Designing Sustainable Rural Wireless Networks for Developing Regions

Sonesh Surana

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2009-188
http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-188.html

December 20, 2009

Designing Sustainable Rural Wireless Networks for Developing Regions

by

Sonesh Surana

B.S. (Carnegie Mellon University) 2001
M.S. (University of California at Berkeley) 2004

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Computer Science

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Eric Brewer, Chair
Professor John Canny
Professor Eva Harris

Spring 2009

The dissertation of Sonesh Surana is approved:

Chair Date

Date

Date

University of California, Berkeley

Spring 2009

Designing Sustainable Rural Wireless Networks for Developing Regions

Abstract

Designing Sustainable Rural Wireless Networks for Developing Regions

by

Sonesh Surana

Doctor of Philosophy in Computer Science

University of California, Berkeley

Professor Eric Brewer, Chair

We design, implement and deploy a rural telemedicine network, connecting several villages to the Aravind Eye Hospitals in India, enabling over 100,000 (and counting) remote eye examinations, via high-quality video-conferencing between doctors and remote rural patients who otherwise had *no access* to any health care services. Of these, close to 20,000 patients have regained their sight, only due to early diagnoses enabled by our network. This dissertation describes the technologies we have created across several layers to enable new options for rural connectivity, more viable than the current state of the art.

Long-distance point-to-point links, up to hundreds of kilometers long, are best suited to introduce connectivity into rural geographies, typically characterized by sparsely-spread clusters of lower-income users. Compared to fiber, cellular, satellite and WiMAX technologies, WiFi offers more exciting possibilities as it is very inexpensive, offers operational freedom due to its use of unlicensed frequencies, and is less complex to deploy and

manage by rural organizations. However, WiFi performs very poorly in long-distance settings due to low channel utilization, intra-link packet collisions, inter-link interference at relay nodes, and asymetrically lossy channels. In addition, managing systems in rural areas is very challenging due to frequent failures resulting from poor-quality power, operator errors by inexperienced staff, and limited opportunities for remote management. Taken together, these challenges prevent financial and operational sustainability, a critical goal rural networks need to achieve in order to scale beyond small pilot deployments, and have any real impact at all.

We treat sustainability as a critical systems design goal and take an end-to-end systems perspective by asking the question: *How can we design and build financially and operationally sustainable WiFi networks that provide high-throughput in long-distance settings, in the face of lossy environments, and in the presence of sytemic link or node failures?*

At the *network layer*, we design and implement WiLDNet, a new TDMA-based MAC-layer that increases link utilization, eliminates most packet collisions in single- and multi-hop settings, and combines FEC and ARQ for link-level loss recovery. Compared to the standard WiFi MAC, WiLDNet enables 2-5 fold improvements in UDP and TCP throughputs. At the *management layer*, we build a range of tools for system-monitoring over intermittent links, backchannels for fault diagnosis, and mechanisms for hardware and software failure recovery. At the *lowest layer*, after careful investigation of the effects of poor-quality grid power, we build low-cost solutions that make wireless nodes more resistant to power-related damage. We also build off-grid power solutions that can extend the life of battery backups, further reducing costs and enabling connectivity in remote areas.

These solutions, across all layers, have enabled financial and operational sustainability of the Aravind telemedicine network, resulting in both *scaling* and *replication* of our work. Aravind is scaling our network to connect to 50 villages, targeting general access to eye care services for 2.5 million people and 500,000 remote eye examinations annually. Other groups such as the Lumbini Eye Institute, the Pakistan Institute for Community Ophthalmology and Inveneo are using our work to replicate similar networks in Nepal, Pakistan and Uganda respectively.

Professor Eric Brewer
Dissertation Committee Chair

For my wonderful parents

Santosh and Subhadra Surana

and my loving Nanisa and Nanasa

Shrimati Bowribai and Shri Mohanlalji Khariwal

in memory of my beautiful Dadisa and Dadasa

Swargya Shrimati Badambai and Swargya Shri Pandit Jodhraj Surana

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation would not have been possible without the work of some amazing collaborators. I feel privileged to have been a part of TIER, a vibrant research group, full of intelligent, generous and fun people, committed to making a real difference in the world. In many ways, TIER reflects the qualities of its leader, Eric Brewer.

I have been fortunate to have Eric as an advisor. Many years ago, it was his vision that a new agenda for systems research was needed to address higher-level problems of health care, education, governance, business and agriculture in the developing world. Back then, it was unclear if this would be an area viable enough for computer science graduate students to bet their careers on. But Eric provided leadership, and continues to do so, in picking the right problems to solve, that are research-worthy, socially relevant and deeply fulfilling.

Personally, I've been inspired by Eric's sense of optimism, readiness for a big challenge, openness in collaboration, and deep values of integrity. Although he's played a major role in guiding me through research, he's played an even more significant role in guiding me to be a good person.

I am grateful to my committee members, John Canny and Eva Harris. Both have been very supportive of my dissertation and have provided excellent advice at various points in time. I am also thankful to Ion Stoica, my advisor when I first got to Berkeley. He encouraged me to explore research opportunities closer to my heart, which is how I found TIER.

Rabin has been a constant partner-in-crime, always ready for any adventure any time, and always willing to help. He is also one of the smartest people I know. Sergiu has been trying to make a bad boy out of me since the day we met. I hope I give in soon. He is the best-balanced person I know, deftly able to manage a hectic work schedule with a hectic party schedule.

Lakshmi brought a sense of urgency to our research, taught us how to write compelling papers, and has always had great insights on job searches (and salsa clubs). Yahel has kept us honest. His unbounded enthusiasm has pushed us to do more and more, and is the major reason why I have come this far. One day, I hope I will install a *pukka* link for him. Manuel helped me understand power electronics when he was at TIER, and was instrumental in the design of the charge controller. Melissa has been a great friend, housemate and colleague. Someday I hope to be able to do as many things as she does. Michael Rosenblum's help was instrumental in deploying the early Aravind links; he managed much of the logistics. The rest of the TIER group, Joyojeet, RJ, Bowei, Demmer, Matt, Omar, Jen, Renee, Assane, Kurtis, have been a great bunch to hang out with.

Overall the TIER group has given me three wonderful gifts. On the academic front, I completed a socially impactful and deeply fulfilling PhD. On the professional front, field work experiences led me to start a company, an exciting phase of life right now. On the personal front, I met my my fiancée, Divya, at TIER. As Dr. Aravind puts it, he's "glad to know that the (Aravind) project connected two urban people apart from connecting 100,000 rural people."

I could not have asked for an organization better than the Aravind Eye Care System (Aravind) as a partner in this work. Most significantly, Aravind transformed me internally. I have learned so much about compassion, hard work, contentment and sense of purpose from the Aravind staff across all levels. The late Dr. V has left an incredible mark on my life, living his life with integrity, fully aligned with his core values of service. I may never meet his standard of values, but working with him has inspired me to align my work with my own values. The current chairman, Dr. Nam has been an incredible support throughout. His efforts in leading Aravind on its mission to eradicate needless blindess have been tireless. He took a big risk by placing trust in me, then an unknown, untested, inexperienced student, for such a big task. Thulsi Uncle, Chitra Aunty, Dr. Aravind, Dr. Kim, Vara Aunty, Bala Uncle and Deepa have become my extended family in Madurai.

The real stars of the show at Aravind are Sashi, Saravanan and Rajan; they put in long hours with me in the hospital and on the field. We've missed buses, lunches, chai and sleep, but its all been worth it, and in the process I learned some Tamil to woo Divya.

Pavi Mehta probably does not know how influential she has been for this work. On extremely short notice, one day in August 2004, she arranged for me to visit Theni, where the seeds of this work were sown. Since then she's been a source of constant support.

My UCB friends Ambuj Tewari, Yatish Patel, Yury Markovsky, Steve Sinha, Mark Whitney, Nemanja Isailovic made the days in Soda Hall fun. Life in Berkeley would not have been the same without the Cal Squash Club. Playing hard against Dror Matalon, Kim Ben-Khoo, Ivo Plsek, Joe Farrell, Tobey Reiner, and Umesh Vazirani recharged me after long days at work.

Charity Focus and Karma Kitchen hold a special place in my heart. Viral Mehta, Pavi Mehta, Mark Jacobs, Yoo-mi Lee, Rahul Brown, Ragu Padmanabhan, Nisha Srinivasan, Nipun Mehta, Guri Mehta, Yaniv Cohen, Neerav Mehta, Trushna Majmundar, Roshni Kasad have all roped me into their many adventures.

Dan Belov and David Seidman have been great friends right from the CMU days, and have helped this work in their own quiet way, coming out all the way to presentations and practice talks. Ameet Jain, Sundeep Kapila, Arjunan Rameswaran, Susmit Sarkar, Sanket Malde, Ashwin Bharambe have been great friends since my IIT Powai days. I may have transfered to CMU after one semester, but it still feels like I never left IIT.

My parents, Santosh and Subhadra, have helped me more than they can ever imagine, and more than I can ever adequately describe. They showed tremendous courage and perseverance in rebuilding our lives after the invasion of Kuwait. I cannot remember a time when they made a decision that was not centered around my benefit; I can only hope to be as selfless as them. Their beautiful traits can be traced to both sets of my wonderful grandparents, who emphasized values of humility, non-violence and service, above everything else.

Every time I think about my chance encounter with Divya at the bus stop on Bancroft and Telegraph, I cannot thank my lucky stars enough. She has brought joy, peace and music into my life, and has always been tremendously supportive. She makes bad days good and good days even better by encouraging me when I need it, and by celebrating life's small victories with me, as they happen.

Chapter 1

Introduction

It is widely accepted that technology is, and has historically been, one of the great enablers for an improved quality of life. Despite the tremendous technological advances leading up to the 21st century, the world still faces many challenges, chief among which is widespread poverty. A key reason for this is that the majority of technology's benefits have been concentrated in industrialized nations and therefore limited to only a small fraction of the world's population. There is a vast disparity between industrialized nations and developing countries in all key indicators of development such as average income, mortality rates, education levels and life expectancies. For example, Table 1.1 shows that the 27 richest countries in the world overwhelmingly outperform the 50 poorest countries across several measures of development.

Among the various technologies, Information and Communications Technologies (ICTs), have been key drivers of the digital revolution that many of us in the industrialized world now take for granted. Network connectivity technolgies such as fiber, cellular and satellite to name a few, and applications such as telephony, e-mail and Internet have

Development indicator	High income OECD countries	Least developed countries
Population (million)	965	799
GDP per capita in US$ (PPP)	36,657	1173
Life expectancy (years)	79	55
Child mortality rate, under 5 years (per 1000)	6	141
Internet users (per 100 people	65.5	1.8
Phone penetration (per 100 people)	145	10

Table 1.1: Disparity in key developmental indicators between the 27 high income OECD countries and the 50 least developed countries in the world. Source: World Development Indicators 2008, World Bank [145]

transformed lives by improving access to information and services, enabling collaboration, and enhancing productivity.

It is with similar hope that the potential for positive impact of ICT [1] is seen in the context of development. Although ICT is not a panacea for the complex problems facing developing nations, it can be a *key enabler* for development when coupled with a broad understanding and a multi-disciplinary approach. As an example, the *Millennium Development Goals*, a set of targets defined by the United Nations to focus efforts on improving socio-economic conditions in the world's poorest countries, explicitly include the spread of ICTs to bridge the digital divide as an aim: "*Make available the benefits of new technologies - especially information and communication technologies*" [139].

A direct transfer of technology is simply not enough. "First world" technologies, originally designed for use in industrialized nations, are typically a bad fit for developing

[1] A very brief history of ICT for development, particularly in the context of networking, can be found in Appendix A for the interested reader.

countries. Attempts to transplant such technologies into developing world scenarios often fail because they ignore the stark differences with ground realities in developing countries: dramatically low purchasing power, low rural user densities, different usage models (higher emphasis on shared versus personal use of technology), limited supporting infrastructure (bad roads, lower quality electricity), and an initial lack of familiarity with the technology. Therefore, after initial deployment, such technologies have a greater risk of failing due to either financial or operational reasons.

The main theme of this thesis is to *develop appropriate communications technology that can overcome these barriers of cost, user densities, usage models and poor quality power, and that are financially and operationally sustainable to have real lasting impact*. In particular we have developed WiLDNet, a system that can provide very long distance, low-cost high-bandwidth point-to-point WiFi links and is particularly suited for targeted rural coverage. We have also developed various system-level mechanisms for fault mitigation, recovery and prevention, and solutions to address the poor quality of power that these networks frequently experience. We have deployed this system in partnership with the Aravind Eye Hospitals to enable live, high-quality, rural, video-based telemedicine for eye care. In the last three years, this system has enabled over 100,000 live remote consultations for patients in areas without any eye doctors; close to 20,000 of these patients recovered their eye sight following early diagnoses enabled by the WiLDNet network. The network is financially and operationally sustainable, and is now managed entirely by Aravind. It is being scaled by the hospital to 50 centers targeting 500,000 remote consultations annually and general access to eye care for 2.5 million people.

1.1 Defining sustainability

Sustainability is generally a poorly understand concept, in part because it has several different connotations in different contexts. Since the end of World War II, substantial mechanisms for providing multilateral aid to underdeveloped nations have emerged including the World Bank and many regional and national development programs. A primary concern for projects supported by these aid programs is sustainability. Although there is a growing belief that ICT can significantly impact development [41], in practice creating sustainable ICT projects is extremely difficult. A typical project consists of a pilot stage that aims to demonstrate the basic goals, followed by a deployment stage that aims for both scale and sustainability. Few projects even get to the latter stage, and many that do eventually fail.

The reasons for ICT failures vary, but at the core is an underappreciation of the many obstacles that limit the transition from a successful pilot to a truly sustainable system. In addition to financial constraints, there are operational issues such as power-related equipment failures, difficulties in local system maintenance, and an ongoing need for trained local staff, who often move on to better jobs. Further, researchers tend to focus on a deployment's sexier aspects, such as a novel technology, performance improvements, or a highly publicized pilot. But real impact requires a sustained presence in both deploying and maintaining ICTs, and this implies that every aspect of the system must be designed to be financially and operationally sustainable within the local context.

We argue that sustainability is a critically important systems design goal and in the rest of this section develop a more comprehensive and actionable view of sustainability.

At a high level, sustainability implies pushing for low-cost, low-power and easy-to-use and maintain solutions. Although worthy goals, they are too high level to be very helpful, and we refine them in this discussion.

Our experiences from the work described in this dissertation indicate that any ICT system must be designed to exhibit three important principles to be sustainable: *optimization of an existing system*, *financial self-sufficiency*, and *operational self-sufficiency*. Throughout this thesis, we describe our efforts in meeting these principles, and in particular discuss our impact on sustainability in Chapter 8. In the next three sections, we define these three principles.

1.1.1 Optimization of an Existing System

Development projects are not deployed in a vacuum; there is almost always an existing approach for a given task, even if it is a poor one. We have found it extremely valuable to view our work as optimizing the existing system rather than deploying something from scratch. There are three key reasons for this:

- the local community understands the needs and motivation for the existing system, so variations on it require less communication and education;
- achieving community buy-in, which is fundamental to sustainability, is easier; and
- it is rarely clear to outsiders why the current system is designed the way it is, which means larger changes might fail due to unforeseen circumstances.

For example, users in Ghana rejected a new medical record system in part because it brought in too much transparency; lab workers secretly depended on performing lab tests on the side for extra income [114]. The optimization principle promotes a kind of humility that admits we cannot fully understand the existing problem space, and thereby prevents "blank slate" approaches that are unlikely to succeed in practice.

1.1.2 Financial Self-sufficiency

Any deployment aiming for sustainability must be cash-flow positive. Projects that fail to recover at least operating costs tend to fail overall, especially when outside funding diminishes at the end of a pilot. The goal of positive monthly cash flow is easier to achieve than profitability, which implies enough income to recover costs for capital investments. Although profitability is more desirable, maintaining positive cash flow is hard enough in rural areas, making it a reasonable target for development projects. Under this view, aid can be used for start-up costs but not for ongoing operations. However, we can frame certain capital expenditures such as PCs as a monthly cost in addition to the operating costs, by dividing the up-front cost into equal monthly fees over the item's expected lifetime. This calculation ignores the interest that money could have earned, which we estimate to be 8 percent. For example, a PC that costs \$600 and lasts five years translates to \$10 per month without interest and $\$600 * (1.08)^5/60 = \$14.70$ per month with interest. But in developing regions, where old components are reused, the PC would not have zero value at the end of five years. Factoring in a salvage value of 20 percent reduces the monthly fee proportionally to \$11.76 per month.

1.1.3 Operational Self-Sufficiency

A major operational issue is ongoing system maintenance and support. This includes power, hardware, and software as well as expansion and new installations. In practice, local groups do not start out with the ability to handle these tasks well. Also, the remoteness of deployment locations makes it hard for experts to help. Thus, three basic design goals for operational sustainability are:

- increased component robustness,
- easy-to-use management tools for local staff, and
- tools for remote management by experts

In a sense, these goals can be addressed by higher operating budgets that can afford robust high-end equipment, stable power sources and highly trained staff; however this is not viable for most grassroots rural organizations.

We have developed various low-cost techniques across all layers of the system to achieve these goals. As a complementary approach, we have also added a middle tier for operational management, between local staff and remote experts (in this case, us), which uses local IT vendors as a second level of support. We train both local staff and local vendors, but the latter have more IT background and can thus handle more issues and ultimately reduce the need for remote management. The vendors charge per-visit for their services, but the system-level tools limit the need for this intervention and therefore, further reduce operational costs. We include these costs in our financial sustainability accounting.

1.2 Thesis Contributions

In this thesis, we address the research challenges that arise in building *low-cost*, *long-range* wireless networks for *low-density* rural areas. Our overall goal in designing such networks is achieving *sustainability*, enabling these networks to scale beyond small pilot deployments and to have large-scale real-world impact in developing regions. We primarily focus on WiFi-based networks given that WiFi is much cheaper than other wireless technologies and that it also operates in unlicensed spectrum, both of which allow for deployment by grassroots organizations. Some of the early works in this space by Bhagwat *et al.* [36] and Raman *et al.* [118] focus on tailoring the 802.11 MAC protocol to achieve high-bandwidth in such settings. Although this is indeed relevant and we build on these efforts, they represent a small part of a much larger puzzle.

In our work, we take an end-to-end systems perspective of the overall challenge: *How can we engineer a sustainable WiFi network that provides high-throughput in long-distance settings, in the face of lossy environments, and in the presence of systemic link or node failures?* Answering this question involves tackling research challenges at various layers. We have addressed problems ranging from poor quality grid power at the base level, to the unsuitability of the standard WiFi MAC for high-throughput long-distance links, to the need for a set of system management tools to allow local rural staff to administer the network.

In particular, we present 6 contributions that fall under three main categories: *enabling rural connectivity*, *enabling rural IT management*, and *a real world deployment.*

1.2.1 Enabling Rural Connectivity

Our *first contribution* is a systematic study investigating the commonly observed drop in performance seen when the standard WiFi Medium Access Control (MAC) protocol is used for long-distance links [129]. In particular, we study losses induced by inefficiencies in the standard MAC protocol, and by loss characteristics of the channel. We show that the presence of external WiFi interference leads to a significant level of packet loss in WiFi-based long-distance (WiLD) links. This is in sharp contrast to the work done in the Roofnet project where the authors observed that multipath was the most significant source of packet loss [26]. Since our work, a more recent study [54] has reexamined the Roofnet data and has shown that external interference and not multipath was the major source of loss.

In this study, we also explore the solution space of loss recovery mechanisms and propose a range of MAC- and network-layer adaptation algorithms to mitigate the protocol- and channel-induced losses. Key lessons from this effort led to our *second contribution*: the design of a Time Division Multiple Access (TDMA) MAC protocol for high performance WiLD links, namely *WiLDNet* [111]. With WiLDNet, we make essential MAC-layer modifications compatible with standard low-cost network cards; thus it is *easily deployable*. We replace the original packet collision avoidance scheme of CSMA/CA [68] with TDMA, and improve channel utilization and loss recovery with bulk acknowledgments. Using WiLDNet, we were able to achieve a throughput of 6 Mbps on a 382 km single-hop link in Venezuela, which is a *new world record* for the longest WiFi link achieved without any custom antenna design or active amplification. The previous record at 279 km achieved a throughput of only 50 – 60 Kbps using the standard MAC. Other researchers have also verified the high-performance achieved by WiLDNet over a 133-kilometer link in Italy [50]. These are a significant results since they imply that, in practice, we are now no longer limited by poor performance of WiFi over long distances in rural networks.

1.2.2 Enabling Rural IT Management

Armed with a solution for a low-cost high throughput network, we deployed it at the Aravind Eye Care System (Aravind) [4] in rural South India to create a rural telemedicine network. During the deployment, we encountered several operational challenges that threatened the viability of the network. Many of these challenges, such as the *extent* of low-quality grid power, were very surprising. Thus, our *third contribution* is a broad systemic view of the problem, documenting operational challenges in an unfamiliar rural context. These include network faults resulting from components failing with high frequency due to poor-quality grid power, difficulties with diagnosis in the absence of local expertise and reliable connectivity for remote experts, and challenges in predicting faults.

This analysis leads to our *fourth contribution*, which is a range of very low-cost solutions for several aspects of the network including monitoring, power, alternate network entry points for backchannel access, automatic recovery mechanisms, and safe software. Through this suite of tools, we have improved the availability and therefore the operational sustainability of the network to a point where we are no longer involved in administering the deployment. In addition to our mechanisms that enable automatic recovery, local rural staff at Aravind use our tools to perform daily management and maintenance of the network.

1.2.3 Large Deployment with Real-world Impact

Our *fifth contribution* is a real-world telemdicine network that addresses a very real and pressing problem in the developing world: *the acute shortage of doctors*. Using WiLDNet links we connect Aravind Eye Hospital to ten *doctor-less* rural villages as far as 40 km away. In three years of operation, our network has enabled over *100,000* remote consultations for rural eye patients who had no other access to an eye doctor. Of these patients, close to *20,000* were diagnosed with severe eye problems and were able to regain their sight due to early diagnosis. With our emphasis on inexpensive WiFi technology supported by system-level mechanisms for management and fault mitigation, the network at Aravind has achieved financial and operational sustainability as defined in section 1.1. This has led to both *scaling* and *replication* of our work.

Aravind is currently scaling the network to connect 50 centers targeting $500,000$ remote consultations a year and access to eye care for around 2.5 million people. Given this successful transition, the network is being replicated by other organizations such as Lumbini Eye Institute [93] in Nepal and the Pakistan Insitute of Community Ophthalmology [107]. In addition, Inveneo [77], a non-profit technology group, has deployed WiLDNet in production systems for Internet and voice-over-IP services in several Ugandan villages.

Finally, our *sixth contribution* is a broad set of lessons for rural ICT research and deployment. Ours is one of the few early works in this space; we have faced several obstacles on the ground and have learned many lessons during our journey. We distill our experiences into four main lessons. First, *sustainability has to be a key systems design goal* if the ICT deployment is to have any real chance of success. Throughout this dissertation, we demonstrate how various design choices were influenced by considerations of sustainability. Second, ICT systems designers must design their systems to have a *reduced* need for trained staff. Although training rural staff well is important and necessary, once trained they eventually leave for better opportunities. [2] Therefore the system must cope with this turnover via easy-to-use system-management mechanisms for periodically incoming new staff with zero or limited training. Third, *simple redesign* of standard components can go a long way in achieving high performance or operational gains at a lower costs; we describe our approaches to both software and hardware in this regard throughout the dissertation. And fourth, the real cost of power for rural IT projects is not the price paid to the utility for grid electricity, but is the cost of *overcoming poor quality grid electricity*, which can increase the cost by 24 times. This is a *major* finding; most groups neither budget for it nor even realize they need low-cost power solutions similar to those described in this dissertation. We expect these lessons are broadly applicable to other ICT for rural development projects, and will save other researchers major effort.

1.3 WiLDNet Architecture

For rural areas in developing countries, wireless infrastructure appears to be the first kind of infrastructure that is affordable, and can arguably lead to sufficient increase in rural incomes, making other infrastructure investments in water and power more viable. Given the low densities of users in rural areas, approaches that provide blanket wireless coverage are not feasible. We propose the more viable approach of targeted coverage to populated areas using long-distance point-to-point links. Such links can rely on WiFi,

[2] We and Aravind view this as a good outcome

point-to-point WiMAX or other long-distance wireless technologies.

For practical and cost-related reasons, we have chosen WiFi-based links. WiFi radio cards are inexpensive and widely available. Since WiFi generally operates in unlicensed spectrum, it has no spectrum fees and is easy to deploy. For around $800, WiFi cards can be combined with off-the-shelf passive directional antennas on single-board computers to create line-of-sight links over several hundreds of kilometers. In addition, there are many open-source drivers available for the radios, allowing us to redesign the standard WiFi MAC to achieve high bandwidth over such long distances.

In the rest of this section, we provide a very basic overview of WiLDNet by first describing the principles that drive its design and then by presenting the various architectural components comprising it. We propose this architecture as an alternative to current solutions such as cellular and wireline technologies in rural areas.

Although WiFi is an inexpensive option for long distance links, it was originally designed for "hotspot" usage in short-range indoor broadcast environments, and therefore fails to perform adequately in long distance rural settings. For example, some of our early long distance links achieved negligible TCP throughput despite the absence of interference. Adapting WiFi to long distance settings by redesigning the MAC protocol to be able to continue leveraging existing inexpensive hardware is one of our main research challenges. Other challenges include building system level-mechanisms to increase the reliability of these links and also low-cost stable-quality power.

1.3.1 Design Principles

Our overall focus in designing low-cost rural connectivity solutions is to achieve sustainability, both financial and operational. This view drives the the following four main design principles of the WiLD architecture.

- **Low cost**: The solution needs to have both low capital cost and low operational cost. Capital costs include the cost of equipment and the cost of initial installation. Operational costs include power and maintenance, but are often not anticipated as it is difficult to predict the overhead in providing good quality stable power or to appreciate the challenges of network maintenance in rural areas. Given the low purchasing power of the small number of users, there is even more pressure on the network to be very low-cost. As an example, rural clinics that we have connected to central hospitals for telemedicine typically have an annual cash flow of $1500 ignoring all network costs [137]. Therefore, the solution needs to cost less than this amount per year.

- **Support grassroots deployments**: We target solutions that can enable small organizations such as rural non-profits, hospitals, and schools to start small-sized network deployments that can later be scaled as demand and revenues increase. Technologies that require expensive base stations are not suitable as they require a large user base to amortize costs. Technologies that operate in unlicensed frequencies are preferable as they can be set up by the grassroots organizations as needed, without dependence on telecom carriers or government clearance.

- **Intranet connectivity**: The solution must support local connectivity between centers as much as, if not more than, the need for Internet connectivity. Internet content, mostly in English, may not always be relevant for the local population due to language reasons. Instead, applications such as voice and video using high-bandwidth intranet connectivity among villages in a larger area can enable many local services using local resources such as telemedicine, agricultural consultations, market pricing information and remote education. Such local applications, if implemented over the Internet, would incur unnecessary international bandwidth charges.

- **Local maintenance**: Any system for rural connectivity should have mechanisms that support local maintenance. A key challenge for longer term sustainability of rural networks is a shortage of trained staff. Reliance on remote experts to fix network issues is very expensive. Training of local staff only helps some, as trained staff leave for better jobs. Poor power damages equipment often and isolated locations make timely maintenance a challenge. Therefore, along with training of staff, the system itself needs to have support for easy monitoring, diagnostics, fault recovery and stable power.

1.3.2 Architectural Components

With these design principles in mind, we propose a network architecture comprising long distance point-to-point links in multi-hop configurations connecting sparsely spread groups of villages together. The network architecture also comprises system monitoring mechanisms at various levels, and has solutions to address poor-quality power.

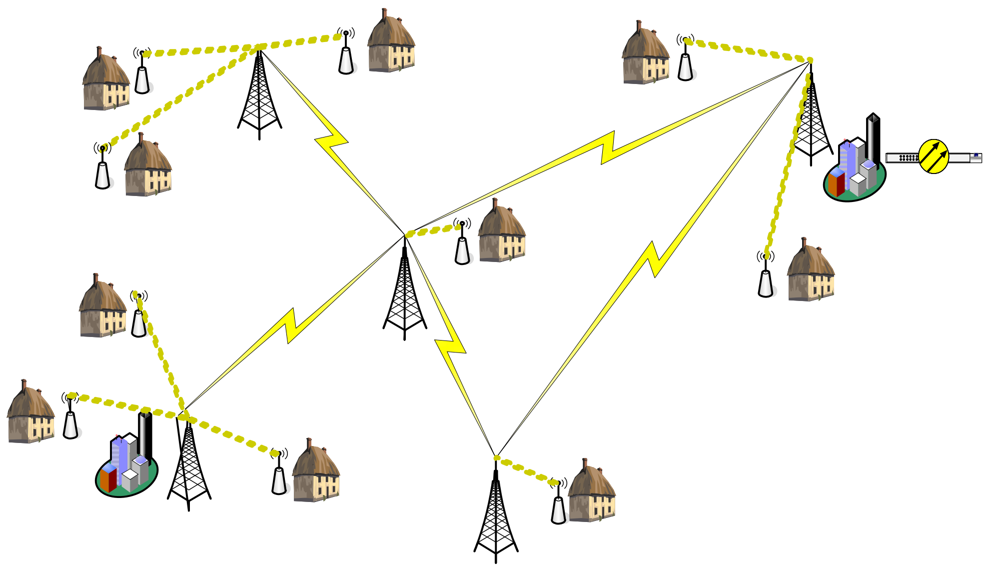

Figure 1.1: WiLD point-to-point network connecting several points over multiple hops.

Point-to-Point Links: Long distance point-to-point links links, ranging in distance from 10 to 100 kilometers, provide high-bandwidth connectivity. Each network node has multiple high-power radios, each with its own directional antenna. Two end points of a link need to have direct line-of-sight and are therefore usually mounted on towers or masts on top of existing structures. The main feature of these links is a novel MAC-layer protocol with better loss recovery mechanisms, and receive/transmission synchronization resulting in high end-to-end bandwidth, even in multi-hop settings.

Shorter range access links to villages can be shorter range point-to-point links. However, these links can also be improved dramatically via steerable antennas and additional MAC modifications; we discuss these future directions in Chapter 8.

System Monitoring: System management poses a steep learning curve to anyone not experienced at it. Early on, we created a tiered model support, in which we trained local IT vendors to provide a middle tier of support, between the the local staff, the first tier of support and the remote management team (us), the third tier of support. Gradually we built our tools around this ecosystem, with specific tools targeted for specific tiers, with a view to eventually migrate all responsibility to the first tier of support, the local staff. To support local and remote management, we need to support the following types of operations: a) running active and passive network measurements, b) collecting the readings of these measurements and other logs, and c) distribution of remote software and configuration updates. To enable the tiered model of operation, our system needs to provide appropriate configuration, monitoring and diagnosis tools for each tier. Figure 1.2 describes our management system in terms of the three tiers.

At the base level, we provide a suite of simple automatic recovery mechanisms such that the system can recover from basic and localized errors, such as a hung OS or disk full errors. An important feature of automatic recovery is triggered reboot (using a watchdog processor) of the system on detecting a software malfunction.

Next, we provide some core utilities that enable remote management for both the second and third tiers. First, after having years of trouble reliably collecting log information from remote nodes, we developed an asynchronous file transfer system based on delay-tolerant networking (DTN). Second, we created a more reliable way to SSH into remote nodes, called PhoneHome, that initiates management connections from inside the network.

Power: Although the problems of power outages in rural areas are well known, what is surprising is the *extent* of poor quality electricity when power is available. It is not

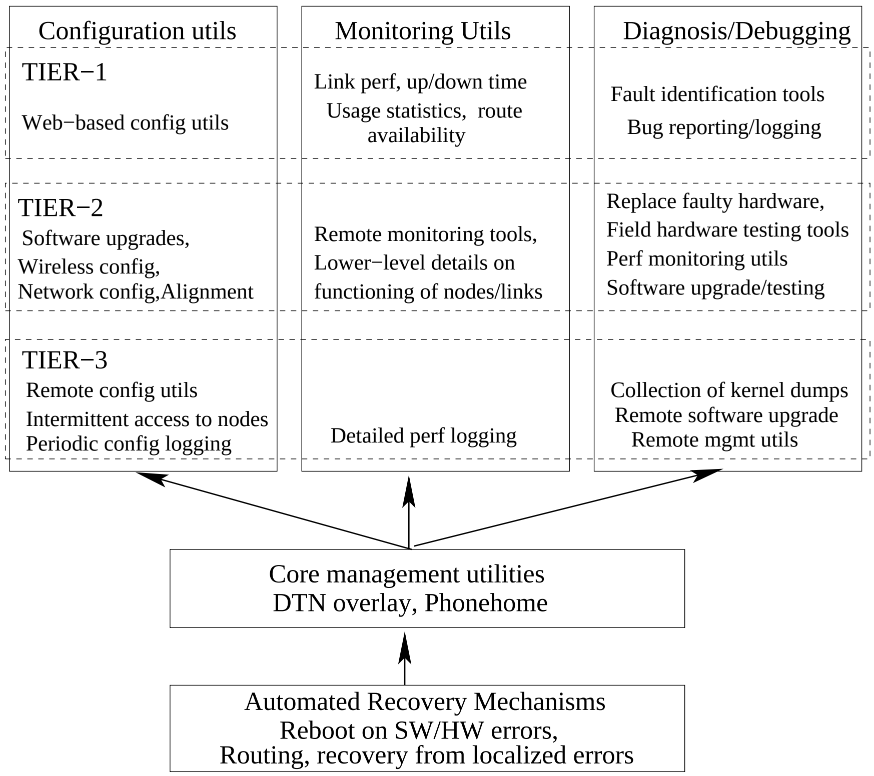

Figure 1.2: Description of System Monitoring

uncommon to see 1000-volt spikes and extremely low voltages from the grid. Such voltage characteristics either damage equipment or cause it to hang, an expensive problem to fix since it requires traveling to the isolated location to reset it. At the lowest level of the system, we provide inexpensive low-voltage disconnect circuits and low-cost charge controllers for battery-connected off-grid solar systems with maximum power point tracking that provide stable power to the equipment. These have been designed after careful study of the electrical characterstics from the grid in our rural deployment locations.

1.4 Thesis organization

In the next chapter, we argue that WiLD networks can enable practical rural connectivity for grassroots organizations. We first introduce our main rural partner organization, the Aravind Eye Care System. We discuss its connectivity needs and how prior approaches were limited in impact. Second, we explore the tradeoffs between various connectivity options and make the case for WiLD deployments in rural areas. Third, we provide a very brief description of some of our early experiences operating WiLD links at Aravind. Specifically, we motivate the need for addressing challenges not just at the network layer, but also from an end-to-end systems perspective to improve the overall availability of the system in the face of various failures typically seen in rural deployments. Finally, we provide several examples of grassroots organizations attempting similar deployments to illustrate that this is indeed a general problem, and therefore our work is widely applicable.

In Chapter 3, we first provide an overview of the physical (PHY) and medium access (MAC) layers of the legacy 802.11 family of standards, originally designed for wireless LANs in indoor broadcasts. Second, we illustrate that 802.11 has outgrown its intitial intended use and has become a popular choice for outdoor wireless mesh and WiLD networks. Third, we provide an overview of three techniques that can be used to improve the range and throughput in WiLD links: 1) extending the range of WiLD links via the use of high-power radios, high-gain atennas and careful link path planning, 2) extending the SNR (and therefore range) and throughput via PHY-layer optimizations to 802.11, including the use of variable channels and multiple paths, and finally, 3) extending throughput via MAC-layer optimizations such as changing timing parameters and enforcing synchronization among multiple co-located radios. We show that MAC-layer parameter optimizations are not enough to improve throughput in WiLD links and that new MAC-layer designs that incorporate simultaneous synchronized transmit and receive are necessary.

In Chapter 4, we characterize, via a detailed measurement study, the throughput and loss seen when the standard 802.11 MAC protocol is used in WiLD links. First, we identify the fundamental shortcomings of the 802.11 MAC protocol itself in long-distance settings: *inefficiency link-layer recovery*, *collisions at long distances*, and *inter-link interference*. Second, we analyze three well-known causes of channel losses: *external WiFi interference*, *non-WiFi interference*, and *multipath interference*. Among these we show that external WiFi interference is the most significant source of packet loss in WiLD environments while the effects of non-WiFi and multipath interference are not significant. Third, we analyze the variability in packet loss over WiLD links in both urban and rural settings. Fourth, we measure the effect of protocol- and channel-induced losses on end-to-end TCP performance, and how various link parameters affect the TCP throughput performance. Finally, having identified external interference as the primary source of losses in WiLD links, we evaluate three potential remedies to mitigate these losses: *802.11 frequency channel adaptation*, *802.11 PHY datarate adaptation* and *adaptive forward error correction (FEC)*.

In Chapter 5, we describe the design of WiLDNet and elaborate on how it addresses the 802.11 protocol shortcomings as well as achieves good performance in high-loss environments. In the previous chapter, we identified three basic problems with 802.11; (a) low utilization, (b) collisions at long distances, and (c) inter-link interference. To address the problem of low utilization, we propose the use of bulk packet acknowledgments (Section 5.1). To mitigate loss from collisions at long distances as well as inter-link interference, we replace the standard CSMA medium access method with a TDMA-based approach. We build upon the TDMA protocol design of 2P [118], and make the necessary changes to adapt 2P to high-loss environments. Additionally, to handle the challenge of high and variable packet losses, we design adaptive loss recovery mechanisms that use a combination of FEC and retransmissions with bulk acknowledgments.

In Chapter 6, our overall objective is to characterize the operational faults, many of which were surprising and unexpected, that we experienced in hitherto unfamiliar rural settings. We start by first discussing why faults seem more likely in rural settings. We then describe our rural deployments at the Aravind Eye Hospital and the AirJaldi community network. Next, we present our operational experiences with these two networks, particularly the types of faults seen and the difficulties posed to system management from decreased component robustness and difficulties in diagnosing and predicting faults.

In Chapter 7 we first discuss our tiered support model and then present five key aspects of our network management system: monitoring, power, backchannels, independent recovery mechanisms, and software. Each aspect has been designed to specifically address the following goals: *increasing component robustness*, *enabling fault diagnosis*, and *supporting fault prediction*. For each aspect, wherever appropriate, we also discuss tradeoffs affecting our design choices.

In Chapter 8, we present the real-world impact of our WiLDNet deployment at the Aravind Eye Care System for telemedicine between the Aravind hospitals and various rural vision centers. We first look at the number of patients that have benefited. Next, we demonstrate that the networking infrastructure contributes to the financial self-sufficiency of the overall telemedicine setup, resulting in its continued operation and patient benefit. Finally, we look at operational self-sufficiency, particularly in terms of the migration of operational responsibility to local staff as an indicator of local capacity building.

In Chapter 9, we first discuss the limitations of our work. Next, we propose a few promising areas for future work. We then provide a detailed summary of our contributions and close with the lessons learned during this effort.

Chapter 2

Motivation for WiFi-based Long-Distance (WiLD) Networks

Historically, profitable options for rural network connectivity in developing countries have been limited. High deployment costs turn operators away from willingly providing service in low-density lower-income rural areas where it is hard to recoup initial costs. This has long hindered the possibilities of connecting rural areas to urban areas for the provision of much-needed services such as telemedicine, distance education, e-governance, and general access to information. In this chapter, we make the case for adapting inexpensive WiFi technology to network topologies comprising point-to-point WiFi-based long-distance (WiLD) links as a viable initial choice to introduce targeted connectivity into rural areas from urban centers.

The nature of computer systems research specifically targeted for rural development is such that it requires deep partnership with rural organizations on the ground to understand their technology needs and constraints, enabling researchers to design appropriate systems. [1] Therefore, we first introduce our main rural partner organization, the Aravind Eye Care System (Section 2.1). We discuss its connectivity needs and how prior approaches were limited in impact. Second, we explore the tradeoffs between various connectivity options and make the case for WiLD deployments in rural areas (Section 2.2). Third, we provide a very brief description of some of our early experiences operating WiLD links at Aravind. Specifically, we motivate the need for addressing challenges not just at the network layer, but also from an end-to-end systems perspective to improve the overall availability of the system in the face of various failures typically seen in rural deployments (Section 2.3). Finally, we provide several examples of grassroots organizations attempting similar deployments to illustrate that this is indeed a general problem, and therefore our work is widely applicable (Section 2.4).

2.1 Aravind Eye Care System

The Aravind Eye Care System (Aravind) [4] is a group of 5 eye hospitals in the state of Tamil Nadu in southern India. These hospitals are located in Madurai, Theni, Tirunelveli, Coimbatore and Pondicherry, all of which are urban or semi-urban centers (figure 2.1). By volume, it is the largest eye care service provider in the world. In 2005 alone, Aravind saw 1.7 million patients and performed 245,000 eye surgeries, most of which were cataract surgeries [29]. Aravind's stated mission is to eradicate preventable blindness. Preventable blindess, also commonly known as "needless blindness", is defined as blindness for which simple treatments exist but may not be available to patients for reasons such

[1] This has been particularly important and useful for us as most of our technology perspectives were largely shaped by experiences in industrialized countries

Figure 2.1: The Aravind Eye Care System is a group of 5 hospitals in the state of Tamil Nadu in southern India. The hospitals are located in the districts of Madurai, Theni, Tirunelveli, Coimbatore and Pondicherry. The figure shows the current bed capacity (for both paying and free patients) and year of founding for each hospital

as a lack of accessibility or awareness. The most common causes of needless blindness are refractive errors and cataracts, which can be treated with prescription glasses and cataract surgeries respectively. India shares the largest burden of preventable blindness globally; of the 45 million people needlessly blind globally, India has 15 million [89]. About 70% of India's population is rural, where the risks of blindess are highest and the access to eye care services is the lowest [106].

Aravind places major emphasis on providing rural eye care to address this disparity. Known for its high volume and high quality of care, Aravind uses income from higher-income patients to heavily subsidize costs of the same quality of treatment for its lower-income patients; 40% of Aravind's patients are paying patients while the rest receive free or heavily subsidized treatment [119]. Having significantly brought down the costs of eye care, Aravind is mainly limited by its ability to reach rural patients from its urban centers.

2.1.1 A Shortage of Doctors

The primary limitation to eye care is a severe shortage of trained doctors and nurses in developing countries, particularly in rural areas [83]. For example, India has only about 10,000 ophthalmologists serving a population of about 1 billion, or roughly one ophthalmologist for every 100,000 people [58], a very low overall ratio. This distribution is further biased against rural areas. With 90% of Indian ophthalmologists based in urban areas [58] and 70% of India's total population residing in rural areas, there is only one ophthalmologist for every 700,000 people in rural areas. Therefore, rural patients typically must travel much longer distances to clinics or hospitals to get care. Travel expenses, even if not large in absolute terms, can be significant fractions of rural patients' incomes, and as a result they are unable to get or simply decline treatment. For example, one patient in need of cataract surgery had gone seven years without treatment until a local rural Aravind center enabled by our work opened nearby; until then he was effectively blind.

The long-term solution is to increase the number of doctors but that will take decades, involving policy changes and heavy investments. In the short-term, a practical approach is to improve the utilization of the existing set of doctors by extending their reach into rural areas via services such as telemedicine.

