University of California
Santa Barbara

Robust, Resilient Networked Communication in Challenged Environments

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

Doctor of Philosophy
in
Computer Science

by

Michael S. Nekrasov

Committee in charge:

Professor Elizabeth Belding, Chair
Professor Chandra Krintz
Professor Rich Wolski

March 2020

The Dissertation of Michael S. Nekrasov is approved.

Professor Chandra Krintz

Professor Rich Wolski

Professor Elizabeth Belding, Committee Chair

March 2020

Robust, Resilient Networked Communication in Challenged Environments

by

Michael S. Nekrasov

Acknowledgements

This work is based on co-authored publications with Elizabeth Belding, Ryan Allen, Vivek Adarsh, Udit Paul, Max Ginier, Miriam Metzger, Lisa Parks, Esther Showalter, Ellen Zegura, Daniel Iland, Morgan Vigil-Hayes, Irina Artamonova, and Ben Zhao.

I would like to express my greatest gratitude to my committee Elizabeth Belding, Rich Wolski, and Chandra Krintz for guiding me throughout my degree. You have invested countless hours in my education and helped shape me as a computer scientist. Elizabeth thank you in particular for years of meetings, advice, and support.

Thanks to my coworkers and lab-mates, Nevena Golubovic, Morgan Vigil-Hayes, Daniel Iland, Paul Schmitt, Esther Showalter, and Mai EL-Hussein for support and the endless hours of fruitful discussion. Thank you for your professionalism and friendship. A special thanks to Vivek Adarsh and Udit Paul for the great company, fantastic ideas, and endless hours of field work. You made the hours of work go by in a flash. I am looking forward to all the amazing research you will produce in the coming years. In addition, thank you to local UCSB collaborators Lisa Han, Hannah Goodwin, Kristi Hocevar, and Pritha Narayanappa for collaborative work on freedom of Speech.

None of this research would have been possible without many US and international collaborators. In particular, thank you to Jerrold Baca, Joseph Peralta, Kiss Abraham, and Zaya Nara for your time and resource investment. It was a pleasure to work with you. Thanks to many local partners, including the Southern California Tribal Chairmen's Association's Tribal Digital Village program, the Iipay Nation of Santa Ysabel, Santa Clara Pueblo, and Ohkay Owingeh. Thanks to Coal Oil Point Reserve for allowing us to utilize your space for our aerial testbed.

Thanks to Kris Jaroensutasinee and Mullica Jaroensutasinee for hosting me at the Center for Excellence in Ecoinformatics at Walailak University during my Fulbright tenure

and prior. Without your support I would not be where I am today. In addition, I would like to express gratitude to my PhD colleagues at Walailak who have worked with me, including Sirilak Chumkiew, Premrudee Noonsang, Siriwan Wongkoon, Wittaya Pheera, Uthai Kuhapong, and Puangrat Jinpon.

Thank you to the Taiwan Forestry Research Institute, including Chau Chin Lin, Sheng-Shan Lu, and Yu-Huang Wang for a decade of collaborations on terrestrial sensor networks. Your group has been incredibly patient and supporting of my work from a fledgling undergraduate to a professional researcher.

Thanks to Peter Arzberger for providing my first research opportunities and having faith in my abilities to grow into the person I am today. You initially pushed me to pursue a PhD, and taught me the skills needed to succeed. Thank you to Tony Fountain for advising me at CALIT2 during the transition period from undergraduate to graduate school. You taught me how to conduct research and grow collaborations.

Thanks to Irina Artamonova for supporting me emotionally, financially and intellectually. In addition to a great mom, she is a talented statistician who has been instrumental in my work. Thank you for always being there for me to catch me when I falter.

Finally, and most importantly thank you to my life partner Sherri Lynn Conklin. She has been patient, supporting, encouraging, and absolutely key to all my work, including this dissertation. An accomplished academic herself, and a far better writer, she has spent hundreds of hours reading through drafts of my work and grant proposals. Thank you so much for always being there when I need you most.

This work was funded in part through multiple University of California Santa Barbara PhD scholarships, a United States National Science Foundation (NSF) Graduate Research Fellowship, the U.S. Fulbright program, the U.S. Department of State's Bureau of Democracy, Human Rights and Labor (DRL), and the NSF Smart & Connected Communities award NSF-1831698.

Michael S. Nekrasov

https://michaelnekrasov.com

mikrasov@gmail.com

(858) 754-9676

EDUCATION

Ph.D. Computer Science 2020
University of California, Santa Barbara

M.S. Computer Science 2018
University of California, Santa Barbara

B.S. Computer Science 2011
University of California, San Diego (cum laude)
Minor in Mathematics

