It Rains, Ask Grandma to Disconnect the Nano: Maintenance & Care in Havana's StreetNet". In: *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW Nov. 7, 2019), 187:1–187:27.
DOI: 10.1145/3359289.

[51] Nico Echaniz. *The right to co-create the Internet*. Altermundi: Portada. Dec. 13, 2017.
URL: https://altermundi.net/2017/12/13/the-right-to-co-create-the-internet/ (visited on 07/06/2022).

[52] Nicholas Economides and Joacim Tåg. "Network Neutrality on the Internet: A Two-Sided Market Analysis". In: *Information Economics and Policy* 24.2 (June 1, 2012), pp. 91–104.
DOI: 10.1016/j.infoecopol.2012.01.001.

[53] Matthias Falkner, Michael Devetsikiotis, and Ioannis Lambadaris. "An Overview of Pricing Concepts for Broadband IP Networks". In: *IEEE Communications Surveys Tutorials* 3.2 (2000), pp. 2–13.
DOI: 10.1109/COMST.2000.5340798.

[54] Roderick Fanou. "On the State of Interdomain Routing in Africa". MA thesis. Universidad Carlos III de Madrid, Spain, Sept. 29, 2014. 13 pp.
URL: https://eprints.networks.imdea.org/949/ (visited on 03/15/2021).

[55] Paul Feldman. "A Practical Scheme for Non-Interactive Verifiable Secret Sharing". In: *28th Annual Symposium on Foundations of Computer Science (Sfcs 1987)*. 28th Annual Symposium on Foundations of Computer Science (Sfcs 1987). Oct. 1987, pp. 427–438.
DOI: 10.1109/SFCS.1987.4.

[56] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Eric Pujol, Igmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapiador, Narseo Vallina-Rodriguez, Oliver Hohlfeld, and Georgios Smaragdakis. "Implications of the COVID-19 Pandemic on the Internet Traffic". In: *Broadband Coverage in Germany; 15th ITG-Symposium*. Broadband Coverage in Germany; 15th ITG-Symposium. Mar. 2021, pp. 1–5.
ISBN: 978-3-8007-5474-8.

[57] Adrienne Porter Felt, Richard Barnes, April King, Chris Palmer, Chris Bentzel, and Parisa Tabriz. "Measuring HTTPS Adoption on the Web". In: *USENIX Security Symposium*. USENIX Security '17. Vancouver, BC, Canada: USENIX, Aug. 16, 2017, p. 17.
ISBN: 978-1-931971-40-9.

[58] Agustin Formoso, Josiah Chavula, Amreesh Phokeer, Arjuna Sathiaseelan, and Gareth Tyson. "Deep Diving into Africa's Inter-Country Latencies". In: *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*. IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. Honolulu, HI, USA: IEEE, Apr. 2018, pp. 2231–2239.
DOI: 10.1109/INFOCOM.2018.8486024.

[59] Pantelis A. Frangoudis, George C. Polyzos, and Vasileios P. Kemerlis. "Wireless Community Networks: An Alternative Approach for Nomadic Broadband Network Access". In: *IEEE Communications Magazine* 49.5 (May 2011), pp. 206–213.
DOI: 10.1109/MCOM.2011.5762819.

[60] Batya Friedman, Peter H. Kahn, and Alan Borning. "Value Sensitive Design and Information Systems". In: *The Handbook of Information and Computer Ethics*. In collab. with Herman T. Tavani, Kenneth E. Himma, Kenneth Einar Himma, and Herman T. Tavani. Hoboken, NJ, USA: Wiley, John Wiley & Sons, Inc, 2009, pp. 69–101.
DOI: 10.1002/9780470281819.ch4.

[61] Amitabha Ghosh, Andreas Maeder, Matthew Baker, and Devaki Chandramouli. "5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15". In: *IEEE Access* 7 (2019), pp. 127639–127651.
DOI: 10.1109/ACCESS.2019.2939938.

[62] Richard J. Gibbens and Frank P. Kelly. "Resource Pricing and the Evolution of Congestion Control". In: *Automatica* 35 (1999), pp. 1969–1985.
DOI: 10.1016/S0005-1098(99)00135-1.

[63] Oana Goga and Renata Teixeira. "Speed Measurements of Residential Internet Access". In: *Passive and Active Measurement*. International Conference on Passive and Active Network Measurement. Ed. by Nina Taft and Fabio Ricciato. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 168–178.
DOI: 10.1007/978-3-642-28537-0_17.

[64] Ben Greenstein. "Delivering the Mobile Web to the Next Billion Users". In: *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications* (Tempe, Arizona, USA). HotMobile '18. New York, NY, USA: ACM, 2018, pp. 99–99. DOI: 10.1145/3177102.3180277.

[65] Rebecca E. Grinter, W. Keith Edwards, Marshini Chetty, Erika S. Poole, Ja-Young Sung, Jeonghwa Yang, Andy Crabtree, Peter Tolmie, Tom Rodden, Chris Greenhalgh, and Steve Benford. "The Ins and Outs of Home Networking: The Case for Useful and Usable Domestic Networking". In: *ACM Trans. Comput.-Hum. Interact.* 16.2 (June 2009), 8:1–8:28. DOI: 10.1145/1534903.1534905.

[66] Sarthak Grover, Mi Seon Park, Srikanth Sundaresan, Sam Burnett, Hyojoon Kim, Bharath Ravi, and Nick Feamster. "Peeking behind the NAT: An Empirical Study of Home Networks". In: *Proceedings of the 2013 Conference on Internet Measurement Conference*. IMC '13. New York, NY, USA: Association for Computing Machinery, Oct. 23, 2013, pp. 377–390. DOI: 10.1145/2504730.2504736.

[67] Ali Güngör. *Aligungr/UERANSIM*. June 29, 2022. URL: https://github.com/aligungr/UERANSIM (visited on 06/29/2022).

[68] Leon Tinashe Gwaka, Julian May, and William Tucker. "Towards Low-Cost Community Networks in Rural Communities: The Impact of Context Using the Case Study of Beitbridge, Zimbabwe". In: *The Electronic Journal of Information Systems in Developing Countries* 84.3 (2018), e12029. DOI: 10.1002/isd2.12029.

[69] Sangtae Ha, Carlee Joe-Wong, Soumya Sen, and Mung Chiang. "Pricing by Timing: Innovating Broadband Data Plans". In: SPIE OPTO. Ed. by Benjamin Dingel, Raj Jain, and Katsutoshi Tsukamoto. San Francisco, California, USA: International Society for Optics and Photonics, Jan. 21, 2012, p. 82820D. DOI: 10.1117/12.914462.

[70] Sangtae Ha, Injong Rhee, and Lisong Xu. "CUBIC: A New TCP-friendly High-speed TCP Variant". In: *SIGOPS Oper. Syst. Rev.* 42.5 (July 2008), pp. 64–74. DOI: 10.1145/1400097.1400105.

[71] Sangtae Ha, Soumya Sen, Carlee Joe-Wong, Youngbin Im, and Mung Chiang. "TUBE: Time-Dependent Pricing for Mobile Data". In: *Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication - SIGCOMM '12*. The ACM SIGCOMM 2012 Conference. Helsinki, Finland: ACM Press, 2012, p. 247. DOI: 10.1145/2342356.2342402.

[72] Amir Haleem, Andrew Allen, Andrew Thompson, Marc Nijdam, and Rahul Garg. *Helium: A Decentralized Wireless Network*. Nov. 4, 2018. URL: http://whitepaper.helium.com/ (visited on 08/08/2022).

[73] Jean Hardy, Susan Wyche, and Tiffany Veinot. “Rural HCI Research: Definitions, Distinctions, Methods, and Opportunities”. In: *Proc. ACM Hum.-Comput. Interact.* 3 (CSCW Nov. 2019).
DOI: 10.1145/3359298.

[74] Shaddi Hasan, Mary Claire Barela, Matthew Johnson, Eric Brewer, and Kurtis Heimerl. “Scaling Community Cellular Networks with CommunityCellularManager”. In: *16th USENIX Symposium on Networked Systems Design and Implementation (NSDI 19)*. NSDI 2019 (Boston, MA). NSDI’19. Boston, MA, USA: USENIX Association, 2019, pp. 735–750.
DOI: 10.5555/3323234.3323294.