2.1.2 Prior Approaches to Extending Rural Reach

The first approach taken by Aravind to provide rural eye care was to periodically conduct "eye camps" in rural areas. A team of Aravind doctors would set out to a remote area and conduct comprehensive eye exams for an entire day. For example, in 2005 Aravind conducted 1335 eye camps with an average of 327 patients per camp [29]. Although clearly

Figure 2.2: Each vision center is a rented room in a rural home and is staffed by a technician and a counselor, both trained specifically by Aravind. PC with webcam for videoconferencing enabled by our work can be seen on the left

an improvement, this approach reached at best only 7% of the target population [30]. The primary reason for the low turn out was the difficulty in raising awareness. The transient nature of the camps also led to delayed treatments as patients would wait for the camp to arrive to save on travel costs, rather than actively seek care. It also resulted in very limited opportunities for patient follow-up as the eye camps would visit each village only a handful of times a year. Finally, physically moving the doctors around did not really address the underlying issue of the low utilization of doctors.

To increase the utilization of doctors, Aravind has adopted the *vision center* (VC) model, in which doctors remain at the 5 base hospitals, but interact with rural patients over a communication network. As shown in figure 2.2, a vision center is typically a room rented by Aravind in a rural family's home in the village. It is equipped with some basic ophthalmic equipment (slit lamp, refractive lens kit) and a PC with a webcam. The center is staffed by two people: a technician that operates the ophthalmic equipment and PC, and a counselor that follows up with patients based on the diagnosis of the remote doctor. Generally, neither staff member has a degree or even a broad technical skill set: they are specifically trained for 6 months by Aravind for their roles.

At a vision center, a patient walks in and gets some basic tests (refraction, cataract examination) done by the technician. The counselor presents the results by talking to the doctor at the base hospital via telemedicine, after which the patient talks with the doctor. The counselor then implements the doctor's advice, such as handing out prescriptions, filling out referral forms or creating glasses. If advised by the doctor, the patient is referred to the base hospital for further examinations or treatments such as cataract surgery. The cost to the patient of a consultation visit to the vision center is 25 cents. Cataract surgery, if required, can be up to \$60 at the hospital.[2] However, about 60% of the patients are low-income and are not able to pay; they receive surgeries for free and are subsidized by paying patients. Each vision center is expected to cover a target population of 50,000. A typical VC makes between \$200 to \$2000 in annual operating income, not including the operating cost of the network [137].

By 2005, Aravind had set up 3 vision centers based on CorDECT wireless local-loop (WLL) technology [33]. This technology was supplied by a for-profit local operator focused on rural connectivity. CorDECT operates in a licensed frequency range of 1880 – 1930 Mhz. It comprises a base station on the infrastructure side and customer premises equipment (CPE) units on the customer side. CorDECT has a range of up to 8 kilometers. It provides 32 kbps voice channels in addition to simultaneous 35 kbps Internet. In many

[2]A surgery of comparable quality in the US costs about \$2000.

ways, corDECT was a technology breakthrough for rural connectivity in terms of cost-effectiveness. The per-line cost, or capital cost in other words, for an operator to provide a rural corDECT connection was $450, a 50% reduction compared to the $900 per-line cost to provide a telephone line [82]. As a result, CorDECT enabled connectivity for Aravind to a certain extent, but it still did not enable financial and operational sustainability for Aravind for several reasons that we discuss next.

- **Performance**: With a total bandwidth of 35 kbps per site, it was not surprising that the quality of the video-conferencing was poor, although the audio did prove useful for communication. As a result of the poor quality of video-conferencing, patients would still prefer to travel to the base hospital but would wait until their condition worsened considerably to delay their travel expenditures. This led to higher costs of treatment for the patients, and also lower vision center revenues for Aravind.

- **Costs**: The subscriber unit cost to Aravind was around $700 [1]. In addition, the use of licensed frequencies resulted in the operator charging service fees of $500 per year per vision center at Aravind. This was a significant fraction of the centers' operating income, ranging from 25% to 100%.

- **Operational Freedom**: Operating through a carrier limited Aravind's ability to start centers in areas in dire need of eye care. Around Theni in Bodi and Chinnamanur, for example, despite being ready with trained personnel for the centers, Aravind was unable to start centers in several locations because the carrier did not consider those areas profitable enough to deploy a base station; Aravind would have been the sole paying customer. This prevented viable service for rural patients in those areas.

2.2 The Case for WiFi-Based Long-Distance (WiLD) Networks

In this section, we look at alternative technologies for high-bandwidth (5 to 50 Mbps) network connectivity in developing regions. Three key features of rural areas constrain the deployment model and choice of technology: *low population density*, *low purchasing power* and *uncertain initial demand*.

2.2.1 Other Technology Options: Fiber, Cellular and Satellite

Although there are a number of technologies that can be used, some technologies such as fiber or cellular cannot amortize costs over a small number of low-income customers. Other technologies such as CorDECT [33] and ham radios are inexpensive, but provide low bandwidth, inadequate for sharing among any users or for providing high-bandwidth services such as telemedicine. Having already discussed corDECT, we discuss fiber, cellular, WiMAX and satellite technologies next.

Fiber: The default choice for high-bandwidth connectivity when there is very high demand for bandwidth (gigabits per second) is fiber as it provides nearly unlimited bandwidth. However, when demand is low, it is not the right choice due to high costs. For example, the cost of laying fiber in India is quite high at $1000 per kilometers [40], and cannot be recouped from low-density low-income areas with low initial demand. In such cases, we argue that it is more feasible to build wireless backbones, and when high demand gets established, these links can be upgraded to fiber. Wireless infrastructure seems to be the first viable networking infratstructure that can be deployed in rural areas.

Cellular: The most visible communication technology in developing countries has been cellular telephony, which has seen tremendous growth in the last few years. With 1.15 billion mobile phones sold worldwide in 2008, penetration has increased fastest in India, Brazil, China and sub-Saharan Africa. The flagship example of the impact of cellular telephony is Grameen Telecom [55] in Bangladesh, where women in villages use micro-credit to buy phones as franchisees and then operate it as a pay phone by renting it to others in nearby locations. At its peak, Grameen Telecom had 95,000 franchisees that provided coverage to around 60 million people in 50,000 of Bangladesh's 68,000 villages.

Networks with expensive base station models need to amortize the high deployment cost over a large number of users. In low density regions, such base stations simply do not cover enough users to be economically viable.

Therefore the expectation that cellular solves the connectivity problem for developing regions is somewhat of a myth. Cellular success in developing regions is still largely an urban phenomenon, with few exceptions. For example, even the best known "rural" cellular system, Grameen Telecom, avoids rural-only base stations. Instead, by exploiting the high population density of Bangladesh, Grameen places base stations such that they cover both higher income urban users and lower income rural users, thereby subsidizing their costs of rural coverage. Typically, there is no coverage for areas that are not near an urban base station. China is one country with good universal coverage, but it is dictated by strong government policy despite the economic difficulties in providing such coverage.

Other countries either subsidize rural users via taxation, much like the US universal coverage tax, or mandate rural coverage as part of urban spectrum allocation. Thus, many cellular providers incur losses in low user-density regions and partially recoup these losses by either charging very high usage rates or imposing a universal service charge on all users.

WiMAX: The WiMAX standards [25, 74] are still in flux, and most of the current effort appears to be on enabling high-bandwidth mobile urban users, leading to expensive higher-capacity and higher-power base stations. This reduces the chances that WiMAX equipment may be used for rural solutions. In addition, WiMAX, with the intended deployment model of expensive base stations covering many users, shares the shortcomings of cellular technologies.

Satellite: A Very Small Aperture Terminal (VSAT) station is a two-way satellite ground station with a dish antenna that relays data to other earth stations via satellites in geosynchronous orbits. VSAT data rates vary depending on technology, but they typically range from narrowband to at most 4 Mbps. The upload and download channels are asymmetric, with higher download rates. VSAT networks can be configured as star topology in which a large, central hub station connects to small terminals, or as a mesh toplogy where VSAT terminals communicate directly with other VSAT terminals.

The biggest advantage of VSAT is great coverage; VSAT networks can be deployed almost anywhere, independent of wireline infrastructure. However, VSAT systems are very expensive, with installation costs as high as $10,000 and recurring costs as high as $2000 for 1 Mbps. Therefore, VSAT is mostly available for well-to-do private businesses in rural areas, or is donated by governments to rural organizations. In fact, Aravind Theni is connected to

Aravind Madurai via VSAT donated by the Indian Space Research Organization. Since the equipment is donated, there is no clear responsibility for maintenance, and the link suffers frequent downtimes; we have measured downtimes of 30%.

2.2.2 WiFi

In 2008, 387 million WiFi chipsets were sold globally. The widespread popularity of WiFi, its rapid adoption all over the world, and continuing double digit growth in chipset sales volumes have been primarily driven by commoditization and standardization. High volume production resulting from commoditization has brought the prices down, making WiFi-based technologies very attractive for developing world markets. The standardization of protocols and use of unlicensed frequencies have enabled interoperability between different vendors, further driving adoption.

It is therefore not surprising that the use of WiFi has far outgrown its originally intended use of wirelessly connecting devices in indoor broadcast environements. [3] WiFi has now become the most popular choice for building outdoor community networks and long distance rural networks where the distances are orders of magnitude longer (one to hundreds of kilometers) than distances within offices.

Inexpensive WiFi radios can be combined with directional antennas to create point-to-point links of distances as long as 100 kilometers. The disadvantage is that the WiFi protocol is ill-suited to operate in such settings, and performs very poorly. However, we can redesign the protocol to provide high bandwidths independent of the distance of the link, and this is a major focus of this dissertation (chapters 4 and 5).

[3] The earliest use of WiFi was to connect cash registers at check out lanes

The key advantages of WiFi are that it is inexpensive, consumes very little power, provides operational freedom by using unlicensed frequencies, and can be modified to work very well in point-to-point long-distance settings suitable for targeted rural coverage.

2.3 The Need for an End-to-end Systems Perspective

In this section, we demonstrate the need for an end-to-end sytems perspective that addresses not just the challenge of making WiFi work well over long-distances, but also addresses various non-WiFi system-level issues ranging from power, monitoring, backchannel access and failure recovery.

2.3.1 Our WiLD Deployment at Aravind: A Very Brief History

In January of 2005, we established our first WiLD link from the Aravind hospital in Theni to the Bodi vision center. This link used a novel TDMA MAC-layer (chapter 5) and achieved around 5 Mbps over a two-hop 15 kilometer link to Bodi. The existing corDECT link continued to be used for consultations by Aravind while we sent dummy traffic over the WiLD link, mostly for operational experience.

Our first year of operating the link saw several failure modes that we discuss in more detail in chapter 6. The network was down for close to one-third of the full year for various reasons. For example, 50% of the downtime incidents were caused by poor quality power. We had to work through these issues to develop a system that would not only be high-performance but also robust to failures in rural areas.

By the beginning of 2006, satisfied with the high performance (5–6 Mbps per link), the freedom due to unlicensed spectrum, and their improving ease of managing the

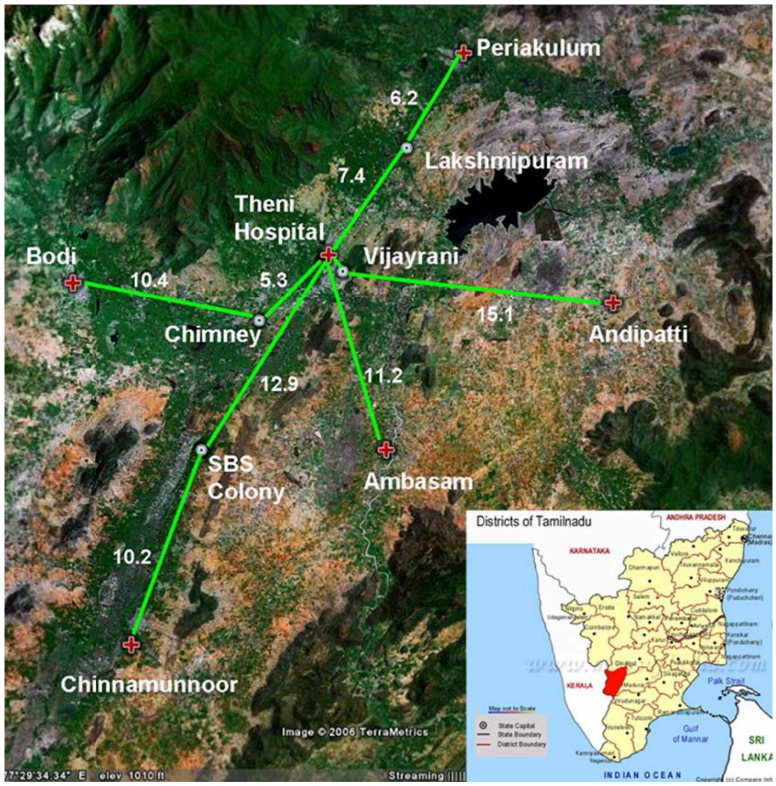

Figure 2.3: WiLD Deployment at Aravind Hospital in Theni. All link distances are in kilometers. Inset marks the district of Theni in the state of Tamil Nadu

system, Aravind phased out the CorDECT links, converting the existing three centers and completing the others that had been impeded. Figure 2.3 shows the WiLD network deployed at Aravind Hospital in Theni. Since then Aravind established a five more vision centers using WiLD links near the Tirunelveli and Madurai hospitals. Most recently, two more centers were set up in early 2009 connecting the Theni Hospital.

This process reflects the optimization principle discussed in Section 1.1.1: the doctors were already using a form of telemedicine; we just needed to improve their access to rural areas by working with the hospital to address their full range of concerns.

2.3.2 Operational Challenges

Software or hardware errors, power issues, or other transient environmental issues such as wireless interference caused outages leading to poor experiences for users and administrators in our deployments. As these new systems strive for acceptance, outages can chase away potential users and jeopardize viability.

There are several specific reasons why maintenance in rural areas is hard. First, local staff tend to start with limited knowledge about wireless networking. This leads to limited diagnostic capabilities, inadvertent equipment misuse, and misconfiguration. Thus management tools need to help with diagnosis and must be educational in nature. Training helps as well, but high IT turnover limits the effectiveness, so education must be an ongoing process and the system should be equipped with capabilities for monitoring and fault detection.

Second, the chances of hardware failures are higher as a result of poor power quality. Although we have not conclusively inferred the failure rate of equipment for power reasons in rural areas, we have lost far more routers and adapters for power reasons in rural India than we have lost in our Bay Area testbed. This calls for a solution that provides stable and quality power to equipment in the field.

Third, many locations with wireless nodes, especially relays, are quite remote. It is important to avoid unnecessary visits to remote locations. Also, we should enable preventive maintenance during the visits that do occur. For example, gradual signal strength degradation could imply cable replacement or antenna realignment during a normal visit.

Fourth, the wireless deployment may not be accessible remotely or through the

Internet. The failure of a single link might make parts of the network unreachable although the nodes themselves might be functional. This makes it very hard for remote experts in other towns or even local administrators to diagnose or resolve the problem. This points to a need for a low-cost alternative or “back channel” (e.g. SMS) that allows remote access to the nodes even in the event of a failure of the primary link.

Troubleshooting can be hard even for experienced users or experts. We argue that these troubles exist in a large part because the research community has not tried very hard: these systems are not designed or deployed with support for easy diagnosis built in right from the start.

We therefore need an end-to-end systems framework with support for diagnosis even in the presence of primary faults; this can improve the operational sustainability of the network. In addition to a data collection and monitoring infrastructure that can operate over intermittent networks or alternate channels, we need to build hardware and software modules that can be queried independently from the primary ways in which they are normally accessible. This allows for diagnosis of subsystems that would not be available when the primary link is down. Note that this redundancy is typically used for diagnosis rather than for failure recovery which is usually much more expensive. However, whenever possible, we are able to exploit this redundancy for failure recovery at low cost.

2.4 Other Groups

Outside of the United States, groups like AirJaldi [3] that runs a wireless mesh network in Dharamsala (India), Enlace HispanoAmericano de Salud (EHAS) Foundation [48] that uses WiFi to build wireless networks for health applications in Colombia [121] and Peru [131], CRCNet that helps connect rural communities in New Zealand [8] and Nepal Wireless [101] that connects a number of villages in the foothills of the Himalayas have pioneered the spread of WiFi in long-distance settings. The Digital Gangetic Plains project [9] consisted of a dozen point-to-point links, the longest of which was 39 kilometers.

The common theme among all these attempts is that standard WiFi is being used in point-to-point long-distance configurations with mixed levels of success, depending on network topology, interference levels and system management issues. Our solution of WiLDNet coupled with our suite of system management tools is widely applicable to all these groups.

2.5 Summary

In this chapter, we motivated the need for WiLD connectivity in rural areas using the specific example of our rural partner organization, the Aravind Eye Care System. Specifically, we described the technology needs for Aravind, and discussed how its initial approach of using corDECT WLL technology was insufficient. We compared alternative technologies and demonstrated that adapting WiFi for use in point-to-point long-distance topologies is the most promising option to introduce connectivity into rural areas. Next, we used examples from our early deployment experiences at Aravind to motivate the need for an end-to-end systems perspective on WiLD links in order to address the operational challenges of running WiLD networks in rural areas. Finally, we showed that this is a general problem by providing examples of several groups attempting to do the same, thus showing the generality of our work.

Chapter 3

Overview of WiFi-based Long Distance (WiLD) Networks

The IEEE 802.11 family of standards, commonly known as *WiFi* [1], is one of the most popular wireless technologies in use today. In 2008, 387 million WiFi chipsets were sold globally. The widespread popularity of WiFi, its rapid adoption all over the world, and continuing double digit growth in chipset sales volumes have been primarily driven by commoditization and standardization. High volume production resulting from commoditization has brought the prices down, making WiFi-based technologies very attractive for developing world markets. The standardization of protocols and use of unlicensed frequencies have enabled interoperability between different vendors, further driving adoption.

It is therefore not surprising that the use of WiFi has far outgrown its originally intended use of wirelessly connecting devices in indoor broadcast environments. [2] WiFi

[1] The term "WiFi, first used commercially in August 1999, was coined by a brand consulting firm hired by the WiFi Alliance (then called the Wireless Ethernet Compatibility Alliance) to create a name catchier than IEEE 802.11. The term WiFi was created as a play on words with HiFi (High Fidelity), but does not mean Wireless Fidelity (or anything else), as is widely misunderstood

[2] The earliest use of WiFi was to connect cash registers at check out lanes

has now become the most popular choice for building outdoor community networks and long distance rural networks where the distances are orders of magnitude longer (one to hundreds of kilometers) than distances within offices.

In this chapter, we first provide an overview of the physical (PHY) and medium access (MAC) layers of the legacy 802.11 family of standards, originally designed for wireless LANs in indoor broadcasts (Section 3.1). Second, we illustrate that 802.11 has outgrown its intitial intended use and has become a popular choice for outdoor wireless mesh and WiLD networks (Section 3.2). Third, we provide an overview of three techniques that can be used to improve the range and throughput in WiLD links: 1) extending the range of WiLD links via the use of high-power radios, high-gain atennas and careful link path planning (Section 3.3), 2) extending the SNR (and therefore range) and throughput via PHY-layer optimizations to 802.11, including the use of variable channels and multiple paths (Section 3.4), and finally, 3) extending throughput via MAC-layer optimizations such as changing timing parameters and enforcing synchronization among multiple co-located radios (Section 3.5). We show that MAC-layer parameter optimizations are not enough to improve throughput in WiLD links and that new MAC-layer designs that incorporate simultaneous synchronized transmit and receive are necessary.

3.1 Overview of 802.11

IEEE 802.11 [68] is a set of standards originally created for wireless local area network communication in the 2.4, 3.6 and 5 GHz frequency bands. It defines several air interface modulation techniques for the physical (PHY) layer, and a medium access control

(MAC) layer used by the physical layer. We provide a brief overview of the PHY- and MAC-layers in this section.

3.1.1 Physical Layers

At the physical layer, IEEE 802.11 defines various air interface modulation techniques such as Direct Sequence Spread Spectrum (DSSS), Frequency Hopping Spread Spectrum (FHSS) and Orthogonal Frequency Division Multiplexing (OFDM).

The original 802.11 standard specified two net data rates of 1 Mbps or 2 Mbps, and three alternative physical layer technologies: diffuse infrared at 1 Mbps, Frequency Hopping Spread Spectrum (FHSS) at 1 Mbps or 2 Mbps, and Direct Sequence Spread Spectrum (DSSS) at 1 Mbps or 2 Mbps. The latter two physical layers use microwave transmission in the 2.4 GHz ISM band. The legacy standard is now obsolete and has been amended by several extensions.

Maintaining backwards compatibility with the original specification, the 802.11b [70] extension uses DSSS in the $2.4 - 2.495$ GHz ISM band but in addition to data rates of 1 and 2 Mbps, it supports data rates of 5.5 and 11 Mbps. At net data rates of 11 Mbps, it is possible to achieve up to 7 Mbps of application throughput.

The 802.11a [69] extension uses OFDM to split the signal across 52 separate subcarriers to enable much higher net data rates ranging from 6 Mbps to 54 Mbps. It operates in the $5.725 - 5.825$ GHz ISM band and also in the $5.15 - 5.32$ GHz UNII band. These relatively unused 5 Ghz bands provide little interference compared to the crowded 2.4 GHz band in which several devices such as cordless phones, microwave ovens, baby monitors, bluetooth devices and other 802.11b routers operate. However, the higher frequency makes

802.11a incompatible with 802.11b/g and reduces the achievable range compared to that of 802.11b/g at the same transmit power.

The 802.11g [71] extension uses the same OFDM modulation technique as 802.11a, but works in the 2.4 GHz ISM range of 802.11b. Therefore, it is backwards compatible with 802.11b while achieving net data rates up to 54 Mbps.

802.11n, a pending amendment as of this writing, adds Multiple-Input Multiple-Output (MIMO) and channel bonding to the PHY-layer. MIMO allows the use of multiple antennas per station for signal diversity and spatial multiplexing, enabling higher data rates. Channel bonding allows the use of two non-overlapping channels to transmit data, increasing the amount of data that can be transmitted at the PHY-layer. It is projected that 802.11n will increase the current maximum data rates of 54 Mbps to a maximum of 600 Mbps.

3.1.2 Medium Access Control

Unlike wired ethernet, WiFi transceivers are half-duplex, and hence cannot receive while transmitting. Therefore, mechanisms for collision avoidance rather than collision detection are necessary. The 802.11 standard for wireless LANs specifies two different ways to coordinate transmissions among stations: a mandatory Distributed Coordination Function (DCF) and an optional Point Coordination Function (PCF). DCF allows *distributed contention-based* channel access using Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), while PCF provides *centralized contention-free* channel access based on a poll-and-response mechanism.

The 802.11e [72] enhances the DCF and PCF modes through a new coordination function called the Hybrid Coordination Function (HCF), which in turn has two methods of channel access: Enhanced Distributed Channel Access (EDCA) and HCF Controlled Channel Access (HCCA).

We briefly describe the basics of both DCF and PCF modes of operation; full details can be found in the standards specifications [68, 69, 70]. We also summarize the key aspects of EDCA and HCCA.

DCF mode

In DCF mode, any station (client or access point) with a packet ready for transmission first senses the channel. If the channel is busy, the station waits until the channel becomes idle, and then waits for an additional time interval known as the DCF Inter Frame Space (DIFS) interval. If the channel stays idle throughout the DIFS interval, the station starts a random backoff process. It views time in units of discrete interval called slots and selects a random backoff counter. For each slot during which the channel continues to remain idle, the random backoff counter is decremented. The packet is transmitted when the counter reaches zero. However, if the channel becomes busy during backoff, the backoff is suspended. It is resumed with the most recent counter value after the channel stays idle for a full DIFS interval. This method of random backoff prevents multiple stations that sense the channel idle simultaneously, from transmitting packets at the same time, thus reducing the likelihood of packet collisions.

Each station maintains a contention window (CW) to select the random backoff count. The backoff count is selected uniformly at random in the interval $[0, CW]$. A simple *stop-and-wait* protocol is used between the transmitting and receiving stations. Upon suc-

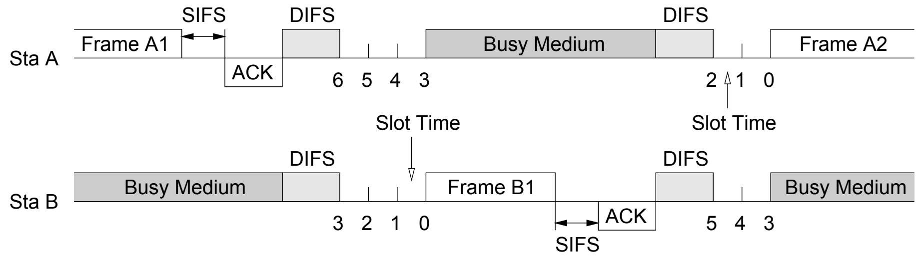

Figure 3.1: After successful transmission and acknowledgment of frame A1, station A waits for DIFS interval and selects a backoff interval of 6 slots before attempting to transmit the next frame A2. Meanwhile, station B which has a smaller backoff interval of 3 senses the channel is idle, and transmits frame B1. Station A senses the channel is busy during its backoff, freezes its counter at slot 3 and continues to decrement later when it senses the channel is idle for DIFS interval.

cessful backoff, the frame is transmitted. For each successfully received frame, the receiving station sends an acknowledgment (ACK) frame within a Short Inter Frame Space (SIFS) interval, which is much shorter than the DIFS interval. This protects the ACK frame from contention with data frames of other stations. If an ACK frame is not received by the transmitter within an interval called the *ACKTimeout* [3], the frame is retransmitted after another random backoff based on a new value of CW. Initally, CW is assigned to CW_{min}. After every unsuccessful frame transmission, the value of CW is doubled up to an upper bound of CW_{max}. Upon successful transmission, CW is reset to CW_{min} and the station waits for another DIFS interval and random backoff period before it sends the next consecutive frame. This prevents a station with multiple frames ready for transmission from monopolizing a shared channel.

Wireless networks can suffer from the hidden terminal problem [28] where two stations, out of range from each other, can independently transmit to (and receive from) a

[3] Several protocol implementations commonly define the *ACKTimeout* as SIFS + ACK Transmission Time + Slot Time

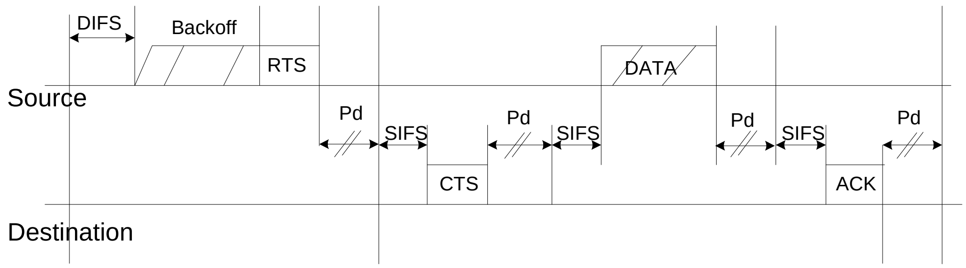

Figure 3.2: Timeline for DCF operation with RTS/CTS. Pd is the propagation delay. The source and destination participate in the RTS/CTS exchange while other stations (not shown) wait for the NAV timer to count down to zero before contending for the channel.

common third station, leading to increased packet collisions at the third station. This can lead to severely degraded performance. To mitigate the hidden terminal problem, DCF has an optional Request-to-Send/Clear-to-Send (RTS/CTS) mechanism. A short RTS frame is transmitted by the sending station, followed by a short CTS frame by the receiving station before the actual exchange of a data frame and ACK between the stations. The RTS and CTS frames include a network allocation vector (NAV) that specifies the expected duration of transmission of the data frame and its corresponding ACK. Other stations, hearing the transmitting station and the hidden stations, do not transmit any frames. They start their NAV timer and do not contend for the medium until it counts down to zero. Between every two consecutive frames in the sequence of RTS, CTS, Data and ACK frames, the stations wait for a SIFS interval. Figure 3.2 shows the time line of this sequence.

PCF mode

PCF is a contention-free mode of operation that is available only in “infrastructure mode” where stations are connected to the network via an access point (AP). The AP sends beacon frames at regular intervals, usually every 0.1 seconds. Between beacon frames, PCF defines two periods: a Contention Free Period (CFP) and a Contention Period (CP). In CFP, the AP sends poll messages to each station, one at a time, asking for data. If the polled station has a frame to send to the AP, it is included in the response message. If the AP itself has data to send to the station being polled, this data can be included in the poll message. In CP, DCF is used. The PCF mode of operation is intended to support Quality-of-Service by enabling the AP to regulate the bandwidth allotted to each of the client stations. However, as PCF is optional, very few WiFi devices actually implement it.

EDCA Mode

The Enhanced Distributed Channel Access (EDCA), specified by 802.11e [72] within its Hybrid Coordination Function (HCF), enhances the original DCF to provide quality-of-service (QoS). EDCA introduces traffic priority classes. Each priority level is assigned Transmit Opportunity (TXOP) which is a bounded interval of time during which a station can send as many frames as possible. In practice, different transmit opportunities are achieved by varying two DCF parameters: the size of the contention window CW, and the size of defer time, the DIFS interval. Lower CW and DIFS give higher priority for channel access. EDCA defines that a channel has up to four access categories, each with its own defer time interval called the Arbitrary Distributed Inter Frame Space (AIFS) interval, and its own CW value. This provides higher priority traffic a better chance to contend for the channel. WiFi Multimedia (WMM) certified access points are enabled with EDCA and TXOP.

HCCA mode

Also specified by 802.11e within its HCF, the HCF Controlled Channel Access (HCCA) mode is very similar to PCF, but with some important differences. Unlike PCF, in which the interval between two beacon frames is divided into a Contention Free Period (CFP) and a Contention Period (CP), HCCA allows for CFPs being initiated at almost any time during a CP. HCCA is generally considered to be the most advanced coordination function, allowing QoS to be configured with precision. However, HCCA support is not mandatory for access points and we are not aware of any that are enabled for it.

3.2 802.11 in Long Distance Networks

The availability of high-performance WiFi radios at low cost [4] and the opportunities provided by unlicensed spectrum have spurred several researchers, non-profit groups and enthusiasts to experiment with outdoor WiFi deployments, either as unplanned mesh networks or as planned point-to-point networks.

In the United States, some research efforts include MIT's Roofnet project [37] and Harvard University's CitySense project [7], and some non-profit efforts include the Champaign-Urbana Community Wireless Network (CuWin) [6], NYCWireless [15] and Seattle Wireless [19].

Internationally, several groups have pioneered the use of WiFi in long-distance settings. For example, AirJaldi [3] runs a wireless mesh network with several point-to-point long-distance links in Dharamsala (India), Enlace HispanoAmericano de Salud (EHAS)

[4]Ubiquiti Networks [23] currently offers high-power 2.4 GHz radios for less than $50.

Foundation [48] builds long-distance wireless networks for health applications in Colombia [121] and Peru [131], CRCNet [8] connects isolated rural communities in New Zealand and Nepal Wireless [101] links villages together the foothills of the Himalayas. The Digital Gangetic Plains [9] in India consists of a dozen point-to-point links, the longest of which was 39 kilometers. The Berkeley TIER group has also been involved in similar links deploying links ranging from 40 to over 300 kilometers in the Bay Area, Ghana, Uganda, India and Venezuela.

The common challenge facing all these efforts attempting to use standard off-the-shelf hardware is that the standard WiFi protocol is not designed to work in these environments [36], and therefore often performs quite poorly [129].

We next discuss techniques for adapting WiFi for high-throughput long-distance settings. Specifically, we discuss several range extension techniques orthogonal to the WiFi protocol, followed by techniques involving PHY and MAC layer modifications.

3.3 Extending Range

A point-to-point link consists of two radios, each with its own antenna, separated by the physical path traveled by the wireless signal. In order to establish communication between the two endpoints a certain minimum radio frequency (RF) signal needs to be seen by the receiving radio. Determining whether the link is feasible requires a link budget calculation. Link feasibility depends on the power of the transmit radio, the gain by the antennas, and energy losses in cables, connectors and along the wireless path.

3.3.1 Link Budget Calculation

The parameters involved in link budget calculation are transmit power, antenna gain, receiver sensitivity, losses in cables and connectors, and path loss.

Transmit power: Transmit power is usually expressed either in milliwatts (mW) or in decibels relative to a milliwatt (dBm), and for typical commercially available 802.11 radios ranges from 30 to 200 mW. High-power radios can go up to 600 mW (28 dBm) [23].

Antenna gain: Antennas are passive devices that amplify electromagnetic waves by virtue of their shape. Antenna gain is defined by the ratio of the radiation intensity of an antenna in a given direction to the intensity produced by hypothetical ideal (isotropic) antenna that distributes energy uniformly in all directions and has no losses. It is measured in dBi. Antenna gain is symmetric, amplifying the signal both at the transmit end and at the receive end. Typical gains for commercially available antennas are around 8 dBi for omnidirectional antennas, 12 dBi for sector antennas, and 19-24 dBi for parabolic directional antennas [133].

Receiver sensitivity: A receiver's sensitivity is a measure of the receiver's ability to discern a faint signal. It is defined as the minimum input signal required to produce a specified signal-to-noise ratio (SNR) and is measured in dBm. A lower value implies better sensitivity. For example, a receive sensitivity of -98 dBm is better than a receive sensitivity of -95 dBm by 3 dB, or a factor of two. Therefore, at a specified data rate, a reciever with -98 dBm sensitivity can discern signals that are half the power of those heard by a receiver with a sensitivity of -95 dBm.

Losses in cables and connectors: Additional RF energy is lost due to attenuation along cables connecting the radios to their antennas, and also due to the attenuation in the antenna connectors themselves. Cable attenuation depends on cable quality, signal frequency and cable length. For example, at 2.4 GHz, typical attenuation ranges from 0.05 dB to 0.8 dB per foot of the cable. Connectors can attenuate the signal by 2 or 3 dB each.

Path Loss: The most significant loss in RF energy occurs over the air. There are several factors contibuting to this attenuation: free space loss, attenuation by the environment and multipath. Free space loss occurs as the power attenuates with the geometric spreading of the wave. This factor increases with distance since the RF wave spreads more the farther it travels. Since it is independent of the environment and other factors, it is not possible to mitigate this loss.

Another factor contributing to path loss is the attenuation resulting from signal absorption as the wave passes through foliage, walls and other obstacles. This attenuation is difficult to quantify and can be as much as 20 dB for walls or trees. This can be mitigated by ensuring clear line of sight between the transmit and receive antennas such that the Fresnel zone [65] (discussed shortly) is unobstructed.

Finally signals are also attenuated due to multipath effects. RF energy spreads out as it radiates from the antenna. Part of this energy travels to the receiving antenna along the direct path, while other parts reflect off the ground and other obstacles. Since the reflected signal travels a longer path, it arrives at the receiver later than the direct signal. The receiver combines signals coming on different paths, but due to the associated delay spreads, these signals can either sum up or cancel each other out. In practice, multipath is a very location-specific phenomenon, more pronounced in areas with many possibilities for reflections, such as urban areas. Multiple paths do not generally pose trouble for point-to-point long-distance rural links, as the delay spreads between reflected signals decrease with distance (section 4.3.3)

If we consider the combined effects of free space loss, attenuation, and multipath, the overall path loss can be expressed in dB as

$$Path\ Loss = 92.45 + 20 * log_{10}(F) + 10 * n * log_{10}(D) + L_{Attenuation} \tag{3.1}$$

where F is the radio frequency in GHz, D is the distance between radios in kilometers, n is the free space loss exponent depending on the environment, and $L_{attenuation}$ is the attenuation estimated for absorption losses due to obstacles along the path; $n = 2$ represents essentially no multipath, $n = 3$ is typical for outdoor environments, while $n = 4$ is representative of indoor links.

The Link Budget: Taking into account all the factors contributing to the loss or gain in signal, the received signal strength can be computed as

$$\begin{aligned} RxPower &= TxPower + TxAntennaGain - TxCableLoss \\ &\quad -PathLoss + RXAntennaGain - RxCableLoss \end{aligned} \tag{3.2}$$

Therefore, the key challenge in operating a long-distance link is to ensure that the received signal strength exceeds the receiver's sensitivity by a reasonable margin. As an example, let us consider an outdoor point-to-point link that is 20 kilometers long with the following assumptions. There is perfect line-of-sight and no multipath ($L_{Attenuation} = 0$ and $n = 2$). The link uses high-power radios with trasmit powers of 200 mW (23 dBm) and receive sensitivies of -90 dB. The antenna gains are 19 dBi, and losses due to cables and connectors are 4 dB on each end. Using equation 3.1, we have $Path\ Loss = 92.45 + 20 * log_{10}(2.4) + 10 * 2 * log_{10}(20) + 0 \approx 126\ dB$. Therefore, from equation 3.2, we have $RX\ Power = 23 + 29 - 4 - 126 + 19 - 4 \approx -73\,dB$. Since the network card has sensitivity of $-90dB$, the link is feasible and we have about $17dB$ of margin to accommodate additional losses due to obstacles or multipath.

3.3.2 Range Extension Techniques

Considering all the factors relevant for the link budget computation, two obvious ways to ensure feasible links at longer distances are to *increase gain* or *decrease loss*.

Increasing gain

The easiest method to increase signal strength is to simply use higher power radios, such as Ubiquiti's XR [23] series of radios that can transmit up to 600 mW. It is possible to achieve even higher transmit power levels via the use of external power amplifiers, but these are expensive and rarely needed.

A second method is to increase the Effective Isotropically Radiated Power (EIRP) in a particular direction by using highly directional or sector antennas. For example, parabolic directional antennas with gains of 24 dBi in the 2.4 GHz band are commercially available today for less than $100.

Decreasing loss

The most important factor in reducing path loss is ensuring clear line-of-sight between the end points of the links to prevent absorption or multipath losses. RF line-of-

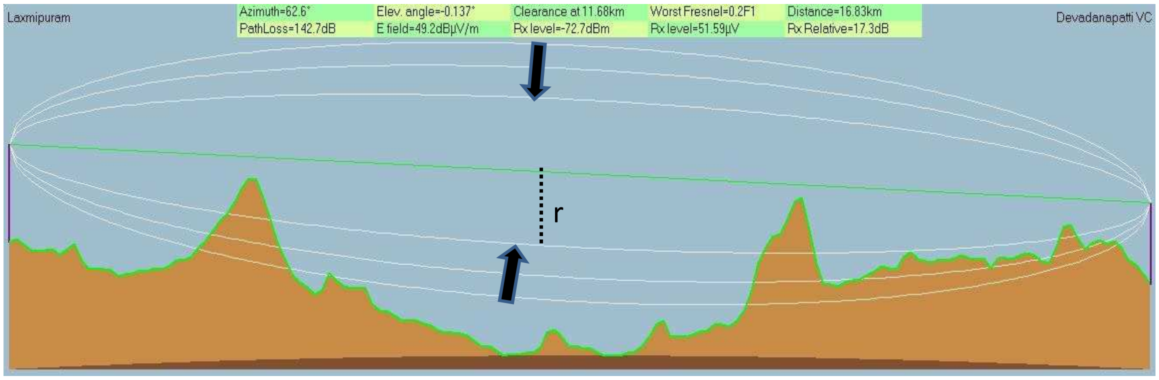

Figure 3.3: RF propagation profile showing the fresnel zones between Laxmipuram and Devadanapatti, a 16 kilometer path. Three fresnel zones are shown. The radius of the first fresnel zone at a given point is r. Although visual line-of-sight is clear, the first fresnel zone is greatly obstructed by the peaks

sight does not simply imply visual line-of-sight. RF waves spread out as they travel, and in order to preserve energy, an elliptical zone around the visual line-of-sight needs to remain largely unobstructed. This zone is called the first Fresnel zone [65]. There are many fresnel zones, but only the first one is important for RF propagation, as that is the zone in which reflections of the RF wave will interfere constructively.

Figure 3.3 shows an RF path profile with three fresnel zones for a WiLD link between Laxmipuram and Devadanapatti in the Theni district where Aravind is located. We see that despite clear visual line-of-sight, the first fresnel zone is greatly obstructed by the two peaks as a result of which this link is infeasible.

At any point along the path we can compute the radius r of fresnel zone using

$$r = 17.31\sqrt{\frac{n(d1 * d2)}{f * d}} \tag{3.3}$$

where r is the radius in meters, $d1$ and $d2$ are the distances to the link end-points in meters,

f is the frequency in MHz, $d = d1 + d2$ is the total link distance in meters, and n is the zone we are interested in (first, second and so on). In practice, at least 60% of the first fresnel zone should be kept clear; therefore the minimum value used for n is 0.6.

In order to ensure clear RF line-of-sight, one must consider elevation profiles and other possible obstacles. High-elevation locations, such as rooftops, hills, or communication towers are usually good choices to place link end-points.

3.4 PHY-layer Optimizations

PHY-layer optimizations such as the use of variable width channels and the use of multiple paths can improve signal-to-noise-ratio (SNR). This improves the link budget and therefore link distance, and can also improve link throughputs. We discuss these two approaches in this section.

3.4.1 Using variable width channels

In the United States, the FCC allows 802.11b to use 11 channels in the 2.412-2.462 GHz frequency range. Each channel is separated by 5 MHz but the channel widths are 22 MHz wide, implying that adjacent channels overlap and interfere with each other. Effectively, only three orthogonal channels can be used without overlap to avoid any interference.

Fundamentally, maximum bitrate on a channel is roughly proportional to the channel width. Therefore, changing the channel width offers the option to trade off increased simultaneous transmissions on orthogonal channels with higher throughput on a single wider channel. A detailed analysis of the tradeoffs for single links is presented by Chandra et al [43].

Using narrower channel widths has two main advantages for WiLD links: *higher range* from better SNR and resilience to multipath, and *higher cumulative throughput* by using more orthogonal channels.

Higher range from better SNR

If we assume that the transmit power of a wireless radio is constant regardless of the channel width, then at narrower widths, the radio can transmit higher energy per unit Hz. If the noise per unit Hz is assumed to be constant at all channel widths, then the SNR is higher for narrower widths, improving the link budget and therefore enabling a longer link range. Theoretically, with ideal radios, if we use half the width of a channel, we should get a 3 dB increase in SNR, and if we use a quarter of the channel width, we should see a 6 dB increase in SNR, and so on.

More orthogonal transmissions

We can also achieve more simultaneous transmissions if we use narrower channels, thereby increasing the total cumulative throughput of a network. It is known that the most optimal throughput on a channel can be achieved by using successive interference cancellation [138]. However, this requires complex signal processing on specialized hardware, Instead, it can be proven that when stations that are continuously backlogged, optimal throughput can always be achieved by allocating the right spectrum width to each channel [57]. Although the maximum bitrate of the channel is also reduced with narrower channels, the cumulative throughput of all the narrower channels can equal the optimal throughput, if the channel widths are allocated proportionally to the received SNR on the

Channel width	5 MHz	10 MHz	20 MHz	40 MHz
Symbol Duration	16 μs	8 μs	4 μs	2 μs
SIFS	40 μs	20 μs	10 μs	5 μs
Slot Duration	20 μs	20 μs	20 μs	20 μs
Guard Interval	3.2 μs	1.6 μs	0.8 μs	0.4 μs

Table 3.1: 802.11 timing parameters for different channel widths [43].

channel [57].

Implementing narrower channel widths

Today, channel widths can be configured on several off-the-shelf radios. Atheros chipsets support "turbo mode" where two adjacent 20 MHz channels can be bonded together to form a 40 MHz channel that can achieve data rates up to 108 Mbps.