SKILLS

Programming:	Python; Java; JavaScript; C; Assembly Languages (SPARC)
Web:	HTML 5; CSS 3; NodeJS; React; Gatsby; JQuery; XML; SQL; JSON; AJAX; Twitter Bootstrap; API Development; REST;
Technologies:	Pandas; NumPy; SciPy; Android Development; Django; TCP/IP Stack; 802.11; 802.15.4; Cloud Computing (AWS); OpenCV;
Tools:	Linux/Windows; PyCharm; IntelliJ; Eclipse; Adobe Photoshop/-Lightroom; LaTeX; Microsoft Office; UAS Pilot (Part 107);
Social:	Teaching; Communicating between disciplines; Multi-national collaborations; Engage with stakeholders; Grant writing
Other:	US Citizen; Fieldwork; Sensor Network Deployment; Photography; PADI Scuba Certified; International Travel (40+ Countries); Fluent in Russian;

HONORS & AWARDS

ACM DroNet Best Paper	2019	UCSD Graduated Cum Laude	2011
Dissertation Year Fellowship	2018	UCSD Returnee of the Year	2010
MobiSys Best Poster	2013	UCSD Robins Scholarship	2010
Fulbright Scholar	2012	Boeing Scholarship	2008
NSF Graduate Research Fellowship	2012	BAE Systems Scholarship	2009
UCSB Doctoral Scholars Fellowship	2012	UCSD I-Center Scholarship	2007

WORK EXPERIENCE

UC Santa Barbara
Research Assistant - MOMENT lab 2014 - 2020

◇ **Flying Ubiquitous Sensor Networks**: Optimizing UAS 802.15.4 for data delivery for wireless sensor networks.

◇ **Cellular Congestion Monitoring**: SDR to detect network congestion via LTE broadcast messages.

◇ **Tribal Internet Access**: Automated aerial network assessment to rapidly plan new network deployments.

◇ **Voxel Based Aproximation**: Novel algorithm for locating devices from Unmanned Aircraft Systems.

◇ **Verifiable Group Anonymity**: Android application using cryptography and proxy to anonymous social media accounts while protecting group identity and message integrity.

Walailak University - Center of Excellence in Ecoinformatics, Thailand
Fulbright Scholar 2013

◇ **Early Warning Flood Detection** : Real-time event detection of flooding in Bandon Bay, Thailand.

UC San Diego
Research Assistant - CALIT2 2009 - 2012

◇ **Coral Sensor Networks**: Middleware for coral reef observatories streaming and sharing data.

◇ **Bee Counting**: Computer vision for automatic counting of bee population changes.

◇ **Coral Spawning**: Automated coral spawning detection using computer vision coupled with coral fluorescence.

Front-End Web Developer - International Center 2008 - 2010

MobileTrac, San Diego
Front-End Web Developer 2009 - 2012

Mikrasov Design (Independent Contractor)
Web Design, Photography, and Computer Consulting 2003 - Present

TEACHING EXPERIENCE

Instructor of Record
UC Santa Barbara - School for Scientific Thought
- ◇ Computer Ethics: Reshaping Society Through Technology — Winter 2003

Teaching Assistant
UC Santa Barbara
- ◇ Introduction to Computer Communication Networks — Fall 2019
- ◇ Translation of Programming Languages — Spring 2018
- ◇ Computer Science Bootcamp — Winter 2018
- ◇ Foundations of Computer Science — Fall 2012

FORMAL TRAINING

Graduate: Computational Geometry; Distributed Computing; Computer Networks; Mobile Computing; Cloud Computing; Mobile Networks; Mobile Imaging; Data Intensive Computing; Java Distributed Computing; Sociology & Biology Networks; Education: Blended Learning Course Design

Undergraduate: Artificial Intelligence; Programming Languages; Assembly; Compilers; Computer Design; Computer Architecture; Operating Systems; Computer Security; Computer Communication and Networks; Basic Data Structures & Object Oriented Design; Advanced Data Structures; Software Engineering; Cognitive Science; Computability & Intractability; Mathematical Reasoning; Advanced Calculus; Advanced Linear Algebra; Graph Theory; Statistics; Combinatorics; Number Theory; Physics (Electromagnetic, Optics, Mechanical, Thermodynamics); Analog Design; Circuits and Systems

PUBLICATIONS

Peeking through Cellular Walled Gardens to Estimate Congestion 2020
Adarsh, V., Nekrasov, M., Belding, E.
Undetermined. (In preparation)

Real-Time Multilateral RSSI Localization from UASs for Disaster Response 2020
Nekrasov M, Belding E.
Undetermined. (In preparation)

The Past 110 Years: Historical Data on the Underrepresentation of Women in Philosophy Journals 2020
Hassoun, N., Conklin, S., Nekrasov, M., West, J.
The Journal of Ethics. (Under Submission)

Impact of 802.15.4 Radio Antenna Orientation on UAS Aerial Data Collection 2020
Nekrasov, M., Maxton, G., Allen, R., Artamonova, I., Belding, E.
Proceedings of the 29th International Conference on Computer, Communication and Networks (ICCCN). (Under Submission)

#Outage: Detecting Power and Communication Outages from Social Networks 2020
Paul, U., Alexander, E., Nekrasov, M., Adrash, V., Belding, E.
Proceedings of the Web Conference (WWW).

Evaluating LTE Coverage and Quality from an Unmanned Aircraft System 2019
Nekrasov M, Adarsh V, Paul U, Showalter E, Zegura E, Vigil-Hayes M, Belding E.
Proceedings of the 16th IEEE International Conference on Mobile Ad-Hoc and Smart Systems (MASS).