[75] Shaddi Hasan, Yahel Ben-David, Max Bittman, and Barath Raghavan. “The Challenges of Scaling WISPs”. In: *Proceedings of the 2015 Annual Symposium on Computing for Development*. DEV ’15. New York, NY, USA: ACM, 2015, pp. 3–11.
DOI: 10.1145/2830629.2830637.

[76] Shaddi Hasan, Kurtis Heimerl, Kate Harrison, Kashif Ali, Sean Roberts, Anant Sahai, and Eric Brewer. “GSM Whitespaces: An Opportunity for Rural Cellular Service”. In: *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*. 2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN). Apr. 2014, pp. 271–282.
DOI: 10.1109/DySPAN.2014.6817804.

[77] Jetmir Haxhibeqiri, Eli De Poorter, Ingrid Moerman, and Jeroen Hoebeke. “A Survey of LoRaWAN for IoT: From Technology to Application”. In: *Sensors* 18.11 (11 Nov. 2018), p. 3995.
DOI: 10.3390/s18113995.

[78] Huan He, Ke Xu, and Ying Liu. “Internet Resource Pricing Models, Mechanisms, and Methods”. In: *Networking Science* 1 (2011), pp. 48–66.
DOI: 10.1007/s13119-011-0004-5.

[79] Kurtis Heimerl, Shaddi Hasan, Kashif Ali, Eric Brewer, and Tapan Parikh. “Local, Sustainable, Small-scale Cellular Networks”. In: *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers - Volume 1* (Cape Town, South Africa). ICTD ’13. New York, NY, USA: ACM, 2013, pp. 2–12.
DOI: 10.1145/2516604.2516616.

[80] Kurtis Heimerl, Shaddi Hasan, Kashif Ali, Tapan Parikh, and Eric Brewer. “A Longitudinal Study of Local, Sustainable, Small-Scale Cellular Networks”. In: *Information Technologies & International Development* 11.1 (2015), p. 20.
URL: http://itidjournal.org/index.php/itid/article/view/1359.

[81] Kurtis Heimerl, Anuvind Menon, Shaddi Hasan, Kashif Ali, Eric Brewer, and Tapan Parikh. "Analysis of Smartphone Adoption and Usage in a Rural Community Cellular Network". In: *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ICTD '15. New York, NY, USA: Association for Computing Machinery, May 15, 2015, pp. 1–4.
DOI: 10.1145/2737856.2737880.

[82] Charlotte Hess. "The Virtual CPR: The Internet as a Local and Global Common Pool Resource". In: *Fifth Annual Meeting of the International Association for the Study of Common*. Annual Meeting of the International Association for the Study of Common. Bodoe, Norway: Indiana University Press, May 24–28, 1995, p. 37.
URL: http://hdl.handle.net/10535/234.

[83] Calvin Artemies G. Hilario, Mary Claire Barela, Mar Francis D. De Guzman, Rizza T. Loquias, Ramon Vann Cleff B. Raro, Jean Jay J. Quitayen, and Joel Joseph S. Marciano. "LokaLTE: 600 MHz Community LTE Networks for Rural Areas in the Philippines". In: *2020 IEEE Global Humanitarian Technology Conference (GHTC)*. 2020 IEEE Global Humanitarian Technology Conference (GHTC). Oct. 2020, pp. 1–8.
DOI: 10.1109/GHTC46280.2020.9342849.

[84] Prieto-Egido Ignacio, Aragon Valladares Joel, Muñoz-Medina Olga, Cordoba Bernuy Cesar, Simo-Reigadas Javier, Auccapuri Quispetupa Darwin, Bravo Fernández Alejandro, and Martinez-Fernandez Andrés. "Small Rural Operators Techno-Economic Analysis to Bring Mobile Services to Isolated Communities: The Case of Peru Amazon Rainforest". In: *Telecommunications Policy* 44.10 (Nov. 1, 2020), p. 102039.
DOI: 10.1016/j.telpol.2020.102039.

[85] *ILO Monitor: COVID-19 and the World of Work. 7th Edition*. Briefing note. International Labour Organization, Jan. 25, 2021, p. 35.
URL: http://www.ilo.org/global/topics/coronavirus/impacts-and-responses/WCMS_767028/lang--en/index.htm (visited on 05/28/2021).

[86] Youngbin Im, Carlee Joe-Wong, Sangtae Ha, Soumya Sen, Ted "Taekyoung" Kwon, and Mung Chiang. "AMUSE: Empowering Users for Cost-Aware Offloading with Throughput-Delay Tradeoffs". In: *IEEE Transactions on Mobile Computing* 15.5 (May 2016), pp. 1062–1076.
DOI: 10.1109/TMC.2015.2456881.

[87] Raynell A. Inojosa, Philip A. Martinez, Ramon Vann Cleff B. Raro, Riza Carmela M. Pineda, Jerome Dylan S. Villamater, Kenneth Rey L. Sumalinog, Maria Aya Lei P. Banzuela, Kerry C. Hiponia, Kieth Joshua M. Manato, and Peter Antonio B. Banzon. "Towards the Development and Deployment of Community LTE Networks in Rural Areas". In: *2022 International Conference for Advancement in Technology (ICONAT)*. 2022 International Conference for Advancement in Technology (ICONAT). Jan. 2022, pp. 1–6.
DOI: 10.1109/ICONAT53423.2022.9725850.

[88] *Internet Performance during the COVID-19 Emergency*. The Cloudflare Blog. Apr. 23, 2020. URL: https://blog.cloudflare.com/recent-trends-in-internet-traffic/ (visited on 05/27/2021).

[89] Vikram Iyer, Vamsi Talla, Bryce Kellogg, Shyamnath Gollakota, and Joshua Smith. "Inter-Technology Backscatter: Towards Internet Connectivity for Implanted Devices". In: *Proceedings of the 2016 ACM SIGCOMM Conference*. SIGCOMM '16. New York, NY, USA: Association for Computing Machinery, Aug. 22, 2016, pp. 356–369. DOI: 10.1145/2934872.2934894.

[90] Esther Jang, Mary Claire Barela, Matt Johnson, Philip Martinez, Cedric Festin, Margaret Lynn, Josephine Dionisio, and Kurtis Heimerl. "Crowdsourcing Rural Network Maintenance and Repair via Network Messaging". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada). CHI '18. Montreal QC, Canada: Association for Computing Machinery, Apr. 19, 2018, pp. 1–12. DOI: 10.1145/3173574.3173641.

[91] Esther Han Beol Jang, Philip Garrison, Ronel Vincent Vistal, Maria Theresa D. Cunanan, Maria Theresa Perez, Philip Martinez, Matthew William Johnson, John Andrew Evangelista, Syed Ishtiaque Ahmed, Josephine Dionisio, Mary Claire Aguilar Barela, and Kurtis Heimerl. "Trust and Technology Repair Infrastructures in the Remote Rural Philippines: Navigating Urban-Rural Seams". In: *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW Nov. 7, 2019), 99:1–99:25. DOI: 10.1145/3359201.

[92] Jiongkuan Hou, Jie Yang, and S. Papavassiliou. "Integration of Pricing with Call Admission Control to Meet QoS Requirements in Cellular Networks". In: *IEEE Transactions on Parallel and Distributed Systems* 13.9 (Sept. 2002), pp. 898–910. DOI: 10.1109/TPDS.2002.1036064.

[93] Carlee Joe-Wong, Sangtae Ha, and Mung Chiang. "Time-Dependent Broadband Pricing: Feasibility and Benefits". In: *2011 31st International Conference on Distributed Computing Systems*. 2011 31st International Conference on Distributed Computing Systems (ICDCS). Minneapolis, MN, USA: IEEE, June 2011, pp. 288–298. DOI: 10.1109/ICDCS.2011.81.

[94] David L. Johnson, Elizabeth M. Belding, Kevin Almeroth, and Gertjan van Stam. "Internet Usage and Performance Analysis of a Rural Wireless Network in Macha, Zambia". In: *Proceedings of the 4th ACM Workshop on Networked Systems for Developing Regions*. NSDR '10. New York, NY, USA: Association for Computing Machinery, June 15, 2010, pp. 1–6. DOI: 10.1145/1836001.1836008.