Narrower channel widths can be achieved by changing the frequency of the reference clock that drives the phased-locked loops (PLL) inside the radio, via hardware registers. However, changing the clock rate also affects several 802.11 timing parameters as summarized in Table 3.1. As symbol lengths are different across channel widths, the same modulation scheme that achieves 24 Mbps with the default 20 MHz channel width now achieves 6 Mbps and 12 Mbps at 5 MHz and 10 MHz channel widths.

3.4.2 Using multiple paths

The arrival of multiple reflections of a signal at the receiver can cause significant performance degradation for two reasons: Inter Symbol Interference (ISI), where successive symbols within a packet overlap with each other, or destructive interference, where reflections of the signals interfere with each other (signal fading).

The traditional method to deal with multipath is to use antenna diversity at the receiver. The receiver uses multiple antennas to increase the chances of receiving a better signal on one of the antennas by choosing the signal with the best SNR.

However, recent enhancements such as Multiple-Input Multiple-Output (MIMO), proposed as part of the 802.11n draft standard, can now leverage multiple paths to improve SNR and data rates by using *spatial multiplexing*. Both receivers and transmitters employ multiple antennas and radios. The sending station splits a high-rate signal into multiple lower-rate signals, and each signal is transmitted via a different antenna on the same frequency. If these signals are received with sufficiently different spatial signatures by the receiving station, the receiving station can separate these signals, effectively creating parallel channels. Theoretically, MIMO can dramatically improve SNR and throughput to upto 600 Mbps using four antenna and radio pairs at both the sending and and receiving stations [73].

In practice, multiple paths do not cause a lot of performance degradation in WiLD links. As discussed in section 4.3.3 the path differences between the primary line-of-sight path and any ground reflected path become inconsequential with distance.

3.5 MAC-Layer Optimizations

It is not surprising that the original 802.11 MAC-layer protocol, originally designed for wireless LANs in indoor broadcast environments, performs very poorly when used in long-distance settings. In this section, we show that the standard coordination functions, DCF and PCF, are ill-suited for long distances. We then show how the performance of DCF can be improved to some extent by tweaking specific 802.11 parameters. Finally, we present recent work on designing completely new MAC protocols with synchronization techniques specifically for optimizing performance and spatial reuse of multihop WiLD networks.

3.5.1 PCF in long-distance links

Timing requirements make the 802.11 PCF mechanism ill-suited for long-distance links. In particular, the acknowledgment (ACK) sent by the polled station must be received by the access point (AP) within the SIFS interval, which is only 10μs. This corresponds to a round trip of approximately 3 kilometers, limiting links to 1.5 kilometers or less. We are not aware of any chipset that allows modifying the SIFS interval to arbitrary values.

3.5.2 DCF at long distances

Contention resolution in DCF is based on the idea that stations can always detect each other's transmissions and then back off because it is assumed that the maximum distance between any two stations is always less than the distance implied by the slot time.

But this assumption is not true in the case of WiLD links where the propagation delays can be much larger. If two stations start transmissions within an interval that is shorter than the propagation delay, which is very likely in long links, the transmissions will collide.

3.5.3 Optimizing DCF parameters

Recent work estimating the impact of distance on the throughput of DCF suggests that losses from collisions can be decreased if we modify certain timing parameters [130]. These parameters are available in the Enhanced Distributed Channel Access (EDCA) mode specified as a part of the set of amendments for 802.11e. As discussed in section 3.1.2, EDCA allows for quality-of-service for different traffic priority classes.

This traffic differentiation is created by providing the following set of parameters for the i_{th} class of traffic: Aribitrary Inter Frame Space Number ($ASFN_i$), the contention window (CW_i), and the Transmission Opportunity ($TXOP_i$). In addition, we can also control some non-standard parameters such as $SlotTime$ and $ACKTimeout$.

It has been shown that the $SlotTime$ must be increased to twice the propagation delay for distances longer than 3 kilometers (for a slot time equal to 20 μs) or 1.35 kilometers (for a slot time equal to 9 μs) in order to guarantee that two stations that listen to each other may only collide if they transmit in the same slot [125]. These adjustments are a slight violation of the standard but they are feasible with many available commercial WiFi systems, and compatibility with legacy WiFi stations can still be maintained.

Although these parameter adjustments improve DCF's performance at long distances to some extent by eliminating almost all collisions, the final throughput is still less than optimal, mainly because of the inefficient stop-and-wait acknowledgment protocol used by 802.11 [125].

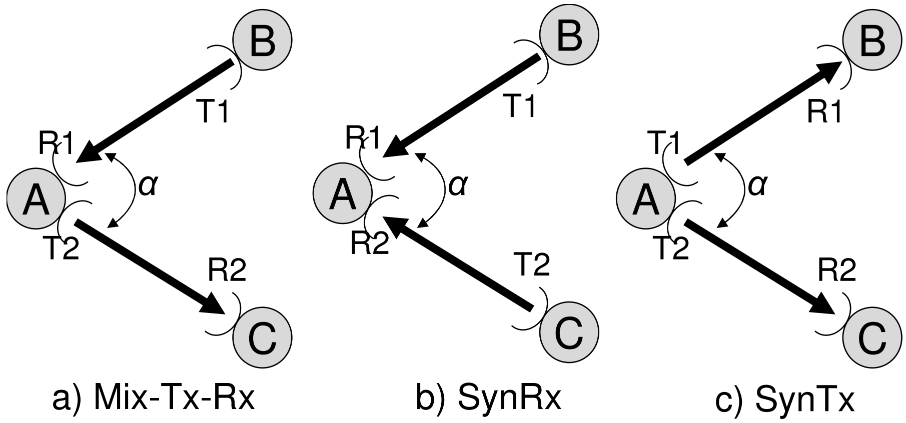

Figure 3.4: SynOp scenarios for interference

3.5.4 Synchronization for spatial reuse

Multihop WiLD wireless networks often have nodes equipped with multiple co-located radios. Although operation on different channels avoids inter-radio interference, there are several practical constraints that may require the operation of co-located radios on the same channel. This can happen due to country-specific regulatory restrictions [135], high spectrum costs, or the limited number of available channels.

Long-distance links with high-gain directional antennas interfere with each other in a very specific manner. More precisely, as first observed by Raman *et al.* [117], and reemphasized in [118, 111], co-located radios (sharing the same physical location) operating on the same wireless channel interfere with each other if one transmits while the other receives. However, two adjacent directional links that either transmit simultaneously (*SynTx*), or receive simultaneously (*SynRx*), will be largely *interference-free*, a mode of operation termed as Simultaneous Synchronized Operation (*SynOp*).

We briefly explain the reason for this behavior. Consider the adjacent directional point-to-point links depicted in Figure 3.4, separated by an angle α. Now consider the following three potential interference scenarios:

1. **Mix-Tx-Rx:** In this scenario, depicted in Figure 3.4(a), T_2's transmissions interfere with R_1's reception due to the physical proximity between the radios and the presence of antenna side-lobes. Therefore, operating the links in this mode is not feasible.

2. **SynRx:** During simultaneous receive, shown in Figure 3.4(b), T_2's transmissions cause interference at R_1, and T_1's transmissions cause interference at R_2. For the interfering signal to be ignored, the difference between the intended signal and the interfering signal must be larger than a certain threshold $Th_{isolation}$, which depends on modulation and data-rate; for example, with 802.11b at 11 Mbps, $Th_{isolation} \approx 12\ dB$ [117]. Fortunately, this isolation can usually be ensured through the differences in gain provided by the directional antennas, if the links are separated by a sufficiently large angle. If we denote the difference between the antenna gain of the main lobe and the gain at an angle α away from the main lobe by S_{alpha} (also called the rejection level at angle α), then adjacent links are interference-free under the following condition [117]:

$$|P_{R1} - P_{R2}| < S_\alpha - Th_{isolation} \tag{3.4}$$

   where P_{R1} and P_{R2} are the receive power levels at R_1 and R_2 respectively.

   For example, if links use typical $24dBi$ grid antennas (also used in our deployments) in horizontal polarization, an angular separation of more than 10° (half the width of the antenna main lobe) translates into an isolation of at least $25dB$. This means that 802.11b links receiving simultaneously are interference-free if $|P_{R1} - P_{R2}| < 15\ dB$.

This can be easily satisfied by a large range of values (e.g. $P_{R1} = P_{R2}$), and even if the path loss of the two links is very different, the condition can be satisfied by adjusting the radio transmit power accordingly (by reducing the TX power on the stronger link).

3. **SynTx:** With simultaneous transmissions, as in Figure 3.4(c), interference may occur at nodes B and C, but not at node A. Once again, R_1 may see interference from T_2, and R_2 may see interference from T_1. Given the symmetry of the two links, ensuring no interference during *SynTx* can be done by enforcing a similar condition to that in Equation 3.4.

   We note that simultaneous transmission is infeasible using a carrier-sensing MAC, such as 802.11, since radios can hear each others' transmission, causing one of the radios to backoff.

In summary, simultaneous synchronized operation (*SynOp*) can allow multiple adjacent WiLD links to simultaneously use the same wireless channel *provided* the links are separated by a sufficiently large angle α and the radio transmit powers are chosen to satisfy the constraint from Equation 3.4. Given the gain pattern of typical grid directional antennas [66], an angular separation α larger than $30°$ provides generous isolation between adjacent links; this has also been demonstrated experimentally [117, 118] and validated in our deployments [111, 136].

3.6 Summary

In this chapter, we first provided an overview of the PHY- and MAC-layers specified by the legacy 802.11 family of standards and its several extensions. Next, we provided a summary of the range of techniques that can be used to extend range and throughput to enable the use of 802.11 in WiLD links.

These techniques ranged from methods to improve the range via the use of higher power radios, higher gain antennas, better path planning, PHY-layer optimizations that increase both SNR and throughput, and MAC-layer optimizations that improve throughput.

We demonstrated that simple modifications to the existing MAC-layer are not sufficient to achieve high throughput, and that new MAC-layer designs that enable synchronous transmit and receive - but disallow mixed transmit and receive - are necessary.

Chapter 4

Characterization of WiLD Links

It is possible to achieve good signal strength in WiLD links by using inexpensive off-the-shelf directional antennas in line-of-sight conditions. Driven by this promise of low-cost long-distance connectivity, many efforts around the world, including our own, have focused on deploying 802.11 links in WiLD settings. However, the performance of such WiLD links has been abysmal. This has been our experience from links we deployed in the Bay Area, India and Ghana, as well as the experience of several other groups such as Digital Gangetic Plains [9] and Akshaya [98].

In this chapter, we characterize the poor performance in terms of throughput and losses seen when the standard 802.11 protocol is used in WiFi-based long-distance (WiLD) links. First, we describe our experimental setup comprising both an outdoor testbed and em-

The work presented in this chapter is joint work with TIER group members Anmol Sheth, Rabin Patra, Sergiu Nedevschi, Lakshminarayanan Subramanian and Eric Brewer. Most of the material on loss characterization and channel-induced losses has been previously published as "Packet Loss Characterization in WiFi-Based Long Distance Networks" [129], and the material on protocol-induced losses has been previously published as "WiLDNet: Design and Implementation of High-Performance WiFi Based Networks" [111].

ulated links indoors (Section 4.1). Second, we demonstrate that the standard 802.11 MAC protocol is ill-suited for WiLD links due to inefficient packet transmission and the breakdown of CSMA over long distances and propagation delays (Section 4.2). We specifically identify three fundamental reasons for these protocol-induced losses: *inefficient link-layer recovery*, *collisions at long distances*, and *inter-link interference*.

Third, we analyze the three causes of channel-induced losses in WiLD links: *external WiFi interference*, *non-WiFi interference* and *multipath interference* (Section 4.3). Among these, we show that external WiFi interference is the most significant source of packet loss in WiLD environments while the effects of multipath and non-WiFi interference are not significant. This is in contrast to the results of the Roofnet mesh network [26], where the authors observed multipath to be the most significant source of packet loss. Since our work, a more recent study [54] has reexamined the Roofnet data and has shown that external interference and not multipath was the major source of loss.

Fourth, we classify the loss patterns over time into two basic categories: *bursts* and *residual* losses (Section 4.4). We further classify bursts into short and long bursts. We make three important observations: (a) Although the burst arrival patterns can be approximately modeled based on a Poisson process, the duration and magnitude of a burst are harder to predict; (b) The residual loss characteristics over certain links are stationary, while some others exhibit non-stationary behavior even over daily timescales; (c) The loss variability observed in our urban links significantly differs from that under rural settings as observed in prior work [44]. Next, we measure the effect of protocol- and channel-induced losses on end-to-end TCP performance, and how various link parameters affect the TCP throughput (Section 4.5). Finally, having identified external WiFi interference as the primary source of losses in WiLD links, we discuss three potential remedies to mitigate these losses (Section 4.6): (a) 802.11 frequency channel adapation, (b) 802.11 PHY datarate adaptation, and (c) adaptive forward error correct (FEC).

Our measurement study is significantly different from other WiFi loss measurements in mesh networks [26] or hotspot settings [124]. The study by Chebrolu et. al. [44] is the only other measurement-based study of WiLD deployments that we are aware of. However, these two studies are orthogonal. We focus on determining the various sources of loss, understanding loss variability, and studying potential remedies to alleviate loss. Their work focused on the effect of parameters such as weather, SNR, payload size and datarate on loss.

4.1 Experimental Setup

We use three different experimental setups to conduct our loss characterization measurements.

1. **Real testbed**: We perform our packet loss characterization measurements on a WiLD network testbed comprising of links in both rural and urban environments. Table 4.1 summarizes some of the urban and rural links in our deployments. The links range from 2 to 20 kilometers in length.

2. **Wireless Channel Emulator**: In addition to the testbed, we also use a wireless channel emulator (Spirent 5500 [21]) to study each source of packet loss in isolation. The emulator allows us to place the two ends of the link in separate RF-isolation boxes

Link	Distance (km)	Environment	Antenna height(m)
K-P	20	Urban	50
B-R	8	Urban	30
M-P	2	Urban	40
T-A	11	Rural	20
T-S	13	Rural	25
W-N	15	Rural	20

Table 4.1: List of our urban and rural WiLD testbed links.

(80 dB of isolation) and then emulate in real time the RF channel between them. The Spirent 5500 accurately emulates radio channel characteristics with channel loss, fast and slow fading and delay spreads. This enables us to emulate links of any length or loss profile with repeatable results. We perform tests by connecting the channel emulator to the same radios used in our WiLD deployments.

3. **Indoor multi-hop testbed**: We perform controlled multi-hop experiments on an indoor multi-hop testbed consisting of 4 nodes placed in RF-isolation boxes. The setup was designed to recreate conditions similar to long-distance outdoor links where transmissions from local co-located radios interfere with each other, but which also enable simultaneous reception on multiple co-located radio interfaces. We can also control the amount of external interference by placing an additional wireless node in each isolation box just to transmit packets mimicking a real interferer. The amount of interference is controlled by the rate of the CBR traffic sent by this node.

We use Atheros 802.11a/b/g radios for all our experiments. The wireless nodes are 266 MHz x86 Geode single board computers [17] running Linux 2.4.26. We use *iperf* [10] to measure throughput. All our results are based on CBR UDP traffic streams. Unless otherwise stated, we turn off MAC-layer ACKs and set the maximum retry limit to zero for all our experiments. This allows us to measure the real underlying channel loss rate in absence of any MAC-layer acknowledgments and retries.

We instrument the standard Atheros *madwifi* driver to log fine-grained information for each frame received and transmitted. In addition to capturing all the frames on the link, we also capture and log frames being transmitted by external WiFi sources. This is achieved by creating a virtual network interface set in "monitor mode" on the same channel as the primary interface. This technique is equivalent to using two physical network interfaces, one being the primary interface and the other a passive monitor. To summarize, we collect the following information for every frame: complete 802.11 MAC header and IP payload, received signal strength, PHY layer transmit datarate, timestamp, PHY and CRC errors, and the noise floor immediately after the frame is received. We also modify the Atheros driver to pass up frames with CRC and PHY errors.

Using the WiLD testbed and the channel emulator, we explore two categories of loss: *protocol-induced losses* by the standard 802.11 MAC protocol and *channel-induced losses* due to the wireless channel itself. Specifically, for 802.11 protocol-induced losses, we investigate 1) *timeouts due to propagation delay* and 2) *the breakdown of CSMA over long distances*. For channel-induced losses we investigate 1) *external WiFi interference*, 2) *external non-WiFi interference*, and 3) *Multipath interference*.

4.2 802.11 Protocol Shortcomings

In this section, we study three main limitations of the standard 802.11 protocol: *inefficient link-layer recovery*, *collisions in long distance links*, and *inter-link interference*. These limitations make 802.11 ill-suited even in the case of a single WiLD link. Based on extensive experiments, we also show that just modifying the driver-level parameters of 802.11 is insufficient to achieve good performance.

4.2.1 Inefficient Link-Layer Recovery

The 802.11 MAC uses a simple stop-and-wait protocol, with each packet independently acknowledged. Upon successfully receiving a packet, the receiver node is required to send an acknowledgement within a tight time bound (ACKTimeout), or the sender has to retransmit. This mechanism has two drawbacks:

1. As the link distance increases, propagation delay increases as well, and the sender waits for a longer time for the ACK to return. This decreases channel utilization.
2. If the time it takes for the ACK to return exceeds the ACKTimeout parameter, the sender will retransmit unnecessarily and waste bandwidth.

We illustrate these problems by performing experiments using the wireless channel emulator. To emulate long distances, we configure the emulator to introduce a delay to emulate links ranging from 0–200 kilometers. Figure 4.1 shows the performance of the 802.11 stop-and-wait link recovery mechanism over increasing link distances. With the MAC-layer ACKs turned off (No ACKs), we achieve a throughput of 7.6 Mbps at the PHY layer data rate of 11 Mbps. When MAC ACKs are enabled, we adjust the ACK timeout as

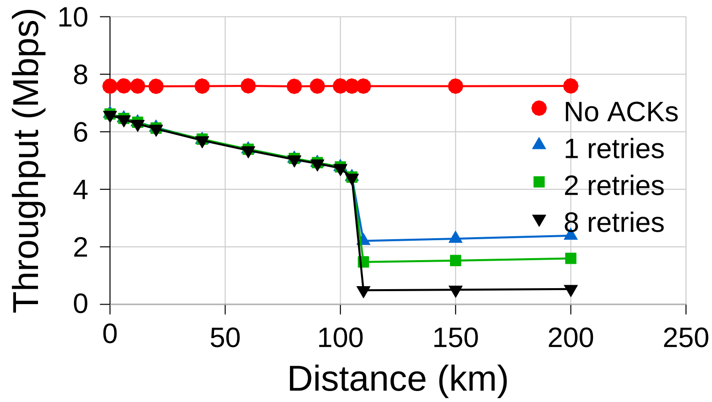

Figure 4.1: Unidirectional UDP throughput for standard 802.11 with CSMA on single emulated link. ACK timeouts were adjusted with increasing distance (on Atheros cards). Traffic is 1440 byte CBR UDP packets in 802.11b at PHY layer datarate of 11 Mbps.

the distance increases. In this case, the sender waits for an ACK after each transmission, and we observe decreasing channel utilization as the propagation delay increases. At 110 km, the propagation delay exceeds the maximum ACK timeout (746 μs for Atheros; this is smaller for Prism 2.5 chipsets and cannot be modified) and the sender always times out before the ACKs can arrive. We notice a sharp decrease in received throughput, as the sender retries to send the packet repeatedly even though the packets were most likely received, until the maximum number of retries is reached (this happens because if an ACK is late, it is ignored). This causes the received throughput to stabilize at $BW_{110km}/(no\_of\_retries + 1)$.

4.2.2 Collisions on long-distance links

The 802.11 protocol uses a CSMA/CA channel-access mechanism, in which nodes listen to the medium for a specified time period (DIFS) before transmitting a packet, thus ensuring that the channel is idle before transmission. This translates to a maximum allow-

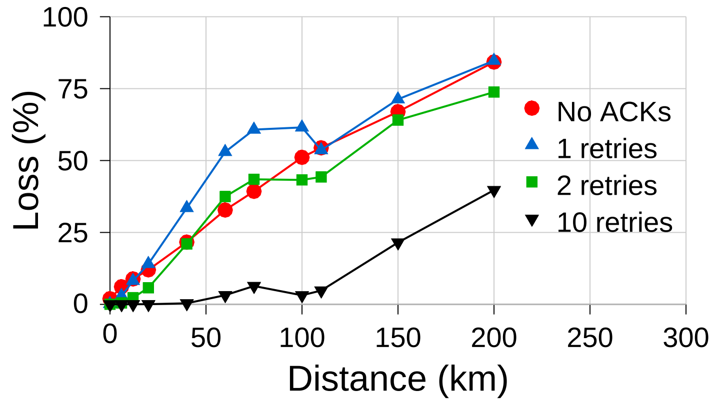

Figure 4.2: Bidirectional UDP loss for standard 802.11 with CSMA. ACK timeouts were adjusted with increasing distance (on Atheros cards). Traffic is 1440 byte CBR UDP packets in 802.11b at PHY layer datarate of 11Mbps

able distance at which collisions can be avoided of about 15km for 802.11b (DIFS is 50μs), 10.2 km for 802.11a and 8.4 km for 802.11g. For longer links it is possible that a node starts transmitting a packet unaware of another packet transmission at the other end. As the propagation delay increases, this probability of loss due to collisions increases.

We illustrate the above-mentioned effect by using a simple experiment: we send bidirectional UDP traffic at the maximum possible sending rate on the emulated link and measure the percentage of packets successfully received at each end. Figure 4.2 shows how the packet loss rate increases with distance. Figure 4.3 shows the sum of the throughputs achieved at both ends for bidirectional UDP traffic as we increase the distance for a link. Note that there are no losses due to attenuation or outside interference in this controlled experiment; all of the losses are due to collisions.

A possible solution to this issue would be to increase the DIFS time interval in order to permit longer propagation delays. However, just as in the case of the ACK timeout,

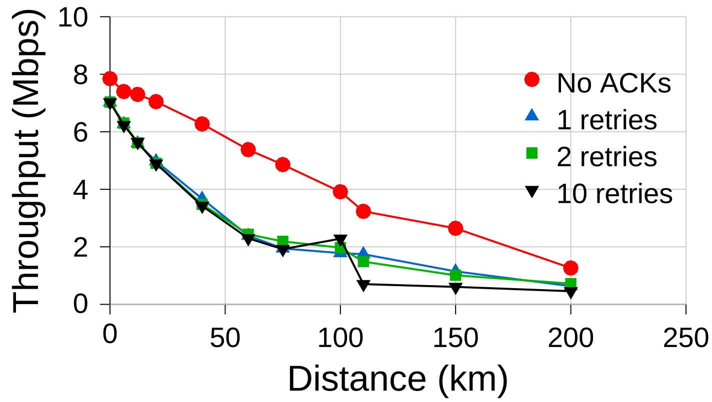

Figure 4.3: Bidirectional UDP throughput for standard 802.11 with CSMA on single emulated link. ACK timeouts were adjusted with incresaing distance (on Atheros cards). Traffic is 1440 byte CBR UDP packets in 802.11b at PHY layer datarate of 11Mbps

this measure would decrease channel utilization substantially for longer links. Furthermore, we are not aware of any 802.11 card chipsets that allow the DIFS interval to be configured via the driver.

4.2.3 Multiple Link Interference

Another important source of errors is the interference between adjacent 802.11 links operating in the same channel or in overlapping channels. While interference between adjacent links can be avoided by using non-overlapping channels, there are numerous reasons, as described by Raman *et al* [118], that make it advantageous to operate adjacent links on the same frequency channel. Moreover, there are WiLD topologies such as the Akshaya network [98] where different channels cannot be allocated to all the pairs of adjacent links, given the high connectivity degree of several nodes.

Inter-link interference occurs because the high-power radios create a strong RF field in the vicinity of the radios, enough to interfere with the receptions at nearby co-located radios. Directional antennas also have sufficiently high gain (4–8 dBi) side lobes [36] in addition to the main lobes.

The first type of problem occurs when multiple radios attached to the same node attempt to transmit at the same time. As soon as one radio starts transmitting after sensing the carrier to be idle, all other radios in the vicinity find the carrier to be busy and backoff. This is desirable in a broadcast network to avoid collisions between two senders at any receiver node. However, in our network where each of these radios transmits over point-to-point long-distance links to independent receivers, this backoff leads to suboptimal throughput. A second problem occurs when packets being received at one link collide with packets being simultaneously transmitted on some other link on the same node. The signal strength of packets transmitted locally on a node overpowers any packet reception on other radios.

In order to illustrate and quantify the magnitude of these effects, we perform experiments on two adjacent links from our testbed (Table 4.1), K-P and P-M, using CBR UDP traffic, with the radios configured to operate in 802.11b mode at a datarate of 11Mpbs.

First, we transmit UDP traffic from P to both nodes K and M simultaneously. We gradually increase the packet sending rate and also vary the channel separation between the adjacent links. We observe that the total cumulative send-throughput on both links can be as high as 14.20 Mbps when they are on non-overlapping channels (separation $\geq$ 4), but this drops by 50% to only 7.88 Mbps when the links operate on the same channel. Next we perform a similar experiment in which we send UDP traffic from node M to node

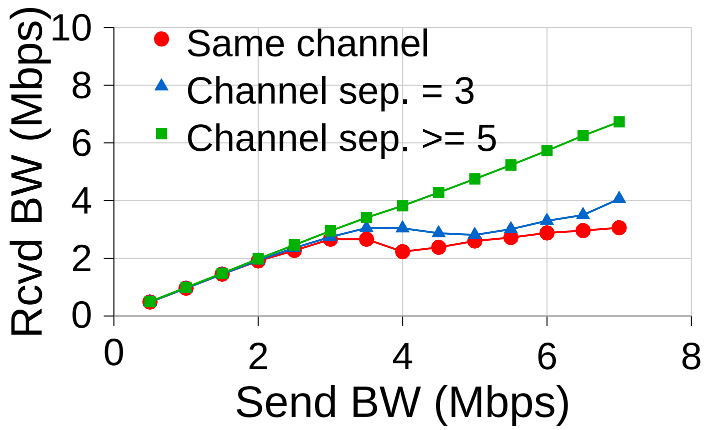

Figure 4.4: Effect of interference on received UDP throughput when sending from M to K through relay node P (see Table 4.1). Channel separation is the number of 802.11b channels by which the two links, M-P and P-K,are separated by. Traffic is 1440 byte CBR UDP packets in 802.11b at PHY layer datarate of 11Mbps

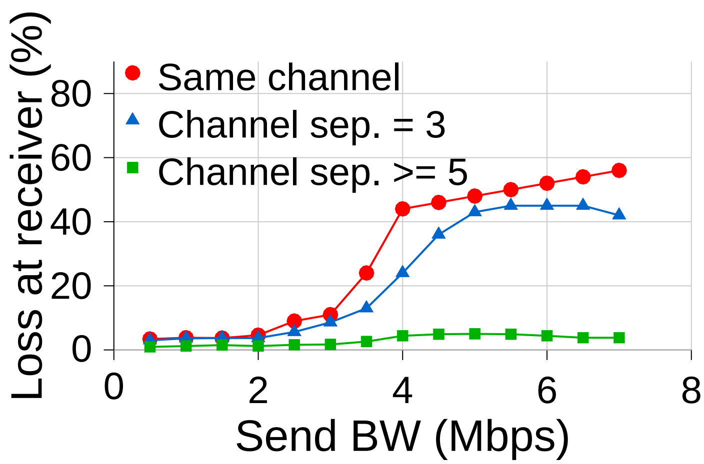

Figure 4.5: Effect of interference on received UDP error rate when sending from M to K through relay node P (see Table 4.1). Channel separation is the number of 802.11b channels by which the two links, M-P and P-K, are separated by. Traffic is 1440 byte CBR UDP packets in 802.11b at PHY layer datarate of 11Mbps

K, relayed through node P. We then measure the received throughput and packet loss rate for various channel separations between the two adjacent links, as presented in Figures 4.4

and 4.5. We observe that interference does reduce the utilization of the individual links and significantly increases the link loss rate (even in the case of partially overlapping channels). For links operating on the same channel, loss rates can be as high as 60%, and even for a channel spacing of 4 the loss rate due to interference can exceed 40%.

Therefore, the maximum channel diversity that one can simultaneously use at a single node in the case of 802.11b is restricted to 3 (channels 1, 6 and 11) which may not be sufficient for many WiLD networks. This motivates the need for a scheme that allows the efficient operation of same-channel adjacent links, as proposed in Chapter 5.

4.2.4 Implications

1. Per-packet acknowledgments and a stop-and-wait protocol are extremely inefficient in long distance links. These problems can be alleviated by disabling per-packet acknowledgments and switching to a more efficient flow-control mechanism that permits batched transfer of packets to achieve high throughput despite the large propagation delays.

2. Due to long propagation delays, the probability of collisions between packets sent by link ends is very high. This indicates that more appropriate medium-access policies that synchronize between link endpoints and that do not simply rely on carrier sensing are required.

3. Inter-link synchronization between adjacent links operating on overlapping channels is a significant problem. This could be alleviated by employing mechanisms that synchronize transmissions across multiple adjacent links, avoiding inter-link interference.

We address all these issues by designing and implementing WiLDNet, a new MAC for long-distance WiFi links (Chapter 5).

4.3 Channel-induced Loss

Apart from protocol shortcomings, another cause for poor performance is high packet loss rates in the underlying channel due to external factors, independent of the protocol used by the links. We refer to these as *channel-induced losses*.

4.3.1 External WiFi Interference

In this section, we investigate external WiFi interference as a potential source of packet loss in WiLD links. Any WiFi traffic that is not a part of the primary WiLD link is categorized as external WiFi interference. Based on the measurements performed on our WiLD testbed and on the wireless channel emulator, we show three key results:

1. In the presence of external WiFi interference, the loss rate is strongly correlated with the amount of external traffic received on the same and adjacent channels. In contrast, there was no such strong correlation observed in Roofnet [37].

2. Packet loss due to external WiFi interference is far more significant in WiLD deployments than local mesh networks.

3. The loss due to external WiFi interference depends on the relative power level between the primary and external traffic, their channel separation, and the rate of external interference.

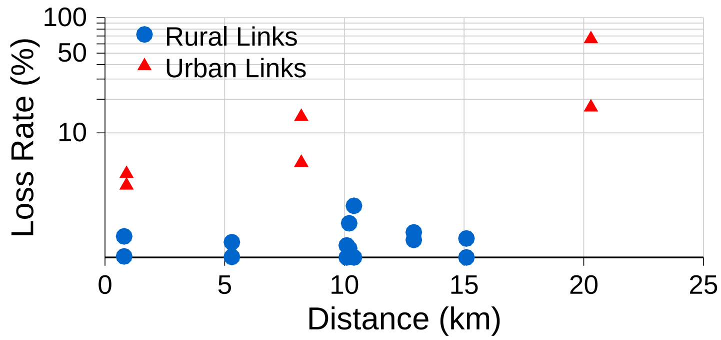

Figure 4.6: Scatter plot of loss rates observed in links deployed in urban and rural areas. Loss rate is plotted on log scale.

In the rest of this section we investigate the correlation between loss rate and external WiFi traffic, the higher impact of external interference due to the exacerbated hidden terminal effect in WiLD settings, and the effects of the relative power and rate of external interference.

Correlation of loss rate and external WiFi traffic

To measure the effect of external WiFi traffic interference on loss in our WiLD links we create a virtual interface in monitor mode as described in Section 4.1. A CBR traffic source of 1 Mbps is used to generate traffic on the WiLD link and the loss rate is averaged every minute.

Figure 4.6 shows the loss rate across all (rural and urban) our WiLD links. We observe that the loss rate of the urban links vary across a wide range (4–70%). In contrast, all the rural WiLD links have a very small loss rate. The maximum loss rate observed in all our rural WiLD links was 2%.

To study this contrast between the rural and urban links, we collected detailed

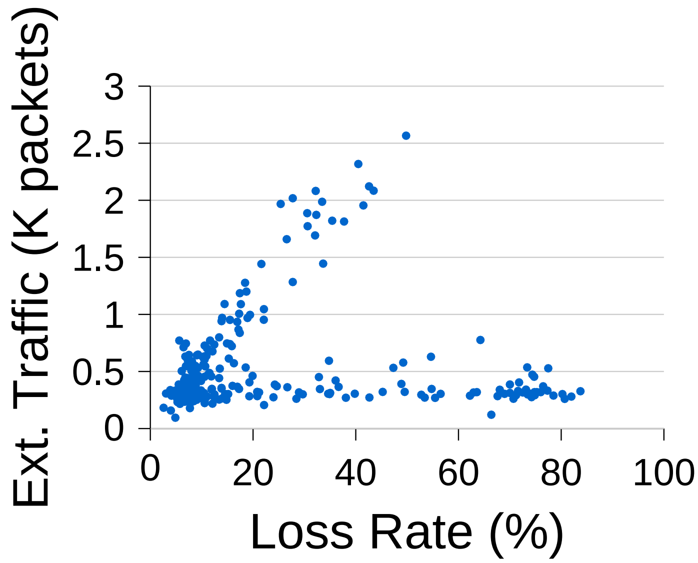

Figure 4.7: Correlation between loss rate and external traffic observed on a WiLD link P-K (see Table 4.1). Traffic is 5Mbps UDP CBR packets of 1440 bytes each at 802.11b PHY datarate of 11Mbps.

packet level MAC traces. By parsing the MAC header source and destination fields, we were able to count the number of frames received from external WiFi sources (interference). In the traces collected over all our rural links we do not capture any external WiFi traffic. However, significant amount of external WiFi traffic was captured from the traces collected in the urban WiLD deployment.

Figure 4.7 shows a scatter plot between the loss rate and the absolute number of external WiFi traffic frames received on an urban link (K-P) for a period of 6 hours. The figure shows that a subset of the loss rate samples are strongly correlated with the external traffic. For the other subset of the samples, the loss rate increases even when there is no significant increase in WiFi traffic on the same channel.

To investigate this further, we performed a controlled experiment using the wireless

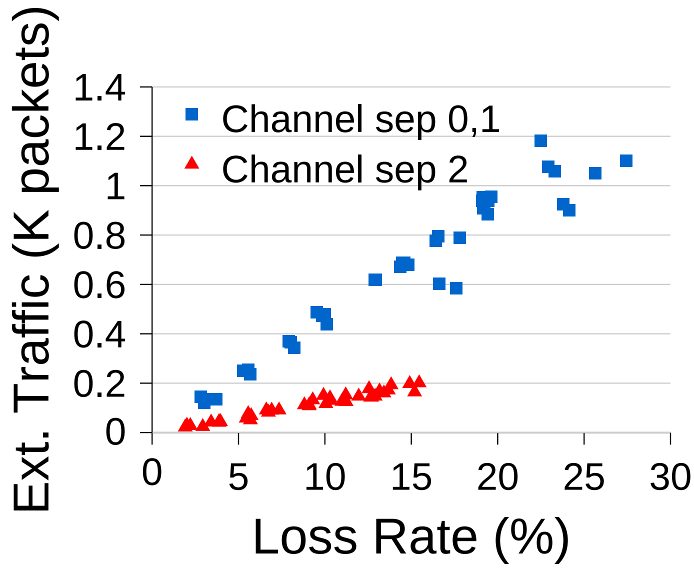

Figure 4.8: Correlation between loss rate and external traffic observed in the wireless emulator. Traffic is 5Mbps CBR UDP packets of 1440 bytes each at 802.11b PHY datarate of 11Mbps.

channel emulator. To model interference from an external traffic source, we introduced a controlled interfering source at the receiver. The traffic rate of the interfering source was varied from 0.1 to 1 Mbps and the traffic rate on the primary link was fixed at 5 Mbps. The data rate was fixed at 11 Mbps on both links. Figure 4.8 shows a scatter plot of the loss rate and the total number of frames received from the external interference source. We observe that for a given loss rate, the amount of external traffic captured by the monitor device depends on the channel separation of the primary and interference source.

The above observed trend is the same as that in Figure 4.7. At a channel separation of 0 and 1, the receiver can receive both the primary link traffic as well as the frames from the interference source. Hence, the loss rate is directly correlated with the amount of external WiFi traffic captured by the monitor interface. At a channel separation of 2, the receiver is not able to receive the frames from the external interference source. However, the signal spillage of the interference source in the primary channel is sufficient to cause frame corruption. This was validated by collecting detailed packet level logs at the MAC layer. From these traces we observed that almost 100% of the lost frames contained CRC errors.

Effect of hidden terminals in WiLD networks

Unlike outdoor mesh network deployments such as Roofnet [37], WiLD deployments show significant correlation between loss rate and external interference. In a mesh-network deployment, an external interference source (I) that is within range of the omni-directional transmitter (Tx) could sense the medium to be free and backoff its transmission. However in WiLD links, the long distance between the two end-points increases the propagation delay, and highly directional antennas with narrow beamwidths are used for transmission. These factors in combination exacerbate the *hidden terminal* problem in WiLD networks, increasing the likelihood of collisions.

Collisions at the receiver can manifest in two different situations: a) when I does not hear Tx, and initiates a transmission when the medium is busy with an ongoing packet transmission from Tx, and b) When Tx does not hear I, and causes a collision by interrupting an ongoing packet transmission from I.

To isolate these two cases and measure the performance degradation due to each case, we perform controlled experiments using two WiFi links. We simultaneously send packets from both Tx at 512 Kbps and I at 3Mbps, and measure the packet loss rate on the primary link ($Tx \rightarrow Rx$) with MAC-layer ACKs disabled.

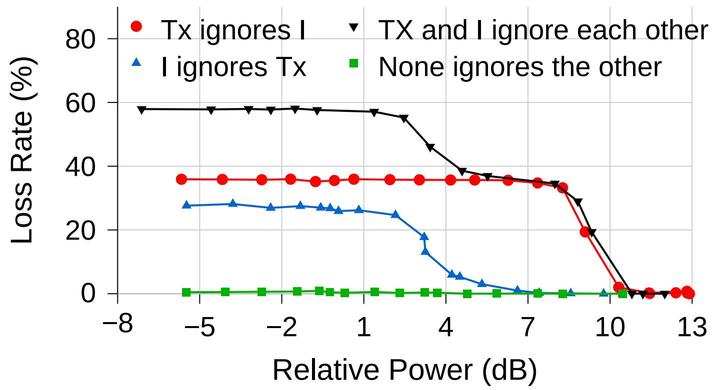

Figure 4.9: Losses due to hidden terminal effects. Both main and interfering traffic is 1440 byte CBR UDP packets at 802.11 PHY datarate of 11Mbps.

To create the situation where Tx cannot hear I, we disable the Clear Channel Assessment (CCA) at Tx, which simply causes Tx to ignore I. We also eliminate propagation delay between Tx and I so that I's CCA works perfectly. We reverse the operations to create the situation in which I cannot hear Tx, but Tx hears I perfectly.

We then run four experiments, mirroring the losses in four situations: when Tx cannot hear I, when I cannot hear Tx, when neither can hear each other (representative of cases in WiLD networks), and when both Tx and I hear each other (representative of most cases in urban mesh networks).

Figure 4.9 shows the loss rate for each of the above four cases. In the case where I ignores Tx, to overcome the interferer completely (achieve 0% loss), packet transmissions from the Tx have to be 7dB stronger than the interfering transmissions. This threshold, at which the primary link is loss free, is higher at 12 dB in the case where Tx ignores I. We can explain this in terms of the capture effect.

When Rx begins receiving a packet P_I from the the interfering source I, Rx's radio calibrates to the channel parameters based on the packet preamble, and continues receiving the packet data without searching for another packet preamble until P_I is fully received. If a subsequent packet P_{Tx} is sent by Tx, this packet can be properly received by Rx only if its signal strength is large enough to cause the receiver to drop packet P_I. Morever, Rx needs to identify P_{Tx}'s preamble quickly. This can happen only if the signal received from Tx is stronger by at least 12 dB than the signal received from I. If instead Tx is the first one to transmit packet P_{Tx}, it is easier for Rx to successfully finish the receipt of this packet. Even if the interferer sends a subsequent packet P_I, the receipt of P_{Tx} is successful as long as the signal received from Tx is 7 dB stronger than that of the interfering source I.

When neither Tx and I can hear each other, both these kinds of collisions are possible. Hence the loss rate is a summation of the losses generated by the above two collision types. However, when both Tx and I are in range of each other, resembling a mesh-network, losses due to collisions are close to zero. In this case, CSMA ensures that the two transmitters, Tx and I, share the medium well.

From the above experiment we conclude that the effect of hidden terminals, causing collisions at the receiver, are greatly exacerbated in WiLD networks compared to urban mesh networks.

Effect of relative power and rate of external interference

To study the effect of relative power and rate of the external WiFi traffic on the loss of the primary link, we perform two experiments using the wireless channel emulator.

In the first experiment, we fix the relative power between the interference source and primary WiLD link, and vary the rate of the external interference source. The received

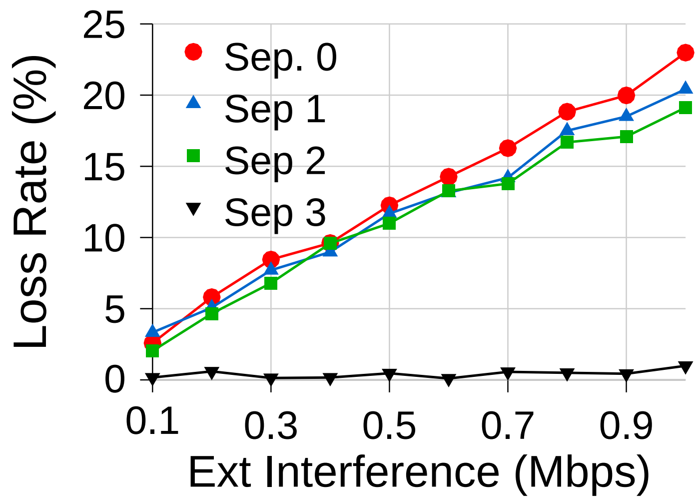

Figure 4.10: Loss rate at different channel separations while varying the interference rates

signal strength of the interfering source was approximately 6dB higher than the primary link traffic. In Figure 4.10, we observe that for channel separations of 0, 1 and 2, the loss rate increases as the rate of the external interference increases. Also, the loss rate is almost the same for all the above channel separations. However, beyond a channel separation of 2, there is no significant interference from the external WiFi traffic source and the loss rate is almost zero.

Figure 4.11 shows the variation in loss rate for different relative power levels of the interference source and WiLD link. In this experiment, we fix the signal strength of the primary WiLD link traffic at 42 dBm and vary the relative power of the interference source from -15 dBm to +13 dBm as shown in the figure. The primary link transmits CBR traffic at 512 Kbps, while the interferer transmits at a rate of 3 Mbps.

We observe that when the interference source is on the same channel, even an

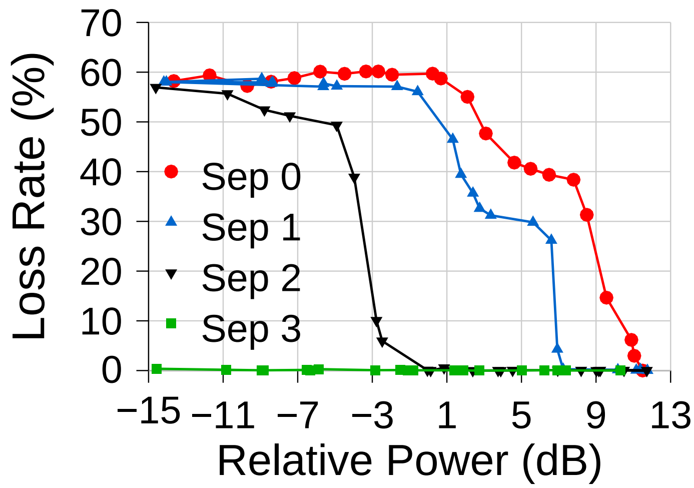

Figure 4.11: Loss rate at different channel separations and different transmit powers of the interfering sources

interference signal which is 12dB lower than the primary WiLD link could lead to packet loss on the primary WiLD link. When the interference source is significantly higher than the WiLD link(-6dB and beyond), the loss rate is very high ($\geq$50%) for channel separations 0, 1 and 2. This corresponds to the situation where any collision results in the capture of the packet on the primary link. At a channel separation of 2, the WiLD link is affected only when the interference source has a higher power. Beyond a channel separation of 2, we do not observe any loss on the primary link.