[95] Matthew Johnson, Jenny Liang, Michelle X. Lin, Sudheesh Singanamalla, and Kurtis Heimerl. "Whale Watching in Inland Indonesia: Analyzing aSmall, Remote, Internet-Based Community Cellular Network". In: *Proceedingsof the Web Conference 2021*. WWW '21. Ljubljana, Slovenia: ACM, Apr. 2021, p. 12. DOI: 10.1145/3442381.3449996.

[96] Matthew Johnson, Spencer Sevilla, Esther Jang, and Kurtis Heimerl. "dLTE: Building a More WiFi-like Cellular Network: (Instead of the Other Way Around)". In: *Proceedings of the 17th ACM Workshop on Hot Topics in Networks* (Redmond, WA, USA). HotNets '18. New York, NY, USA: ACM, 2018, pp. 8–14.
DOI: 10.1145/3286062.3286064.

[97] R. P. Jover and J. Lackey. "dHSS - Distributed Peer-to-Peer Implementation of the LTE HSS Based on the Bitcoin/Namecoin Architecture". In: *2016 IEEE International Conference on Communications Workshops (ICC)*. 2016 IEEE International Conference on Communications Workshops (ICC). May 2016, pp. 354–359.
DOI: 10.1109/ICCW.2016.7503813.

[98] Anne Kadet. "New York City Neighbors Build Cheaper Way to Connect to Web". In: *Wall Street Journal. US* (Aug. 6, 2019).
URL: https://www.wsj.com/articles/new-york-city-neighbors-build-cheaper-way-to-connect-to-web-11565100000 (visited on 03/12/2021).

[99] Kanchana Kanchanasut, Apinun Tunpan, Mohammad Awal, Dwijendra Das, Thirapon Wongsaardsakul, and Yasuo Tsuchimoto. "DUMBONET: A Multimedia Communication System for Collaborative Emergency Response Operations in Disaster-Affected Areas". In: *International Journal of Emergency Management* 4 (Jan. 1, 2007).
DOI: 10.1504/IJEM.2007.015736.

[100] Nabin Kumar Karn, Zhang Hongli, and Muhammad Shafiq. "Measuring Broadband Internet Performance in Nepal: A Comparative Study". In: *Procedia Computer Science* 107 (Jan. 1, 2017), pp. 64–69.
DOI: 10.1016/j.procs.2017.03.057.

[101] Mohamed M. Kassem, Mahesh K. Marina, and Bozidar Radunovic. "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*. COMPASS '18. Menlo Park and San Jose, CA, USA: Association for Computing Machinery, June 20, 2018, pp. 1–5.
DOI: 10.1145/3209811.3212703.

[102] Antonios Katsarakis, Zhaowei Tan, Matthew Balkwill, Bozidar Radunovic, Andrew Bainbridge, Aleksandar Dragojevic, Boris Grot, and Yongguang Zhang. "rVNF: Reliable, Scalable and Performant Cellular VNFs in the Cloud". In: (Apr. 2, 2021).
URL: https://www.microsoft.com/en-us/research/publication/rvnf-reliable-scalable-and-performant-cellular-vnfs-in-the-cloud/ (visited on 08/12/2022).

[103] Finn Kensing and Jeanette Blomberg. "Participatory Design: Issues and Concerns". In: *Computer Supported Cooperative Work* 7 (Sept. 1, 1998), pp. 167–185.
DOI: 10.1023/A:1008689307411.

[104] Tobias Klenze, Giacomo Giuliari, Christos Pappas, Adrian Perrig, and David Basin. "Networking in Heaven as on Earth". In: *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*. HotNets '18. New York, NY, USA: Association for Computing Machinery, Nov. 15, 2018, pp. 22–28.
DOI: 10.1145/3286062.3286066.

[105] Deepti Kumar, David Martin, and Jacki O'Neill. "The Times They Are A-Changin': Mobile Payments in India". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. New York, NY, USA: Association for Computing Machinery, May 7, 2011, pp. 1413–1422.
DOI: 10.1145/1978942.1979150.

[106] Adam Langley, Alistair Riddoch, Alyssa Wilk, Antonio Vicente, Charles Krasic, Dan Zhang, Fan Yang, Fedor Kouranov, Ian Swett, Janardhan Iyengar, Jeff Bailey, Jeremy Dorfman, Jim Roskind, Joanna Kulik, Patrik Westin, Raman Tenneti, Robbie Shade, Ryan Hamilton, Victor Vasiliev, Wan-Teh Chang, and Zhongyi Shi. "The QUIC Transport Protocol: Design and Internet-Scale Deployment". In: *Proceedings of the Conference of the ACM Special Interest Group on Data Communication* (Los Angeles, CA, USA). SIGCOMM '17. New York, NY, USA: ACM, 2017, pp. 183–196.
DOI: 10.1145/3098822.3098842.

[107] David Lawrence and Sean Turner. *Web Packaging (Wpack)* -. Web Packaging (wpack): About. 2021.
URL: https://datatracker.ietf.org/wg/wpack/about/ (visited on 02/13/2021).

[108] Adisorn Lertsinsrubtavee, Liang Wang, Arjuna Sathiaseelan, Jon Crowcroft, Nunthaphat Weshsuwannarugs, Apinun Tunpan, and Kanchana Kanchanasut. "Understanding Internet Usage and Network Locality in a Rural Community Wireless Mesh Network". In: *Proceedings of the Asian Internet Engineering Conference*. AINTEC '15. Bangkok, Thailand: Association for Computing Machinery, Nov. 18, 2015, pp. 17–24.
DOI: 10.1145/2837030.2837033.

[109] Ioana Livadariu, Ahmed Elmokashfi, and Amogh Dhamdhere. "Measuring IPv6 Adoption in Africa". In: *E-Infrastructure and e-Services for Developing Countries*. AFRICOMM 2017. Ed. by Victor Odumuyiwa, Ojo Adegboyega, and Charles Uwadia. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Dakar, Senegal: Springer International Publishing, 2017, pp. 345–351.
DOI: 10.1007/978-3-319-98827-6_32.

[110] Jialu Lun, Pål Frenger, Anders Furuskar, and Elmar Trojer. "5G New Radio for Rural Broadband: How to Achieve Long-Range Coverage on the 3.5 GHz Band". In: *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall). Sept. 2019, pp. 1–6.
DOI: 10.1109/VTCFall.2019.8891556.

[111] Zhihong Luo, Silvery Fu, Mark Theis, Shaddi Hasan, Sylvia Ratnasamy, and Scott Shenker. "Democratizing Cellular Access with CellBricks". In: *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*. SIGCOMM '21. New York, NY, USA: Association for Computing Machinery, Aug. 9, 2021, pp. 626–640.
DOI: 10.1145/3452296.3473336.

[112] Andra Lutu, Diego Perino, Marcelo Bagnulo, Enrique Frias-Martinez, and Javad Khangosstar. "A Characterization of the COVID-19 Pandemic Impact on a Mobile Network Operator Traffic". In: *Proceedings of the ACM Internet Measurement Conference*. IMC '20. New York, NY, USA: Association for Computing Machinery, Oct. 27, 2020, pp. 19–33.
DOI: 10.1145/3419394.3423655.

[113] L. Maccari, M. Karaliopoulos, I. Koutsopoulos, L. Navarro, F. Freitag, and R. LoCigno. "5G and the Internet of EveryOne: Motivation, Enablers, and Research Agenda". In: *2018 European Conference on Networks and Communications (EuCNC)*. 2018 European Conference on Networks and Communications (EuCNC). June 2018, pp. 429–433.
DOI: 10.1109/EuCNC.2018.8443200.

[114] Jeffrey K. MacKie-Mason and Hal R. Varian. "Pricing Congestible Network Resources (Invited Paper)". In: *IEEE Journal on Selected Areas in Communications* 13 (1995), pp. 1141–1149.
DOI: 10.1109/49.414634.

[115] Yale Maguire. *High Altitude Connectivity: The next Chapter*. Facebook Engineering. June 27, 2018.
URL: https://engineering.fb.com/2018/06/27/connectivity/high-altitude-connectivity-the-next-chapter/ (visited on 03/11/2021).

[116] Carleen Maitland and Ying Xu. *A Social Informatics Analysis of Refugee Mobile Phone Use: A Case Study of Za'atari Syrian Refugee Camp*. SSRN Scholarly Paper ID 2588300. Rochester, NY: Social Science Research Network, Mar. 31, 2015.
DOI: 10.2139/ssrn.2588300.

[117] Anna Maria Mandalari, Andra Lutu, Ana Custura, Ali Safari Khatouni, Özgü Alay, Marcelo Bagnulo, Vaibhav Bajpai, Anna Brunstrom, Jörg Ott, Marco Mellia, and Gorry Fairhurst. "Experience: Implications of Roaming in Europe". In: *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking* (New Delhi, India). MobiCom '18. New York, NY, USA: ACM, 2018, pp. 179–189.
DOI: 10.1145/3241539.3241577.

[118] Andrés Martinez-Fernandez, Josep Vidal, Javier Simo-Reigadas, Ignacio Prieto-Egido, Adrián Agustin, Juan Paco, and Álvaro Rendon. "The TUCAN3G Project: Wireless Technologies for Isolated Rural Communities in Developing Countries Based on 3G Small Cell Deployments". In: *IEEE Communications Magazine* 54.7 (July 2016), pp. 36–43.
DOI: 10.1109/MCOM.2016.7509376.

[119] Marcela S. Melara, Aaron Blankstein, Joseph Bonneau, Edward W. Felten, and Michael J. Freedman. "{CONIKS}: Bringing Key Transparency to End Users". In: 24th {USENIX} Security Symposium ({USENIX} Security 15). 2015, pp. 383–398.
ISBN: 978-1-939133-11-3.

[120] Panagiota Micholia, Merkouris Karaliopoulos, Iordanis Koutsopoulos, Leandro Navarro, Roger Baig Vias, Dimitris Boucas, Maria Michalis, and Panayotis Antoniadis. "Community Networks and Sustainability: A Survey of Perceptions, Practices, and Proposed Solutions". In: *IEEE Communications Surveys Tutorials* 20.4 (2018), pp. 3581–3606.
DOI: 10.1109/COMST.2018.2817686.

[121] Ali Mohammadkhan, K. K. Ramakrishnan, and Vivek A. Jain. "CleanG—Improving the Architecture and Protocols for Future Cellular Networks With NFV". In: *IEEE/ACM Transactions on Networking* 28.6 (Dec. 2020), pp. 2559–2572.
DOI: 10.1109/TNET.2020.3015946.

[122] Mehrdad Moradi, Yikai Lin, Z. Morley Mao, Subhabrata Sen, and Oliver Spatscheck. "SoftBox: A Customizable, Low-Latency, and Scalable 5G Core Network Architecture". In: *IEEE Journal on Selected Areas in Communications* 36.3 (Mar. 2018), pp. 438–456.
DOI: 10.1109/JSAC.2018.2815429.

[123] Mehrdad Moradi, Karthikeyan Sundaresan, Eugene Chai, Sampath Rangarajan, and Z. Morley Mao. "SkyCore: Moving Core to the Edge for Untethered and Reliable UAV-based LTE Networks". In: *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking* (New Delhi, India). MobiCom '18. New York, NY, USA: ACM, 2018, pp. 35–49.
DOI: 10.1145/3241539.3241549.

[124] Mehrdad Moradi, Wenfei Wu, Li Erran Li, and Zhuoqing Morley Mao. "SoftMoW: Recursive and Reconfigurable Cellular WAN Architecture". In: *Proceedings of the 10th ACM International on Conference on Emerging Networking Experiments and Technologies*. CoNEXT '14. New York, NY, USA: Association for Computing Machinery, Dec. 2, 2014, pp. 377–390.
DOI: 10.1145/2674005.2674981.

[125] Richard Mortier, Tom Rodden, Peter Tolmie, Tom Lodge, Robert Spencer, Andy crabtree, Joe Sventek, and Alexandros Koliousis. "Homework: Putting Interaction into the Infrastructure". In: *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA). UIST '12. New York, NY, USA: ACM, 2012, pp. 197–206.
DOI: 10.1145/2380116.2380143.

[126] nancymharvey. *AST & Science*. AST & Science. 2021.
URL: https://ast-science.com/ (visited on 03/09/2021).

[127] Binh Nguyen, Tian Zhang, Bozidar Radunovic, Ryan Stutsman, Thomas Karagiannis, Jakub Kocur, and Jacobus Van der Merwe. "ECHO: A Reliable Distributed Cellular Core Network for Hyper-scale Public Clouds". In: *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. MobiCom '18. New York, NY, USA: Association for Computing Machinery, Oct. 15, 2018, pp. 163–178.
DOI: 10.1145/3241539.3241564.

[128] O3B. *O3b mPOWER*. SES. 2021.
URL: https://www.ses.com/networks/networks-and-platforms/o3b-mpower (visited on 03/11/2021).

[129] Andrew Odlyzko. "Internet Pricing and the History of Communications". In: *Computer Networks*. Theme Issue: The Economics of Networking 36.5 (Aug. 1, 2001), pp. 493–517.
DOI: 10.1016/S1389-1286(01)00188-8.

[130] Andrew Odlyzko. "Paris Metro Pricing for the Internet". In: *Proceedings of the 1st ACM Conference on Electronic Commerce* (Denver, Colorado, USA). EC '99. New York, NY, USA: ACM, 1999, pp. 140–147.
DOI: 10.1145/336992.337030.

[131] Shree Om, Carlos Rey-Moreno, and William David Tucker. "Towards a Scalability Model for Wireless Mesh Networks". In: Telkom, 2015.
ISBN: 978-0-620-67151-4.

[132] *Open Street Map: Bokondini*. 2022.
URL: openstreetmap.org/copyright.

[133] Elinor Ostrom. "A General Framework for Analyzing Sustainability of Social-Ecological Systems". In: *Science* 325.5939 (July 24, 2009), pp. 419–422.
DOI: 10.1126/science.1172133.

[134] Elinor Ostrom. "Beyond Markets and States: Polycentric Governance of Complex Economic Systems". In: *American Economic Review* 100.3 (June 2010), pp. 641–672.
DOI: 10.1257/aer.100.3.641.

[135] Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Political Economy of Institutions and Decisions. Cambridge ; New York: Cambridge University Press, 1990.
ISBN: 978-0-521-37101-8.

[136] Elinor Ostrom, Joanna Burger, Christopher B. Field, Richard B. Norgaard, and David Policansky. "Revisiting the Commons: Local Lessons, Global Challenges". In: *Science* 284.5412 (1999), pp. 278–282.
DOI: 10.1126/science.284.5412.278.

[137] I.Ch. Paschalidis and J.N. Tsitsiklis. "Congestion-Dependent Pricing of Network Services". In: *IEEE/ACM Transactions on Networking* 8.2 (Apr. 2000), pp. 171–184.
DOI: 10.1109/90.842140.

[138] Torben Pryds Pedersen. “Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing”. In: *Advances in Cryptology — CRYPTO ’91*. Ed. by Joan Feigenbaum. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1992, pp. 129–140. DOI: 10.1007/3-540-46766-1_9.

[139] Avery Pennarun. *How Tailscale Works*. Tailscale. Mar. 20, 2020. URL: https://tailscale.com/blog/how-tailscale-works/ (visited on 06/29/2022).

[140] A. Pentland, R. Fletcher, and A. Hasson. “DakNet: Rethinking Connectivity in Developing Nations”. In: *Computer* 37.1 (Jan. 2004), pp. 78–83. DOI: 10.1109/MC.2004.1260729.

[141] Gregers Petersen. “Freifunk: When Technology and Politics Assemble into Subversion”. In: *Subversion, Conversion, Development: Cross-Cultural Knowledge Exchange and the Politics of Design*. Boston, MA, USA: The MIT Press, Apr. 25, 2014, pp. 39–56. ISBN: 978-0-262-32249-2.

[142] *Policy Brief: Spectrum Approaches for Community Networks*. Internet Society. URL: https://www.internetsociety.org/policybriefs/spectrum/ (visited on 11/13/2018).

[143] Cynthia Putnam, Beth Kolko, and Siri Wood. “Communicating About Users in ICTD: Leveraging HCI Personas”. In: *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (Atlanta, Georgia, USA). ICTD ’12. New York, NY, USA: ACM, 2012, pp. 338–349. DOI: 10.1145/2160673.2160714.