4.3.2 Non-WiFi Interference

The 802.11b communication protocol operates in the 2.4GHz shared ISM band. This frequency band is shared with a host of other non-802.11 devices, such as microwave ovens, cordless phones, baby monitors, etc. Most of these non-802.11 devices do not follow a channel-access protocol. The lack of a common channel-access protocol could lead to a significant amount of interference caused by these devices.

In Sheth et al. [128], the authors were able to detect and measure non-WiFi interference by sampling the noise floor of the Atheros chipset. The authors observed that in presence of external non-WiFi noise, the noise floor increases linearly with noise. We performed the same experiment on our WiLD testbed, where we sample the noise floor for every packet received. In the presence of external noise causing high loss, we would expect the noise floor to be correlated with the loss rate. However, based on extensive measurements carried out on the urban links we do not see any correlation between noise floor and loss rate. In fact, the noise floor remains mostly constant with only minor 1–2 dB variations.

In addition to the above test, we also check for wide-band non-WiFi noise. A wide-band noise source would cause interference across the entire 802.11 spectrum. Ideally, this can be measured using a spectrum analyzer and detecting a rise in power across the entire spectrum. However, using a spectrum analyzer is infeasible on the outdoor WiLD links. Thus, to detect wide band noise in our WiLD deployment we synchronize the two ends of a link to rotate across channel 1, 6 and 11 periodically. The sender generates 1 Mbps UDP CBR traffic on each channel and the receiver measures the loss rate on each channel. In the presence of any wide-band noise, we would expect to observe a correlated increase in loss rate across all three channels. However, based on long-term experiments performed on three urban links, we determined that there was no statistically significant correlation, and thus no significant broadband noise.

4.3.3 Multipath Interference

Multipath interference is a well known source of packet loss in WiFi networks [26, 45]. It occurs when the RF signal takes different paths from a source to a destination node. Hence, along with the primary line-of-sight signal, the receiver also receives multiple secondary reflections that distort the primary signal. Although it is difficult to measure the exact delay between the primary and secondary paths on our WiLD deployments using commodity off-the-shelf equipment, based on the experiments using the wireless channel emulator we conclude the following:

1. The order-of-magnitude lower delay spreads in WiLD deployments significantly reduce the interference due to multipath in WiLD deployments.

2. If WiLD links are deployed in dense urban deployments with non-line-of-sight, multipath interference could lead to significant loss at the higher 802.11b data rates of 5.5 and 11 Mbps.

Multipath interference in Roofnet and WiLD deployments

The authors of the Roofnet deployment [26] concluded that multipath interference was a significant source of packet loss in Roofnet. However unlike urban 802.11 mesh deployments, multipath interference is significantly lower in WiLD network deployments due to the order-of-magnitude lower delay spreads. The two factors contributing to lower delay spreads in WiLD networks are the long distance between the two end hosts and the line-of-sight deployment of the nodes. The strong line-of-sight component in WiLD deployments ensures that the attenuation of the primary signal is only due to path loss,

Dist. (km)	Delay spread (μsec)
0.5	(4.75, 3.59)
1.0	(2.4, 1.80)
8.0	(0.3, 0.22)
16.0	(0.15, 0.11)
100.0	(0.02, 0.01)

Table 4.2: Delays between a primary and secondary reflection

and most of the secondary paths are due to reflections from the ground. In comparison to our WiLD deployment, the Roofnet deployment has shorter links and non-LOS deployments. The median link length is 0.5 km and the longest link is 2.5 km, and links are rarely line-of-sight.

Table 4.2 shows the delay between the primary path and secondary path assuming the antenna is mounted at a height of 30 meters and reflection is only from the ground. The two delays are computed for a secondary path reflecting at the quarter point and at the mid-way point between the transmitter and the receiver. Although multipath reflections arriving at the receiver are not constrained to these distances, the table provides the relative difference in delay spreads observed in the Roofnet deployment and WiLD deployment. As the length of the link increases, the primary and the secondary path travel almost the same distance, and hence the delay between the primary and secondary reflection reduces. As seen from the table, there is an order-of-magnitude difference between the delay in WiLD links and medium range Roofnet links. Aguayo et al. [26] also observed that the RAKE receiver is able to tolerate delay spreads upto 0.3–0.4 μsec.

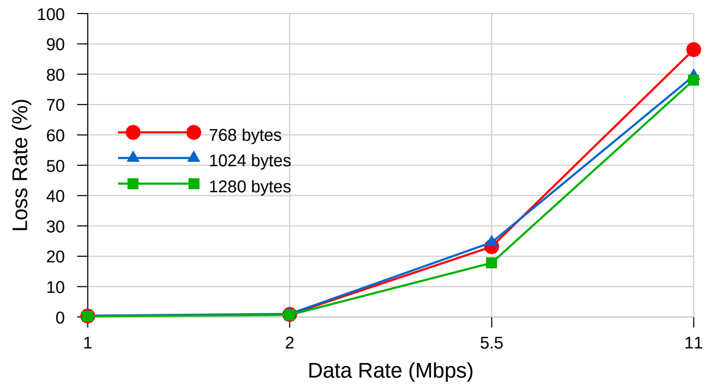

Figure 4.12: Higher data rates are more susceptible to multipath interference

To further validate that multipath interference is not a significant source of packet loss, we perform measurements over WiLD links deployed in rural environments. Our hypothesis was that most of the loss in our urban deployment was due to external WiFi interference. Hence, in absence of external interference the WiLd links deployed in the rural areas should not have any loss. Figure 4.6 validates our hypothesis, which shows loss rates observed across three such rural links. From the table we observe that the maximum loss rate observed was 1.7% with low variance.

Effect of non-line-of-sight dense urban deployments

To study the effect of multipath when WiLD links are deployed in absence of line-of-sight, we perform controlled experiments in the wireless channel emulator. Due to the lack of analytical models, we build an artificial model consisting of 12 reflected paths with the path delay increasing in steps of 0.18 μsec and the power exponentially decaying. Hence the maximum delay between the primary and the longest secondary path is 2.16μsec.

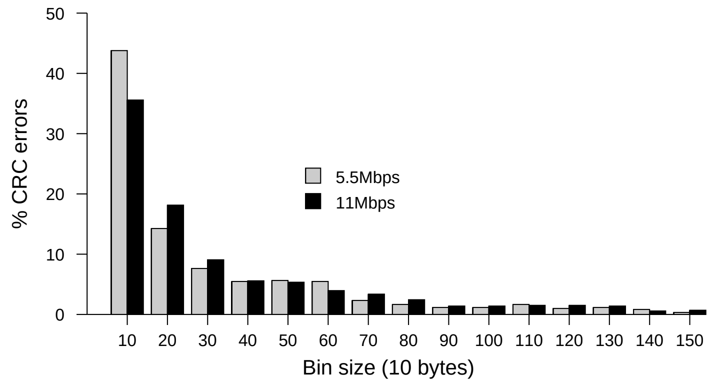

Figure 4.13: Effect of multipath in non-line-of-sight dense urban deployments. Distriution of number of bytes being corrupted for payload of 768 bytes

Figure 4.12 shows the loss rate for payloads of size 768, 1024 and 1280 bytes and the four data rates of 802.11b. From the figure we observe that the lower data rates are resilient to multipath and the length of a frame does not affect the loss rate.

Looking closer at the packet traces collected at the receiver node, we observed that almost 100% of the errors were caused due to CRC errors. Figure 4.13 shows a histogram of the number of bytes corrupted as a percentage of total number of CRC error frames received. Since the loss rates were almost zero for 1 and 2 Mbps, we present the histogram only for the 5.5 Mbps and 11 Mbps data rates. From the figure we observe that the distribution has a heavy tail and almost 90% of the CRC error frames have less than 10% of the bytes corrupted in the payload.

4.3.4 Implications

In the previous sections we have measured the effects of external interferece and multipath in WiLD networks. Here we discuss the implications of these effects for WiLD protocol and network design.

External Interference:

1. We conclude that external WiFi interference is a significant source of packet loss in WiLD networks. Any deployment of WiLD networks in dense urban deployments has to take into account external WiFi interference. In rural links, if one end of the link is in an urban area, external interference on the urban side can significantly affect the overall throughput of the link

2. When calculating the link budget for urban links, it is beneficial to over-provision the received power. A high received signal strength relative to that of external interference can potentially isolate the WiLD link from external WiFi traffic.

3. MAC layer adaptation algorithms like adaptive channel switching, rate adaptation, and adaptive FEC could significantly reduce the loss due to external WiFi interference. In section 4.6 we evaluate each one of these as potential remedies to mitigate external WiFi interference.

Multipath:

1. Multipath interference is not a significant problem in WiLD links in line-of-sight conditions as a result of very low delay spreads. But in dense urban non-line-of-sight deployments, the higher data rates of 11 Mbps and 5.5 Mbps are much more sensitive to multipath interference as compared to the lower data rates of 1 Mbps and 2 Mbps. This has an implication on the rate selection algorithm for applications that can trade-off bandwidth for a loss free channel. For example, in the presence of sig-

nificant multipath interference, a rate selection algorithm for such applications could directly move from 11 Mbps to 1 or 2 Mbps instead of stepping down to 5.5 Mbps.

2. In Figure 4.13 we observe that 90% of the corrupted frames have less than 10% of bytes corrupted. Also from Figure 4.12 we observe that increasing the length of the payload does not affect the loss rate. Hence, an alternate approach to rate adaptation could be to divide the payload into smaller blocks and encode these blocks to add in sufficient redundancy to tolerate the CRC errors due to multipath interference.

4.4 Loss Variability

In this section, we analyze the variability of packet loss over time on the WiLD links. We first propose a simple mechanism we use to classify loss periods as either bursts or residual losses and then individually describe the loss characteristics for bursts and residual losses.

4.4.1 Burst-Residual Separation

We observe that all the links in our testbed exhibit a bi-modal loss variation over time where the loss-rate at any given time can be classified into into two categories: *bursts* and *residual* losses. While bursts refer to time-periods with sharp spikes in the loss rate, residual losses refer to the losses that constantly occur in the underlying channel over time. Unlike previous studies on WiLD links in rural environments [44], we observe a non-zero residual loss-rate in most of our links in urban environments.

Given the bimodal loss variation property, we use a simple mechanism to separate bursts and residual losses. To classify each time-period into either a bursty or residual loss-period, we determine a *demarcation region* for the loss distribution on a given link. We estimate parameters p_1 and p_2 ($> p_1$) such that a significant majority ($> 99\%$) of the loss samples fall in the regions $[0, p_1)$ and $(p_2, 1]$. All loss periods with the loss-rate in the range $[0, p_1)$ are classified residual and those in the range $(p_2, 1]$ are classified bursty. The remaining samples are considered transition phases. If adjacent loss periods of a transition period are bursty, then the transition phase is also classified as bursty.

4.4.2 Burst characteristics

To analyze burst characteristics, we need to measure the variability of three parameters associated with bursts: duration, arrival pattern and magnitude.

Burst duration and arrival

Based on the duration of bursts, one can classify a burst as either as a *short burst* or a *long burst*. Across our links, we observe a majority of the bursts to be short bursts that last for less than 0.3s. The median loss rate is less than 1s across most links. However, in certain links, especially those in urban environments, we observe a continuous burst period that can last up to 70s. The characteristic arrival pattern that we observed for long bursts is that a single long burst is followed by a string of other long bursts separated by short time-periods (on the order of a few seconds). Overall, the entire string of long bursts that occur together in time lasts for several minutes representing elongated time periods where the underlying channel experiences very high loss rates. Based on the results in Section 4.3.1, we conclude that these elongated bursts occur due to interference from external WiFi traffic

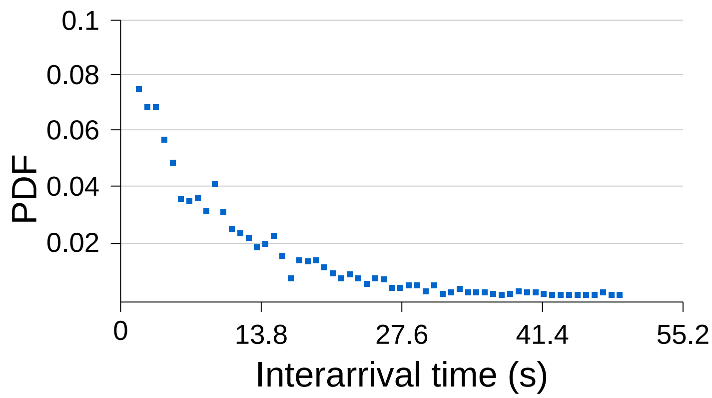

Figure 4.14: Probability distribution of inter-arrival time of bursts (R-B link)

sources.

We next focus on the arrival pattern of short bursts. Figure 4.14 shows the distribution of inter-arrival times between bursts with short durations for the R-B link in our testbed. For this link, we observe that the underlying distribution of inter arrival time resembles an exponential distribution with a mean inter-arrival time of 15s. In addition, we observe that the inter-arrival time distribution is stationary across various time-periods. These observations suggest that the underlying arrival process can potentially be modeled based on a Poisson arrival process. We observe a similar behavior across all the links in our testbed.

Burst-loss magnitude

We found burst magnitudes to be very hard to predict. For both short spikes and long-duration bursts, the loss-rate varied across the entire spectrum between $10-60\%$. Even within a single burst, we observed the loss-rate across episodes to fluctuate rapidly. Given that our links operate in static environments, such wild fluctuations in very short periods appear to be triggered due to external WiFi interference as opposed to multi-path fading channel conditions.

4.4.3 Residual loss characteristics

Every link in an urban environment in our testbed exhibits a non-trivial residual loss rate where packet losses occur at regular intervals as opposed to bursts. The residual loss rate varies between $1 - 10\%$ in our urban links in the testbed. However, residual loss-rates are negligible in our rural links. Upon analyzing the loss distributions over different timescales for different links, we make two observations. First, except for one specific link (K-P), we observed that the residual loss distribution is stationary over hourly time scales while on the K-P links, the distribution is time-varying. Second, we observe that the residual loss rate on any link remains roughly constant over a few minutes even in the presence in short bursts during such periods.

4.4.4 Implications

In summary, we make three observations. First, we can classify the loss sample at any time period into three categories: short burst, long burst or residual. Second, while the arrival of short bursts can be approximately modeled based on a Poisson arrival process, the arrival of long bursts are highly correlated in time and not memoryless. Finally, unlike rural links which exhibit negligible residual losses, we observe a non-negligible residual loss-rate in urban environments.

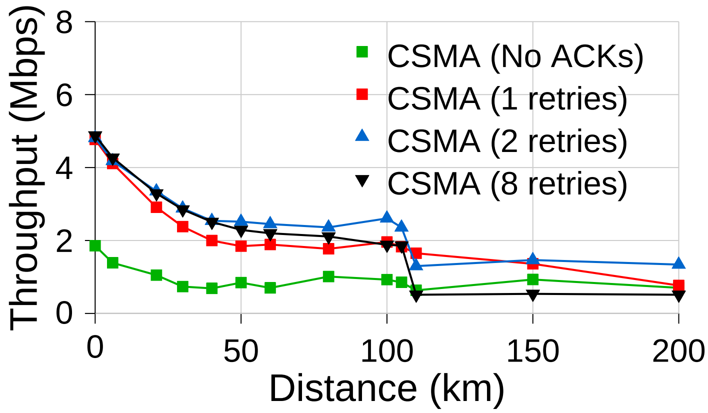

Figure 4.15: Cumulative throughput for TCP in both directions simultaneously over standard CSMA with 10% channel loss on an emulated link. Traffic is 802.11b at PHY layer datarate of 11Mbps

4.5 Impact on TCP

Taken together, the protocol shortcomings of 802.11 and channel induced losses significantly lower end-to-end TCP performance. The use of stop-and-wait over long distances reduces channel utilization. In addition, we see correlated bursty collision losses due to interference from unsynchronized transmissions (over both single-link and multi-hop scenarios) as well as from external WiFi sources. Under these conditions, TCP flows often timeout resulting in very poor performance. To handle these losses, the only knob available in the driver is to tune the number of packet retries. Setting a higher value on the number of retries decreases the loss rate, but at the cost of lower throughput resulting from lower channel utilization.

To better understand this trade-off, we measure the aggregate throughput of TCP flows in both directions on an emulated link while varying distance and introducing a channel packet loss rate of 10%. Figure 4.15 presents the aggregate TCP throughput with various number of MAC retries of the standard 802.11 MAC. Due to increased collisions and larger ACK turnaround times, throughput degrades gradually with increasing distances.

4.6 Remedies

Having identified external WiFi interfernce as the main source of packet loss in WiLD networks, in this section we outline the potential remedies to mitigate external WiFi interference. We evaluate adaptive frequency selection, rate adaptation and adaptive forward error correction (FEC) algorithms as the potential remedies. For each, we simulate the adaptation algorithms and measure the improvements gained for real loss traces from our testbed and experiments performed on the wireless channel emulator.

4.6.1 Frequency Channel Adaptation

A simple solution to mitigate external WiFi interference could be to select an alternate less congested channel and switch to that channel. To motivate this simple technique we perform a channel switching experiment on our WiLD deployment on the K-P link. The source and destination switch between channel 1 and 11 synchronously every 30 seconds. Figure 4.16 shows the variability of loss rate across the two channels for a period of about 2 hours. We can observe that both channel 1 and 11 show bursts that stretch upto a few minutes. It is important to note that by averaging the loss rate over 30 seconds we are not capturing the transient changes in the channel conditions.

Given the above loss trace across the two channels, table 4.3 compares different channel switching algorithms by the achieved loss rate and the no of channel switches

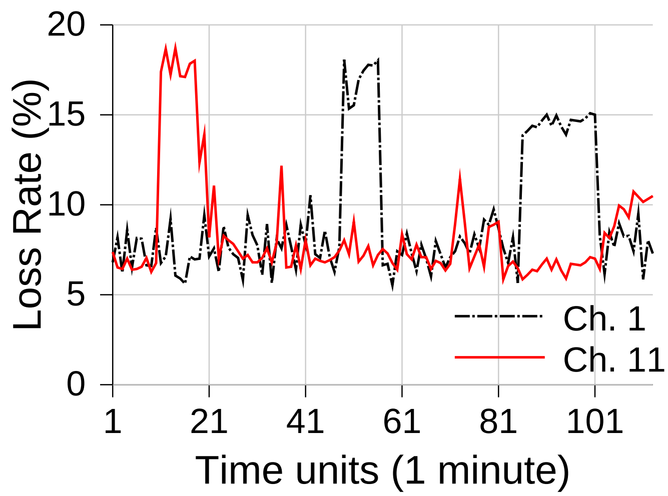

Figure 4.16: Loss variation over time across channels 1 and 11. Loss rate averaged every 1 minute.

required. In the base case (No adapt), where the channel is fixed at either channel 1 or 11, the average loss rate across the entire trace is either 9.2 or 8.3%. If the receiver has complete knowledge of the loss rate on both channels 1 and 11 at the beginning of a time interval (Oracle), then switching to the least lossy channel at any given time achieves the lowest loss rate (at 6.8%); but this comes at a cost of frequent switches of the channel. Adding a small hysteresis of 5% (Oracle 5%) for channel switching reduces the number of switches from 40 to 26 without increasing the average loss rate significantly. In absence of knowledge of loss rates on other channels, we can use the simple approach of jumping to the alternate channel when the loss rate on the current channel exceeds a threshold (e.g. 10% in Change $\geq$ 10%).

Although the reduction in loss rate shown in Table 4.3 by the different algorithms is only of the order of 1-2%, the advantages of channel switching could be significant in

	Loss	Number of Switches
No adapt	(9.2, 8.3)	0
Lowest rate	6.8	40
Oracle (5%)	7.01	26
Change $\geq$ 10%	7.76	8

Table 4.3: Channel switching algorithms for the trace (loss rate and number of switches)

presence of long or high-loss bursts.

Implications of channel switching

Even though adaptive channel switching seems to be a viable solution, large scale WiLD mesh deployments require careful channel assignment to avoid interference between multiple radios mounted on the same tower [111, 118]. Switching the frequency channel on one link could lead to a cascading effect requiring other links to also change their operating channel. Hence, although it could mitigate interference, it is not always possible to switch a frequency channel in a large scale deployment.

4.6.2 Rate Adaptation

Figure 4.17 shows the variation of loss rates as the relative power of the primary transmitter is increased with respect to that of the interference source for different 802.11 datarates.

We observed that in presence of external WiFi interference, data rate adaptation could either degrade the performance further or cause no effect on the loss rate. From

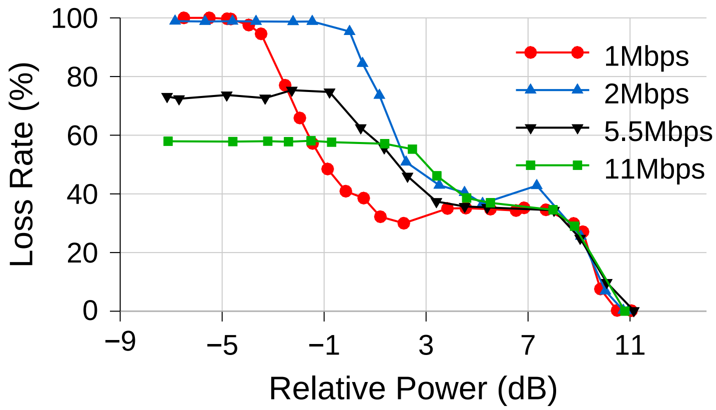

Figure 4.17: Loss rate for 802.11b encoding rates at varying power of transmitter relative to interferer. Traffic is 1440 byte UDP CBR packets at 11Mbps PHY datarate of 802.11b.

figure 4.17 we see that when the received signal strength of the primary transmitter is higher than that of the interference source (from 0 to 12 dB), there is no difference in the loss rate for different 802.11b datarates. Whereas, when the interferer is stronger than the transmitter, reducing the data rate actually exacerbates the performance. This is because the increased transmission time of the frame increases the probability of a collision with the external traffic.

Implications for datarate selection

Most of the 802.11 radios have built in rate-adaptation algorithms which selects a lower rate with resilient encoding on experiencing high loss. However, the above analysis shows that in the presence of loss due to external WiFi interference, it is not worthwhile to adapt the data rate. Rather, we propose using other techniques such as adaptive FEC and link-layer retransmissions to mitigate the loss.

4.6.3 Adaptive Forward Error Correction

As discussed in the previous two sections, both channel and rate adaptation may not be feasible in large-scale WiLD mesh deployments. Furthermore, they only provide coarse-grain adaptations, which may not be suitable for QoS specific applications like video streaming. In this section we propose adaptive FEC as a solution to achieve fine-grained control. With an estimate of the channel loss variability, adaptive FEC allows addition of the “right” amount of redundancy to cope with the channel losses.

We evaluate a simple Reed-Solomon based adaptive FEC mechanism. Time is divided into slots (25 ms) and at the end of each slot the receiver informs the transmitter of the loss observed in the previous slot. Based on this link information, the transmitter adjusts the redundancy for the next round. To deal with transient spikes in loss rate, the sender maintains a moving window average of the loss rate (WinSize = 10). The application traffic is assumed to be a CBR traffic source (1.8 Mbps) and consuming only half the available bandwidth (3.8 Mbps at 11 Mbps); there is sufficient bandwidth per slot to introduce 100% redundancy. To differentiate between new incoming packets and backlogged packets that were dropped in the previous slot, two separate queues are maintained. New incoming packets are queued at the beginning of every slot in the *application queue*. A *backlog queue* (2 x application queue length) maintains a queue of packets that have been transmitted, but not yet successfully acknowledged. Packets in the application queue are given higher priority as compared to the backlogged packets. Packets are dequeued off the backlog queue for re-transmission only when there is additional bandwidth remaining in the current slot. In absence of the backlog queue, packets transmitted with inadequate redundancy would

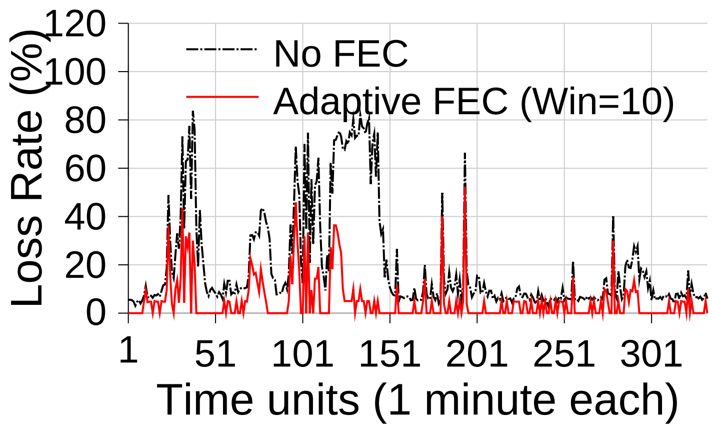

Figure 4.18: Comparison of loss rate observed with and without adaptive FEC. Adaptive FEC can significantly reduce the loss rate during periods of long bursts.

be dropped.

Figure 4.18 shows a loss trace on the M-P link. The traffic source was a 1.5 Mbps UDP CBR traffic generator and the loss rate was averaged over 1 minute for a duration of approximately 6 hours. Here again, the MAC-layer ACKs were turned off and retries set to zero. From the above figure we observe that the link was extremely bursty with bursts as high as 70–80% lasting for 20–30 mins. Table 4.4 shows the performance comparison of the adaptive FEC algorithms. We measure the average loss rate and percentage of packets delayed. A packet is marked as delayed if it is received in a slot later than its scheduled slot. The baseline case is when there is no FEC being applied (No FEC). In this case, the average loss rate across the entire 6 hour period is 19.98% and approximately 12.67% packets were delayed. If the exact loss rate could be predicted for each slot (Oracle), then the loss rate and delayed packets are zero. However, in practice the channel loss rate cannot be predicted accurately, especially since the loss rate is determined by the external WiFi interference. A

	Loss Rate	% of packets delayed
No FEC	19.98	12.67
Oracle FEC	0	0
Adapt FEC (Win = 10)	4.78	2.86

Table 4.4: Summary of the effectiveness of adaptive FEC over the trace shown in Figure 4.18
.

simple approach is to maintain a moving average window of the loss rate and for every time slot encode the frames to transmit additional redundant packets determined by the current value of the moving average.

Table 4.4 shows that this simple approach (Adapt FEC) significantly reduces the loss rate to 4.78% and the packets delayed to 2.86%. Figure 4.18 shows the loss rate along with the original loss rate. From the figure we observe that the above simple approach can tolerate long bursts of high loss rate. However, FEC cannot adapt to transient high bursts.

4.7 Related Work

There have been several studies on packet loss characteristics. Here we focus on works closely related to ours.

The study by Chebrolu et. al. [44] is the only other measurement-based study of WiLD deployments that we are aware of. However, these two studies are orthogonal. We focus on determining the various sources of loss, understanding loss variability, and studying potential remedies to alleviate loss. Their work focused on the effect of parameters such as weather, SNR, payload size and datarate on loss.

Aguayo et al. [26] present a detailed link layer measurement for Roofnet, an outdoor 802.11 mesh deployment, in which they identify the sources of packet loss. Our study indicates that WiLD deployments are faced with a different set of problems as compared to an outdoor 802.11 mesh deployment. The authors of Roofnet conclude that the main source of packet loss was due to multipath interference. The authors observed only a weak correlation between loss rates and parameters such as SNR and link distances. They did not observe any correlation between loss and external interference. Since our work, a more recent study [54] has reexamined the Roofnet data and has shown that external interference and not multipath was the major source of loss.

Jamieson et.al. [79] experimentally evaluate the limitations of carrier-sense with respect to achieving high throughput in multi-hop environments. Garretto et. al. [53] show that CSMA performs very badly in multihop wireless networks, and that this is not due to any particular implementation of a CSMA protocol, but is indeed a general coordination problem in multihop wireless networks.

A large number of measurement based studies have also been carried out to study the source of packet loss in indoor large scale 802.11 deployments [63, 79, 80, 124]. The authors in [80, 124] study the performance of 802.11 in a conference setting, where a large number of clients are using the wireless network. The authors observed both short- and long-term variability in link quality and performance degradation under heavy usage of the wireless network. The authors also point out that rate fallback exacerbates the link quality, leading to a higher number of retransmissions and dropped frames.

4.8 Summary

The performance of standard 802.11 in long-distance links is quite abysmal. We performed a detailed study of protocol-induced losses due to inefficient transmission and breakdown of CSMA in WiLD settings. We also showed that 802.11b protocol limitations make it unsuitable not just for point-to-multipoint links, as claimed in prior work, but also unsuitable for simple point-to-point links.

We then performed a detailed study of channel-induced losses due to external WiFi interference, non-WiFi interference Our main result is that most of the losses arise due to external WiFi interference on same and adjacent channels. This result is in contrast to loss studies of urban mesh networks, where multipath is reported to be the most significant source of loss.

In addition, we analyzed the loss variability in both urban and rural links and showed that urban links suffer from a higher degree of residual loss. Finally, we proposed and analyzed the effectiveness of three remedial strategies to mitigate the losses caused by external WiFi interference.

Chapter 5

WiLDNet Design

In this chapter, we describe the design of WiLDNet and elaborate on how it addresses the 802.11 protocol shortcomings as well as achieves good performance in high-loss environments. In the previous chapter, we identified three basic problems with 802.11; (a) low utilization, (b) collisions at long distances, and (c) inter-link interference. To address the problem of low utilization, we propose the use of bulk packet acknowledgments (Section 5.1). To mitigate loss from collisions at long distances as well as inter-link interference, we replace the standard CSMA medium access method with a TDMA-based approach. We build upon the TDMA protocol design of 2P [118], and make the necessary changes to adapt 2P to high-loss environments (Section 5.2). Additionally, to handle the challenge of high and variable packet losses, we design adaptive loss recovery mechanisms that use a combination of FEC and retransmissions with bulk acknowledgments (Section 5.3). We

The work presented in this chapter is joint work with TIER group members Rabin Patra, Sergiu Nedevschi, Anmol Sheth, Lakshminarayanan Subramanian and Eric Brewer. The material on protocol design and evaluation has been previously published as "WiLDNet: Design and Implementation of High-Performance WiFi Based Networks"[111].

then describe the implementation of WiLDNet (Section 5.4), present evaluation results (Section 5.5) and discuss the tradeoffs in throughput, delay and error rates achieved by varying several WiLDNet parameters (Section 5.6).

WiLDNet follows three main design principles. Firstly, the system should not be narrowly focused to a single set of application types. It should be configurable to provide a broad trade-off spectrum across different end-to-end properties like delay, bandwidth, loss, reliability and jitter. Secondly, all mechanisms proposed should be implementable on commodity off the shelf 802.11 cards. And thirdly, the design should be lightweight, such that it can be implemented on the resource constrained single board computers (266 MHz CPU and 128 MB memory) used in our testbed.

5.1 Bulk Acknowledgments

We begin with the simple case of a single WiLD link, with each node having a half duplex radio. As we have seen in the previous chapter, when propagation delays become longer, the default CSMA mechanism cannot determine whether the remote peer is sending a packet in time to back-off its own transmission and avoid collisions. Moreover, such a contention-based mechanism is an overkill, since we know precisely that only two hosts share the channel for a directional link.

Thus, a simple and efficient solution to avoid these collisions would be to have an echo protocol between the sender and the receiver, allowing the two end-points to take turns sending and then receiving packets. Hence, from a node's perspective, we divide time into send and receive time slots, with a burst of several packets being send from one host to its peer in each of the slots. Having these bursty transmissions, we are presented with the opportunity to introduce a more efficient flow control mechanism.

Consequently, to improve link utilization, we replace the stock 802.11 stop-and-wait protocol with a sliding-window based flow-control approach in which we transmit a *bulk acknowledgment* from the receiver for a window of packets. We generate a bulk acknowledgment as an aggregated acknowledgment for all the packets received within the previous slot. In this way, a sender can rapidly transmit a burst of packets rather than transmit each frame and wait for an acknowledgment after each.

The bulk acknowledgment can be either piggybacked on data packets sent in the reverse direction, or sent as stand-alone packets if no data packets are available. Each bulk ACK contains the sequence number of the last packet received in order and a variable-length bit vector ACK for all packets following the in-order sequence. Here, the sequence number of a packet is locally defined between a pair of end-points of a WiLD link.

Like 802.11, the bulk acknowledgment mechanism is not designed to guarantee perfect reliability. 802.11 has a maximum number of retries for every packet. Similarly, upon receiving a bulk acknowledgment, the sender can choose to advance the sliding window skipping unacknowledged packets if the retry limit is exceeded. In practice, we can support different retry limits for packets of different flows. The bulk ACK mechanism introduces packet reordering at the link layer, which may not be acceptable for TCP traffic. To handle this, we provide in-order packet delivery at the link layer either for the entire traffic or at a per-flow level.

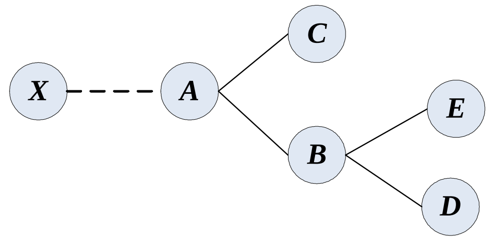

Figure 5.1: Example topology

5.2 Designing TDMA in Lossy Environments

To address the inappropriateness of CSMA for WiLD networks, 2P [118] proposes a contention-free TDMA based channel access mechanism. 2P eliminates inter-link interference by synchronizing all the packet transmissions at a given node (along all links which operate on the same channel or adjacent overlapping channels). In 2P, a node in transmission mode simultaneously transmits on all its links for a globally known specific period, and then explicitly notifies the end of its transmission period to each of its neighbors using marker packets. A receiving node waits for the marker packets from all its neighbors before switching over to transmission mode. In the event of a loss of a marker packet, a receiving node uses a timeout to switch into the transmission mode.

The design of 2P, while functional, is not well suited for lossy environments. Consider the simple example illustrated in Figure 5.1, where all links operate on the same or adjacent overlapping channels. Consider the case where (X, A) is the link experiencing high packet loss-rate. Let T denote the value of the time-slot. Whenever a marker packet transmitted by X is lost, A begins transmission only after a timeout period T_0 $(\geq T)$. This, in turn, delays the next set of transmissions from nodes B and C to their other neighbors by a time period that equals $T_0 - T$. Unfortunately, this propagation of delay does not end here. In the time slot that follows, D's transmission to its neighbors is delayed by $T_0 - T$. Hence, what we observe is that the loss of marker packets has a "ripple effect" in the entire network creating an idle period of $T_0 - T$ along every link. When markers along different links are dropped, the ripples from multiple links can interact with each other and cause more complex behavior.

Ideally, one would want $T_0 - T$ to be very small. If all nodes are perfectly time synchronized, we can set $T_0 = T$. However, in the absence of global time synchronization, one needs to set a conservative value for T_0. 2P chooses $T_0 = 1.25 \times T$. Hence, the loss of marker packets leads to an idle period of $0.25 \times T$ throughout the network (in 2P, this is 5 ms for $T = 20$ ms). In bursty losses, transmitting multiple marker packets may not also suffice.

Given that many of the links in our network experience sustained loss-rates over 5–40%, in WiLDNet, we use an implicit synchronization approach that aims to reduce the value of $T_0 - T$. In WiLDNet, we use a simple loose time synchronization mechanism similar to the basic linear time synchronization protocol NTP [95], where during each time slot along each link, the sender acts as the master and the receiver as the slave. Consider a link (A, B) where A is the sender and B is the receiver at a given time. Let t_{send_A} and t_{recv_B} denote the start times of the slot as maintained by A and B. All the packets sent by A are timestamped with the time difference (δ) between the moment the packet has been sent (t_1) and the beginning of the send slot(t_{send_A}). When a packet is received by B at time t_2, the beginning of B's receiving slot is adjusted accordingly: $t_{recv\_B} = t_2 - \delta$. As soon as B's receive slot is over, and $t_{send\_B} = t_{recv\_B} + T$ is reached, B starts sending for a period T.

Due to the propagation delay between A and B, the send and corresponding receive slots are slightly skewed. The end-effect of this loose synchronization is that the value of $T_0 - T$ is limited by the propagation delay across the link even in the face of packet losses (assuming clock speeds are roughly comparable). Hence, an implicit synchronization approach significantly reduces the value of $T_0 - T$ thereby reducing the overall number of idle periods in the network.

5.3 Adaptive Loss Recovery

Handling high and variable loss-rates primarily induced by external WiFi interference is a challenging problem. However, to achieve predictable end-to-end performance, it is essential to have a loss recovery mechanism that can hide the loss variability in the underlying channel and provide a bound on the loss-rate perceived by higher level applications along a single link. More specifically, the loss recovery mechanism should provide a loss-rate bound q independent of the underlying link loss-rate.

Achieving such a bound is not easy in our setting due to two factors. First, it is hard to predict the arrival and duration of bursts; also, bursts occur very frequently in some of our links. Second, the loss distribution that we observed on our links is non-stationary even on long time scales (hourly and daily basis). Hence, it is not easy to use a simple model to capture the channel loss characteristics. In WiLDNet, we can either use retransmissions or Forward Error Correction (FEC) to deal with losses (or a combination of both). A retransmission based approach can achieve the loss-bound q with minimal throughput overhead but at the expense of increased delay. A FEC based approach incurs additional throughput overhead but does not incur a delay penalty especially since it is used in combination with TDMA on a per-slot basis. However, a FEC approach cannot achieve arbitrarily low loss-bounds mainly due to the unpredictability of the channel.

5.3.1 Tuning the Number of Retransmissions

To achieve a loss bound q independent of underlying channel loss rate $p(t)$, we need to tune the number of retransmissions. One can adjust the number of retransmissions $n(t)$ for a channel loss-rate $p(t)$ such that $(1 - p(t))^{n(t)} = q$. Given that our WiLD links support in-order link-layer delivery (or in-order delivery on a per-flow basis, a larger $n(t)$ also means a larger maximum delay, equal to $n(t) * T$ for a slot period T. One can set different values of $n(t)$ for different flows. We found that estimating $p(t)$ using an exponentially weighted average is sufficient in our links to achieve the target loss estimate q. A purely retransmission based recovery mechanism has minimal throughput overhead as only the lost packets are retransmitted but this comes at a cost of high delay due to the long round trip times over WiLD links.

5.3.2 Adaptive FEC-Based Recovery

Designing a good FEC mechanism in highly variable lossy conditions requires accurate estimation of the underlying channel loss. When the loss is underestimated, the redundant packets cannot be decoded at all making them useless, but overestimating the loss rate leads to unnecessary transmission of FEC packets. In our environment the loss

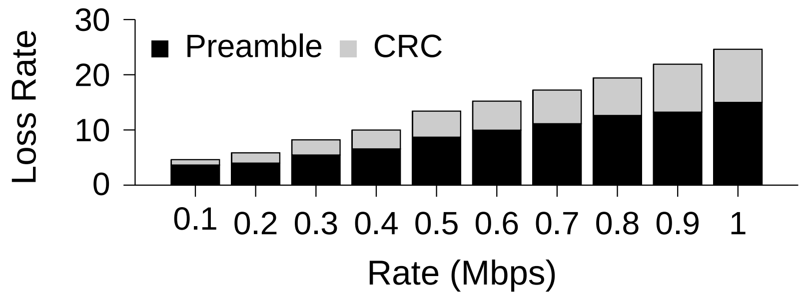

Figure 5.2: Proportion of CRC and preamble errors during channel loss. Traffic is at UDP CBR packets of 1440 bytes each at 802.11b PHY datarate of 11Mbps. Main link is sending at 2Mbps. The sending rate of the interferer increases from 0.1Mbps to 1Mbps.

distribution is non-stationary over large time scales, making it difficult to determine an accurate loss estimator.

Motivating inter-packet FEC: We can perform two types of FEC: inter-packet FEC (coding across packets) or intra-packet FEC (coding redundant blocks within a packet). In WiLD environments, we found that intra-packet FEC is not beneficial. Based on extensive measurements on a wireless channel emulator we observe that in presence of external WiFi interference, lost packets can be categorized into either CRC errors or preamble errors. A CRC error packet is still received by the driver with a checksum error. But an error in the preamble leads to the entire packet being dropped completely. Figure 5.2 shows the breakup of the loss rate with increasing external interference. We observe that although the proportion of preamble errors decreases as external interference increases, it still causes at least 50% of all errors. Moreover, a substantial number of the CRC error packets were truncated. Intra-packet FEC on the other hand can help recover only those packets that have CRC errors. Hence, we chose to perform only inter-packet FEC.

Estimating the level of redundancy: We apply FEC in combination with TDMA. For every time slot of N packets, we add $N - K$ redundant packets for every K packets. To estimate the redundancy factor, $r = (N - K)/K$, we choose a simple but not perfect estimation policy based on a weighted average of the losses observed in the previous M time slots. Here, we specifically chose a small value of $M = 10$. First, predicting the start of a burst is very hard; hence, we do not even attempt to predict it. Second, a small value of M, can quickly adapt to both the start and the end of a loss burst saving unnecessary redundant FEC packets. For a time slot of $T = 10ms$, $M = 10$ corresponds to 200 ms to adapt to a change in the loss behavior. Third, due to non-stationary loss distributions, the added reduction that we observed in the perceived loss rate obtained by applying more complicated distribution based estimation approaches [134] is marginal. FEC is best suited for handling residual losses and bursts that are longer than the time required for weighted loss average estimation mechanism to adapt.

5.4 Implementation

In this section, we describe the implementation details of WiLDNet. Our implementation comprises two parts: (a) driver-level modifications to control or disable features implemented in hardware (Section 5.4.1); (b) a *shim* layer that sits above the 802.11 MAC (Section 5.4.2) and uses the Click [88] modular router software to implement the functionalities described earlier.

5.4.1 Driver Modifications

The wireless cards we use in our implementation are the high power (200-400 mW) Atheros-based chipsets. To implement WiLDNet, we have to disable the following 802.11 MAC mechanisms:

•We disable **link-layer association** in Atheros chipsets using the *AdHoc-demo* mode.

•We disable **link layer retransmissions and automatic ACKs** by using 802.11 QoS frames with WMM extensions set the no-ACK policy.

•We disable **CSMA** by turning off the Clear Channel Assessment (CCA) in Atheros chipsets. With CCA turned off, the radio card can transmit packets right away without waiting for a clear channel.

5.4.2 Software Architecture Modifications

In order to implement single-link and inter-link synchronization using TDMA, the various loss recovery mechanisms, sliding window flow control, and packet reordering for in-order delivery, we use the Click modular router [88] framework. We use Click because it enables us to quickly prototype a modular MAC layer by composing different click elements together. It is also reasonably efficient for packet processing especially if loaded as a kernel module. Using kernel taps, Click creates fake network interfaces, such as *fake0* in Figure 5.3 and the kernel communicates with these virtual interfaces. Click allows us to intercept packets sent to this virtual interface and modify them before sending on the real wireless interface and vice versa.