[144] Zafar Ayyub Qazi, Melvin Walls, Aurojit Panda, Vyas Sekar, Sylvia Ratnasamy, and Scott Shenker. “A High Performance Packet Core for Next Generation Cellular Networks”. In: *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*. SIGCOMM ’17. New York, NY, USA: ACM, 2017, pp. 348–361. DOI: 10.1145/3098822.3098848.

[145] Ali Raza, Yasir Zaki, Thomas Pötsch, Jay Chen, and Lakshmi Subramanian. “xCache: Rethinking Edge Caching for Developing Regions”. In: *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*. ICTD ’17. New York, NY, USA: Association for Computing Machinery, Nov. 16, 2017, pp. 1–11. DOI: 10.1145/3136560.3136577.

[146] *Remote Work, Regional Lockdowns and Migration of Internet Usage*. The Cloudflare Blog. Apr. 11, 2020. URL: https://blog.cloudflare.com/remote-work-regional-lockdowns-and-migration-of-internet-usage/ (visited on 05/27/2021).

[147] Rethink Technology Research. *IoT Should Drive LTE into Underused Sub-500 MHz Spectrum*. Volume 16, issue 35. 450 MHz Alliance, Mar. 29, 2019. URL: https://450alliance.org/iot-should-drive-lte-into-underused-sub-500-mhz-spectrum/ (visited on 03/11/2021).

[148] Carlos Rey-Moreno. *Supporting the Creation and Scalability of Affordable Access Solutions: Understanding Community Networks in Africa*. 1st ed. Galerie Jean-Malbuisson 15, CH–1204 Geneva, Switzerland: Internet Society, May 1, 2017. 40 pp.
ISBN: 978-0-692-89777-5.

[149] Carlos Rey-Moreno, Anriette Esterhuysen, Mike Jensen, Peter Bloom, Erick Huerta, and Steve Song. "Can the Unconnected Connect Themselves? Towards an Action Research Agenda for Local Access Networks". In: *Community Networks: The Internet by the People, for the People*. Ed. by Luca Belli. FGV Direito Rio, Nov. 2017, pp. 103–118.
ISBN: 978-85-959701-0-6.

[150] Carlos Rey-Moreno, Amalia Sabiescu, and Masbulele Siya. "Towards Self-Sustaining Community Networks in Rural Areas of Developing Countries: Understanding Local Ownership". In: *ICTs for Inclusive Communities in Developing Societies*. 8th International Development Informatics Association Conference. Port Elizabeth, South Africa: International Development Informatics Association, Nov. 3, 2014, pp. 63–77.
ISBN: 978-0-620-63498-4.

[151] Carlos Rey-Moreno, William B. Tucker, D. Cull, and R. Blom. "Making a Community Network Legal Within the South African Regulatory Framework". In: *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. ICTD'15 (Singapore, Singapore). ICTD '15. New York, NY, USA: ACM, 2015, 57:1–57:4.
DOI: 10.1145/2737856.2737867.

[152] J. W. Roberts. "Internet Traffic, QoS, and Pricing". In: *Proceedings of the IEEE* 92.9 (Aug. 2004), pp. 1389–1399.
DOI: 10.1109/JPROC.2004.832959.

[153] Nithya Sambasivan, Paul Lee, Greg Hecht, Paul M. Aoki, Maria-Ines Carrera, Jenny Chen, Michael Youssefmir, David Cohn, Pete Kruskall, Everett Wetchler, and Astrid Twenebowa Larssen. "SmartBrowse: Design and Evaluation of a Mobile Data Price Transparency Tool for Mobile Web Use". In: *Information Technology and International Development* 11(1) (2015), pp. 21–40.
URL: http://www.itidjournal.org/index.php/itid/article/view/1360/507 (visited on 01/11/2020).

[154] Nithya Sambasivan and Thomas Smyth. "The Human Infrastructure of ICTD". In: *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development* (London, United Kingdom). ICTD '10. New York, NY, USA: ACM, 2010, 40:1–40:9.
DOI: 10.1145/2369220.2369258.

[155] Christian Sandvig. "Connection at Ewiiaapaayp Mountain: Indigenous Internet Infrastructure". In: *Race After the Internet*. New York: Routledge, 2012, pp. 168–200.
ISBN: 978-0-415-80236-9.

[156] Srivatsan Sankaranarayanan, Panagiotis Papadimitratos, Amitabh Mishra, and Steven Hershey. "A Bandwidth Sharing Approach to Improve Licensed Spectrum Utilization". In: *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.* First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. Nov. 2005, pp. 279–288.
DOI: 10.1109/DYSPAN.2005.1542644.

[157] Jithesh Sathyan. *Fundamentals of EMS, NMS, and OSS/BSS*. 1st ed. Boca Raton, Fla.: CRC Press : Auerbach Publications, 2010. 586 pp.
ISBN: 978-1-4200-8574-7.

[158] Brandon Schlinker, Italo Cunha, Yi-Ching Chiu, Srikanth Sundaresan, and Ethan Katz-Bassett. "Internet Performance from Facebook's Edge". In: *Proceedings of the Internet Measurement Conference*. IMC '19. New York, NY, USA: Association for Computing Machinery, Oct. 21, 2019, pp. 179–194.
DOI: 10.1145/3355369.3355567.

[159] Paul Schmitt and Barath Raghavan. "Pretty Good Phone Privacy". In: 30th USENIX Security Symposium (USENIX Security 21). 2021, pp. 1737–1754.
ISBN: 978-1-939133-24-3.

[160] *SD-Core*. Open Networking Foundation. Jan. 2022.
URL: https://opennetworking.org/sd-core/ (visited on 07/10/2022).

[161] Mennan Selimi, Amin M. Khan, Emmanouil Dimogerontakis, Felix Freitag, and Roger Pueyo Centelles. "Cloud Services in the Guifi.Net Community Network". In: *Computer Networks: The International Journal of Computer and Telecommunications Networking* 93.P2 (Dec. 24, 2015), pp. 373–388.
DOI: 10.1016/j.comnet.2015.09.007.

[162] Mennan Selimi, Leandro Navarro, Bart Braem, Felix Freitag, and Adisorn Lertsinsrubtavee. "Towards Information-Centric Edge Platform for Mesh Networks: The Case of CityLab Testbed". In: *2020 IEEE International Conference on Fog Computing (ICFC)*. 2020 IEEE International Conference on Fog Computing (ICFC). Apr. 2020, pp. 50–55.
DOI: 10.1109/ICFC49376.2020.00016.

[163] Soumya Sen, Carlee Joe-Wong, Sangtae Ha, and Mung Chiang. "A Survey of Smart Data Pricing: Past Proposals, Current Plans, and Future Trends". In: *ACM Comput. Surv.* 46.2 (Nov. 2013), 15:1–15:37.
DOI: 10.1145/2543581.2543582.

[164] Soumya Sen, Carlee Joe-Wong, Sangtae Ha, and Mung Chiang. "Smart Data Pricing: Using Economics to Manage Network Congestion". In: *Communications of the ACM* 58.12 (Nov. 23, 2015), pp. 86–93.
DOI: 10.1145/2756543.

[165] Soumya Sen, Carlee Joe-Wong, Sangtae Ha, and Mung Chiang. "Time-Dependent Pricing for Multimedia Data Traffic: Analysis, Systems, and Trials". In: *IEEE Journal on Selected Areas in Communications* 37.7 (July 2019), pp. 1504–1517.
DOI: 10.1109/JSAC.2019.2916490.

[166] Spencer Sevilla. *CoLTE: The Community LTE Project.* Github: UW ICTD Lab, Oct. 12, 2018.
URL: https://github.com/uw-ictd/colte (visited on 10/13/2018).

[167] Spencer Sevilla, Matthew Johnson, Pat Kosakanchit, Jenny Liang, and Kurtis Heimerl. "Experiences: Design, Implementation, and Deployment of CoLTE, a Community LTE Solution". In: *The 25th Annual International Conference on Mobile Computing and Networking*. MobiCom '19. Los Cabos, Mexico: Association for Computing Machinery, Oct. 11, 2019, pp. 1–16.
DOI: 10.1145/3300061.3345446.

[168] Kushal Shah, Philip Martinez, Emre Tepedelenlioglu, Shaddi Hasan, Cedric Festin, Joshua Blumenstock, Josephine Dionisio, and Kurtis Heimerl. "An Investigation of Phone Upgrades in Remote Community Cellular Networks". In: *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*. ICTD '17. New York, NY, USA: Association for Computing Machinery, Nov. 16, 2017, pp. 1–12.
DOI: 10.1145/3136560.3136569.

[169] Adi Shamir. "How to Share a Secret". In: *Communications of the ACM* 22.11 (Nov. 1, 1979), pp. 612–613.
DOI: 10.1145/359168.359176.

[170] Asheesh Sharma, Manveen Kaur, Zahir Koradia, Rahul Nishant, Sameer Pandit, Aravindh Raman, and Aaditeshwar Seth. "Revisiting the State of Cellular Data Connectivity in India". In: *Proceedings of the 2015 Annual Symposium on Computing for Development*. DEV '15. New York, NY, USA: Association for Computing Machinery, Dec. 1, 2015, pp. 149–157.
DOI: 10.1145/2830629.2830649.

[171] Savannah Wei Shi, Mu Xia, and Yun Huang. "From Minnows to Whales: An Empirical Study of Purchase Behavior in Freemium Social Games". In: *International Journal of Electronic Commerce* 20.2 (2015), pp. 177–207.
DOI: 10.1080/10864415.2016.1087820.

[172] Javier Simo-Reigadas, Esteban Municio, Eduardo Morgado, Eva M. Castro, Andres Martinez, Luis F. Solorzano, and Ignacio Prieto-Egido. "Sharing Low-Cost Wireless Infrastructures with Telecommunications Operators to Bring 3G Services to Rural Communities". In: *Computer Networks*. Community Networks 93 (Dec. 24, 2015), pp. 245–259.
DOI: 10.1016/j.comnet.2015.09.006.

[173] Sameer Kumar Singh, Rohit Singh, and Brijesh Kumbhani. "The Evolution of Radio Access Network Towards Open-RAN: Challenges and Opportunities". In: *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. 2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW). Apr. 2020, pp. 1–6. DOI: 10.1109/WCNCW48565.2020.9124820.

[174] *srsLTE: Open Source SDR LTE Software Suite*. srsLTE, July 16, 2018. URL: https://github.com/srsLTE/srsLTE (visited on 07/18/2018).

[175] Reuters Staff. "Motorola Settles Iridium Bankruptcy Cases". In: *Reuters* (May 20, 2008). URL: https://www.reuters.com/article/us-motorola-iridium-idUSN2033916620080520 (visited on 03/11/2021).

[176] RIPE NCC Staff. "Ripe Atlas: A Global Internet Measurement Network". In: *Internet Protocol Journal* 18.3 (2015). URL: http://ipj.dreamhosters.com/wp-content/uploads/2015/10/ipj18.3.pdf.

[177] Statista Research Department. *Nonprofit Organizations in the U.S.* Statista. URL: https://www.statista.com/topics/1390/nonprofit-organizations-in-the-us/ (visited on 06/29/2022).

[178] Samuel Sudar, Matt Welsh, and Richard Anderson. "Siskin: Leveraging the Browser to Share Web Content in Disconnected Environments". In: *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. COMPASS '18. New York, NY, USA: Association for Computing Machinery, June 20, 2018, pp. 1–7. DOI: 10.1145/3209811.3209820.

[179] Srikanth Sundaresan, Walter de Donato, Nick Feamster, Renata Teixeira, Sam Crawford, and Antonio Pescapè. "Broadband Internet Performance: A View from the Gateway". In: *ACM SIGCOMM Computer Communication Review* 41.4 (Aug. 15, 2011), pp. 134–145. DOI: 10.1145/2043164.2018452.

[180] Sonesh Surana, Rabin K. Patra, Sergiu Nedevschi, Manuel Ramos, Lakshminarayanan Subramanian, Yahel Ben-David, and Eric A. Brewer. "Beyond Pilots: Keeping Rural Wireless Networks Alive." In: *NSDI*. USENIX Symposium on Networked Systems Design and Implementation. Vol. 8. San Francisco, California, USA: USENIX Association, 2008, pp. 119–132. URL: https://www.usenix.org/legacy/events/nsdi08/tech/full_papers/surana/surana_html (visited on 09/09/2017).

[181] The Magma Authors. *About Magma*. Magma Documentation. 2022. URL: https://magmacore.org/about-magma/ (visited on 08/12/2022).

[182] Nussara Tieanklin, Matthew Johnson, and Kurtis Heimerl. "Poster: The Low Impact of COVID-19 on Rural Community Network Traffic". In: *ACM SIGCAS Conference on Computing and Sustainable Societies*. COMPASS '21. New York, NY, USA: Association for Computing Machinery, June 28, 2021, pp. 417–422. DOI: 10.1145/3460112.3472311.

[183] Mariah Timms. "AT&T Outage: Internet, 911 Disrupted, Planes Grounded after Nashville Explosion. Get the Latest Updates". In: *The Tennessean* (Dec. 25, 2020).
URL: https://www.tennessean.com/story/news/local/2020/12/25/att-outage-internet-down-hours-after-nashville-explosion/4045278001/ (visited on 06/21/2022).

[184] Guillaume Touchard. *Unlocking Rural Coverage: Enablers for Commercially Sustainable Mobile Network Expansion*. London: GSMA Intelligence, 2017.
URL: https://www.gsma.com/mobilefordevelopment/resources/unlocking-rural-coverage-enablers-commercially-sustainable-mobile-network-expansion/ (visited on 09/13/2019).

[185] Deepak Vasisht, Zerina Kapetanovic, Jongho Won, Xinxin Jin, Ranveer Chandra, Sudipta Sinha, Ashish Kapoor, Madhusudhan Sudarshan, and Sean Stratman. "FarmBeats: An IoT Platform for Data-Driven Agriculture". In: 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17). Boston, MA, USA: USENIX, 2017, pp. 515–529.
ISBN: 978-1-931971-37-9.

[186] Davide Vega, Roger Baig, Llorenç Cerdà-Alabern, Esunly Medina, Roc Meseguer, and Leandro Navarro. "A Technological Overview of the Guifi.Net Community Network". In: *Computer Networks*. Community Networks 93 (Dec. 24, 2015), pp. 260–278.
DOI: 10.1016/j.comnet.2015.09.023.

[187] Davide Vega, Llorenç Cerdà-Alabern, Leandro Navarro, and Roc Meseguer. "Topology Patterns of a Community Network: Guifi.Net". In: *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. 2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). Barcelona, Spain: IEEE, Oct. 2012, pp. 612–619.
DOI: 10.1109/WiMOB.2012.6379139.

[188] Benny Vejlgaard, Mads Lauridsen, Huan Nguyen, Istvan Z. Kovacs, Preben Mogensen, and Mads Sorensen. "Coverage and Capacity Analysis of Sigfox, LoRa, GPRS, and NB-IoT". In: *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. 2017 IEEE 85th Vehicular Technology Conference (VTC Spring). IEEE, June 2017, pp. 1–5.
DOI: 10.1109/VTCSpring.2017.8108666.

[189] *Verizon 5G Standalone Core Trial Paves Way for Robust 5G Consumer and Enterprise Solutions*. July 9, 2020.
URL: https://www.verizon.com/about/news/verizon-5g-standalone-core (visited on 07/10/2022).

[190] Morgan Vigil, Matthew Rantanen, and Elizabeth Belding. "A First Look at Tribal Web Traffic". In: *Proceedings of the 24th International Conference on World Wide Web*. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, May 18, 2015, pp. 1155–1165.
ISBN: 978-1-4503-3469-3.

[191] Morgan Vigil-Hayes, Md Nazmul Hossain, Alexander K Elliott, Elizabeth M. Belding, and Ellen Zegura. “LoRaX: Repurposing LoRa as a Low Data Rate Messaging System to Extend Internet Boundaries”. In: *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*. COMPASS ’22. New York, NY, USA: Association for Computing Machinery, June 29, 2022, pp. 195–213.
DOI: 10.1145/3530190.3534807.