Figure 5.3 presents the structure of the Click elements of our layered system, with different functionality (and corresponding packet header processing) at various layers, which

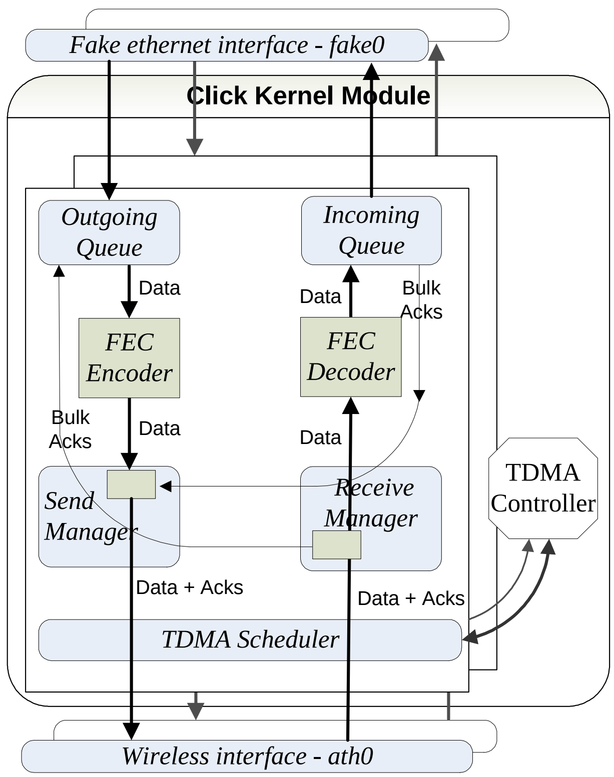

Figure 5.3: Click Module Data Flow

we describe next.

Incoming/Outgoing Queues: The mechanisms supporting sliding window packet flow, bulk acknowledgments, selective retransmission and reordering for in-order delivery are implemented by the incoming/outgoing queue pair. Packet buffering at the sender is necessary for retransmissions, and buffering at the receiver ensures reordering. In-order delivery and packet retransmission are optional, and the maximum number of retries can be set on a per-packet basis.

FEC Encoder/Decoder: An optional layer is responsible for inter-packet forward error correction encoding and decoding. For our implementation we modify a FEC library [123] that uses erasure codes based on Vandermonde matrices computed over $GF(2^m)$. This FEC method uses a (K, N) scheme, where the first K packets are sent in their original form, and $N - K$ redundant packets are generated, for a total of N packets sent. At the receiver, the reception of any K out of the N packets enables the recovery of the original packets. We choose this scheme because, in loss-less situations, it introduces very low latency: the original K packets can be immediately sent by the encoder (without undergoing encoding), and immediately delivered to the application by the decoder (without undergoing decoding).

Send and Receive Managers: These elements are responsible for managing the bulk ACKs and the encoded data packets. They combine the bulk ACKs generated by the receive queue with outgoing packets and also direct the incoming ACKs to the send queue.

TDMA Scheduler: This element ensures that packets are being sent only during the designated send slots, and manages packet timestamps as part of the synchronization mechanism.

TDMA Controller: This element is common for all the interfaces supported by the click module. It implements synchronization among the wireless radios on the same channel, by enforcing synchronous transmit and receive operation (all the radios on the same channel have a common send slot, followed by a common receive slot).

5.4.3 Timing issues

We do not use Click timers to implement time synchronization because the underlying kernel timers are not precise at the granularity of our time slots (10ms-40ms) on our hardware platform (266MHz CPU). Additionally, packet queuing in the wireless interface causes variability in the time between the moment Click emits a packet and the time the packet is actually sent on the air interface. Thus, the propagation delay between the sending and the receiving click modules on the two hosts is not constant, affecting time slot calculations. Fortunately, this propagation delay is predictable for the first packet in the send slot, when the hardware interface queues are empty. Thus, in our implementation, we only timestamp the first packet in a slot, and use it for adjusting the receive slot at the peer. If this packet is lost, the receiver's slot is not adjusted in the current slot, but since the drift is slow this does not have a significant impact.

Another timing complication is related to estimating whether we have time to send a new packet in the current send slot. Since the packets are queued in the interface, the time when the packet leaves Click cannot be used to estimate this. To overcome this aspect, we use the notion of *virtual time*. At the beginning of a send slot, the virtual time t_v is same as current (system) time t_c. Every time we send a packet, we estimate the transmission time of the packet on the channel and recompute the virtual time: $t_v = max(t_c, t_v) + duration(packet)$. And a packet is sent only after checking that the virtual time after sending this packet will not exceed the end of the send slot. Otherwise, we postpone the packet until the next slot. Although our synchronization scheme works reasonably well, we intend to move this part of the system into the Atheros firmware for increased accuracy.

5.5 Experimental Evaluation

The main goals of WiLDNet are to increase link utilization and to eliminate the various sources of packet loss observed in a typical multi-hop WiLD deployment, while simultaneously providing flexibility to meet different end-to-end application requirements.

We believe these are the first actual implementation results over an outdoor multi-hop WiLD network deployment.

Raman *et al.* [118] show the improvements gained by the 2P-MAC protocol in simulation and in an indoor environment. However, a multi-hop outdoor deployment also has to deal with high losses from external interference. 2P in its current form does not have any built in recovery mechanism. The authors of 2P also do not propose how any recovery mechanism can be combined with their marker-based synchronization protocol. Hence, we do not have any direct comparison results with 2P on our outdoor wireless links. Secondly, the proof-of-concept implementation of 2P was for the Prism 2.5 wireless chipset. WiLDNet relies on firmware features in the Atheros chipset and it is not clear how we can implement the modifications proposed for 2P with the Atheros driver.

Our evaluation has three main parts:

• We analyze the ability of WiLDNet to maintain high performance (high link utilization) over long-distance WiLD links. At long distances, we demonstrate 2–5x improvements in throughput with TCP flows in both directions.

• Next, we evaluate the ability of WiLDNet to scale to multiple hops and eliminate inter-link interference. WiLDNet yields a 2.5x improvement in TCP throughput on our real-world multi-hop setup.

• Finally, we evaluate the effectiveness of the two link recovery mechanisms of WiLDNet: Bulk Acks and FEC.

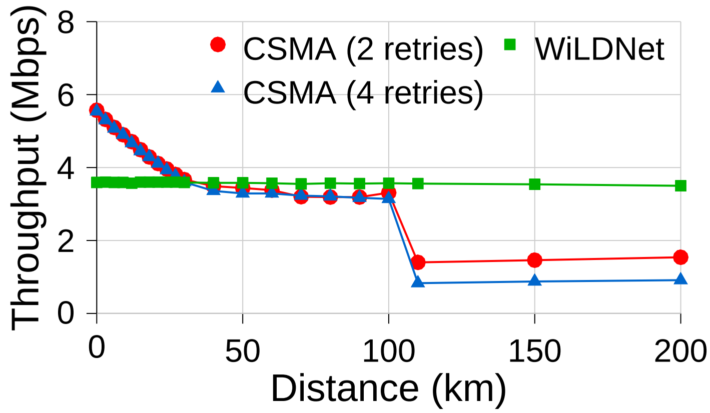

Figure 5.4: TCP flow in one direction

5.5.1 Single Link Without Channel Losses

In this section we demonstrate the ability of WiLDNet to eliminate link under-utilization and packet collisions over a single WiLD link. We compare the performance of WiLDNet (slot size of 20ms) with the standard 802.11 CSMA (2 retries) base case.

The first set of results show the improvement of WiLDNet on the emulator link with the distance going up to 200km. Figure 5.4 compares the performance of TCP flowing only in one direction. The lower throughput of WiLDNet, approximately 50% of channel capacity, is due to symmetric slot allocation between the two end points of the link. However, over longer links (>50 km), the TDMA-based channel allocation avoids the under-utilization of the link as experienced by CSMA. Also, beyond 110 km (the maximum possible ACK timeout), the throughput with CSMA drops rapidly because of unnecessary retransmits (as pointed out earlier in Section 4.2.1).

Figure 5.5 shows the cumulative throughput of TCP flowing simultaneously in

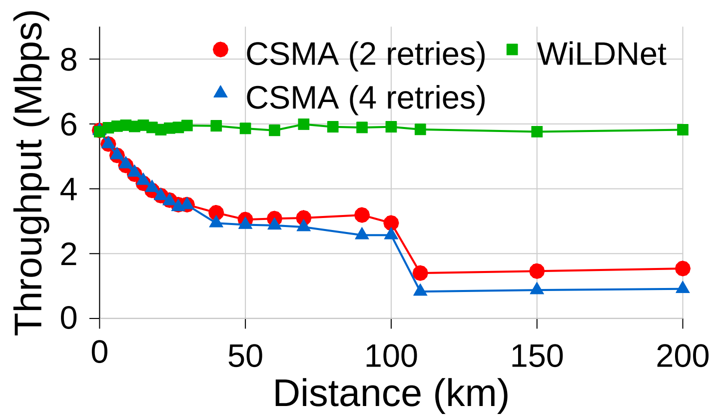

Figure 5.5: TCP flow in both directions

both directions. In this case, WiLDNet effectively eliminates all collisions occurring in presence of bidirectional traffic. TCP throughput of 6 Mbps is maintained for all distances.

Table 5.1 compares WiLDNet and CSMA for all our outdoor wireless links. We show TCP throughput for flows in one direction and the cumulative throughput for TCP flowing in both directions. Since these are outdoor measurements, there is significant variation over time and we show both the mean and standard deviation for the measurements. We can see that as the link distance increases, the improvement of WiLDNet is more substantial.

5.5.2 Multiple Hops

This section validates that WiLDNet eliminates inter-link interference by synchronizing receive and transmit slots in TDMA resulting in up to 2x TCP throughput improvements over standard 802.11 CSMA in multi-hop settings.

The first set of measurements were performed on our indoor setup (Section 4.1)

Link	Distance	Loss rates	802.11 CSMA		WiLDNet	
	(km)	(%)	One dir	Both dir	One dir	Both dir
B-R	8	3.4	5.03 (0.02)	**4.95** (0.03)	3.65 (0.01)	**5.86** (0.05)
P-I	15	40	3.0	4.0	4.0	5.0
P-S	45	2.6	3.62 (0.20)	**3.52** (0.17)	3.10 (0.05)	**4.91** (0.05)
Ghana	65	1.0	2.80 (0.20)	**0.68** (0.39)	2.98 (0.19)	**5.51** (0.07)

Table 5.1: Mean TCP throughput (flow in one direction and cumulative for both directions at same time) for WiLDNet and CSMA for various outdoor links (distance and loss rates). The standard deviation is shown in parenthesis for 10 measurements. Each measurement is for TCP flow of 30s at 802.11b PHY layer datarate of 11Mbps.

where we recreated the conditions of a linear outdoor multi-hop topology using the RF isolation boxes. Thus transmissions from local radios interfere with each other but multiple local radio interfaces can receive simultaneously. We then measure TCP throughput of flows in the one direction and then both directions for both standard 802.11 CSMA and WiLDNet (with slot size of 20ms). All the links were operating on the same channel. As we see in Table 5.2, as the number of hops increases, standard 802.11's TCP throughput drops substantially when transmissions from a radio collide with packet reception on a nearby local radio on the same node. WiLDNet avoids these collision and maintains a much higher cumulative TCP throughput (up to 2x for the 3-hop setup) by proper synchronization of send and receive slots.

Linear setup	802.11 CSMA			WiLDNet		
	Dir 1	Dir 2	Both	Dir 1	Dir 2	Both
2 nodes	5.74 (0.01)	5.74 (0.01)	**6.00** (0.01)	3.56 (0.03)	3.53 (0.02)	**5.85** (0.07)
3 nodes	2.60 (0.01)	2.48 (0.01)	**2.62** (0.01)	3.12 (0.01)	3.12 (0.01)	**5.12** (0.03)
4 nodes	2.23 (0.01)	2.10 (0.01)	**1.99** (0.02)	2.95 (0.05)	2.98 (0.04)	**4.64** (0.24)

Table 5.2: Mean TCP throughput (flow in each direction and cumulative for both directions simultaneously) for WiLDNet and standard 802.11 CSMA. Measurements are for linear 2,3 and 4 node indoor setups recreating outdoor links running on the same channel. The standard deviation is shown in parenthesis for 10 measurements of flow of 60s each at 802.11b PHY layer datarate of 11Mbps.

We can also see that although WiLDNet has more than 2x improvement over standard 802.11, the final throughput (4.6Mbps) is still much smaller than the raw throughput of the link (6-7Mbps). This can be attributed to the overhead of synchronization, as well as to inefficient packet processing in Click, when run on our low-frequency single board relays. A more efficient synchronization mechanism implemented in the firmware (rather than Click) would deliver even more of an improvement.

We also measure this improvement on our outdoor testbed in Table 5.3 between the nodes K and M relayed through node P. We again compare the TCP throughput for WiLDNet and standard 802.11 CSMA with links operating on the same channel. In order to quantify the effect of inter-link interference, we also perform the same experiments with the links operating on different, non-overlapping channels, in which case the inter-link interference is almost zero, as previously shown.

Description	One direction	Both directions
Standard TCP: same channel	**2.17**	**2.11**
Standard TCP: diff channels	3.95	4.50
WiLD TCP: same channel	**3.12**	**4.86**
WiLD TCP: diff channels	3.14	4.90

Table 5.3: Mean TCP throughput (flow in single direction and cumulative for both directions simultaneously) comparison for WiLDNet and standard 802.11 CSMA over a 2-hop outdoor setup ($K \leftrightarrow P \leftrightarrow M$). Averaged over 10 measurements of TCP flow for 60s at 802.11b PHY layer datarate of 11Mbps.

We can see from Table 5.3 that, for same channel operation, the cumulative TCP throughput in both directions with WiLDNet (4.86 Mbps) is more than twice the throughput observed over standard 802.11 (2.11 Mbps). The improvement is substantially lower for the unidirectional case (3.14 Mbps versus 2.17 Mbps), because the WiLD links are constrained to send in one direction only roughly half of the time.

Another key observation is that WiLDNet is capable of eliminating almost all inter-link interference. This is shown by the fact that the throughput achieved by WiLDNet is almost the same, whether the links operate on the same channel or on non-overlapping channels. This result is very important, as it makes channel allocation a non-issue.

5.5.3 WiLDNet Link-Recovery Mechanisms

Our next set of experiments evaluate WiLDNet's adaptive link recovery mechanisms in conditions closer to the real world, where errors are generated by a combination of collisions and external interference. We evaluate both the bulk ACK and FEC recovery

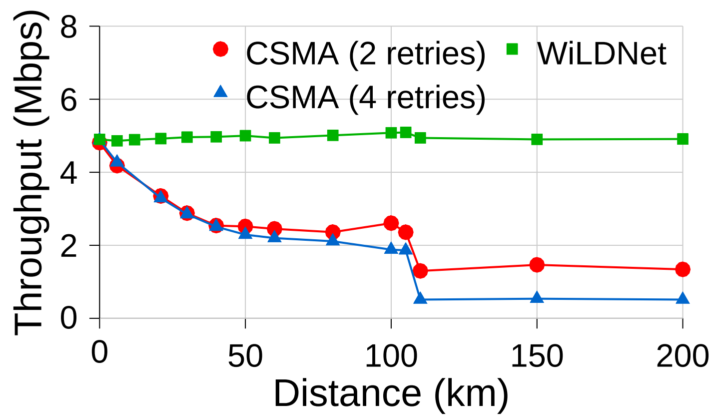

Figure 5.6: TCP in both directions, 10% channel loss. TCP throughput for WiLDNet vs 802.11 CSMA. Each measurement is for a TCP flow of 60s, 802.11b PHY, 11Mbps.

mechanisms.

Bulk ACK Recovery Mechanism

For our first experiment, presented in Figure 5.6, we uniformly vary the link length on the emulator, and we introduce a 10% error rate through external interference. We again measure the cumulative throughput TCP flows in both directions for both WiLDNet and standard 802.11 CSMA. Again, WiLDNet maintains a constant throughput with increasing distance as opposed to the 802.11 CSMA. Due to the 10% error, WiLD incurs a constant throughput penalty of approximately 1 Mbps compared to the no-loss case in Figure 5.5.

In our second experiment we fix the distance to 80 km, and vary external interference rates. The measurement results, presented in Figure 5.7, show that WiLDNet maintains roughly a 2x improvement over standard CSMA's recovery mechanism (with four retries), for packet loss rates up to 30%.

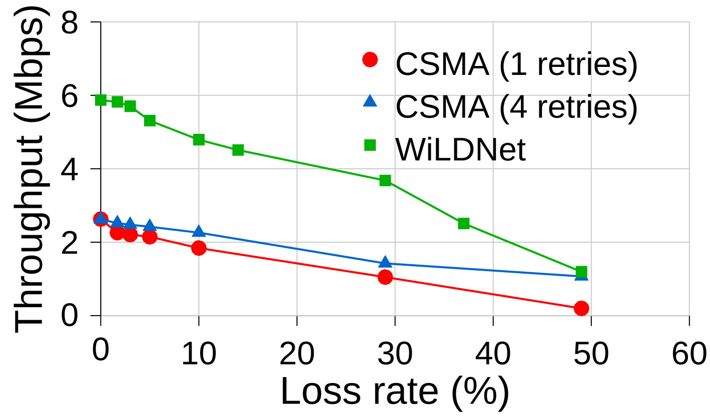

Figure 5.7: Cumulative TCP throughput on both directions for increasing loss. Each measurement was for 60s TCP flows of 802.11b at 11Mbps PHY datarate.

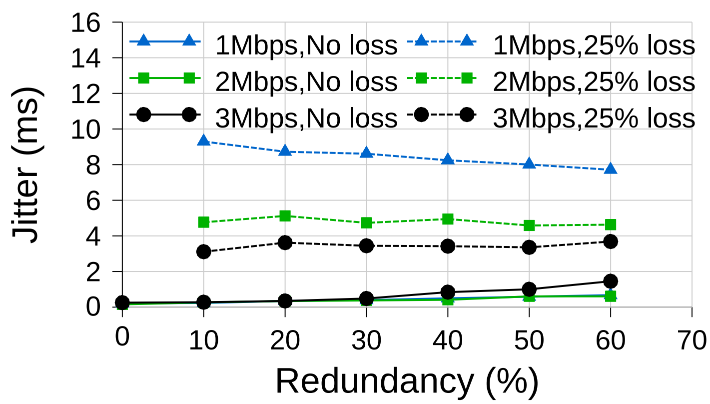

Figure 5.8: Overhead of FEC encoding and decoding. 1440 byte UDP CBR packets at PHY datarate of 11Mbps in 802.11b

Forward Error Correction (FEC)

To measure the jitter introduced by the FEC mechanism, we performed a simple experiment where we measured the jitter of a flow under two conditions: in the absence of any loss and in the presence of a 25% loss. Figure 5.8 illustrates the jitter introduced by

WiLDNet's FEC implementation. We can see that in the absence of any loss , when only encoding occurs, the jitter is minimal. However, in the presence of loss, when decoding also takes place, the measured jitter increases. However, the magnitude of the jitter is very small and well within the acceptable limits of many interactive applications (voice or video), and decreases with higher throughputs (since the decoder waits less for redundant packets to arrive).

Moreover, considering the combination of FEC with TDMA, the delay overheads introduced by these methods overlap, since the slots when the host is not actively sending can be used to perform encoding without incurring any additional delay penalties.

5.6 Tradeoffs

One of the main design principles of WiLDNet is to build a system that can be configured to adapt to different application requirements. In this section we explore the tradeoff space of throughput, delay and delivered error rates by varying the slot size, number of bulk retransmissions and FEC redundancy parameters. We observe that WiLDNet can perform in a wide spectrum of the parameter space, and can easily be configured to meet specific application requirements.

5.6.1 Choosing number of retransmissions

The first tradeoff that we explore is choosing the number of retries to get a desired level of final error rate on a WiLD link. Although retransmission based loss recovery achieves optimal throughput utilization, it comes at a cost of increased delay; the loss rate can be reduced to zero by arbitrarily increasing the number of retransmissions at the cost

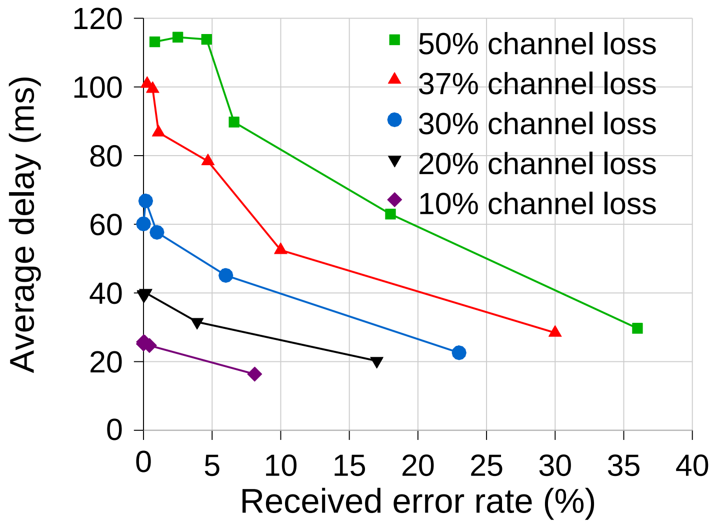

Figure 5.9: Avg delay with decreasing target loss rate (X-axis) for various loss rates.

of increased delay. This tradeoff is illustrated in Figure 5.9 which shows a plot of delay versus error rate for varying channel loss rates (10% to 50%). Retries are increased from 0 to 10 in increments of 1 moving from right to left for a given line in the figure. All the tests are unidirectional UDP tests at 1 Mbps for a fixed slot size of 20ms. For example, at a loss rate of 50%, 3 retries are required to reduce the error rate to 5%, which leads to a delay of approximately 110 ms. We can see that as we try to reduce the final error rate at the receiver, we have to use more retries and this increases the average delay. In addition, we also observe that larger the number of retries, larger the end-to-end jitter (especially at higher loss rates).

This tradeoff has important implications for applications that are more sensitive to delay and jitter (such as real time audio and video) as compared to applications which require high reliability. For such applications, we can achieve a balance between the final

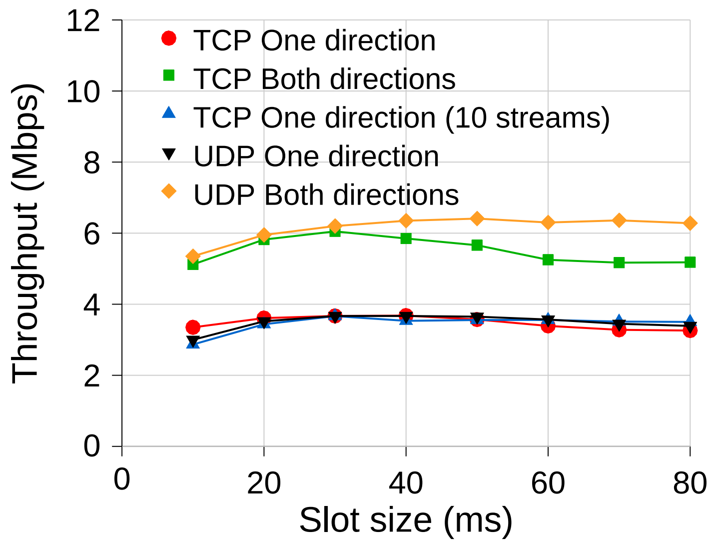

Figure 5.10: Throughput for increasing slot sizes (X-axis).

error rate and the average delay by choosing an appropriate retry limit. For applications that require improved loss characteristics without incurring a delay penalty, we need to use FEC for loss recovery.

5.6.2 Choosing slot size

The second tradeoff that we explore is the effect of slot size. The two factors that affect slot size are the end-to-end delay requirements of the application and the overhead in switching from transmit to receive mode.

We first analyze the effect of slot size on received throughput. Our experiments are performed on a 60-km emulated link with 10% packet loss rate. Figure 5.10 presents the UDP and TCP throughputs for various slot sizes. Ideally, we would not expect the received throughput to change with slot size. However, as discussed in Section 5.2, switching

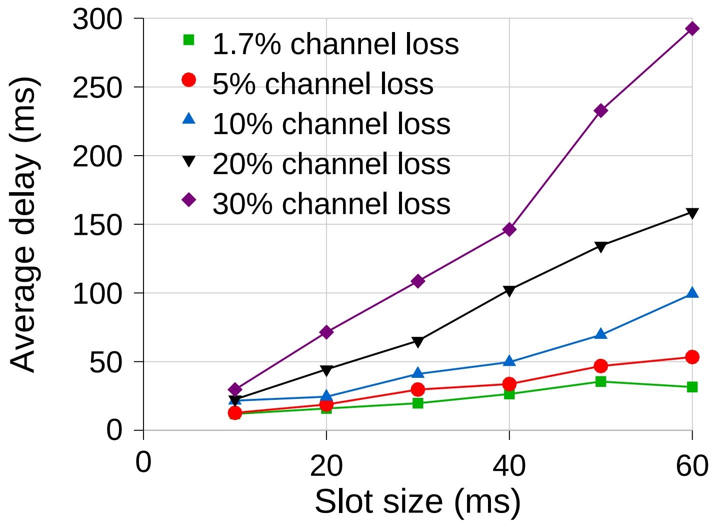

Figure 5.11: Avg delay at increasing slot sizes (X-axis) for various loss rates.

between send and receive slots incurs a fixed overhead. This overhead is non-negligible for the Click based WiLDNet implementation. This overhead although constant for all slot sizes, occupies a higher fraction of the lot for for smaller slots sizes. As a consequence, as seen in figure 5.10, at small slot sizes the achieved throughput is lower. However, the UDP throughput levels off beyond a slot size of 20 ms. We also observe the TCP throughput reducing slightly at higher slot size. This is because the throughput-delay product of the link increases with slot size, but the send TCP window sizes are fixed. UDP throughput is not affected at higher slot sizes.

In the next experiment, we measure the average UDP packet transmission delay while varying the slot size, for several channel error rates. The results are presented in Figure 5.11; each series represents a unidirectional UDP test (1 Mbps CBR) at a particular channel loss rate with WiLDNet using maximum number of retries. Figure 5.11 shows the

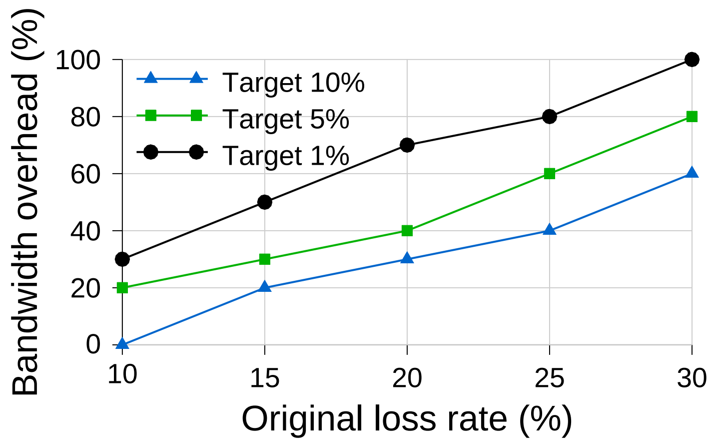

Figure 5.12: FEC throughput overhead versus channel loss rate.

increase in delay with increasing slot size. It is clear that slot sizes beyond 20 ms do not result in substantially higher throughputs, but they do result in much larger delay. Thus, if delay constraints are not too stringent, a good choice for a slot size is 20ms. However, if lower delay is required, smaller slots can be used at the expense of some throughput overhead consumed by the switching between the transmit and receive modes.

5.6.3 Choosing redundancy parameters for FEC

The primary FEC parameter that we can tune is the redundancy factor $r = (N-K)/K$, also referred to as throughput overhead. While FEC incurs a higher throughput overhead than retransmissions, it incurs a smaller delay penalty as illustrated earlier in Section 5.5.3. To analyze the tradeoff between FEC throughput overhead and the target loss-rate, we consider the case of a single WiLD link (in our emulator environment) with a simple Bernoulli loss-model (every packet is dropped with probability p). Here, we set a specific value of r and measure the observed target loss-rate for different values of r.

Figure 5.12 shows the amount of redundancy required to meet three different target loss-rates of 10%, 5% or 1% for different error rates (namely p). The primary observation we make is that in order to achieve very low target loss-rates, FEC needs to expend a lot of throughput overhead (for example, FEC incurs a 100% throughput overhead to reduce the loss-rate from 30% to 1%).

In general, the redundancy factor required to achieve a certain target loss rate, q, is dependent on three factors: (a) the target loss-rate q; (b) the loss characteristics of the channel; (c) the predictability of losses in the channel. In general, when a channel is very bursty and has an unpredictable burst arrival pattern, it is very hard for FEC to achieve arbitrarily low target loss-rates.

In general, for applications that can tolerate one round of retransmissions, one can perceive different combinations of FEC and retransmissions that can provide a tradeoff between overall throughput overhead, delay and target loss-rate. In the case of a channel with a stationary loss distribution, OverQoS [134] shows that the optimal policy to minimize overhead is to not use FEC in the first round but use it in the second round to pad retransmission packets. With unpredictable and highly varying channel loss conditions, an alternative promising strategy is to use FEC in the first round during bursty periods to reduce the perceived loss-rate.

5.7 Related Work

Long Distance WiFi: The use of 802.11 for long distance networking with directional links and multiple radios per node, raises a new set of technical issues that were first illustrated in [36]. Raman *et al.* built upon this work in [118, 117] and proposed the 2P MAC protocol. WiLDNet builds upon 2P to make it robust in high loss environments. Specifically we modify 2P's implicit synchronization mechanism as well as build in adaptive bulk ACK based and FEC based link recovery mechanisms.

Other wireless loss recovery mechanisms: There is a large body of research literature in wireless and wireline networks that have studied the tradeoffs between different forms of loss recovery mechanisms. Many of the classic error control mechanisms are summarized in the book by Lin and Costello [91]. OverQoS [134] performs recovery by analyzing the FEC/ARQ tradeoff in variable channel conditions and the Vandermonde codes are used for reliable multicast in wireless environments [123].

Of particular interest for this work are the Berkeley Snoop protocol [34] which provides transport-aware link-layer recovery mechanisms in wireless environments. To compare the WiLDNet bulk ACK recovery mechanism with recovery at a higher layer, we experimented with a version of the original Snoop protocol [35] that we modified to run on WiLD links. Basically, each WiLD router ran one half of Snoop, the fixed host to mobile host part, for each each outgoing link and integrated all the Snoops on different links into one module.

We measured the performance of modified Snoop as a recovery mechanism over both standard 802.11 (CSMA) and over WiLDNet with no retries. We found that WiLDNet was still 2x better than Snoop. We also saw that Snoop was better than vanilla CSMA only at lower error rates (less than 10%). Thus, this indicates that higher layer recovery mechanisms might be better than stock 802.11 protocol, but only at lower error rates.

Other WiFi-based MAC protocols: Several recent efforts have focused on leveraging off-the-shelf 802.11 hardware to design new MAC protocols. Overlay MAC Layer (OML) [120] provides a deployable approach towards implementing a TDMA style MAC on top of the 802.11 MAC using loosely-synchronized clocks to provide applications and competing nodes better control over the allocation of time-slots. SoftMAC [102] is another platform to build experimental MAC protocols. MultiMAC [47] builds on SoftMac to provide a platform where multiple MAC layers co-exist in the network stack and any one can be chosen on a per-packet basis.

WiMax: An alternative to WiLD networks is WiMax [25]. WiMax does present many strengths over a WiFi: configurable channel spectrum width, better modulation (especially for non-line of sight scenarios), operation in licensed spectrum with higher transmit power, and thus longer distances. On the other hand, WiMax currently is primarily intended for carriers (like cellular) and does not support point-to-point operation. In addition, WiMax base-stations are expensive ($10,000) and the high spectrum license costs in most countries dissuades grassroots style deployments. Currently it is also very difficult to obtain licenses for experimental deployment and we are not aware of open-source drivers for WiMax base-stations and clients. However, most of our work in loss recovery and adaptive FEC would be equally valid for any PHY layer (WiFi or WiMax). With appropriate modifications and cost reductions, WiMax can serve as a more suitable PHY layer for WiLD networks.

Chapter 6

Understanding Faults in Rural WiLD Networks

WiLDNet has greatly benefited from real deployments undertaken in parallel with research. In particular, we have gained valuable insights into the operational challenges of running WiLD networks in rural areas.

Over the last 3 years we have deployed WiLDNet at the Aravind Eye Hospital in southern India. Patients in villages are linked to centrally located Aravind doctors in cities via point-to-point WiLD links. The patients video-conference with doctors over these links for consultations. We have also deployed a few WiLD links for both experimental and production purposes at the AirJaldi community network in the town of Dharamsala in northern India to provide Internet and voice-over-IP services to the Tibetan community in exile.

The work presented in this chapter is joint work with TIER group members Rabin Patra, Sergiu Nedevschi, Manuel Ramos, Lakshminarayanan Subramanian, Yahel Ben-David and Eric Brewer. Some of the material presented in this chapter has been previously published as "Beyond Pilots: Keeping Rural Wireless Networks Alive"[136].

Our overall objective in this chapter is to characterize the faults, many of which were surprising and unexpected, that we experienced in hitherto unfamiliar rural settings. We start by first discussing why faults seem more likely in rural settings (Section 6.1). We then describe our rural deployments at the Aravind Eye Hospital and the AirJaldi community network (Sections 6.2, 6.3 and 6.4). Next, We present our operational experiences with these two networks, particularly the types of faults seen and the difficulties posed to system management from decreased component robustness and difficulties in diagnosing and predicting faults (Sections 6.5, 6.6 and 6.7). Finally, we list the key components for system management in rural WiLD networks (Section 6.8).

6.1 Faults in Rural Settings

Although there has been much active research on wireless networking technologies for developing regions, one important area tends to be overlooked: the ability of local staff to keep the system running over the long term. Software or hardware errors, power issues, or other transient environmental issues such as wireless interference can cause outages leading to poor experiences for users (and administrators). As these new systems strive for acceptance, outages can chase away potential users and jeopardize viability.

In two years of experience with rural wireless deployments, we have seen that the inability of local rural staff to fix errors causes extensive outages and prevents the healthy growth and expansion of the network.

There are several specific reasons why maintenance in rural areas is hard. First, local staff tend to start with limited knowledge about wireless networking. This leads to limited diagnostic capabilities, inadvertent equipment misuse, and misconfiguration. Thus management tools need to help with diagnosis and must be educational in nature. Training helps as well, but high IT turnover limits the effectiveness, so education must be ongoing and part of the process.

Second, the chances of hardware failures are higher as a result of poor-quality power. Although we have not conclusively inferred the failure rate of equipment for power reasons in rural areas, we have lost far more routers and adapters for power reasons in rural India than we have lost in our Bay Area testbed. This calls for a solution that provides stable and quality power to equipment in the field.

Third, many locations with wireless nodes, especially relays, are quite remote. It is important to avoid unnecessary visits to remote locations. Also, we should enable preventive maintenance during the visits that do occur. For example, gradual signal strength degradation could imply cable replacement or antenna realignment during a normal visit.

Fourth, the wireless deployment, although connecting local nodes, may not be accessible remotely or through the Internet. The failure of a single link might make parts of a network unreachable although the nodes themselves might be functional. This makes it very hard for remote experts in another town or even local administrators to resolve or even diagnose the problem. This points to a need for a low-cost alternate or "back channel" (e.g. SMS) that allows remote access to the nodes even in the event of a failure of the primary link.

Overall, troubleshooting is hard even for experienced users or experts. The troubleshooting decision tree is not always obvious, which makes it harder to design guidelines for rural users. Users or administrators have to hunt for data, e.g. run ifconfig, ping reachability scripts, log into remote nodes to isolate the fault. Although some this can be automated and even visualized, we may still not know the root cause of the fault. Some amount of non-trivial domain knowledge will still be needed to perform the actual diagnosis. And many times, even experts resort to hopeful reboots without a clear understanding of what is wrong.

We argue that these troubles exist in a large part because the research community has not tried very hard: these systems are not designed or deployed with support for easy diagnosis built in right from the start.

6.2 The Aravind Network

As shown in Figure 6.1, the Aravind network in Theni consists of five rural vision centers connected to the main hospital in Theni. The network has total of 11 wireless routers or nodes, of which 6 are end points and 5 are relay nodes. There are 9 point-to-point links. The link distances range from just 1 kilometer (Theni-Vijerani link) to 15 kilometers (Vijerani-Andipatti link). Six of the wireless nodes are installed on towers, heights of which range from 24 to 42 meters; the others are mounted on short poles on rooftops or already existing tall structures, such as the chimney of a power plant on the premises of a textile factory. Recently, Aravind has expanded this model to their hospitals in Madurai and Tirunelveli where they have added a total 3 and 2 vision centers respectively, with links as long as 40 kilometers. The network is currently financially viable and a further expansion to 50 clinics around 5 hospitals is being planned to provide 500,000 annual eye examinations.

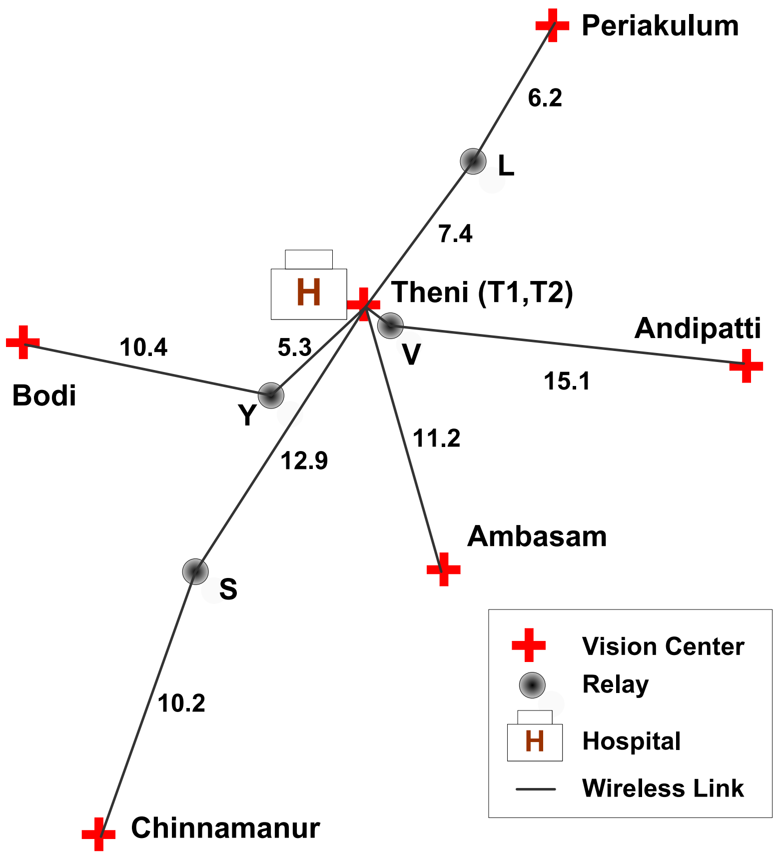

Figure 6.1: Aravind Telemedicine Network. Theni hospital is connected to 5 vision centers. The other nodes are all relays.

6.2.1 Hardware

The wireless nodes are a mix of 266 MHz x86 single board computers such as the WRAP and ALIX boards from PC Engines [17], and the RB230 boards from MikroTik [12]. These routers have up to three 200–400 mW Atheros 802.11 a/b/g radio cards. The longer links use 24 dBi directional antennas. The routers consume about 4.5 watts of power when idle and only 9.5 watts when transmitting at full bandwidth from 2 radios simultaneously; 7 watts is the average power consumption for a node. They run a stripped-down Linux distribution stored on a 512 MB Compact Flash card that includes the 2.4.26 kernel and our software for WiLDNet, monitoring, logging, and remote management.

The routers are placed in small and lightweight waterproof enclosures, and are mounted externally, close to the antennas, to minimize signal losses. They are powered via power-over-ethernet (PoE); therefore, only a single ethernet cable from the ground to the router is sufficient for both power and data. We use uninterruptible power supplies (UPS) or solar panels with battery backup to provide clean power.

6.2.2 Applications

The primary application is video conferencing. We currently use software from Marratech [94]. Although most sessions are between doctors and patients, video conferencing is also used for remote training of staff at the rural vision centers. Typical throughput on the links ranges between 5–7 Mbps with less than 2% channel losses. However, 256 Kbps in each direction is sufficient for very good quality video conferencing. Our network is thus over-provisioned. We also use the network to transmit 4-5 MB-sized retinal images taken by fundus cameras. The hospital has a VSAT link to the Internet, but most applications, except for remote management, require only intranet access within the network.

6.3 The AirJaldi Network

The AirJaldi network provides Internet access and voice-over-IP telephony services to about 10,000 users, primarily members of the Tibetan community in exile, within a radius of 70 km in the rural mountainous terrain of Dharamsala, India. The network has 8 point-to-point WiLDNet links ranging from 10 to 41 kilometers with a total of 10 end points, as shown in Figure 6.2. In addition, the network also has over a hundred low-cost modified

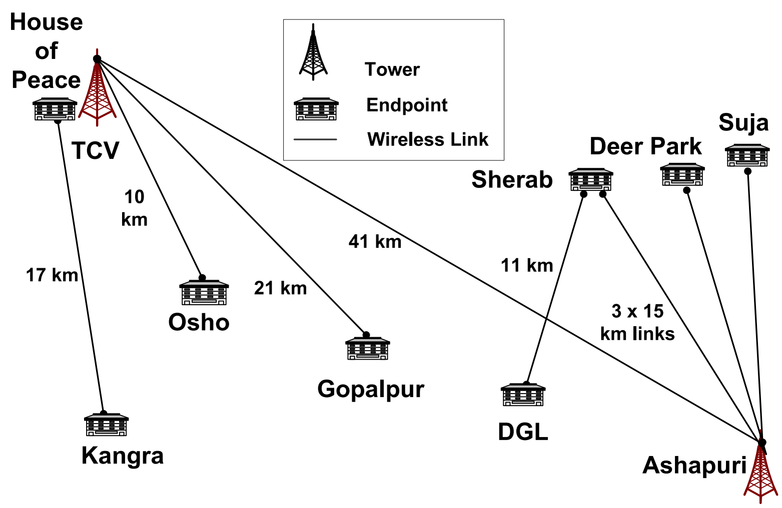

Figure 6.2: The AirJaldi Network. There are 8 long distance links with directional antennas with 10 endpoints.

consumer access points that use a wide variety of outdoor antennas. Three of the nodes are solar-powered relay stations mounted on towers located at remote hills and mountains. All other antennas are installed on low-cost masts less than 5 meters in height; the masts are typically water pipes on the rooftops of subscribers.

6.3.1 Hardware

Most of the routers are modified consumer devices, either Linksys WRT54GL units [11] or units from Buffalo Technology [5], and cost less than $50. These routers are built around a 200MHz MIPS processor with 16 MB of RAM, 4 MB of on-board flash memory, and a low-power Broadcom 802.11b/g radio. We run OpenWRT [16] on these routers, and use open source software for mesh routing, encryption, authentication, QoS, remote management and logging. They are housed inside locally designed and built weatherproof enclosures, and are mounted externally to minimize signal losses. The antennas, power supplies and batteries are all manufactured locally in India. For WiLDNet links and remote relay stations, we use routers similar to the ones used in the Aravind deployment: PCEngines WRAP boards [17], MikroTik router boards [12], and Ubiquiti LiteStation2 units [23], all with Atheros-based radios.

6.3.2 Applications

The Internet bandwidth at AirJaldi consists of 5 ADSL lines ranging from 144 Kbps to 2 Mbps for a total of about 7 Mbps downlink and 1 Mbps uplink bandwidth. The longest link from TCV to Ashapuri (41 km) achieves a throughput of about 4–5 Mbps at 2–5% packet loss, while the link from TCV to Gopalpur (21 km) only gets about 500–700 Kbps at 10–15% packet loss due to the absence of clear line of sight.

This bandwidth is sufficient for applications such as Internet access and VoIP that cater primarily to the needs of the community, namely schools, hospitals, Tibetan monasteries and other non-profit organizations. AirJaldi only provides connectivity to fixed installations and does not offer wireless access to roaming users or mobile devices. A cost-sharing model is used among all network subscribers to recover the operational costs. The network is currently financially sustainable and is also growing rapidly.

6.4 Fault Characterization

The most common description of a fault by our rural partners is that the "link is down". At Aravind, this typically means that the video conference is not working and at AirJaldi, it typically means that the Internet is no longer working. Table 6.1 lists the range of faults, excluding application level errors, that we saw in 2005 and that resulted in

Type	Occurences	Downtime (days)	Fault description
HW	63	63	Router board not powered on (grid outage, battery dead)
	7	10	Router powered but wedged (low voltage, corrupt CF cards)
	21	34	Router powered but not connected to remote LAN (loose ethernet cables, burnt ethernet ports)
	3	2	Router on, but wireless cards not transmitting due to low supplied voltage
	1	45	Router on, but pigtails not connected or other RF connectors gone bad
	1	55	Router on, but antenna misaligned
SW	4	4	No default gateway specified
	3	3	Wrong ESSID, channel, mode
	2	2	Wrong IP address
	2	3	Misconfigured routing

Table 6.1: Various hardware and software errors that resulted in link downtimes at Aravind in 2005. The number of occurrences are an under-estimate since they are based on what we saw and not fully based on what the local staff experienced as well. This is because they did not keep accurate logs of the faults, partly because of the difficulty in diagnosis and tried various solutions till the fault was "fixed".

a "link being down" at Aravind. The table shows whether each fault was a hardware or software issue, the number of times it occurred, the resulting downtime in terms of days, and the root cause of the fault. As seen in the table, there are a variety of reasons for link outages and it is not always easy to diagnose the root cause of the fault. It is important to note that the root causes of the faults listed in the table were only found by extensive after-the-fact investigation, including much traveling to the remote locations of the routers.

We see that over 90% of the faults were related to power (the first four hardware faults), and they caused 50% of the total downtime experienced in 2006, which was roughly a third of a year. Many of these faults were diagnosed by us or rural staff only after traveling to the locations. Had we known the cause beforehand, it would have changed our response. For example, if the fault was due to a general grid power outage, we would not have traveled to the remote location and would instead have waited for power to come back on. If a router was wedged (hung), we would have to go there to reset the router no matter what. If the PoE ports were burned due to high voltage, we would have gone there equipped to replace the boards. This is important as equipment is scarce and spare inventories are limited, making careful planning necessary.

Our operational experiences both at Aravind and at AirJaldi indicated that components were more likely to fail due to unexpected power behavior, diagnosing and anticipating faults was more difficult due to difficulties in monitoring and a lack of alternate backchannels. Next, we describe our experiences in this regard and provide several real examples from Aravind and AirJaldi.

6.5 Decreased Component Robustness

Operating conditions at Aravind and AirJaldi have greatly contributed to a substantial decrease in the robustness of system components that would otherwise work quite reliably. The lack of stable quality power has been a major culprit in this regard. Although issues such as frequent power outages in rural areas are well known, we were surprised by the *degree of poor-quality power* in rural villages, even when power from the grid was avail-

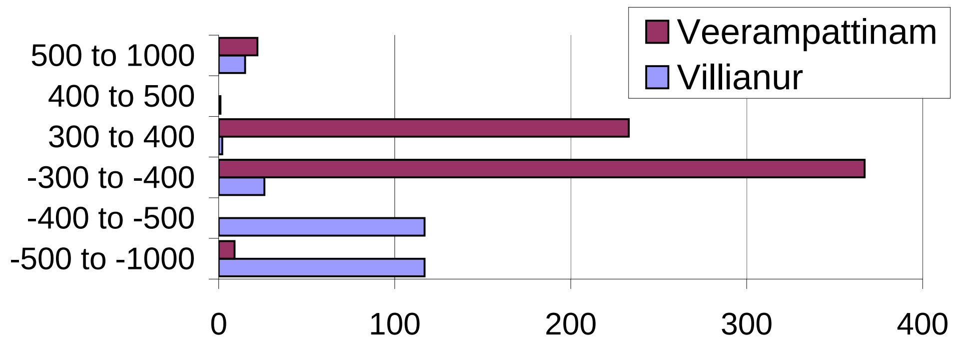

Figure 6.3: Histogram of voltage spikes from two rural villages, Veerampattinam and Villianur. The bins (y axis) are the size range of the spike in volts, while the x axis is the count. Negative bins imply that the power meter was connected with its polarities reversed.

able. Before addressing the power issues (section 7.3), not a single day went by without failures related to low power quality in either network. Any effort that is focused on rural deployments must necessarily fix the power issues. Therefore we describe the quality of rural power in detail, particularly because it has not been previously documented.

Low-quality Power: Figure 6.3 shows data on voltage spikes measured by a power logger placed in two different rural villages in southern India for 6 weeks. We group the spikes based on their magnitude in volts; negative voltages simply indicate that the power logger was connected to the grid with its polarities reversed. We see many spikes above 500 volts, and some even reaching 1000 volts! Clearly such spikes can damage equipment (burn power supplies), and have affected us greatly. We have also seen extended sags below 70 volts and swells above 350 volts (normal voltage in India is 220-240 volts). Although the off-the-shelf power supplies we use function well at a wide range of input voltages (80 to 240 volts), they are not immune to such widely ranging fluctuations. Also, locations far away from transformers are subject to more frequent and extreme power fluctuations. Our

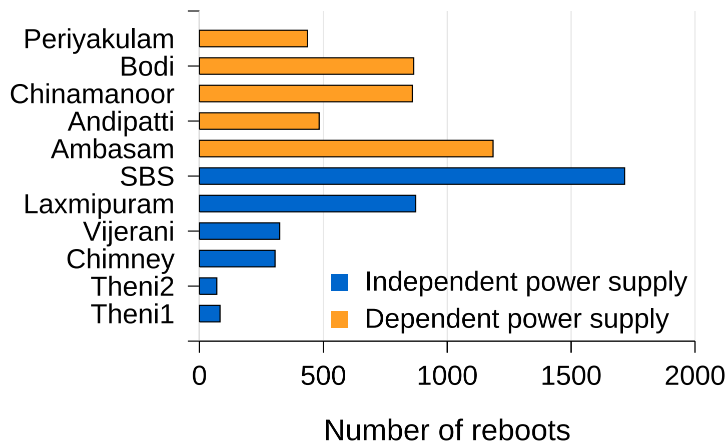

Figure 6.4: Number of reboots estimated per node in the Aravind network for about one year of operation. Nodes with power supplies dependent on the vision center are turned on or off everyday. Nodes with independent power supplies are typically relay nodes or hospital nodes.

first approach was to use Uninterruptible Power Supplies (UPSes) and battery backups. However, affordable UPS systems are only of the "standby" type, in which grid power is passed through to the equipment whenever it is available, regardless of quality. Therefore if the grid power is of poor quality, it is still sent to the equipment without any stabilization. This causes the spikes and surges to pass through to the equipment. Ironically, in such situations we only get "good power during an outage", when the UPS supplies stable battery power to the equipment.

6.5.1 Failures from Bad Power Quality

We have experienced a wide range of failures from bad quality power. First, spikes and surges have damaged our power supplies and router boards. In the AirJaldi network, we have lost more than 50 power supplies, about 30 ethernet ports and 5 boards to power surges, while in the Aravind network, we have lost 4 boards, more than 5 power supplies and have seen several burned ethernet ports as well.

Second, voltage sags have caused brown outs. Low voltages leave routers in a wedged state, unable to boot completely. The on-board hardware watchdog, whose job is to reboot the router, is also often rendered useless because of the low voltages, thus leaving the router in a hung state indefinitely. Third, fluctuating voltages cause frequent reboots, which corrupt and occasionally damage the CF cards through writes during the reboots.

As a typical example, the router at SBS in Aravind rebooted at least 1700 times in a period of 12 months (Figure 6.4), close to 5 times per day, going up to 10 times for some days. In contrast, another router at Aravind deployed on top of the chimney of a power plant from where it derives reasonably stable power has shown uptimes for several months at a stretch. In practice, we have observed that routers with more frequent reboots are more likely to get their flash memory corrupted over time. We had at least 3 such cases at nodes co-located with the vision centers (Figure 6.4), which experienced more reboots since staff at these locations shut down and boot up the routers everyday. Finally, frequently fluctuating voltage also prevents optimal charging of the battery backup and halves its overall lifetime. We've seen batteries rated for 2 years die withing 1 year on the field after experiencing suboptimal charging conditions due to the various power problems.

Lack of quality power increases not only downtime but also maintenance costs. Traveling to remote relay locations just to reboot the node or replace the flash memory is expensive and sometimes has taken us several days, especially in Dharamsala where the terrain is rough.

6.5.2 Other Power-related Problems

In Dharamsala, one of the stormiest locations in India, lightning strikes have often damaged our radios. We have learned the hard way that whenever we deployed a mix of omni and directional antennas, the radios connected to the omni antennas were much more likely to get damaged during lightning storms compared to the radios connected to directional antennas.

It turns out that omni-directional antennas attract lightning more as they are usually mounted on top of masts and have a sharper tip, while directional antennas are typically mounted below the maximum height of the mast. To mitigate this problem, we install omni antennas about 50 cm below the top of the mast. However, this creates dead zones behind the mast where the signal from the antenna is blocked. To reduce these dead zones, we sometimes use an arm to extend the omni antenna away from the mast. After lowering the omni antennas, we have not lost any radios during storms.

6.6 Challenges of Fault Diagnosis

Accurate diagnosis of the problem can greatly reduce response time and thus downtime. The most common description of a fault by our rural partners is that the "link is down." There are a wide variety of reasons for network outages and it is not always easy to diagnose the root cause. The lack of appropriate monitoring tools for inexperienced staff, combined with unreliable connectivity, prevents accurate diagnosis.

For example, a remote host might be running properly, yet is unreachable when an intermediate wireless link goes down. The non-functional link makes it impossible to query

Figure 6.5: The Theni to Vijerani link in the Aravind network was completely obstructed by a newly constructed elevator shaft. This problem was not resolved until we visited Theni after 2 months.

the remote host for diagnosis. In fact, there have been many instances where rural staff have traveled to the remote site with great difficulty only to realize that it was a regular power shutdown from the grid (in which case nothing could be done anyway), or that it was a software problem which could have been fixed if there were an alternate backchannel to the router. Accurate diagnosis of such problems can save considerable time and effort, and prevent unnecessary travel. Furthermore, our own ability to help the local staff by logging in remotely to diagnose the problem is limited by connectivity. For instance, we use the VSAT link at Theni (in the Aravind network) to aid the local staff in monitoring and managing the network, but the VSAT backchannel has worked for only 65% of the time in the last year.

Sometimes local misunderstandings of equipment usage make it even harder to diagnose problems. For example, as shown in Figure 6.5, an elevator shaft was constructed right in front of the directional antenna at Aravind Theni hospital, completely obstructing the line of sight to the remote end. Whenever we remotely logged in to the Theni end of the link from Berkeley, everything seemed fine except that we could not communicate with the remote end. We had no other network access to the remote host so local staff kept physically checking the remote end, but did not (ourselves included) think of checking the roof at Theni. The resulting downtime lasted for two months until we flew there and saw the problem! To be clear, this kind of problem is somewhat rare and unlikely to be repeated. But it does illustrate the difficulties in diagnosis. Next we describe several kinds of more common faults that we found hard to diagnose in the absence of any system level support.

6.6.1 Packet Loss due to Interference

In the AirJaldi network, a decrease in VoIP performance was reported for a particular link at very regular intervals. However without any additional information to diagnose the problem, no action could be taken and this behavior persisted for three months. Finally, after some detailed monitoring by us (and not the rural staff), we saw a regular pattern of packet loss between 8am to 9am every day except Sundays. But scanning the channels showed no external WiFi interference. We were finally able to attribute the problem to a poorly installed water pump that was acting like a powerful spark generator, interfering with wireless signals in the vicinity. Without packet loss information, both the rural staff and we would have had a lot of trouble solving this problem.

6.6.2 Signal Strength Decrease

In a few links in the Aravind Theni network, we noticed a drop in signal strength of about 5 to 10 dB that persisted for 10 to 30 days. Without further information it was hard to tell whether the antennas were misaligned, or the pigtail connectors were damaged, or the radio cards were no longer working well. In the end, several different attempts were made by local staff over multiple trips; the radio cards, the connectors and even the antennas were replaced, and the signal strength bumped back up without it being fully clear what finally helped.

6.6.3 Network Partition

We experienced network partitions many times, but for several different reasons. For example, at Aravind, staff misconfigured the routing and added static routes while dynamic routing was already enabled. This created a routing loop partitioning the network. In another instance of operator error, the default gateway of one of the routers was incorrectly configured. There were also a few instances when operators changed the IP addresses of the endpoints of a link incorrectly, such that the link was non-functional even though it showed up as being associated. And as mentioned earlier, the construction of the elevator shaft left the network partitioned for two months.

6.6.4 "Fixing" by users

A recurring problem is that well-meaning rural staff often attempt to fix problems locally when the actual root cause is not local. For example, at AirJaldi we have seen that when an upstream ISP goes down, rural staff tend to change local settings in the hope of fixing the problem. These attempts typically create new problems, such as misconfiguration, and in a few cases have even resulted in damage to equipment. In all these cases, the network remained non-functional, but now for a different reason, even after the ISP resumed normal connectivity. Thus we need mechanisms to indicate when a link is having problems at the remote end, so as to prevent local attempts at repair.

In general, if the link appears to be down with no additional information or alternate connectivity into the wireless router, it is hard for even experienced administrators to resolve the problem.

6.7 Challenges of Fault Prediction

Some of the node locations in our networks, especially relays, are quite remote. Site maintenance visits are expensive, time consuming, and require careful planning around the availability of staff, tools, and other spare equipment. Therefore, visits are generally scheduled well in advance, typically once every six months. In this scenario, it is especially important to be able to anticipate failures so that they can be addressed during the scheduled visits, or if a catastrophic failure is expected, then a convincing case can be made for an unscheduled visit for timely action. But without an appropriate monitoring and reporting system that includes backchannels, it is difficult to prepare for impending faults. We next discuss the significance of predicting three common faults we have observed in our networks.

6.7.1 Predicting Battery Uptime

At both Aravind and AirJaldi we use battery backups. Loss of grid power at the nodes causes their batteries to start discharging. It is generally not known when the batteries will finally run out. If this information is somehow provided to the staff, they can prevent downtime of the link by taking corrective measures such as replacement of the the battery in time. Such feedback would also suggest if the problem were regional (as other routers would also suffer loss of grid power) or site-specific such as a circuit breaker trip.

6.7.2 Predicting Battery Lifetime

Battery life is limited by the number of deep cycle operations that are permitted. This lifetime degrades sharply because of fluctuating voltages seen in our deployments that do not charge the battery optimally. At Aravind, batteries rated with a lifetime of two years last for roughly three to six months. Information about remaining battery life can also enable prevention of catastrophic failures.

6.7.3 Predicting Disk Failure

We have observed that with frequent reboots over time, the disk partition used to store system logs accumulates bad *ext2* blocks. Unless we run *fsck* periodically to recover the bad blocks, the partition becomes completely unusable very soon. We have also seen that many flash disks show hardware errors, and it is important to keep track of disk errors and replace them before they cause routers to completely fail.

6.8 Summary of Requirements for System Management

First, all aspects of system management require some level of network monitoring. The monitoring system needs to work over intermittently connected links or via alternate backchannels when the main wireless link is down. Second, power quality and availability have been our biggest concerns at both Aravind and AirJaldi. Any rural networking project must necessarily solve the problem of decreased component reliability due to low quality power.

Third, it is challenging to debug a network in the presence of faults, as the only options for connectivity to remote nodes are the failed wireless links themselves. As a result we need to build alternate mechanisms or backchannels to reach remote nodes or query systems inaccessible by primary links. Fourth, failure-independent recovery mechanisms are essential for managing systems remotely. The best solution is to have fully redundant systems (e.g., parallel links between the same pair of wireless routers), but they are often too expensive. An intermediate solution, more viable for rural areas, is to have some independpent modules that enable diagnosis and some recovery. Finally, many faults such as network configuration errors and CF card failures from excessive writes can be easily prevented in software through various safety mechanisms. We therefore propose a management system that has the following key features:

- Support for Monitoring
- Support for Stable Quality Power
- Alternate Network Backchannels
- Independent Recovery Mechanisms
- Safe Software

We discuss the design and implemenation of each of these features in the next chapter.

Chapter 7

System Management for Rural WiLD Networks

The most important operational issue in rural ICTD deployments is the ongoing maintenance and support of the system. This includes handling hardware, software and power failures, in addition to expansion and new installations. In practice, local grassroots groups do not start out with the ability to handle these tasks well. System management tools, no matter how effective, pose a steep learning curve to rural staff with no prior network management experience. To supplement our system management tools, we have found it very useful to create a tiered support model by inserting a complementary layer of support, comprising local IT vendors, between local staff (the first level of support) and the remote management experts (us). Our system management tools limit the need for

---

The work presented in this chapter is joint work with TIER group members Rabin Patra, Sergiu Nedevschi, Manuel Ramos, Lakshminarayanan Subramanian, Yahel Ben-David and Eric Brewer. Some of the material presented in this chapter has been previously published as "Beyond Pilots: Keeping Rural Wireless Networks Alive"[136].

intervention by the local IT vendors and the remote management team, and therefore also help keep costs down in addition to building local capacity.

In this chapter, we first discuss our tiered support model (Section 7.1) and then present five key aspects of our network management system: monitoring (Section 7.2), power (Section 7.3), backchannels (Section 7.4), independent recovery mechanisms (Section 7.5), and software (Section 7.6). Each aspect has been designed to specifically address the goals identified in the previous chapter: *increasing component robustness*, *enabling fault diagnosis*, and *supporting fault prediction*. For each aspect, wherever appropriate, we also discuss tradeoffs affecting our design choices.

7.1 Three-tier support model

The three-tier support model recognizes the range of skills and availability of potential support personnel. In *Tier-1*, we have the *local staff* whose role is to take care of the day-to-day network management and maintenance. They have basic computer skills and brief training in network operation. In *Tier-2*, we have *local network integrators*, who are representatives of local vendors trained in installation, configuration and debugging of networking components, with skills varying from system and RF engineers to tower-climbing technicians. They are paid for their services by the local organization where the network is deployed. We train both the Tier-1 and Tier-2 staff, but the latter have more IT background and can thus handle more issues and ultimately reduce the need for remote management. In *Tier-3*, we have the *remote management team*, comprised of highly skilled professionals who are familiar with the design of the hardware and software of all the networking components.

In this particular case, we play the role of the remote management team.

The goal of our tiered support system is to *increase system availability* and *decrease operational expenditure* by enabling people that are physically close to the deployment and whose services are less expensive to be responsible for as much of the maintenance as possible. Sustainable management of the Aravind system requires capabilities for both local management and remote management. Initially we performed the installation of the first two links. During this stage, we identified local IT vendors who could perform the role of network integrators and trained them, and we also trained the local staff. After system installation and the beginning of operation, a large portion of the maintenance tasks were performed by the remote management team. With time, as problems arose, the management tools and training were updated so that subsequent occurrences could be handled locally, either by the local vendor or the staff. The goal is thus to transition support tasks toward Tier-1. In the last year, the remote management team has not been called into action for any troubleshooting of existing links or help with new installations. The local staff aided by our system management tools and limited help from the *Tier 2* support staff, have managed the entire network themselves. We review this migration in more detail in Section 8.3.

7.2 Monitoring

All aspects of system management require some level of monitoring. During the initial deployment at Aravind, we faced three main challenges in designing a monitoring system. First, the Aravind network at Theni only allowed us to initiate connections from within the network. Second, local staff was not familiar with Linux or with configuration of standard monitoring software such as Nagios [14]. Third, in the absence of backchannels and fault recovery mechanisms, the wireless links would operate intermittently, making reliable log collection very challenging. We addressed these challenges via two solutions: *PhoneHome* to initiate management connections from within the network, and *DTNMon*, an asynchronous file transfer system based on Delay Tolerant Networking [46, 49] to transfer logs over intermittent links.

7.2.1 PhoneHome

PhoneHome is a *push-based* monitoring mechanism in which each wireless router pushes status updates upstream to our US-based server. We chose this method over the general *pull-based* architecture in which a daemon running on a local server polls all the routers. The pull-based approach would require constant maintenance via re-configuration of a local server every time a new router would be added to the network. In contrast, the push-based approach enabled us to configure the routers only once, at installation, by specifying the HTTP proxy to be used.

The Aravind network features two remote connectivity options, both of which are slow and unreliable (Section 7.4): (1) a direct CDMA network connection on a laptop at the central hospital node, and (2) a VSAT connection to another hospital, which has a DSL connection to the Internet. PhoneHome is installed on each of the wireless routers. All the routers periodically post various parameters to our US server website. Server-side daemons analyze this data and plot visual trends.

We collect node and link-level information and end-to-end measurements. The comprehensive list of the measured parameters is presented in Table 7.1. Most of these

Scope	Type	Measured Parameter
Node	Passive	CPU, disk and memory utilization, interrupts, voltage, temperature, reboot logs (number & cause), kernel messages, solar controller periodic data
	Active	disk sanity check
Link	Passive	*traffic:*, traffic volume(#bytes, packets) *wireless:* signal strength, noise level, # control packets, # retransmissions, # dropped packets *interference:* # of stations overheard & packet count from each, # corrupted packets
	Active	liveness, packet loss, maximum link bandwidth
System	Passive	route changes, pairwise traffic volume & type
	Active	pairwise end-to-end delay & max throughput

Table 7.1: Parameters collected by PhoneHome and DTNMon.

parameters can be measured passively, without interfering with normal network operation. However, several of these measurements, such as maximum link or path throughput, require active testing. Some of these tests can be performed periodically (e.g. pinging every network host), and some of them are done on demand (e.g. finding the throughput achievable on a particular link at a given time).

We also use the PhoneHome mechanism for remote management. Every time PhoneHome connects to our US server, it opens a reverse SSH tunnel back into the wireless node, enabling interactive SSH access to the Aravind machines. As the VSAT connection only allows access over an HTTP proxy, we are required to run SSH on top of HTTP, and configure PhoneHome with the proxy. In case of a direct connection to the Internet, no such configuration is required. Another option (employed in the remote management of AirJaldi) is to use the OpenVPN software to open VPN tunnels between network routers and remote servers.

7.2.2 DTNMon

Given that the remote connectivity options into Aravind via CDMA and VSAT were unreliable, we created a DTN-based overlay network for reliable log transfers over intermittently connected links. DTN, the architecture for Delay Tolerant Networking [49] is designed for "challenged" network environments that cannot maintain a low-loss and low-latency end-to-end connection at all times. DTN uses store-and-forward on an overlay network, performs fragmentation to increase efficiency, and ensures reliability by persistent state management to handle failures in network connectivity. A detailed description of the DTN implementation can be found in [46].

We create a DTN overlay network by running a DTN daemon called *DTNMon* on each of the network nodes, as well as on our remote server in the United States. This overlay is organized using a tree topology, with each router maintaining an open TCP connection with its upstream and downstream nodes. When a link goes down, the DTN daemon attempts to reconnect until successful. The overlay routing is statically configured, with each node of the network aware of the routes to all the other nodes in the system.

We use DTN for reliable file transfers because it guarantees reliable message delivery by storing messages in persistent data storage. It also performs automatic fragmentation, which is very useful in long-lasting file transfers over slow and unreliable connections, such as our uplink connection. If a link failure or other type of disconnection occurs during the transfer, DTN can resume the file transfer from the partial message buffered in persistent storage. An operator also does not have to wait for successful delivery of the update/file to the whole network as DTN accepts the message right away at the root node and can asynchronously deliver it to all the nodes whenever they are available.

7.2.3 A Few Examples

In this section, we provide some examples of how our monitoring system enabled diagnosis and prediction of certain faults. PhoneHome and DTNMon proved to be helpful in understanding failures, diagnosing and predicting many faults. First, it helps maintain network reachability information, alerting the local staff when the network was down and that action was needed to be taken for link recovery. Earlier, only a phone call from a rural clinic could alert the local administrator, and depending on the awareness of the staff at the rural clinic, this call would not always happen.

Second, kernel logs transferred using PhoneHome and DTNMon helped us diagnose several interesting problems. For example, in certain instances routers configured with two network interfaces reported only one interface as being active. Pairing this information with power data, we realized that a low voltage supply can prevent two radio interfaces from functioning simultaneously. In another instance, kernel logs and system messages allowed us to examine flash disk error messages and predict when disk partitions needed repartitioning or replacement.

Third, by examining the posted routing table and interface parameters, we were able to diagnose routing misconfigurations or badly assigned IP addresses.

Fourth, continuous monitoring of wireless link parameters helped us narrow the

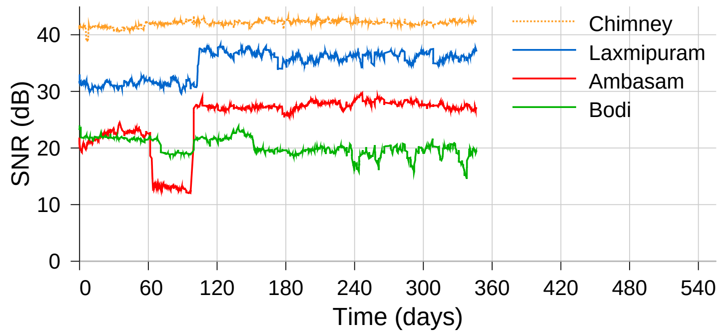

Figure 7.1: Signal strength (shown in dB) variation for all links. Each point is average of measurement over 2 days. The Ambasam link shows a temporary drop in SNR of 10 dB for about 40 days. While the Bodi link is gradually degrading as its SNR has dropped by 4 dB over the last year, the Chimney link's SNR has remained constant.

scope of the problems in many cases. As an example, figure 7.1 shows the signal strength variation in some of our network links at Aravind Theni. While a majority of these links show fairly stable signal strength, some of them show important variations over time. For example, a sudden 10dB signal drop on the link between Ambasam to Theni indicated some kind of a drastic event such as a possible antenna misalignment that needed an immediate visit. On the other hand, a steady decline in signal strength on the Bodi link indicated a gradual degradation of a connector or the RF cable to the antenna that required an eventual visit.

7.2.4 Tradeoffs

We contrast our approach with monitoring at AirJaldi where we use various off-the-shelf tools such as Nagios [14] and SmokePing [20] to collect node, link, and network level parameters. Information is stored at a local data server in Dharamsala and then copied to a US server for detailed analysis. Various graphing toolkits such as MRTG [104] are used to visualize trends and detect anomalies.

The difference in approach compared to Aravind is in part due to the higher experience of the AirJaldi staff, and in part due to the better connectivity we have to AirJaldi. The advantage of having local servers polling for information is that they can be configured by local staff to look for relevant problems, but such an approach is beneficial only if local staff are experienced enough to take advantage of these features.

After three years of operation, the local Aravind staff (some of whom we lost due to turnover after they gained more experience through our training) are more familiar with system configuration, and show less apprehension in taking the initiative and maintaining the system on their own. Therefore, we are now beginning to use a *pull-based* model.

In general, we believe that during the initial phase of a network deployment, minimal configuration *push-based* mechanisms are more appropriate for data collection. However, after building enough local expertise, the monitoring system should be migrated towards a more flexible *pull-based* approach.

7.3 Power

Power quality and availability has been our biggest concern at both Aravind and AirJaldi. Low-quality power damages the networking equipment (boards and power adapters) and sometimes also batteries. As dicussed in Section 6.1, over 90% of the incidents we have experienced have been related to poor quality of power. Thus, designing to increase component reliability in the face of bad power is the most important task. After carefully studying the behaviour of rural grid power (Section 6.5), designing solutions became a much easier task. We have developed two separate approaches to address the effects of low-quality power. The first is a Low-Voltage Disconnect (LVD) solution, which prevents both routers from getting wedged at low voltages and also over-discharge of batteries. The second is a low-cost power controller that supplies stable power to the equipment by combining input from solar panels, batteries, and even the grid.

7.3.1 Low-Voltage Disconnect (LVD)

Excessive discharge of batteries can reduce their lifetimes significantly. Owing to the poor quality of grid power, all AirJaldi routers are on battery backup. LVD circuits, built into battery chargers, prevent excessive discharge of batteries by disconnecting the load (router) when the battery voltage drops below a threshold. As a beneficial side-effect, low-voltage disconnect prevents the router from being powered by a low-voltage source, greatly reducing the likelihood of hung routers. Initially we used off-the-shelf LVD circuits, but found that they oscillated frequently, bringing the router up and down and eventually damaging the on-board flash memory. Every week, there were roughly fifty reboot incidents per router due to hangs caused by low-voltage. A dedicated team of 5 people had to make physical visits to each of these routers every week just to reset them. However, we designed a new LVD circuit [2] with no oscillation and better delay; since then the hangs per week per router have reduced to zero in the AirJaldi network.

7.3.2 Power Controller

Initially we used Uninterruptible Power Supplies (UPS'es) to stabilize the power coming in from the grid. However, affordable UPS systems (around $300) are of the "stand-

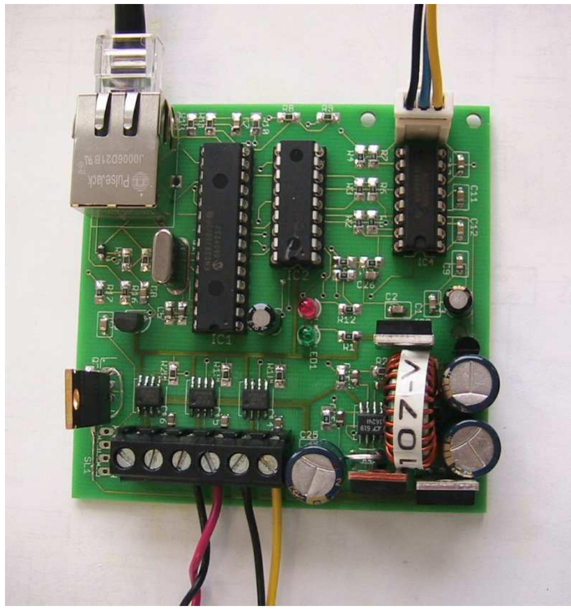

Figure 7.2: Solar controller

by" type. They assume stable grid electricity and pass whatever power comes in from the grid to the load without any correction. As a result, the spikes, surges and sags are passed through to the equipment, thus damaging them. Stable quality power is only provided during grid outages when the UPS routes power to the load from the batteries. Therefore, we had to create our own sources of stable power in remote rural locations with absent or poor quality grid electricity. We focused on small scale solar setups since our wireless routers consume only 7 watts of power on average.

The main components of a small solar power setup such as the solar panel, battery, and charge controller are readily available. But current designs, particularly of the controller which is the most important component of the overall system, are too expensive and too complicated to set up.

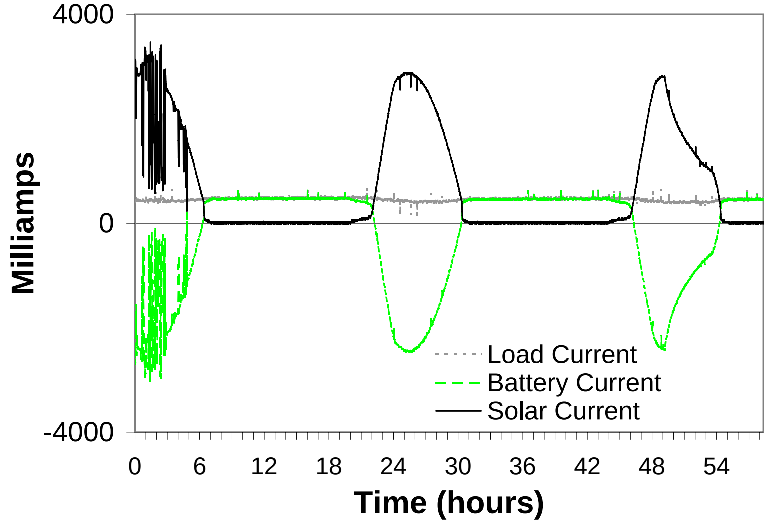

Figure 7.3: Current flow over 60 hours. The load stays even at 7W, while the solar panel and battery shift their relative generation over time. The battery current is negative when it is charging.

We have developed a microcontroller-based solar power charge controller [22], shown in figure 7.2, that provides a stable output of 18V for the routers and intelligently manages the charging and discharging of the battery pack. It has several features such as maximum power point tracking, low-voltage disconnect (LVD), trickle charging and very importantly, support for remote management via ethernet. The setup is trivial as it supplies power to the router using power-over-ethernet. This combination is novel for its price of around $70. Comparable contollers such as Morningstar's SunSaver MPPT [13] cost $250 and do not include remote management.

We use TVS diodes [24] to absorb spikes and surges, and also use a robust voltage regulator to get clean 18-volt power from widely ranging input conditions. Figure 7.3 shows the flow of current through the board over a 60-hour period. We note that power is always available to the router. When enough sunlight is available, the solar panel powers the router and charges the battery. During periods of no sun, the battery takes over powering the router. The frequent swings observed on the left part of the graph are typical for a cloudy day. The graphs also demonstrate how the battery is continually charged when sunlight is available.

LVD and Trickle Charging: Low-Voltage Disconnect (LVD) breaks the connection between the power source and the load in times of low-voltage, thus protecting equipment from brownouts. LVD also prolongs battery life by not allowing the battery to get deeply discharged; it disconnects the load automatically when battery voltage falls below a preset threshold, and powers it back up when the battery voltage is significantly above this threshold. Trickle charging enables the board to charge the battery to maximum capacity while not overcharging it. Using the programmable microcontroller, the charging voltage to the battery can be finely varied thus allowing us to refine the charging plan based on the type of the battery.

Maximum Power Point Tracking: Depending on the load and the intensity of sunlight, at any given snapshot in time the solar panel is operating at a particular power point according to its characteristic I-V curve. For maximum power point tracking (MPPT) [64] we use a high-efficiency step-down DC-DC converter, which functions as the optimal electronic load for the panel, and converts the panel power to a voltage or current level more suitable to the actual load (router or batteries) of the system. This decoupling of the panel from the load allows the operation of the panel at its most efficient power point independent of the effect of the load on the panel at a particular intensity of sunlight. We have measured up to 20% more efficient power draw from 50W panels using our controller via maximum

mode	Battery Voltage (V)	Battery Current (A)	Battery Power (W)	Panel Voltage (V)	Panel Current (A)	Panel Power (W)
direct connect	12.78	2.77	35.39	12.78	2.77	35.39
switch mode 1	12.73	3.20	40.74	16.21	2.65	43.00
switch mode 2	12.90	2.83	36.49	17.45	2.19	38.28
switch mode 3	12.86	2.03	26.11	18.58	1.52	28.24
switch mode 4	12.80	1.73	22.12	18.63	1.25	23.24
switch mode 5	12.77	1.71	21.89	18.76	1.24	23.32
switch mode 6	12.76	1.46	18.62	18.63	1.03	19.26
switch mode 7	12.72	0.78	09.92	19.29	0.54	10.38

Table 7.2: Power Output for Direct Connect and Different Switch Mode Settings

power point tracking.

We use 7 preset conversion ratios or switch modes that determine by what factor the panel voltage is stepped down. The MPPT algorithm determines the best switch mode to use. We use a simple perturb-and-observe algorithm that periodically cycles through all the switch modes and picks the one where the panel output power is the greatest.

Table 7.2 shows a snapshot in time of the amount of power drawn from the panel at different panel voltages. The table also shows the corresponding battery voltages. In direct connect mode, the panel is directly connected to the battery (the controller is shorted). In this case, we see that the panel voltage is driven to the battery voltage, and this is clearly less than optimal. For maximum power at this instant, the panel voltage should be set to the right level, switch mode 1 in this case which is higher than the battery voltage.

Another way to look at the benefit of maximum power point tracking is with respect to the amount of power being sent to the battery. One can alternatively track and maximize the power being sent to the battery. This approach provides a better dynamic range for tracking, especially in situations where power from the panel is barely enough to meet the demands of the load.

Most inexpensive charge controllers do not perform MPPT and hence behave as if they are in the direct connect mode. From table 7.2, we see that 20% more power is extracted from the panel and 15% more power is delivered to the battery for charging, under switch mode 1. Furthermore, we have not lost any routers for power reasons while using our controller.

Remote Management: The controller sends voltage and current information about the panel, load and battery through the power-over-ethernet cable to the wireless router, which then sends it back for diagnosis and prediction of remaining battery uptime.

Advantages: Our charge controller provides four major advantages: simplicity of setup, more efficient power draw, intelligent battery management, and lower cost.

The additional 15% gain in power translates to more power being available to charge the battery faster during a fixed period of sunshine and/or more power available to the equipment being operated. This can factor directly into having longer battery backup times. Additionally, in terms of the overall solar power setup, smaller solar panels can be used, which means lower cost since panels are usually the most expensive part of the solar setup.

Since lead-acid batteries are typically used in solar power setups, it is important that the battery is not deeply discharged nor overcharged to prolong battery life. With low-voltage disconnect and trickle charging, the controller intelligently manages battery charge levels. This leads to longer battery life. The controller senses the battery voltage and switches to trickle charging at a preset threshold voltage, which keeps the battery at near 100% capacity without overcharging.

Also, the board disconnects the load when a minimum battery voltage threshold is reached. The programmed low-voltage disconnect threshold setting allows for an 80% depth-of-discharge (DoD). Without low-voltage disconnect, the battery goes to 100% DoD which shortens battery life significantly. For sealed lead-acid batteries, the total number of charge-discharge cycles for 80% DoD is twice compared to 100% DoD [140]. This translates to twice the battery life in solar setups where the battery goes through a charge-discharge cycle daily. But due to higher levels of battery abuse seen in developing countries, in practice we can triple battery lifetimes. Reducing the churn rate of lead-acid batteries is a huge win in developing countries as spent batteries are discarded by the roadside where the lead seeps into soil, or the batteries are dangerously refilled with acid by hand.

Cost is especially important for deployments in rural developing regions. The controller costs around $70 to construct, which is about the cost of a basic barebones charge controller. Charge controllers with peak power tracking are typically $250 and above. Thus, using the new board incurs no additional cost over a basic solar power setup with a simple charge controller. Furthermore, at no additional cost the new board provides power tracking, low-voltage disconnect, and trickle charging.

Finally, the controller pays for itself. First, peak power tracking draws about 15% more power out of the solar panel (compared to a direct connection). This reduces the overall cost of the system by resulting in longer battery backup times and the use of smaller solar panels. Second, the charge controller roughly doubles battery lifetime at the minimum, by ensuring proper charging and discharging. In addition to the financial benefits, doubling the battery life halves the considerable health and environmental problems of lead-acid batteries.

To summarize, this controller increases panel efficiency, doubles or triples battery life (and in turn provides environmental and healthcare benefits), provides stable power that enables other sustainable ICT projects, reduces appliance damage overall, and enables remote management.

7.3.3 Tradeoffs

One of our major findings was that the real cost of power in rural areas is not just the raw grid electricity costs, but is the cost of overcoming power availability and quality issues through UPS, battery-backups, and and charge controllers. The recurring costs can be quite high, and therefore solar power, although still expensive, becomes more competitve than expected as it can produce clean power directly. Currently we choose to use solar for very remote locations. At less remote and critical sites, we tend to use "dumb" analog chargers to reduce costs even further.

Cost of grid power: The cost of grid power in rural areas is not at all obvious and thus we present data on the operational cost of power. For grid power, the real cost is overcoming the availability and quality issues. In India, a kilowatt-hour (kWh) of grid electricity costs about 5 cents on average [32]. A continuous draw of 7 watts for a wireless router requires 5 kWh per month, and therefore costs around 25 cents per month.

	Capital Cost	Lifetime (months)	Monthly Cost	Monthly Cost 8% interest
7W Rural Grid Power			**$0.25**	
18 AH battery	$45	12	$3.75	$4.05
Small UPS without batteries	$80	24	$3.33	$3.89
UPS with 18 AH battery	$260	24	$5.21	**$6.08**
Power Controller	$70	60	$1.17	$1.71
50W Solar Panel	$275	120	$2.29	$4.95
Controller + Panel + 18AH battery	$390	60/120/36	$4.71	**$8.24**

Table 7.3: Basic costs of power-related equipment. The "Monthly Cost" is just the capital cost divided by the lifetime, while the last column includes interest on the borrowed capital, typically used to pay upfront costs. For solar and UPS, the monthly cost of the battery is calculated over its expected extended lifetime of 36 and 24 months respectively.

Battery costs are driven by the maximum outage length; assuming a 7-watt draw for 10 hours a day, we need 70 watt-hours per day. Using 12V batteries, we thus need about $70/12 = 6$ amp-hours (AH) per day of outage, or about 18 AH to tolerate 3 days without power. Battery life is quite variable and depends significantly on the frequency and depth of discharge. However, in practice, due to battery abuse from incorrect charging we see actual lifetimes of less than one year for batteries used directly, and about two years for batteries used within a UPS.

Table 7.3 shows the costs of the basic components of power management. The capital costs and lifetimes are based on our experience at the Aravind and AirJaldi networks. We compute a monthly cost based on spreading out the capital cost over the lifetime.

Although raw electricity only costs 25 cents per month, a UPS with a sufficient battery for three days of outage costs \$6.08 per month, or about *24 times more*. However, standard UPS systems are only the "standby" type; poor quality power is still passed through from the grid to the router by the UPS. Stable power is delivered only during an outage, when it is routed from the batteries. This represents the *real cost* of rural grid power. Some ICT projects incorrectly view the cost of electricity as zero because it is relatively common to steal electricity in rural India [132]. However, such projects that rely on IT equipment, unlike lighting or heating, generally need consistent and stable power, which comes at a cost.

Cost of Solar Power: As table 7.3 shows, the full cost of solar power remains higher than grid power. We assume a 50-W panel with a lifetime of 10 years and a 5-year lifetime for the controller and the same 18-Ah battery. Due to the panel's high capital cost, the overall monthly cost is \$8.24 or about 35% higher than grid power with UPS. This \$2 difference might be acceptable given the better power, but the difference scales with total power. Therefore, the main value is the ability to obtain reliable power essentially anywhere, which is particularly useful for remote relay points.

7.4 Backchannels

A wide variety of problems at Aravind and AirJaldi have caused link downtimes, leaving remote nodes disconnected. The failure of a single link makes part of the network unreachable although the nodes themselves might be functional. In many cases, if we had alternate access to the nodes, the fixes would have been simple such as correcting a router misconfiguration, or rebooting the router remotely. It is important to have out-of-band access or a *backchannel* to the nodes that is separate from the primary wireless path to it. Backchannel access is also useful in cases where the battery is discharging but the router is already down for other reasons. Information about the battery status from the charge controller via the backchannel would still be helpful. We have tried several approaches to backchannels in both networks. We discuss our experiences with backchannels to the network gateway and to every wireless node within the network.