[192] Christian Vogt, Max Jonas Werner, and Thomas C. Schmidt. “Leveraging WebRTC for P2P Content Distribution in Web Browsers”. In: *2013 21st IEEE International Conference on Network Protocols (ICNP)*. 2013 21st IEEE International Conference on Network Protocols (ICNP). Goettingen, Germany: IEEE, Oct. 2013, pp. 1–2.
DOI: 10.1109/ICNP.2013.6733637.

[193] June Wang and Kin F. Li. “Understanding Internet Pricing: An Objective-Oriented Classification”. In: *CCECE 2003 - Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436)*. CCECE 2003 - Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436). Vol. 2. Montreal Canada: IEEE, May 2003, 703–708 vol.2.
DOI: 10.1109/CCECE.2003.1225992.

[194] Alastair Westgarth. *Saying Goodbye to Loon*. Medium. Jan. 22, 2021.
URL: https://medium.com/loon-for-all/loon-draft-c3fcebc11f3f (visited on 03/11/2021).

[195] Saravut Yaipairoj and Fotios C. Harmantzis. “Dynamic Pricing with "Alternatives" for Mobile Networks”. In: *2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No.04TH8733)*. 2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No.04TH8733). Vol. 2. Atlanta, Georgia, USA: IEEE, Mar. 2004, 671–676 Vol.2.
DOI: 10.1109/WCNC.2004.1311266.

[196] Jeonghwa Yang, W. Keith Edwards, and David Haslem. “Eden: Supporting Home Network Management Through Interactive Visual Tools”. In: *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA). UIST ’10. New York, NY, USA: ACM, 2010, pp. 109–118.
DOI: 10.1145/1866029.1866049.

[197] Sarah Yu and Samia Ibtasam. “A Qualitative Exploration of Mobile Money in Ghana”. In: *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. COMPASS ’18. New York, NY, USA: Association for Computing Machinery, June 20, 2018, pp. 1–10.
DOI: 10.1145/3209811.3209863.

[198] Yasir Zaki, Jay Chen, Thomas Pötsch, Talal Ahmad, and Lakshminarayanan Subramanian. “Dissecting Web Latency in Ghana”. In: *Proceedings of the 2014 Conference on Internet Measurement Conference*. IMC ’14. New York, NY, USA: Association for Computing Machinery, Nov. 5, 2014, pp. 241–248.
DOI: 10.1145/2663716.2663748.

[199] Shoshana Zuboff. “Big Other: Surveillance Capitalism and the Prospects of an Information Civilization”. In: *Journal of Information Technology* 30.1 (Mar. 1, 2015), pp. 75–89. DOI: 10.1057/jit.2015.5.

APPENDIX A

Whale Watching in Inland Indonesia Supplemental Information

This appendix contains additional data collected but omitted from the main manuscript.

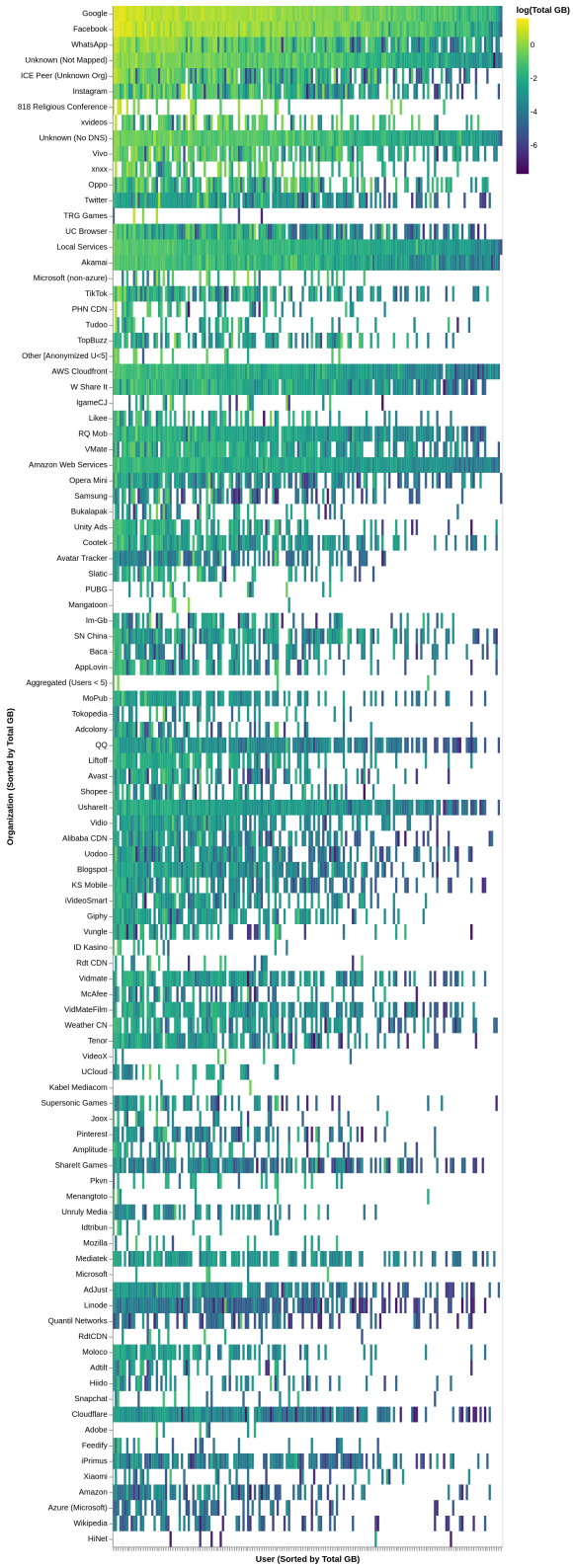


Figure A.1: Heatmap of the amount of bytes transferred between different organizations and the users who communicate with them. Some organizations communicate with most or all users, while others communicate with only a small subset.

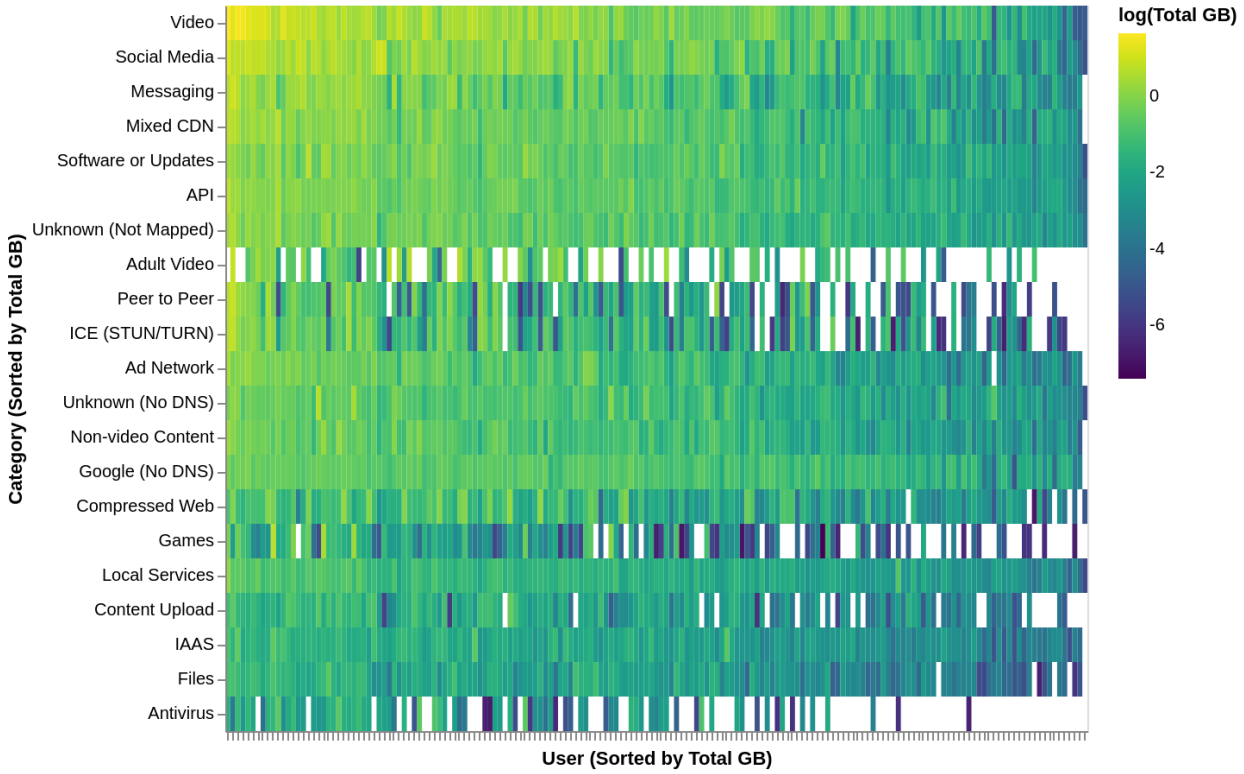


Figure A.2: Heatmap of the amount of bytes in each category transferred by each user. The majority of users have at least some P2P traffic and traffic to a known advertising network. All users interact with the local administration portal. Contrary to prior work, only a small subset of users interact with known antivirus providers, but all receive software and/or application updates from device manufacturer or operating system providers.