Network Backchannel: At the Aravind Theni hospital, we already had some form of backchannel into the Theni network through VSAT. We use PhoneHome to open an SSH tunnel over the VSAT link through an HTTP proxy at the Aravind Madurai hospital. We configure PhoneHome to post monitoring data to our US-based server every 3 hours and also to open a reverse SSH tunnel through which we can log back in for administration purposes. Out of the 2300 posts expected from the router at Theni over 143 days (2 posts every 3 hours), we only received 1510 of them, or about 65%. So this particular backchannel was not very reliable in practice, sometimes not working for long stretches of time. As a result, we used the solitary hospital laptop to connect directly to the Internet using a 1xRTT CDMA card to improve the availability of a backchannel into the network. However, this laptop was used for several other purposes [1] and was mostly unavailable. Furthermore, when the local wireless network was internally partitioned, even a working network backchannel was not useful for diagnosis or recovery.

Node Backchannel: At AirJaldi, we built a node backchannel mechanism using GPRS. In India, GPRS connectivity currently costs roughly $10 per month for unlimited duration

[1] shared hardware is a common feature in rural areas

and bandwidth. We used a Netgear WGT634U router, interfaced through its USB 2.0 port with a mobile phone. The router runs PPP over GPRS and sets up an OpenVPN tunnel to a remote server. To enable remote diagnosis using this link, the backchannel router is connected to the main wireless router using ethernet and optional serial consoles. The backchannel router can also power-cycle the wireless router using a solid-state relay connected to one of its GPIO pins.

This approach has two advantages. First, the cellphone network is completely independent of the wireless link. Second, even though the mobile phone is charged from the same power source, it has its own battery, which allows access via GPRS even if the main power source is down. However, for the Netgear router, we needed additional battery backup, which added to the maintenance complexity.

Tradeoffs: Our experience with the GPRS backchannel in terms of providing real utility for system management has been mixed. Many common problems can be solved by alternative means in simpler ways. In cases of incorrect configuration of routers, we can imagine using the GPRS backchannel to fix problems. But at Aravind, when misconfigurations resulted in routing outages, we used cascaded hop-by-hop logins to move through the network, although this depended on at least the endpoint IP addresses to be set correctly. However, we can also use Link Local IP addressing [18] to have independent hop-by-hop backchannels. Each link gets a local automatic IP address from a pre-assigned subnet; this works even when the system wide routing does not work. This can also be implemented by using virtual interfaces in the Atheros wireless driver [31]. Such virtual link configuration approaches could be permanent and also independent of any network configuration

We have also used the built-in WiFi radio of the backchannel netgear router to remotely scan local air interfaces for interferences or low RF signals from other routers, particularly after storms in Dharamsala. But we found the *most useful* feature of the GPRS backchannel to be console access to the router in case of failed attempts at remote firmware upgrades. But arguably, good practices of testing the upgrade locally on an identical router may suffice. This would mean reducing the variety of router platforms used in the field to standardize testing. However, this can be hard to do practically, especially in initial phases as rural networks move from pilots to scale. In future work we intend to continue exploring the idea of cellphone backchannels.

One idea is that instead of using GPRS as the backchannel, a cheaper mechanism could be using SMS channels. With SMS, console access would need to be implemented from scratch. Instead of console access, one approach would be to just query the remote router over SMS. The reply would have power parameters (grid power, remaining battery, voltage level of power supply), and basic status information from the wireless board if it is up. The phone would be connected to the router within the enclosure over serial. This is often feasible because many places have more ready access to SMS compared to GPRS. For example, all our rural clinics at Aravind, have some degree of SMS coverage provided by 2-3 providers at least.

7.5 Independent Recovery Mechanisms

Failure-independent recovery mechanisms are essential for managing systems remotely. The best solution is to have fully redundant systems, but they are often too expensive. An intermediate solution, more viable for rural areas, is to have some independent modules that enable diagnosis and some recovery (but not full functionality and so cannot do complete failover).

Alternate backchannels can enable independent access to various system components, and we include them in the design of independent recovery mechanisms. However in situations where the main router itself is wedged or is in a non-responsive state, we need components that can reset or reboot the main router for recovery. The components should not be affected by the failure themselves. In this section, we discuss software and hardware recovery.

7.5.1 Software watchdog

Essential software services can enter bad states and crash. For instance, we have seen wireless drivers enter bad states that prevent the wireless card from receiving or transmitting packets even though the OS still keeps running. It is necessary to have a monitoring service that can either restart software services on the router or reboot the router itself.

We have built a software watchdog that is run by `cron` every 4 minutes. A configuration file lists what parameters to monitor such as IP reachability to a set of hosts, channel, SSID and BSSID changes, wireless operation mode as well as a list of processes that need to be running on the node. The configuration file also lists what actions to take upon failure of any of the tests, and how often a test is allowed to fail before an action is taken. Actions range from bringing the wireless interface down and up again, unloading and reloading kernel modules, to rebooting the node. We use this software watchdog in the AirJaldi network currently.

7.5.2 Hardware watchdog

An on-board hardware watchdog will reboot the router periodically unless it gets reset periodically after receiving keep-alive messages from the router. This is a vital feature, but most of the low-cost routers used at AirJaldi do not actually have on-board watchdogs. To address this we have designed for $0.25, a simple external hardware watchdog (a simple delay circuit) that interfaces with the board's GPIO line. We have designed this watchdog to plug into the router's power input port and to also accept PoE-enabled power so it can also power PoE-less routers, which allows us to use lower-cost routers as well. All the boards we use at Aravind have on-board watchdogs, but if the board is wedged due to lower voltage, then the watchdog itself will be rendered useless. However, we can avoid this by using the LVDs we have designed. In some cases, we also use the power controller described in section 7.3 as a form of external hardware watchdog; it monitors the board over ethernet and power-cycles it via PoE if it does not hear a keep-alive message in time.

7.6 Software Design

We have written substantial software for the WiLDNet MAC, monitoring, logging, remote management, fault diagnosis, and fault prediction. In this section we focus on aspects that we have not previously discussed: the boot loader, and configuration and status tools. Both play an important role in reducing failures.

7.6.1 Read-only Operating System

We have seen in both the Aravind and AirJaldi networks that the CF cards used in the wireless routers would often get corrupted because of frequent and unexpected reboots. Writing even a single bit of data can corrupt a flash disk. We discovered that if an oscillating LVD keeps rebooting a router, writes to the CF card during boot up will eventually fail and corrupt the flash memory. Unfortunately, since most boot loaders write to flash memory during the boot up process, we had to replace the boot loader with our own version that does not perform any writes at all.

In addition, it is better to mount the main OS partition read-only so that no write operations occur throughout the normal life cycle of the router. For log collection, we have an extra read-write partition on the CF card.

7.6.2 Configuration and Status Tools

To train local staff in the administration of wireless network without exposing them to the details of underlying Linux configuration files, we designed a web-based GUI for easy configuration and display of simple status information about a particular router.

But to further aid local staff in diagnosing problems we need to build tools that can present an easy to understand view of the problem. For example, a simple mechanism at vision centers can indicate (via something as simple as LEDs) that the local wireless router is up and running, but that reachability to the remote router is down. This will minimize the tendency of *self-fixing* where local staff unnecessarily try to modify the local setup without realizing that the problem might be elsewhere.

Problem description	System Aspects
Component Failures	
Unreliable power supply	P
Bad power causing burnt boards and PoEs	P
CF card corruption: disk full errors	M, P, S
Omni antennas damaged by lightning	P
Fault Diagnosis	
Packet loss from interference	M
Decrease in signal strength	M
Network partitions	M, B
Self fixing by users	S
Routing misconfiguration by users	M, B, S
Failed remote upgrade	B, R
Remote reboot after router crash	B, R, S
Spyware, viruses eating bandwidth	M, S
Anticipating Faults is hard	
Finding battery uptime/status	M, B, P
Predict CF disk replacement	M

Table 7.4: List of the types of faults that we seen in both Aravind and AirJaldi. For each fault, we indicate which aspects of the system, as we have designed it, help mitigate the fault. The different aspects are Monitoring (M), Power (P), Backchannel (B), Independent Recovery Mechanisms (R) and Software (S). The information on faults has been collated from logs and incident reports maintained by the local administrators and remote experts respectively.

7.7 Summary

Ongoing maintenance of the system is the biggest challenge in rural technology deployments. In this chapter we presented a system-wide view of the problems and presented key solutions for monitoring, power, backchannels, independent recovery mechanisms, and safe software. Table 7.4 summarizes the system-faults that are handled by each of these solutions.

Specifically, for monitoring we described our PhoneHome and DTNMon solutions for reliable log collection in the presence of intermittent links. We discussed our approach to overcoming power-related faults via custom low-voltage disconnect circuits and solar charge controllers. We demonstrated GPRS-based router-level backchannels, software watchdogs and hardware watchdogs for recovery, and discussed safe software mechanisms such as read-only operating systems to prevent faults from occuring. We showed concrete examples of how all these mechanisms allowed us to decrease component failures, improve fault diagnosis and prediction, and enable recovery.

We also complemented the existing support model by creating a complementary layer of support, consisting of paid local IT vendors, between local staff and the remote experts. This middle tier handles issues locally and reduces the need for management by remote experts. In addition, the system management tools limit the need for intervention by the middle tier, thus further reducing operational costs.

Chapter 8

Impact

In this chapter, we present the real-world impact of our WiLDNet deployment at the Aravind Eye Care System for telemedicine between the Aravind hospitals and various rural vision centers. We first look at the number of patients that have benefited (Section 8.1). Next, we demonstrate that the networking infrastructure contributes to the financial self-sufficiency of the overall telemedicine setup, resulting in its continued operation and patient benefits (Section 8.2). Finally, we look at operational self-sufficiency, particularly in terms of migration of operational responsibility to local staff as an indicator of local capacity building (Section 8.3).

8.1 Patient Results

Figure 8.1 shows the growth in the total number of patients and rural vision centers at Aravind. The Ambasamudram, Andipatti, and Bodi vision centers were existing

The work presented in this chapter is joint work with TIER group members Rabin Patra, Sergiu Nedevschi and Eric Brewer. Some of the material presented in this chapter has been previously published as "Deploying a Rural Wireless Telemedicine System: Experiences in Sustainability" [137].

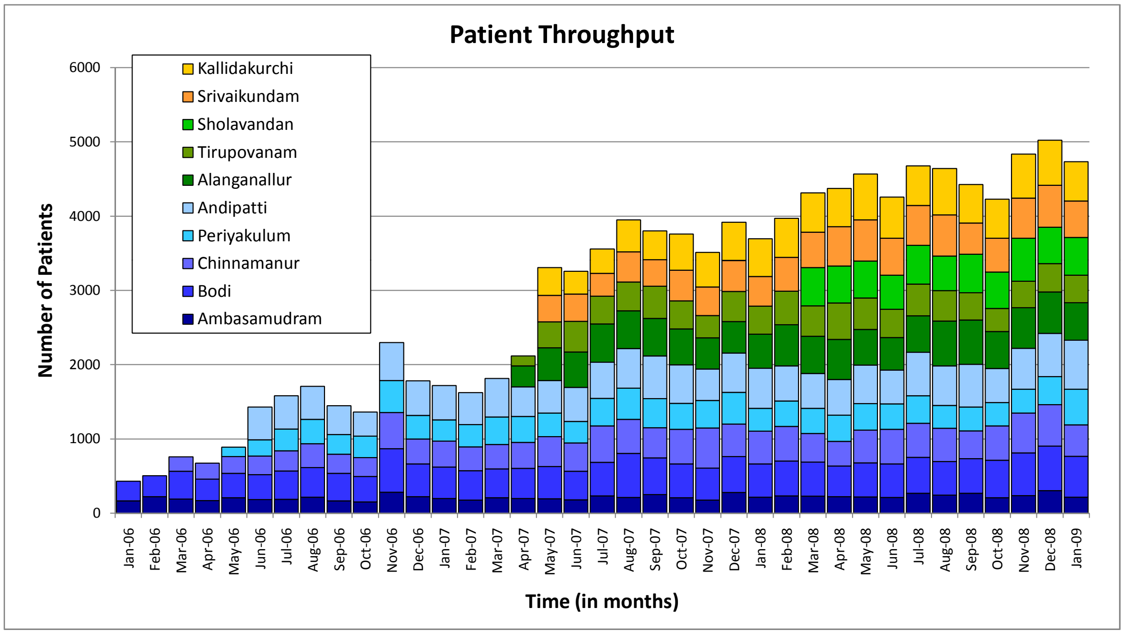

Figure 8.1: Patient Throughput at Aravind

centers that migrated to our high-bandwidth WiLDNet links, while the rest are centers that were enabled by our work. No prior viable technology was available to allow setting up vision centers in those areas. Overall, the network enabled 108,945 remote eye examinations between January 2006 and January 2009. Patient examinations continue to grow.

Since March 2008, when all 10 vision centers were operational, 50,074 patients have been examined at an average of 4,552 patients per month across all 10 vision centers. Of the total patients screened since January 2006, 72% were new patients and 28% were review patients. According to Aravind, this implies two key points: much better penetration of eye care than the eye-camp approach since new patients get treatment every month, and also higher quality of patient eye care with increased patient follow-ups.

Between January 2006 and January 2009, 21,181 patients were diagnosed with

Figure 8.2: Three patients that received their sight back via early diagnosis, resulting in timely cataract surgery.

severe cataract or refractive errors and needed significant vision improvement. Of these, 90% (19,009) got their sight back through cataract surgeries or prescription glasses as advised by the doctor during the video conference. We note again that none of these villages have any ophthalmologists, except for Bodi, which has one doctor in private practice that visits only once a week. We know from interviews that these patients generally would not have received treatment if not for the centers.

The most important impact is economic. In India there is an old saying that a blind person is "just a mouth with no hands", indicating the social burden of blindness. In a study [85] by Aravind, it was shown that 96% of people that get cataract stop working. Among them, of the patients that get surgery as a result of early diagnosis, 85% of men and 58% of women return to income generation activities within a week of surgery. Figure 8.2 shows three of the early patients who received their sight back via early diagnosis, resulting in timely cataract surgeries.

8.2 Financial Sustainability Results

In this section, we show that the centers are cash flow positive, but are probably not yet profitable when considering *all* the capital expenditures. That is, month-to-month they survive, but they do not generate enough income to recover start-up costs or to generate funds to capitalize new centers. We work through these details in part to show the challenges of sustainability and in part to understand how close the project is to true financial sustainability. However, Aravind views the centers as sustainable now, and is expanding to 50 centers. In their view, capital expenditure issues can be handled by a combination of extending the lifetimes of equipment as needed, and by using aid funding for new centers. This appears to match the mode of other development groups.

In our analysis, we consider the capital and operating expenditures of the vision centers and the network separately. We start with the operating finances of the vision centers. Table 8.1 shows the cash-flows of all 5 vision centers in Theni. Income includes consultation fees (25 cents per consultation), medicine sales, test fees, and referral income for patients referred to the Aravind hospital. Expenses include inventories of medicines, diagnostic test kits and other operational costs such as salaries, rent and utilities. They do ont include any capital or operating costs of the networking infrastructure, which we discuss separately. We see that to achieve positive cash flow, centers must get some of the hospital income for referrals. This is reasonable, but it implies that the relatively few expensive procedures at the hospital are part of the sustainability for the centers. If referral income

	Ambasam		**Bodi**		**Chinna**		**Periya**		**Andi**	
Year	2006	2007	2006	2007	2006	2007	2006	2007	2006	2007
Local Income	3,313	3,590	5,526	5,984	3,855	6,381	3,837	5,311	7,110	8,392
Referral Income	1,534	1,890	1,928	2,550	997	2,025	1,018	1,845	2,145	2,865
Expenses	4,676	5,704	6,799	7,377	6,071	6,354	4,530	6,973	8,032	8,451
Cash Flow	171	-224	655	1,157	-1,239	2,052	325	183	1,223	2,806

Table 8.1: Cash-flows, in US dollars, for Aravind-Theni vision centers in the years 2006 and 2007. Local income includes consultation fees, testing fees, and medicine fees. Referral income is income from the main Aravind hospital to which the patient was refered to. Expenses only include the center's operational costs, and do not include any capital or operational costs for the network infrastructure.

is included, then 4 out of 5 centers were cash-flow positive in 2006 and 2007; without it none were positive in 2006, but in 2007 Chinnamanur managed to be positive and Andipatti was borderline. This has encouraged Aravind to set the bar higher and not rely on referral income for sustainability in the future as they believe it is now achievable.

Table 8.2 covers the operating cash flow of one of the centers, Bodi, in more detail. The rest of the income comes from the sale of glasses, medicines (e.g. antibiotics), and glucose tests for diabetes. The actual exam fees are a small part of the income. Overall, this center made about $650 in 2006, excluding the capital and operating costs of the network and the capital cost of the vision center itself. We cover these next.

The local vendor (Tier-2 support) that maintains the links charges about $18

	Income	Expense	Profit
Consulting Fees	807.69		
Medicine Sales	997.55	893.47	104.08
Blood Sugar Fees	138.24	123.81	14.42
Spectacle Sales	3278.51	1517.65	1760.87
Other Income	8.26		
Change in Inventory	295.97		
Referrals to Hospital	1927.60		
Vision Center Rent		630.59	
Office Expenses		342.06	
Telephone Bill		125.54	
Electricity Board		126.52	
Traveling Expenses		55.00	
Staff Salaries		2938.92	
Grand Total	**7453.82**	**6798.55**	**655.27**

Table 8.2: Detailed breakdown of 2006 financials for Bodi vision center

per link per year for maintenance, excluding replacement hardware costs. The reason this cost is low is that our system management tools support local staff and also recover from several faults to limit intervention by the local vendors. Bodi uses two hops with a relay for an annual cost of \$37.70. The total capital cost for those two links was about \$1800 including part of a shared tower, which translates to \$32/month assuming a 5-year lifetime for everything except the tower, 20% salvage value, and 8% interest, or about \$385 per year. This center remained cash-flow positive in 2006 and 2007 when including both the operating and capital costs of the network.

However, if we include all of the capital costs for the center, we get about $11,000. This includes the network but also includes the PC, ophthalmic equipment, furniture, building accessories, etc. If we assume that half the value remains after 10 years, the monthly cost is about $100 with interest, or about $1200 per year. This implies overall net losses of roughly $500 in 2006 and $40 in 2007. Counting these costs, the center is thus not truly profitable, but it is close. For example, assuming a 65% salvage value after 10 years and a 4% interest rate leads to breaking even.

Thus once created, the centers appear to be sustainable and need no extra aid. However, the hospital is unlikely to recoup its initial capital outlay, nor is it earning enough from these centers to pay for the next ones. Since the centers are on the edge of viability, we expect to see a continued emphasis on operating costs and on extending the effective lifetimes of the equipment. On the positive side, patients per month continues to increase, which increases sustainability. From a higher level, this analysis reinforces the need to treat sustainability as a primary goal for systems research in developing regions; we are close only due to considerable effort.

8.3 Operational Sustainability Results

As the system matures it is critical to shift support responsibilities to be entirely local. Thus over time, through training and better tools, we build the capacity to handle more problems locally (either by staff or vendors). Thus an ongoing system design challenge is to examine problems that required remote management and see how they could be handled

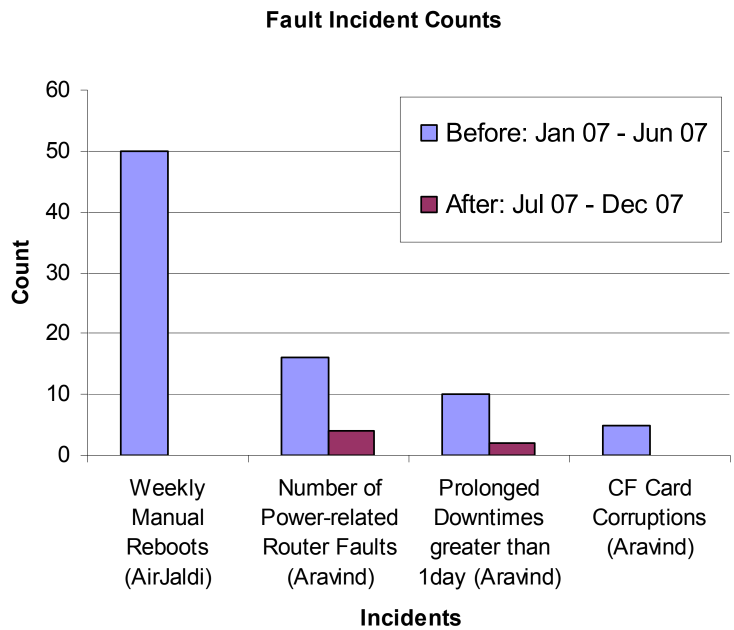

Figure 8.3: Reduction in fault occurrences at Aravind and AirJaldi

locally in the future. We discuss the reduction in faults we have achieved, and also how maintenance responsibility has shifted from remote experts to local staff over the three years of deployment.

Our solutions for power and overall system management have improved the operational sustainability of the networks. As shown in Figure 8.3, after improving the quality of power we have reduced the number of power-related router downtimes. At AirJaldi, we have completely solved the problem of router hangs and reboots associated with poor low voltage disconnects. We have also managed to limit periods of downtimes to under one day; this is primarily due to quicker diagnosis enabled by monitoring and backchannels. Through a combination of quality power and software safety features (e.g., write-protected operating system), we have also completely eliminated compact flash card corruptions.

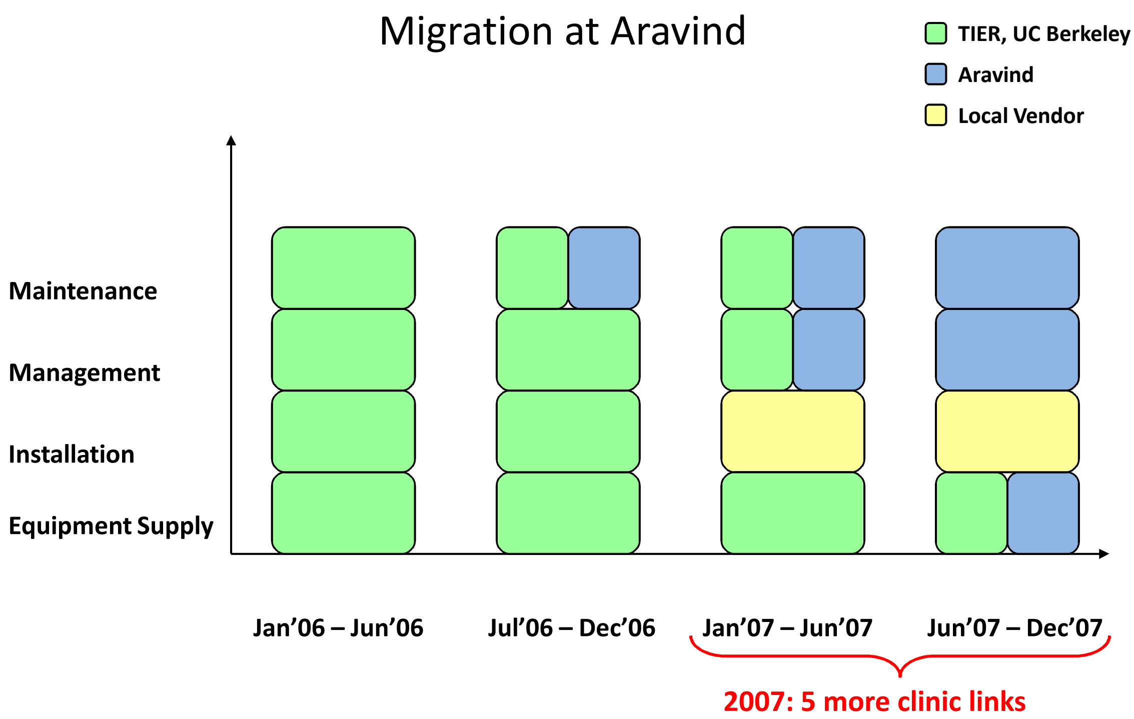

Figure 8.4: Timeline of migration of responsibility over to local staff at Aravind

Over time, operational responsibility for the network has migrated from us (Tier-3) to local staff (Tier-1), and we consider this a big success. As the timeline in Figure 8.4 indicates, we were initially responsible for all aspects of the network in 2006. Through training and development of management tools, local staff has now learned to handle most of the issues. A local vendor (Tier-2) now handles installation issues such as tower construction and antenna alignment. In 2007, our involvement gradually reduced to supplying equipment not easily available in India. However, the equipment was paid for by the hospital, since it was demonstrated that it could recover the network capital and operating costs per vision center. In this time, with our reduced involvement, Aravind managed to set up *five* additional vision centers. Since 2008, our involvement has been virtually zero; we no longer

Type of problem	**Who solved it earlier**	**Who solves it now**
Circuit breaker trip at node locations	**Staff**: Flip the breaker physically at location, added UPS	**Staff**: Monitoring system triggers that node is down
PoE stopped working (transformer explosion)	**Integrators**: Replaced PoE	**Staff**: Replace PoE by checking connectivity and components
Loose ethernet cable jacks	**Experts**, **Staff**: Re-crimp RJ-45 with help from experts, train staff to check for loose cables	**Staff**: Monitoring system triggers that wireless link is up but ethernet is down
Routing misconfiguration: incorrect static routes, absent default gateway	**Experts**: Using reverse SSH tunnel **Integrators**: Using config tool	**Staff/Integrators**: Use config tool for routing
CF card corruption: disk full errors	**Integrators**, **Staff**: Replace CF card **Experts**: Run fsck regularly	**Automatic**: Run fsck on problem **Staff**: Replace CF cards after config.
Wall erected in front of antenna: link went down	**Experts**: After physical verification	**Staff**: Ensure line of sight
Ethernet port on board stopped working	**Integrators**: Replace router board	**Staff/Integrators**: Replace boards

Table 8.3: List of failures that have occurred at various locations in the Aravind network. For each fault, we list who among **staff**, **integrators**, or remote **experts** used to solve the problem, and who solves it now. This information has been collated from logs and incident reports maintained by the local administrators and remote experts respectively. It is an underestimate as not all failures are accounted for in the local logs maintained by local staff.

supply equipment. As recently as early 2009, Aravind has established *two* more vision centers in Theni, and continue to manage the entire network using our system management tools. We have not received any requests for maintenance or management support. Table 8.3 indicates how our system tools have enabled this migration.

Aravind is now expanding the network to include a total of 50 rural vision centers, ultimately providing access to eye care for 2.5 million people.

8.4 Summary

The Aravind vision centers (VCs) have been a profound success for both the hospitals and the patients involved. In three years of operation, the system has enabled over 100,000 remote patient-doctor consultations; the current rate is more than 4,500 patients per month. For about 19,000 patients, the VCs led to restoration of their vision and new economic opportunities. To continue this success, the centers must be both financially and operationally sustainable. We worked to achieve this by reducing deployment and operating costs, increasing deployment flexibility, and developing a long-term support plan that creates a local ecosystem and transitions to local staff and vendors over time. We also developed a range of new tools for system management and to address power issues. Our choices had a large practical impact on making the VCs cash-flow positive, which previous attempts could not do.

Aravind is now scaling the telemedicine network to 50 centers based on the success of this effort. The full network is expected to handle 500,000 exams per year and provide eye care to a rural population of 2.5 million within the next three years. Currently, the VCs break even, but are not quite truly profitable and would have a hard time recovering start-up capital or generating capital for future centers. However, this is still quite favourable as once a small amount of initial capital ($11,000) is raised per center, the centers operate sustainably and need no additional financial assistance.

Although Aravind feels that the network is close enough to sustainability, the process through which we achieved this result highlights the ongoing challenge of sustainability and why it must be a first-class goal for systems research.

Chapter 9

Conclusion

At the beginning of this thesis, we asked the following question: *How can we engineer a sustainable WiFi network that provides high-throughput in long-distance settings, in the face of lossy environments, and in the presence of systemic link or node failures?*

With a focus on achieving sustainability, we answered this question by addressing research challenges across several layers. At the *network-layer*, we designed WiLDNet, a new TDMA-based MAC-layer for point-to-point long-distance WiFi links that increases link utilization, eliminates most packet collisions in single- and multi-hop settings, and combines FEC and ARQ for link-level loss recovery. At the *management layer*, we built a range of tools for system-monitoring over intermittent links, backchannels for fault diagnosis, and mechanisms for hardware and software failure recovery. At the *lowest layer* we designed very low-cost solutions to make wireless nodes more resistant to damage from power fluctuations. We also built off-grid power solutions that extend the life of battery backups, further reducing the costs.

By taking such a broad end-to-end systems perspective, we were able to meet our goals. This work has led to a number of real-world deployments, the most notable of which is the rural telemedicine network at the Aravind Eye Care System, where we have enabled over 100,000 (and counting) high-quality video-based consultations between centrally located doctors and remote rural patients.

In this chapter, we first discuss the limitations of our work (Section 9.1). Next, we propose a few promising areas for future work (Section 9.2). We then provide a detailed summary of our contributions (Section 9.3) and close with a discussion of the lessons learned during this effort (Section 9.4).

9.1 Limitations

Although our work enables new options for rural connectivity, there are some limitations in our approach. Some limitations stem from the choice of technology and the network architecture, while other limitations are specific to the design of the protocol itself.

9.1.1 Line-of-sight constraint

WiLD links require line-of-sight between link end-points, making the network topology highly dependent on the terrain and intervening obstacles. Relays must often be placed at high elevations. When naturally high locations are not available, very tall and expensive towers need to be installed. Network planning can become a sigificant effort [127]. Establishing a link with the required signal strength for a feasible connection implies aligning the directional antennas well to maximize signal strength. However, at very long distances, say 50 kilometers, it becomes challenging to align the links and local rural staff may not be able to do it without specialized equipment.

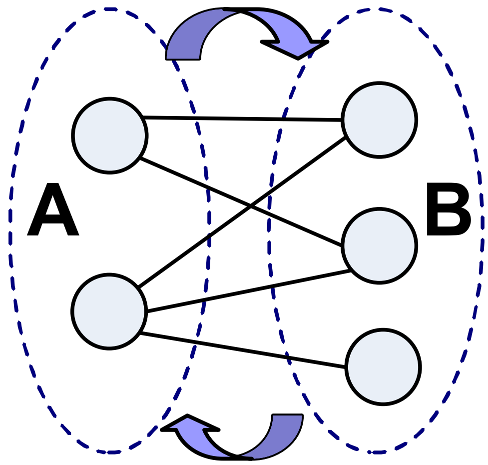

Figure 9.1: SynOp operation in 2P

9.1.2 Bipartite topologies and fixed time slots

Although WiLDNet can operate in multihop long-distance wireless networks using only a single channel, it imposes the constraint that the topology has to be bipartite; for non-bipartite topologies, we can run WiLDNet only on a bipartite subgraph of the overall topology graph.

Figure 9.1 shows an example of such a bipartite network. Using WiLDNet, all nodes in partition A first transmit on all of their links (for a time slot of size $t_{A \to B}$). Following this, all nodes in partition B transmit on all their links (for a time slot of $t_{B \to A}$). The ratio between these slot sizes regulates the bandwidth allocation for every network link between the two partitions. In practice, $t_{A \to B}$ and $t_{B \to A}$ are almost always set to be equal since this maximizes throughput for traffic paths spanning more than two hops [116, 44]. As a result of this constraint of fixed-length slots, currently WiLDNet cannot adapt to dynamic traffic variations.

9.2 Future Work

There are a number of directions that can follow from this work at the network-layer, management-layer, and power-layer. We provide a brief overview of some interesting directions.

9.2.1 Dynamic TDMA slot allocation

Two specific approaches can significantly address the limitations in flexibility and throughput of WiLDNet: 1) *dynamic slot-sizes and* 2) *neighbouring transmissions that can overlap.*

Adapting to Traffic Demand With Dynamic Slot Sizes

WiLDNet features static TDMA slot allocation. This approach is simple, robust, and easy to deploy. However, higher throughputs can be achieved if nodes adapt their slot sizes by using current traffic information. Consider the following examples.

Example 1: Single link: Consider the simplest case of a network with a single link between nodes A and B and assume that the traffic demand only exists from A to B. In this scenario, the highest throughput would be achieved by configuring the link to transmit from A to B for (almost) the entire time. This can be achieved by allocating large transmit slots in the direction $A \rightarrow B$, and very short transmit slots in the reverse direction. If subsequently the direction of traffic flow is reversed, then the optimal slot allocation would correspondingly change, with longer slots from B to A. If we were to use such an adaptive approach, the unidirectional traffic could always be served at close to the full link capacity. Unfortunately, approaches with fixed slot sizes cannot deliver similarly high throughputs.

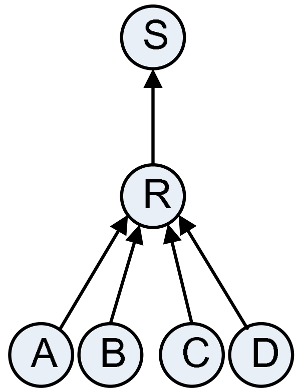

Figure 9.2: Fork Topology

Instead, in these approaches, the link is always scheduled to transmit for $x\%$ of the time in direction $A \rightarrow B$ and $1 - x\%$ in the reverse direction, with a typical setting of $x = 50\%$, to maximize multi-hop throughput [44].

Example 2: A *fork* topology: Figure 9.2 illustrates yet another example. In this scenario, we have a sink node S, and several source nodes A,B,C, and D connected to the sink through relay node R. Let us assume all links have the same data rate. If only one of the sources, say A, sends traffic to the sink, the slot allocation that maximizes throughput is the one in which node R has equally sized transmit and receive slots. In this case, R receives data for 50% of the time, and relays this data for the remainder 50% of the time. Now assume that we have 2 sources sending to S. In this case, the bandwidth-optimal solution would be to have R receive for one-third of the time (from both senders), and then relay this data to S in the remaining two-thirds of the time. Thus, R would have a transmit slot twice as long as the receive slot. Similarly, if all four sources are sending traffic, the best scenario would be the one in which the transmit slot at R is four times longer than its receive slot.

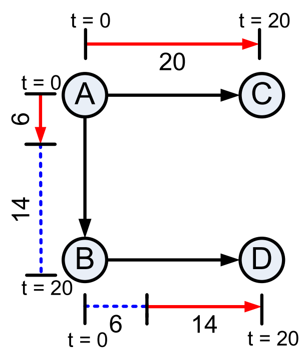

Figure 9.3: Overlap of transmissions

In each of the previous examples, a simple strategy to take advantage of local traffic information would be to monitor the volume of traffic on outgoing links and then adapt the size of TDMA transmit slots to be proportional to the volume of traffic to be transmitted.

Allowing neighboring transmissions that overlap

WiLDNet requires that a node maintains all of its links in transmit mode simultaneously for the same (fixed) time duration. However, there are several situations where this can be needlessly inefficient. For example, consider the topology presented in Figure 9.3, in which traffic flows are represented by arrows. In this topology, since nodes A and B are neighbors, they can never simultaneously operate in transmit mode as per the current WiLDNet protocol.

However, it is possible that the traffic demand is such that A only needs to use a portion of its slot to transmit to B (from say, $t = 0$ to $t = 6$). In this case, we can allow B to start transmitting to a third node (D) at an earlier time ($t = 6$) rather than having to wait until the end of A's transmission slot ($t = 20$). This means that, for a portion of their transmission slots, both A and B can transmit simultaneously while still respecting all the invariants required to avoid interference. Such *neighboring-but-independent* transmissions have the potential to further increase network channel utilization.

These approaches are addressed in the design of JazzyMac [99], a distributed, MAC-layer that came out of the WiLDNet effort. It uses local traffic information to adapt the transmission slot sizes. It uses dynamic slot sizing to negotiate the delay-throughput tradeoff in WiLD networks, and exploits asymmetric traffic, time varying traffic, and non-bipartite topologies. It also enables more efficient use of network capacity by allowing more parallel communication among non-interfering links. Initial simulation results indicate JazzyMac can double the performance of WiLDNet in certain network topologies.

9.2.2 Point-to-Multipoint Networks

Once connectivity is brought into a rural area through WiLDNet, given the density of users it may make sense to provide further omnidirectional coverage over a radius of 10 to 20 kilometers. Current WiFi hardware offers omnidirectonal coverage only up to a few kilometers. The range can be extended using multiple radios and multiple sector antennas.

This offers new research directions in being able to use power control at every base station radio to control the interference among clients. There are also opportunities to save equipment and tower costs by multiplexing several radios across fewer antennas. The key challenge here would be to design a suitable MAC-layer for omnidirectional coverage at longer distances.

9.2.3 Steerable Antennas

The key disadvantage with omni- and directional antennas is the inability to control the radiation pattern, i.e. the direction and gain of the beam.

Antenna designs that allow software control of the radiation pattern are called *electronically steerable antennas* or *smart antennas*. Steerable antennas offers several advantages, and the key research challenge would be to design a MAC-layer than can take advantage of a steerable antenna's capabilities. We briefly describe some of the advantages of steerable antennas.

Automatic alignment: Antenna alignment in long-distance links can be a tedious affair, requiring a person on each making fine adjustments until a reasonable signal strength is achieved. Moreover, the antenna may need to be realigned periodically. Given the beam forming capabilities of steerable antennas, the antennas do not necessarily have to aligned perfectly. The beams can sweep periodically and lock when it achieves the highest link quality.

Dynamic coverage: The ability to change the direction of the beam is a powerful tool to reconfigure links in a network to adapt to changing traffic demands, or reroute traffic when there are failures in the network.

Minimize interference: Steerable antennas can modify their radiation pattern to identify and minimize interference from external sources, by creating nulls in the direction of prominent interfering sources.

9.2.4 System-Management: Enabling Safe Fall Back

In our toolkit of system-management mechanisms, backchannels and independent recovery features can be combined to implement a *safe fallback mechanism* for upgrades. When upgrading the OS on a wireless router, one could use a software watchdog that will be configured to check that the upgrade does not violate any required properties. For example, the board should be able to initialize all the drivers, and ping local interfaces and remote nodes as well. If these conditions are not satisfied, the board should go back to a previously known fail-safe OS state. This can be combined with a hardware watchdog mechanism too that can reboot the router to a fail-safe OS state in cases where the newly installed OS does not even boot.

9.2.5 Power: Stabilizing Grid Power

In this section we describe possible improvements for the power controller controller discussed in Section 7.3.2. An initial direction would be to enable variable-output voltage ranging from 12 to 48 volts to allow the use of the power controller with a wider range of wireless routers. For example, the Mikrotik [12] routers take 48 volts as input.

One very promising direction would be to add an additional feature: the option to take grid power as input, in addition to the current inputs from the solar panel and the battery. This would allow the power controller to fully replace non-performing UPS systems. In other words, the power controller could be used even without solar panels by connecting it to the grid and using power factor correction to stabilize grid power locally. The stabilized grid could then be used to charge the batteries even in the absence of sunlight.

9.3 Summary of Contributions

In this section, we provide a detailed summary of our six major contributions.

9.3.1 Contribution 1: Performance characterization

Our first contribution is a systematic study investigating the drop in performance observed when the standard WiFi MAC is used in long-distance links. Specifically, we showed the following:

- The standard WiFi MAC protocol is ill-suited for long-distance links. In particular, the default stop-and-wait link-layer recovery mechanism results in low channel utilization at long distances. Furthermore, frequent packet collisions occur because of failure of CSMA/CA in long-distance links.

- Although adjusting some 802.11 timing parameters such as ACKTimeout and SlotTime can decrease the number of collisions at long distances, the overall throughput achieved is still much lower than the potential throughput achievable.

- Outdoor links suffer from bursty and variable packet loss due to interference from external WiFi sources, and this variability causes TCP flows to stall and experience timeouts, reducing throughput dramatically.

- In multi-hop settings, co-located radios, operating on the same wireless channel, interfere with each other if one radio transmits while the other receives (Mix-Tx-Rx discussed in Section 3.5.4).

9.3.2 Contribution 2: WiLDNet, a New MAC-Layer

Based on insights from the performance characterization study, we designed, implemented and deployed WiLDNet, a new MAC-layer for use in WiFi-based long-distance point-to-point links. WiLDNet is a TDMA-based MAC-layer that demonstrates signficant performance gains compared to the standard WiFi MAC.

- We used a sliding-window flow control instead of the 802.11's default stop-and-wait packet transmission, and improved channel utilization such that throughput was improved ($\approx$ 6 Mbps) and stayed constant regardless of link distance.

- With the use of TDMA-based slots, we eliminated protocol-induced packet collisions at long distances, that would otherwise occur in the original CSMA/CA scheme.

- In the single link case, we experimentally showed that the higher throughput enabled by WiLDNet is not affected by distance, but only by local interference conditions. Figure 9.4 presents the results of running the WiLDNet MAC on real links from our various deployments in India (Aravind), Venezuela, Ghana and our local Bay Area testbed. Performance of WiLDNet over real links matches the performance seen in emulated links, and it greatly exceeds the performance of the standard WiFi MAC protocol at long distances.

- In collaboration with Ermanno Pietrosemoli from Fundación Escuela Latinoamericano (EsLaRed), we were able to achieve a cumulative bidirectional TCP throughput of 6 Mbps (3 Mbps in each direction simultaneously) over a 382-kilometer single-hop link between Pico Aguila and Platillon in Venezuela (Figure 9.4). To the best of our

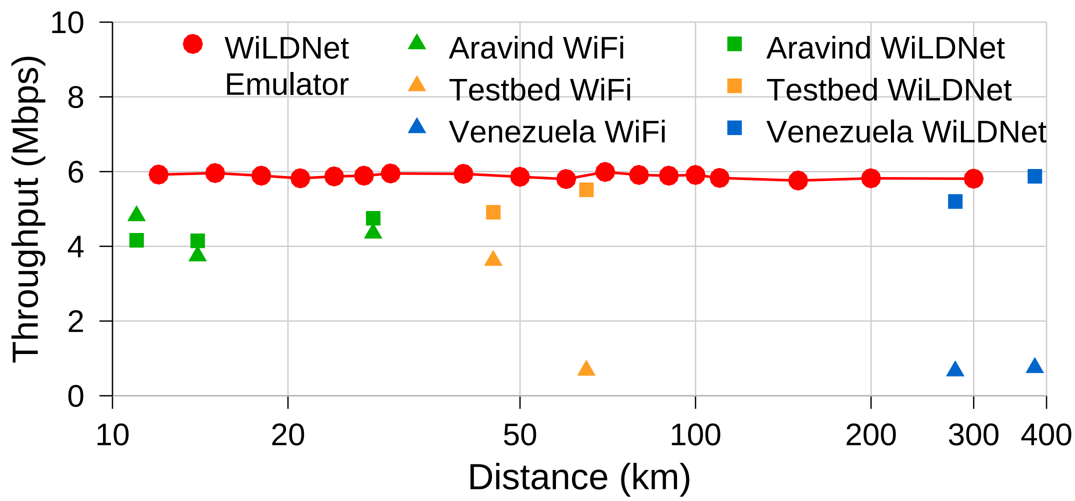

Figure 9.4: Comparison of TCP throughput for WiLDNet MAC (squares) and standard WiFi MAC (triangles) from links in Aravind, Venezuela, Ghana (the 65 km link), and our local testbed in the Bay Area. Most urban links in Aravind had up to 5–10% loss, and so WildNet did not show substantial improvement over standard WiFi. However, WiLDNet's advantage increases with distance. Each measurement is for a TCP flow of 60s, 802.11b PHY, 11 Mbps.

knowledge, this is a new world record for the longest distance at which a stable high-throughput WiFi link has been established without the use of any active amplification or custom antenna design. [1]

- We eliminated inter-link packet collisions by enforcing simultaneous synchronous transmit and receive at all co-located radios on a wireless node. In the multi-hop case, we demonstrated experimentally that the performance achieved by using WiLDNet on identical channels at co-located radios is effectively the same as using the standard WiFi MAC non-overlapping channels.

- We built WiLDNet to run on off-the-shelf Atheros radios running standard firmware,

[1]Each site used a 2.4 GHz 30-dBi reflector grid antenna with 5.3∘ beam-width and a 400 mW Ubiquiti SR2 radio card with the Atheros AR5213 chipset.

enabling *ease of deployment*. In particular, we modified the MadWiFi driver to disable carrier sense, per-packet acknowledgments, and the stop-and-wait flow control mechanism. Then we implemented TDMA slotting and sliding-window flow control in the Click.

9.3.3 Contribution 3: Fault Characterization

Our intial deployment of WiLDNet at the Aravind Hospitals resulted in very valuable operational experiences. We experienced a wide variety of system-level faults, and initially did not know why certain faults kept occuring. We spent a great deal of effort studying the reasons for these faults. Our third contribution is a deeper understanding of how systems can fail in rural operating contexts.

- We showed that poor-quality electricity from the grid was responsible for a majority of the faults. In particular, it caused 90% of the hardware faults at Aravind, resulting in about 50% downtime of the network.

- Further, we demonstrated that the power problem was not just due to grid outages, but mainly due to severe power fluctuations seen when grid power was available. In particular, we performed a careful study of grid voltage and its effects on equipment. We frequently measured voltage spikes over 500 volts, with some even reaching 1000 volts. We also saw extended voltage sags below 70 volts, and extended voltage swells above 350 volts.

- We traced several faults to the effects of fluctuating power. The most obvious were equipment burn-outs caused by voltage spikes. However, extended periods of low voltages also leave the routers in a hung state, unable to boot completely. The onboard hardware watchdog, whose job is to reboot the router into a safe state, is also rendered useless by the low voltage. Therefore the routers stay hung indefinitely. Rapidly fluctuating voltages cause frequent reboots, which corrupt and eventually damage the compact flash (CF) cards in routers by attempting repeated writes to the CF cards during the reboot process.

- We showed that affordable Uninterruptible Power Supplies (UPS'es) do not solve the power issues as they are primarily of the "standby" type. When electricity is available, regardless of its quality, the UPS passes it through to the equipment without stabilizing it. Ironically, stable power is supplied only during outages when the battery backup of a UPS kicked into action.

- One of our major findings is that the real cost of electricity is not just the raw utility price of electricity, but is the cost of stabilizing it. We showed that the approach of using UPS'es and battery backups is up to 24 times more expensive than the raw costs of electricity. Therefore at this point, the use of 50-watt solar panels (normally considered to be expensive) becomes comparable, and has the added advantage of generating stable DC electricity whenever the panel is on.

9.3.4 Contribution 4: System-level mechanisms to mitigate faults

Insights from our initial deployment experiences led us to design system-level tools that included support for monitoring, stabilizing power, alternate network backchannels, independent recovery mechanisms, and safe software. We designed the system with the goal of increasing component robustness, improving fault diagnosis, and enabling fault prediction.

- We implemented two tools for monitoring and remote management, *PhoneHome* and *DTNMon*. PhoneHome is a push-based monitoring system that collects various node, link, and system-level parameters, primarily for diagnosis and prediction of several faults. For example, by pairing network interface kernel logs with power data, we demonstrated that low-voltages can prevent two radio cards from transmitting simultaneously, greatly affecting relays. DTNMon is an asynchronous file transfer daemon that eventually transfers monitoring data reliably in the presence of intermittently operating links.

- We implemented simpler and more effective low-voltage disconnect circuits to prevent routers from hanging or frequently rebooting during low-voltage fluctuations. At AirJaldi we completely eliminated faults due to low voltages. At Aravind, we significantly reduced the effects of low-voltages on the equipment.

- We designed and implemented a microcontroller-based solar-charge controller for use in an off-grid solar-power setup. The controller features maximum power point tracking (MPPT), low-voltage disconnect, trickle charging and power-related remote management.

  The controller can extend battery life by at least 200% via intelligent charging, can draw more power out of the panel via MPPT and is very simple to set up, further reducing the possibilities of faults. Wherever the controller has been used, we have not lost any routers due to poor-quality power.

- We have implemented several backchannel mechanisms that include the use of GPRS and link local IP addressing [18]. At Aravind, quicker fault diagnosis enabled by the use of backchannels has resulted in a 75% decrease in downtimes longer than one day.
- We have completely eliminated CF-card failures due to frequent reboots, by implementing read-only boot loaders.
- We no longer manage any portion of the Aravind network. Local staff, with the help of our management tools and very limited support from the Tier-2 support group, take care of all daily maintenance and management, including expansion of the network.

9.3.5 Contribution 5: Large Deployment with Real-world Impact

WiLDNet has been deployed in several places such as Ghana, Venezuela, the Bay Area and AirJaldi. Here we focus on our most notable deployment, the rural telemedicine network at the Aravind Eye Hospitals.

- We connected ten centers in doctor-less villages to the Aravind hospitals with WiLDNet for high-quality video-based consultations between centrally located doctors and remote rural patients. Over 100,000 (and counting) patients have received remote eye exams in rural areas that had no option for health care before centers enabled by WiLDNet came along. Of these patients, almost 20,000 patients have recovered their sight due to early diagnoses enabled by our network.
- We have enabled the improvement in both the penetration of eye care services and also the quality of care via improved opportunities for follow-up. Every month, 72% of the

patients are new patients and 28% of the patients are follow-up. This is a significant improvement over previous attempts such as eye camps to reach new patients.

- Our previous four research contributions enabled the Aravind network to reach financial and operational sustainability. As a result of this viability, Aravind is scaling the network by adding new links to connect to a total of 50 villages, targeting eye care services for 2.5 million people and 500,000 consultations annually.

- Beyond Aravind, several groups are beginning to use WiLDNet to create similar networks. Examples of such groups include the Lumbini Eye Institute [93] in Nepal, the Pakistan Institute of Community Ophthalmology [107] and Inveneo [77] in Uganda.

9.3.6 Contribution 6: Broad lessons for ICT research

Our final contribution is a broad set of lessons for rural ICT research and deployments. These lessons, listed below, are discussed in detail in the next section:

- Sustainability is a critical systems design goal (Section 9.4.1).
- Prepare for absence of local expertise (Section 9.4.2).
- Simple redesign of components is often enough (Section 9.4.3).
- The real cost of power is in cleaning it (Section 9.4.4).

9.4 Lessons and Closing Remarks

Our work, initiated in 2004, was one of the few early works in the space of Information and Communications Technology (ICT) research specifically focused on achieving positive rural development outcomes. The very nature of this work required deep partnerships and several rounds of deployment with rural development organizations in the field.

We faced several obstacles on the ground during our rural deployments, and have learned many lessons during our journey. We distill our experiences into four main lessons. We believe these lessons are broadly applicable to other rural ICT projects, and sincerely hope they will save other researchers major effort.

9.4.1 Lesson 1: Sustainability is a critical systems design goal

Real impact requires a sustained presence in both deploying and maintaining ICT solutions, and this implies that every aspect of the system must be designed to meet the financial and operational sustainability goals of the partner organizations. In practice, sustainability is a poorly understood concept, as it can have different meanings in different contexts. In this context, at a high level, sustainability implies pushing for low-cost, low-power and easy-to-manage systems. Although these goals are worthy, they are not sufficient by themselves. Our experiences indicate that for an ICT project to be sustainable, it should demonstrate three principles: *optimization of an existing system*, *financial self-sufficiency*, and *operational self-sufficiency*.

Optimization of an existing system

Development projects are not deployed in a vacuum. There is always some existing approach for a given task, even if it is a poor one. We have found it extremely valuable to view our work as optimizing the rural organization's existing system. This approach promotes community buy-in, which is fundamental to sustainability in the long run. This implies a fairly diligent needs assessment of rural organizations in order to frame the overall solution as an optimization of their existing systems. For example, Aravind's previous high-priority efforts at telemedicine were limited by the available technology options at the time. We optimized access to existing doctors for rural patients via low-cost high-performance networks.

Financial self-sufficiency

Any deployment aiming for sustainability must be at least cash-flow positive. Projects that fail to recover at least operating costs tend to fail overall. Under this view, research efforts must focus on technologies with low operating costs. In our case for example, we chose WiFi equipment as it operates in free spectrum, even though it was not suited for our goals initially.

Operational self-sufficiency

As the system matures in usage, it is critical to ensure that maintenance and support be entirely local. Thus an ongoing systems design challenge is to examine problems that require remote management and see how they can be handled locally in the future. In our work, we handled this aspect in several stages. Initially, we (Tier-1 support) managed the network remotely as much as possible. Then we trained local IT vendors (Tier-2 support) to assist Aravind's local staff (Tier-1 support) in system-management, in order to begin the migration of responsibility. Finally, we built several system-management tools to reduce local staff's dependence on the vendors.

9.4.2 Lesson 2: Prepare for absence of local expertise

Most rural ICT projects assume that training will solve the needs of local IT staff, but this is quite difficult in practice. In some sense, better training leads to higher turnover, as local staff apply for better jobs. [2] So instead, we have worked to reduce the dependence on highly trained staff at multiple levels.

Starting at the lowest layers, we have pushed hard on improving the quality of power and the ability of nodes to reboot themselves into a known good state. We have added substantial software for data collection, diagnosis and monitoring. We also developed support for remote management, although it is limited by connectivity issues, especially during faults; in turn, we looked at backchannels to improve the reach for remote management. We also developed GUI tools that are much easier for local staff to use, and are intended to be educational in nature. In the worst case, the Tier-2 network integrators step in to handle issues that local staff cannot solve; earlier local staff would wait until we solved the problem, resulting in extended downtimes. In addition, our suite of management tools further reduce the need for intervention by the Tier-2 support group (Chapter 8).

9.4.3 Lesson 3: Simple redesign of components is often enough

As discussed in Chapter 6, harsh environmental conditions and unreliable power can cause commodity components to fail more often than expected in rural areas. One solution is to use expensive equipment such as military grade routers, big battery backups or diesel generators, as is done with cellular base stations at great cost.

In practice, even simple redesign of selected commodity hardware components can

[2] This is a good outcome

significantly decrease the failure rates without adding much cost. In addition to getting WiFi to work for long distances, we also developed software and hardware changes for low-voltage disconnect, for stable power, and for more reliable automatic reboots, and we developed better deployment techniques to avoid damage due to lightning and power surges.

9.4.4 Lesson 4: The real cost of power is in stabilizing it

It is key to understand that the true cost of electricity in rural areas is not the raw cost of grid electricity, but the cost of stabilizing it using power controllers, batteries and solar-power backup solutions. As shown in Section 7.3.3, this cost of stabilizing electricity can be as high as 24 times the price of grid electricity. Some development projects incorrectly view the cost of electricity as zero, since it is relatively common to steal electricity in rural India.[3] However, the grid cost is irrelevant for IT projects, which generally need stable power anyway (unlike lighting or heating). Due to short lifetimes of batteries and ineffective UPS'es, power stabilization is a recurring cost. Solar power, although still expensive, is thus more competitive than expected as it produces stable power directly. We used solar power for relays or other locations where power was not available, and improved the tolerance for poor-quality elsewhere.

[3]The tolerance of theft is a kind of subsidy for the poor, but it is badly targeted as others steal power too. India loses about 42% of its generated electricity to a combination of theft and transmission losses (vs. 5–10% in the US) [56].

Bibliography

[1] Based on description of expenses by Aravind Eye Care System.

[2] Air Jaldi Charger Hardware Design. http://drupal.airjaldi.com/system/files/Jaldi_Charger_design_1.6.3.pdf.

[3] AirJaldi Wireless Network. http://summit.airjaldi.com.

[4] Aravind Eye Care System. http://www.aravind.org.

[5] Buffalo Technology. http://www.buffalotech.com/home.

[6] Champaign-Urbana Community Wireless Network. http://www.cuwin.net.

[7] CitySense: An Open Urban-Scale Sensor Network. http://www.citysense.net.

[8] CRCNet: Connecting Rural Communities Using WiFi. http://www.crc.net.nz.

[9] Digital Gangetic Plains. http://www.iitk.ac.in/mladgp/.

[10] Iperf. http://sourceforge.net/projects/iperf.

[11] Linksys WRT54G Series. http://www.linksys.com.

[12] MikroTik. http://www.mikrotik.com.

[13] Morningstar sunsaver mppt. http://www.morningstarcorp.com/en/sunsavermppt.

[14] Nagios Wireless Monitoring. http://www.nagios.org.

[15] NY Wireless Network. http://www.nycwireless.net.

[16] OpenWRT. http://www.openwrt.org.

[17] PC Engines. http://www.pcengines.ch.

[18] RFC 3927: Dynamic Configuration of IPv4 Link-Local Addresses . http://www.ietf.org/rfc/rfc3927.txt.

[19] Seattle wireless. http://www.seattlewireless.net.

[20] SmokePing. http://oss.oetiker.ch/smokeping/.

[21] Spirent Communications. http://www.spirentcom.com.

[22] TIER Solar Controller. http://tier.cs.berkeley.edu/wiki/Power.

[23] Ubiquiti Networks. http://www.ubnt.com.

[24] What are tvs diodes? http://www.semtech.com/pc/downloadDocument.do?id=509.

[25] WiMAX forum. http://www.wimaxforum.org.

[26] Daniel Aguayo, John Bicket, Sanjit Biswas, Glenn Judd, and Robert Morris. Link-level Measurements from an 802.11b Mesh Network. In *ACM SIGCOMM*, August 2004.

[27] Airspan Incorporated. Airspan Networks WipLL Solution. http://www.airspan.com/.

[28] D. Allen. Hidden Terminal Problems in Wireless LAN's. *IEEE 802.11 Working Group Paper 802.11/93-xx*.

[29] Aravind Eye Care System. Annual Activities Report 2005. http://www.aravind.org/downloads/reports/activity_report2005.pdf.

[30] Astrid Fletcher, Martine Donoghue, John Devavaram, R D Thulasiraj, Susanna Scott, Mona Abdalla, C A K Shanmugham, P Bala Murugan. Low Uptake of Eye Services in Rural India. *Archives of Ophthalmology*, 117, 1999.

[31] Atheros. MadWiFi driver for Atheros Chipsets. http://sourceforge.net/projects/madwifi/.

[32] P. Audient and F. Verneyre. Electricity in India: Providing Power for the Millions. In *International Engery Agency*, 2002.

[33] A. Balaji, T. A. Gonsalves, and R. Kumar. The corDECT Wireless Local Loop: Capacity Predictions. In *Personal Wireless Communication, year = 1999*.

[34] Hari Balakrishan. *Challenges to Reliable Data Transport over Heterogeneous Wireless Networks*. PhD thesis, University of California at Berkeley, August 1998.

[35] Hari Balakrishnan, Srinivasan Seshan, Elan Amir, and Randy Katz. Improving TCP/IP Performance over Wireless Networks. In *ACM MOBICOM*, November 1995.

[36] Pravin Bhagwat, Bhaskaran Raman, and Dheeraj Sanghi. Turning 802.11 Inside-out. In *Hotnets-III*, 2004.

[37] Sanjit Biswas and Robert Morris. Opportunistic Routing in Multi-Hop Wireless Networks. In *Hotnets-II*, 2003.

[38] B. Bowonder, V. Gupta, and A Singh. Developing a Rural Market e-hub, the Case Study of e-Choupal experience of ITC. *India Planning Commission Report*, 2002.

[39] K. Brauchli, D. O'Mahony, L. Banach, and M. Oberholzer. iPath-a Telemedicine Platform to Support Health Providers in Low Resource Settings. *Studies in Health Technology and Informatics*, 114:11–7, 2005.

[40] Eric Brewer. Technology Insights for Rural Connectivity. October 2005.

[41] Eric Brewer, Michael Demmer, Bowei Du, Melissa Ho, Matthew Kam, Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, Sonesh Surana, and Kevin Fall. The Case for Technology in Developing Regions. *IEEE Computer*, 38(6):25–38, June 2005.

[42] Manuel Castells. *Rise of the Network Society*. Blackwell Publishers, 1996.

[43] Ranveer Chandra, Ratul Mahajan, Thomas Moscibroda, Ramya Raghavendra, and Paramvir Bahl. A Case for Adapting Channel Width in Wireless Networks. In *ACM SIGCOMM*, 2008.

[44] Kameswari Chebrolu, Bhaskaran Raman, and Sayandeep Sen. Long-Distance 802.11b Links: Performance Measurements and Experience. In *ACM MOBICOM*, 2006.

[45] Martin V. Clark, K.K. Leung, B. McNair, and Zoran Kostic. Outdoor IEEE 802.11 Cellular Networks: Radio Link Performance. *IEEE ICC*, 2002.

[46] Michael Demmer, Eric Brewer, Kevin Fall, Sushant Jain, Melissa Ho, and Rabin Patra. Implementing delay tolerant networking. Technical report, 2004.

[47] Christian Doerr, Michael Neufeld, Jeff Filfield, Troy Weingart, Douglas C. Sicker, and Dirk Grunwald. MultiMAC - An Adaptive MAC Framework for Dynamic Radio Networking. In *IEEE DySPAN*, November 2005.

[48] Enlace HispanoAmericano de Salud (EHAS) Foundation. `http://www.ehas.org`.

[49] Kevin Fall. A Delay-Tolerant Network Architecture for Challenged Internets. In *Proceedings of the ACM Symposium on Communications Architectures & Protocols (SIGCOMM)*, 2003.

[50] Rob Flickenger, Steve Okay, Ermanno Pietrosemoli, Marco Zennaro, and Carlo Fonda. Very Long Distance Wi-Fi Networks. In *ACM SIGCOMM Workshop on Networked Systems for Developing Regions (NSDR)*, 2008.

[51] R Gaiha. Decentralization and Poverty Alleviation in Developing Asia. *Economics and Development Resource Center, Asian Development Bank, Manila*, 2001.

[52] R. Gandhi, R. Veeraraghavan, K. Toyama, and V. Ramprasad. Digital Green: A Participatory Digital Framework to Deliver Targeted Agricultural Information to Small and Marginal Farmers. In *Information Technologies and Development (ICTD)*, 2007.

[53] Michele Garetto, Theodoros Salonidis, and Edward Knightly. Modeling Per-Flow Throughput and Capturing Starvation in CSMA Multi-hop Wireless Networks. In *IEEE INFOCOM 2006*, April 2006.

[54] Dattatraya Gokhale, Sayandeep Sen, Kameswari Chebrolu, and Bhaskaran Raman. On the Feasibility of the Link Abstraction in (Rural) Mesh Networks. In *IEEE Infocom*, April 2008.

[55] Grameen Bank. Grameen Telecom Overview. http://www.grameen-info.org/grameen/gtelecom/.

[56] Mark Gregory. India Struggles with Power Theft. http://news.bbc.co.uk/2/hi/business/4802248.stm, 2006.

[57] Ramakrishna Gummadi, Rabin Patra, Hari Balakrishnan, and Eric Brewer. Interference Avoidance and Control. In *HotNets-VII*, 2008.

[58] GV Murthy and SK Gupta and D Bachani and HK Tewari and N John. Human Resources and Infrastructure for Eye Care in India: Current Status. *National Medical Journal of India*, 17, 2004.

[59] N Hanna. Exploiting information technology for development: A case study of India. *World Bank*, 1994.

[60] S Hasan. Introducing E-Government in Bangladesh: Problems and Prospects. *International Social Science Review*, 78(3-4):111–126, 2003.

[61] Richard Heeks. ICT4D 2.0: The Next Phase of Applying ICT for International Development. *IEEE Computer*, pages 26–33, 2008.

[62] C. Heffernan. The Livestock Guru: Demand-led knowledge transfer for poverty alleviation. In *Information and Communication Technologies and Development (ICTD)*, pages 94–100, 2006.

[63] Tristan Henderson, David Kotz, and Ilya Abyzov. The Changing Usage of a Mature Campus-wide Wireless Network. In *MOBICOM*, 2004.

[64] D. P. Hohm and M. E Ropp. Comparative Study of Maximum Power Point Tracking Algorithms. In *Progress in Photovoltaics: Research and Applications*, volume 11, pages 47–62, November 2002.

[65] H.D Hristov. *Fresnel Zones in Wireless Links, Zone Plate Lenses and Antennas*. Artech House, December 1999.

[66] HyperLink. HG2427G 2.4GHz 27dBi High Performance Die Cast Reflector Grid Wireless LAN Antenna. http://www.hyperlinktech.com/multimedia/datasheets/DS_HG2427G.PDF.

[67] M. Ibrahim, B. Bellows, M. Ho, and S. Surana. Improving Healthcare Delivery in Developing Regions: Coupling Output-based Aid and Mobile Technologies. In *University of California Big Ideas Conference*, 2007.

[68] IEEE 802.11 Working Group. IEEE Standard 802.11-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, 1999.

[69] IEEE 802.11 Working Group. IEEE Standard 802.11a-1999, Supplement to Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 5GHz Band, 1999.

[70] IEEE 802.11 Working Group. IEEE Standard 802.11b-1999, Supplement to Part 11:Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High Speed Physical Layer Extension in the 2.4GHz Band, 1999.

[71] IEEE 802.11 Working Group. IEEE Standard 802.11g-2003, Supplement to Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: Amendment 4: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band, 2003.

[72] IEEE 802.11 Working Group. IEEE Standard 802.11e-2005, Supplement to Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: Amendment 8: Medium Access Control Quality, 2005.

[73] IEEE 802.11 Working Group. IEEE 802.11n 1.10 Draft Standard, 2007.

[74] IEEE 802.16. IEEE 802.16 Wireless MAN Standard for Wireless Metropolitan Area Networks.

[75] Indian Institute of Technology, Madras. The TENET Research Group. `http://www.tenet.res.in`.

[76] Intel. Classmate PC. `http://www.intel.com/intel/worldahead/classmatepc`.

[77] Inveneo. `http://www.inveneo.org`.

[78] ITC Limited. The e-Choupal Initiative. `http://www.echoupals.com`.

[79] Kyle Jamieson, Bret Hull, Allen Miu, and Hari Balakrishnan. Understanding the Real-world Performance of Carrier Sense. In *ACM SIGCOMM E-WIND Workshop*, 2005.

[80] Amit P. Jardosh, Krishna N. Ramachandran, Kevin C. Almeroth, and Elizabeth M. Belding-Royer. Understanding Link-layer Behavior in Highly Congested IEEE 802.11b Wireless Networks. In *ACM SIGCOMM E-WIND Workshop*, 2005.

[81] A. Jhunjhunwala. Towards Enabling India through Telecom and Internet Connections. *Variously published essay available online*, 2000. `http://www.tenet.res.in`.

[82] Ashok Jhunjhunwala, Bhaskar Ramamurthi, and Timothy A. Gonsalves. The Role of Technology in Telecom Expansion in India. In *IEEE Communications Magazine*, volume 36, pages 88–94, November 1998.

[83] Joint Learning Initiative. *Human Resources for Health: Overcoming the Crisis*. Harvard University Press, 2004.

[84] M. Kam, D. Ramachandran, A. Raghavan, J. Chiu, U. Sahni, and J. Canny. Practical considerations for participatory design with rural school children in underdeveloped regions: early reflections from the field. *Interaction Design And Children*, pages 25–32, 2006.

[85] Kate Gooding. Poverty and Blindness: A Survey of the Literature. *Sight Savers International*, February 2006.

[86] P. D. Kaushik and N. Singh. Information Technology and Broad-Based Development: Preliminary Lessons from North India. *World Development*, 32(4):591–607, 2004.

[87] G R Kiran. E-governance services through Telecentres - Role of Human Intermediary and issues of Trust. In *Information and Communication Technologies and Development (ICTD)*, pages 131–142, 2006.

[88] Eddie Kohler, Robert Morris, Benjie Chen, John Jannotti, and Frans M. Kaashoek. The Click Modular Router. *ACM Transactions on Computer Systems*, 18(3):263–297, August 2000.

[89] L Dandona and R Dandona and RK John. Estimation of blindness in India from 2000 through 2020: Implications for the blindness control policy. *National Medical Journal of India*, 14, 2000.

[90] F Librero. Uses of the Cell Phone for Education in the Philippines and Mongolia. *Distance Education*, 28(2):231–244, 2007.

[91] Shu Lin and Daniel Costello. *Error Control Coding: Fundamentals and Applications*. Prentice Hall, 1983.

[92] Rowena Luk, Melissa Ho, and Paul M. Aoki. Asynchronous remote medical consultation for Ghana. In *ACM SIGCHI*, 2008.

[93] Lumbini Eye Institute. `http://www.lumbinieyes.org`.

[94] Marratech. Videoconferencing Software. http://www.marratech.com.

[95] David L. Mills. Internet Time Synchronization: The Network Time Protocol. *Global States and Time in Distributed Systems, IEEE Computer Society Press*, 1994.

[96] N Moraveji, U Pawar, and T Kim. A Mouse on Each Desk: An Inexpensive Classroom Interaction Technique for Remote Teaching. *Microsoft Research Working Papers*, 2006.

[97] nanoGSM Picocells. `http://www.ipaccess.com/picocells/nanoGSM.php`.

[98] Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, and Eric Brewer. A Multi-disciplinary Approach to Studying Village Internet Kiosk Initiatives: The case of Akshaya. In *Proceedings of Policy Options and Models for Bridging Digital Divides*, March 2005.

[99] Sergiu Nedevschi, Rabin Patra, Sonesh Surana, Sylvia Ratnasamy, Lakshminarayanan Subramanian, and Eric Brewer. An Adaptive, High Performance MAC for Long-Distance Multihop Wireless Networks. In *ACM MOBICOM*, 2008.

[100] Nicholas Negroponte. *Being Digital*. Random House Inc. New York, 1996.

[101] Nepal Wireless Networking Project. `http://www.nepalwireless.net`.

[102] Michael Neufeld, Jeff Fifield, Christian Doerr, Anmol Sheth, and Dirk Grunwald. SoftMAC - Flexible Wireless Research Platform. In *HotNets-IV*, November 2005.

[103] G Nulens and L Van Audenhove. An Information Society in Africa?: An Analysis of the Information Society Policy of the World Bank, ITU and ECA. *International Communication Gazette*, 1999.

[104] Tobi Oetiker. MRTG: The Multi Router Traffic Grapher. http://oss.oetiker.ch/mrtg/.

[105] Government of India. Community Information Centers. `http://www.cic.nic.in`.

[106] P K Nirmalan, J Katz, A L Robin, R Krishnadas, R Ramakrishnan, R D Thulasiraj, J Tielsch. Utilisation of Eye Care Services in Rural South India: the Aravind Comprehensive Eye Survey. *British Journal of Ophthalmology*, 88, 2004.

[107] Pakistan Institute of Community Ophthalmology. `http://www.pico.org.pk`.

[108] Joyojeet Pal, M Lakshmanan, and Kentaro Toyama. My Child Will be Respected: Parental Perspectives on Computers in Rural India. In *Information Technologies and Development (ICTD)*, volume 2, pages 168–177, Bangalore, 2007.

[109] T Parikh, K Ghosh, and A Chavan. Design Studies for a Financial Management System for Micro-credit groups in Rural India. In *Universal Usability*, pages 15–22, 2003.

[110] Partnership for Higher Education in Africa. Securing the Lynchpin: More Bandwidth at Lower Cost. *Carnegie Corporation and the Ford, MacArthur and Rockefeller Foundations*, 2006.

[111] Rabin Patra, Sergiu Nedevschi, Sonesh Surana, Anmol Sheth, Lakshminarayanan Subramanian, and Eric Brewer. WiLDNet: Design and Implementation of High Performance WiFi Based Long Distance Networks. In *NSDI*, 2007.

[112] Udai Pawar, Joyojeet Pal, R Gupta, and Kentaro Toyama. Multiple Mice for Retention Tasks in Disadvantaged Schools. In *ACM SIGCHI*, 2007.

[113] C K Prahalad and A Hammond. *Serving the World's Poor, Profitably*. Harvard Business Review, 2002.

[114] Private communication with Melissa Ho, PhD Candidate, UC Berkeley School of Information. `http://www.ischool.berkeley.edu/people/students/melissaho`.

[115] One Laptop Per Child Project. `http://www.laptop.org`.

[116] Bhaskaran Raman. Channel Allocation in 802.11-based Mesh Networks. In *IEEE INFOCOM*, April 2006.

[117] Bhaskaran Raman and Kameswari Chebrolu. Revisiting MAC Design for an 802.11-based Mesh Network. In *HotNets-III*, 2004.

[118] Bhaskaran Raman and Kameswari Chebrolu. Design and Evaluation of a new MAC Protocol for Long-Distance 802.11 Mesh Networks. In *ACM MOBICOM*, August 2005.

[119] V. Kasturi Rangan and R.D. Thulasiraj. Making sight affordable (innovations case narrative: The aravind eye care system. *Innovations*, 2:35–49, 2007.

[120] Ananth Rao and Ion Stoica. An Overlay MAC layer for 802.11 Networks. In *MOBISYS*, Seattle,WA, USA, June 2005.

[121] A Rendon, A Martinez, M Dulcey, J Seoane, R Shoemaker, V Villarroel, D Lopez, and J Simo. Rural Telemedicine Infrastructure and Services in the Department of Cauca, Colombia. *Telemedicine Journal and e-Health*, 2005.

[122] C. G. Riley. The Changing Role of the Citizen in the E-Governance and E-Democracy Equation. *Commonwealth Centre for e-Governance*, 2003.

[123] Luigi Rizzo. Effective Erasure codes for Reliable Computer Communication Protocols. *ACM CCR*, 1997.

[124] Maya Rodrig, Charles Reis, Ratul Mahajan, David Wetherall, and John Zahorjan. Measurement-based characterization of 802.11 in a hotspot setting. In *ACM SIGCOMM E-WIND Workshop*, 2005.

[125] Sandra Salmerón, Javier Simó, and Rabin Patra. Comparison of MAC protocols for 802.11 based long distance networks. In *1st Workshop on Wireless Broadband Access for Communities and Rural Developing Regions at International Conference on Mobile Communication Technology for Development (M4D)*, 2008.

[126] L S Satyamurthy and A Bhaskaranarayana. Telemedicine: Indian Space Agency's (ISRO) Initiatives for Speciality Health Care Delivery to Remote and Rural Population. `http://laico.org/v2020resource/files/remote_rural_population.htm`.

[127] Sayandeep Sen and Bhaskaran Raman. Long Distance Wireless Mesh Network Planning: Problem Formulation and Solution. In *WWW*, 2007.

[128] Anmol Sheth, Christian Doerr, Dirk Grunwald, Richard Han, and Douglas Sicker. MOJO: A Distributed Physical Layer Anomaly Detection System for 802.11 WLANs. In *MOBISYS*, New York, NY, USA, 2006. ACM Press.

[129] Anmol Sheth, Sergiu Nedevschi, Rabin Patra, Sonesh Surana, Lakshminarayanan Subramanian, and Eric Brewer. Packet Loss Characterization in WiFi-based Long Distance Networks. In *IEEE INFOCOM*, 2007.

[130] J Simó, C Figuera, J Seoane, and A Martínez. Distance Limits in IEEE 802.11 for Rural Networks in Developing Countries. In *IEEE WRECOM*, 2007.

[131] J Simo, P Osuna, D Espinoza, L Camacho, and R Quispe. Application of IEEE 802.11 Technology for Health in Isolated Rural Environments. In *WCIT*, 2006.

[132] Thomas B. Smith. Electricity Theft: A Comparative Analysis. In *Energy Policy*, volume 32, pages 2067–2076, December 2004.

[133] Streakwave Wireless Inc. `http://www.streakwave.com`.

[134] Lakshminarayanan Subramanian, Ion Stoica, Hari Balakrishnan, and Randy Katz. OverQoS: An Overlay Based Architecture for Enhancing Internet QoS. In *NSDI*, March 2004.

[135] Lakshminarayanan Subramanian, Sonesh Surana, Rabin Patra, Melissa Ho, Anmol Sheth, and Eric Brewer. Rethinking Wireless for the Developing World. In *Hotnets-V*, 2006.

[136] Sonesh Surana, Rabin Patra, Sergiu Nedevschi, Manuel Ramos, Lakshminarayanan Subramanian, Yahel Ben-David, and Eric Brewer. Beyond Pilots: Keeping Rural Wireless Networks Alive. In *Proceedings of the 5th ACM/Usenix Symposium on Networked Systems Design and Implementation (NSDI)*, 2008.

[137] Sonesh Surana, Rabin Patra, Nedevschi Sergiu, and Eric Brewer. Deploying a Rural Wireless Telemedicine System: Experiences in Sustainability. In *IEEE Computer*, volume 41, June 2008.

[138] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[139] United Nations. UN Millennium Development Goals. `http://www.un.org/millenniumgoals/`.

[140] D.G. Vutetakis and H. Wu. The effect of charge rate and depth of discharge on the cycle life of sealed lead-acid aircraft batteries. In *IEEE 35th International Power Sources Symposium*, June 1992.

[141] R. Wang, U. Sahni, S. Sobti, N. Garg, J. P. Singh, M. Kam, A. Krishnamurthy, and T. Anderson. The Digital StudyHall. Technical Report TR-723-05, Computer Science Department, Princeton University, 2005.

[142] U Wehn and R Mansell. *Knowledge Societies*. Oxford University Press, New York on behalf of the United Nations, 1998.

[143] Wi-LAN Incorporated. Wi-LAN Fixed and Mobile Solutions. http://www.wi-lan.com.

[144] World Bank. The Information for Develoment Program (infoDev). http://www.infodev.org/.

[145] World Bank. World Development Indicators, 2008. http://www.worldbank.org/data.

Appendix A

Information and Communications Technology *for* Development (ICTD)

A.1 ICT for Development: A brief history

The link between technology and development has always been a hotbed of academic research. Work in this space includes new ideas of a networked society from both academic communities [42, 100] as well as international development agencies [51, 59, 103, 142].

Much of this interest in developing countries has been driven by the rapid growth of their respective high-tech software industries and their interplay with local social, economic and political challenges. In countries such as India, the highly visible public image of the software industry has created an aspirational value centered around computer literacy among people at all levels of society. Recent studies indicate that computer literacy has acquired immense symbolic value that is more tied to social and economic aspirations than actual functional use [108].

On the business side, much theory has been developed on *bottom-of-the-pyramid* markets. For example, C.K. Prahalad suggested that products and services could be brought to untapped rural markets via profitable business models, not just philantrophic ventures [113].

A combination of technology growth and the possibility of viable rural business models created a strong interest in engineers, scientists, and entrepreneurs around the issues of technology and development. Hundreds of ICT initiatives, ranging from pilots to large-scale initiatives, began to spring from international aid agencies, governments, non-governmental organizations, foundations, corporations and private entrepreneurs. As an example of this surge, the World Bank's *infoDev* website catalogs hundreds of ICT projects [144]

These ICT initiatives target various areas of development ranging from healthcare, agriculture, education, governance and business development. In healthcare, the use of ICTs have been explored to bring telemedicine to remote areas [39], to develop low-cost diagnostic equipment, and for improved information gathering and patient monitoring [67, 92]. In agriculture, ICTs have been proposed for information sharing among farmers [52], to provide price information [62, 86], and for efficient supply chain management [78]. In the field of education, ICTs have been used to improve learning among children through educational games on low-cost computers [115, 112], cellphones [90, 84] and shared computing in schools [96, 141]. ICTs have also been used to improve transparency in governance through e-payments and e-voting [60, 87, 122] and to enable small entrepreneurship through micro-credit [109].

A common theme of many early ICT initiatives was an over reliance of existing off-the-shelf technologies designed specifically for markets in industrialized nations, without proper assessments of what solutions would be appropriate in the rural context and of how off-the-shelf technologies might need to be modified or redesigned. Although well intentioned, a significant number of initiatives remained in the pilot stages of implementation and failed to evolve into sustainable models that could be replicated or scaled to have widespread impact.

Such early experiences led to subsequent academic work revisiting the impact of ICT projects for development, to distill best practices and apply them to future efforts [61]. In the early years of this decade, ICT for Development (ICTD) was proposed as a new area of multi-disciplinary research combining the efforts of social scientists, economists, public health professionals and technology researchers from various fields such as computer science, electrical, mechanical and civil engineering. A number of research groups were formed at universities such as University of California at Berkeley [41], University of Washington [109], Massachussetts Insitute of Technology, Carnegie Mellon, IIT Kanpur [9], IIT Chennai [75], to name a few, to work in the area of technology for development. Around the same time, several research labs of major corporations, such as Hewlett Packard Labs and Microsoft Research set up operations in various developing countries, adding further momentum to the movements thinking about new ideas relevant to the developing world.

A.2 ICT for Development: The challenge

A significant focus in ICTD research has been placed on communications infrastructures. Although reliable communication infrastructure is taken for granted in the industrialized world, similar infrastructure is not widely available in large parts of the developing world, and even in some rural areas of the industrialized world.

The most visible communication technology in developing countries has been cellular telephony, which has seen tremendous growth in the last few years. With 1.15 billion mobile phones sold worldwide in 2008, penetration has increased fastest in India, Brazil, China and sub-Saharan Africa. The flagship example of the impact of cellular telephony is Grameen Telecom [55] in Bangladesh, where women in villages use micro-credit to buy phones as franchisees and then operate it as a pay phone by renting it to others in nearby locations. At its peak, Grameen Telecom had 95,000 franchisees that provided coverage to around 60 million people in 50,000 of Bangladesh's 68,000 villages.

Satellite communication has been a common option, albeit expensive. India's space research organization, ISRO has been providing Very Small Aperture Terminal (VSAT) coverage for over 20 years for remote education and telehealth [126]. ITC's e-Choupal [38] used VSAT terminals to connect its kiosks across villages to provide pricing information to farmers for direct marketing. VSAT is also commonly used to connect telecenters [105].

Although the benefit of various communication technologies is well acknowledged and desired, there are vast areas and sections of populations in developing countries that are under-served and un-served. Although mobile phones have seen fantastic growth rates in developing countries, the overall penetration of mobile phones is still low at around

10% [145]. And within developing countries, the divide between urban areas and the rural areas is even more striking. In India, for example, there are 45 mobile phones for every 100 people in urban areas, but only 4.5 mobile phones for every 100 people in rural areas. In sub-Saharan Africa, rural mobile phone penetration rates are even lower.

Internet usage patterns also exhibit wide gaps. For example, the total number of Internet users in sub-Saharan Africa is only 35 million, which is about the same number of Internet users in South Korea, a country with only $1/16th$ of the population.

The reasons for these disparities are many. Most rural areas in developing countries do not have high enough user density or purchasing power to enable profitable network deployments, whether cellular, VSAT, or wireline. In addition, in many developing countries older regulations allow existing monopolies to restrict access to to network infrastructre or to charge prohibitively exorbitant fees.

The default choice for connectivity in industrialized countries is fiber, but the high costs of laying fiber ($1000 per kilometer in India) cannot be recouped from low-density low-income areas with low initial demand. The total cost of installing a phone in the US is about $500 which is viable given that more than 90% of the households can afford $30 per month on telephone service. In contrast, in India more than 60% of the households can afford to only spend at most $5 on communications every month [81].

Satellite networks provide fantastic coverage, but are very expensive. VSAT equipment installation costs over $10,000 with a recurring monthly cost of over $2000 for a 1 Mbps link. In low user-density areas, VSAT is affordable only for businesses or wealthy users [110].

Networks with a base station model such as WiMAX or cellular networks like

GSM and CDMA have an assymetric deployment model where expensive base stations are amortized by a large number of users. In low density regions, such base stations simply do not cover enough users to be economically viable. Furthermore, in a typical network, 70% of the cost is in the access network, not in the backbone.

Therefore the expectation that cellular solves the connectivity problem for developing regions is somewhat of a myth. Cellular success in developing regions is still largely an urban phenomenon, with few exceptions. For example, even the best known "rural" cellular system, Grameen Telecome [55] in Bangladesh, avoids rural-only base stations. Instead, by exploiting the high population density of Bangladesh, Grameen places base stations such that they cover both higher income urban users and lower income rural users, thereby subsidizing their costs of rural coverage. Typically, there is no coverage for areas that are not near an urban base station. China is one country with good universal coverage, but it is dictacted by strong government policy despite the economic difficulties in providing such coverage. Other countries either subsidize rural users via taxation, much like the US universal service tax or mandate rural coverage as part of urban spectrum allocation. Thus, many cellular providers incur losses in low user-density regions and partially recoup these losses by either charging very high usage rates or imposing a universal service charge on all users. WiMAX, with the intended deployment model of expensive base stations covering many users, shares the shortcomings of cellular technologies.

We do expect that WiMAX and cellular networks, particularly with new developments in lower cost pico and femto cells, will expand over time, but there is immediate need for new rural connectivity options.

A.3 ICT for Development: The time is now

Despite the great challenges in providing low-cost connectivity in rural areas, three trends make now the right time to deploy appropriate ICT solutions in developing regions: *the impact of Moore's law*, *the widespread growth and availability of wireless communication*, and *the emergence of more supportive business environments.*

- *Moore's Law:* The impact of Moore's Law has decreased the cost of computing to fractions of a cent per user. Although this cost reduction applies mostly to shared infrastructure such as servers and datacenters, it also partly applies to personal devices such as PCs, laptops and mobile phones. Today, PCs can be bought for less than $200, small but fully featured laptops such as the Asus Eee PCs and MSI Wind netbooks can be bought for around $300, and mobile phones such as Nokia 1110 are available for $20. Similar scaling advantages have given rise to well known projects such as the Intel Classmate PC [76] and the OLPC laptop [115]. In parallel, models of computing using shared PCs and phones are increasing being adopted. With a focus on shared server infrastructure and low-cost, possibly shared, end-user devices, the cost of computing becomes realistic for even the poorest users.

- *Growth in Wireless Communications:* High volume production of wireless communication equipment, particularly cellular and WiFi, has decreased their costs as well. WiFi radios, available for $50 and generally operating in unlicensed spectrum, have been used with directional antennas for long distance links in sparsely populated rural areas. Integrated radio and antenna solutions such as the NanoStation from Ubiquity Networks are now available for $80. Efforts are also on to manufacture lower cost and lower power base stations to provide highly targeted coverage. For example, GSM pico cells [97] transmit 23bBm (200mW) of power can can provide coverage to small villages. Therefore, for rural areas, a wireless infrastructure appears to be the first kind of infrastructure that is affordable.

- Supportive business environment: There now exist supportive business environments for ICT projects. The diffusion of technology worldwide and the growing access to capital have created a favourable environment for entrepreneurship and experimentation. This environment, combined with the success of franchising as a way to deploy large-scale ICT projects, shows that there is a viable path from research to large-scale impact in the real world.

Leveraging these trends, many early projects for low-cost rural connectivity took shape. The Akshaya network [98] was deployed ty the Indian government in 2004 to connect more than 600 village kiosks in a district of 3.5 million people for the purposes of spreading computer literacy. The network employed a two-tier architecture consisting of backhaul links up to 30 kilometers long and local access links up to 10 kilometers long, using proprietary outdoor wireless technologies developed by WiLAN [143] and AirSpan [27]. Other projects such as AirJaldi [3] in India, CRCNet [8] in New Zealand, and EHAS [48] in Peru have used standard WiFi equipment to build community wireless networks.