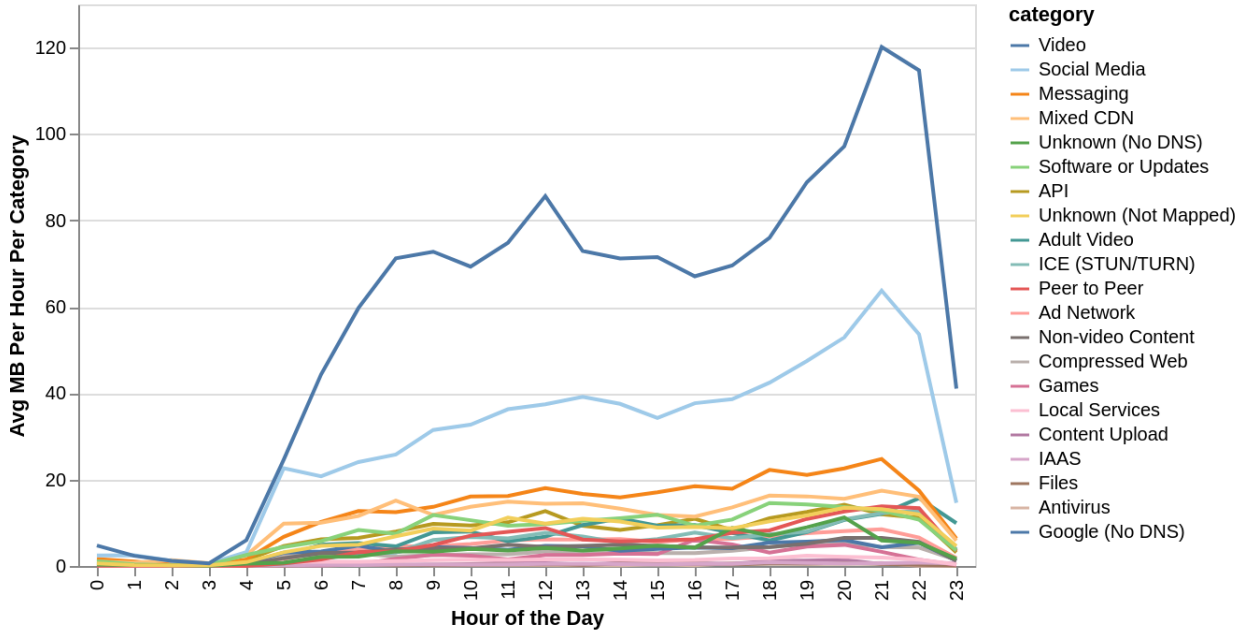


Figure A.3: Video and social media drive most of the evening demand peak. There is a midday video peak as well.

APPENDIX B

CCN Traffic Measurement: COVID-19 Supplemental Information

B.1 Telemedicine Keywords

- corona
- covid
- dokter
- doktor
- health
- infeksi
- kemkes
- kesegaran
- kesehatan
- kewarasan
- korona
- lindungi
- medical
- medicine
- mengobati
- merawat
- oba
- peduli
- pedulilindungi
- pengobat
- penyakit
- telemedicine
- vaccine
- vaksin
- virus
- who

APPENDIX C

dAuth Supplemental Information

C.1 SIM Details

The sequence number, or sqn, is a component of the authentication process that prevents re-authentication using old authentication data.

When considering sqn as a single monotonically increasing value, determining a valid sqn number could be done by choosing any sqn larger than what has already been used. However, SIM cards can maintain a set of independent sqns by dividing the range of possible sqn values into a fixed number of 'slices'. Each slice is the set of all numbers that share the same modulo:

$$\text{slice} = \text{sqn}\%\text{number of total slices}$$

The 3GPP specifies an informative implementation suggestion in TS 33.102 [2] where the SIM card maintains counters for *each slice*. Consider a SIM card that maintains 32 slices. For slice 1 out of 32 (zero indexed), the slice counter would keep track of sqns 1, 33, 65, and so on. Table C.1 below shows this breakdown:

In the SIM Slices table C.1, note that the counters operate independently. It is possible to use a smaller sqn following a larger sqn, provided that the smaller sqn is on a different slice and is the largest seen of that particular slice. For example, a sqn of 33 (slice 1) would be valid, while 64 (slice 0) would be invalid. Table C.2 shows an example valid SIM state in the 3GPP informative implementation.

C.2 Test Network Details

See Table C.3.

SIM Slices				
%32 = 0	**%32 = 1**	**%32 = 2**	**...**	**%32 = 31**
0	1	2	...	31
32	33	34	...	63
64	65	66	...	95
...	...	...	...	...

Table C.1: Table showing sequence of values from 0 and on, separated into slices

SIM Slices				
%32 = 0	**%32 = 1**	**%32 = 2**	**...**	**%32 = 31**
96	1	66	...	31

Table C.2: Table showing an example valid state for internal SIM slice counters

Purpose	Location	Node Type	ISP	Processor	RAM	NIC	Disk
Test	SCN Library	Protectli	Centurylink	Intel Celeron J3160 @ 1.60GHz	8GB DDR3	1Gbps Intel	SATA3 SSD
Test	SCN Community Center	Quotom	Centurylink	Intel i5-4200U @ 1.60GHz	8GB DDR3	1Gbps Intel	SATA3 SSD
Test	UW-Lab	Quotom	UW	Intel Core i3-4005U @ 1.70GHz	8GB DDR3	1Gbps Intel	SATA3 SSD
Test	Cloud Azure US-West-2	F2s v2	Private	Intel Xeon 8370C 2cpu	4GB	?	“Premium” SSD
Test	Cloud AWS US-West-2	C6a.large	Private	AMD EPYC 7R13 Processor 2cpu	4GB	?	Cloud block storage
Test	Cloud Digital Ocean SF	CPU Droplet	Private	Intel Xeon 8168 CPU @ 2.70GHz 2 CPU	4GB	?	Attached SSD
Test	Cloud GCP US-Central-1	c2d standard2	Private	AMD EPYC 7B13 "Milan" 2 cpu	8GB	10Gbps vnic	Balanced persistent cloud storage
Test	UW-Lab	Zotac	UW	Intel Celeron N3160 @ 1.60GHz	8GB DDR3	1Gbps Realtek	SATA2 HDD, 5400RPM
Test	Home A	Zotac	Comcast Business	Intel Celeron N3160 @ 1.60GHz	8GB DDR3	1Gbps Realtek	SATA3 SSD
Test	Home B	Zotac	Comcast Residential Shared	Intel Celeron N3160 @ 1.60GHz	4GB DDR3	1Gbps Realtek	SATA1 HDD, 5400RPM
RAN	Home A	Latitude e7470	Comcast Business	Intel Core i7-6600U @ 2.60GHz	16GB DDR4	1Gbps Intel	SATA3 SSD
RAN	UW-Lab	Dell Precision	UW	Intel Core i7-4790 @ 3.60GHz	8GB DDR3	1Gbps Intel	SATA3 HDD 7200RPM

Table C.3: Details of the nodes in the test network. All nodes are connected via a TailScale Mesh VPN and had direct accessability during the duration of tests. RAN nodes hosted UERANSIM and did not host a dAuth daemon at test time.

Vita

Matt Johnson is a PhD student at the University of Washington. He's passionate about ubiquitious global connectivity, and is excited to work towards a world where everyone can get a performant, reliable, and affordable Internet connection should they desire it.

You can find more info about him and his ongoing work at https://matt9j.net.