Packet-level Congestion Estimation in LTE Networks using Passive Measurements 2019
Adarsh, V., Nekrasov, M., Belding, E.
Proceedings of the Internet Measurement Conference (IMC), pp. 158-164.

Optimizing 802.15. 4 Outdoor IoT Sensor Networks for Aerial Data Collection 2019
Nekrasov, M., Allen, R., Artamonova, I., Belding, E.
MDPI Journal of Sensors, 19(16), p. 3479.

Performance Analysis of Aerial Data Collection from Outdoor IoT Sensor Networks using 2.4GHz 802.15.4 2019
Nekrasov, M., Allen, R., Belding, E.
Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications (DroNet), pp. 33-38.

A User-driven Free Speech Application for Anonymous and Verified Online, Public Group Discourse. 2018
Nekrasov, M., Iland, D., Metzger, M., Parks, L. Belding, E.
Journal of Internet Services and Applications, 9(1).

SecurePost: Verified Group-Anonymity on Social Media 2017
Nekrasov, M., Iland, D., Metzger, M., Zhao, B. Belding, E.
7th USENIX Workshop on Free and Open Communications on the Internet (FOCI).

Limits to Internet Freedoms: Being Heard in an Increasingly Authoritarian World 2017
Nekrasov, M., Parks, L., Belding, E.
Proceedings of the 2017 Workshop on Computing Within Limits (Limits), pp. 119-128.

Sensor Networks Applications for Reefs at Racha Island, Thailand. 2012
Jaroensutasinee, M., Jaroensutasinee, K., Bainbridge, S., Fountain, T., Chumkiew, S., Noonsang, P. Kuhapon, U. Vannarat, S. Poyai, S. Nekrasov, M.
Proceeding of the 12th International Coral Reef Symposium (ICRS), Cairns, Australia, p. 95.

CREON – Integrating Disparate Sources of Remote Coral Reef Sensor Data 2012
Jaroensutasinee, K., Jaroensutasinee, M., Bainbridge, S., Fountain, T., Holbrook, S., Nekrasov, M.
Proceedings of the 12th International Coral Reef Symposium (ICRS), Cairns, Australia, pp. 9-13.

The Open Source DataTurbine Initiative: Empowering the Scientific Community with Streaming Data Middleware. 2012
Fountain, T., Sameer, T., Shin, P., Nekrasov, M.
Bulletin of the Ecological Society of America. 93(3), pp. 242-252.

Coral Sensor Network at Racha Island, Thailand. 2011
Fountain, T., Nekrasov, M., Jaroensutasinee, M., Jaroensutasinee, K., Chumkiew, S., Noonsang, P., Kuhapong, U. Bainbridge, S.
Proceedings of the Environmental Information Management Conference (EIM).

Automatic Analysis of Camera Image Data: An Example of Honey Bee (Apis cerana) Images from the Shanping Wireless Sensor Network. 2011
Lu, S. S., Perry, M., Nekrasov, M. Fountain, T., Arzberger, P., Wang, Y. H., Lin, C. C.
Taiwan Journal of Forest Science, 26(3), pp.305-311.

PROFESSIONAL ACTIVITIES

Technical Program Committees

- ◇ First Annual Android SensorPod Workshop, *TFRI, Taipei, Taiwan* 2013

Invited Talks

- ◇ Remote Lecture: ICT4D, *Brown University* 2016
- ◇ Lecture: Applications of Math and C.S., *Torrey Pines High School, CA* 2016
- ◇ Lecture: Quantitative Methods in Social Sciences, *UC Santa Barbara* 2014
- ◇ TV Interview, *Local news for Nakhon Si Thammarat, Thailand* 2013
- ◇ Sensor Tools Workshop, *LTER Headquarters, Albuquerque, NM* 2012
- ◇ Experimental Learning Conference, *UC San Diego* 2012
- ◇ SensorNIS Workshop, *Hubbard Brook Experimental Forest, NH* 2011

Service

- ◇ Judging SB Hacks (competition), *UC Santa Barbara* 2020
- ◇ Graduate Student Recruitment Committee, *UC Santa Barbara* 2016
- ◇ Student Senate, *UC Santa Barbara* 2015-2016

Abstract

Robust, Resilient Networked Communication in Challenged Environments

by

Michael S. Nekrasov

In challenged environments, digital communication infrastructure may be difficult or even impossible to access. This is especially true in rural and developing regions, as well as in any region during a time of political or environmental crisis. We advance the state of the art in wireless networking and security to design networks and applications that rapidly assess changing networking conditions to restore communication and provide local situational awareness.

This dissertation examines new systems for responding to current and emerging needs for wireless networks. This work looks across the wireless ecosystem of widely deployed standards. We develop new tools to improve network assessment and to provide robust and reliable network communication. By incorporating new technological breakthroughs, such as the wide commercial success of unmanned aircraft systems (UASs), we introduce novel methods and systems for existing wireless standards for these challenged networks.

We assess how existing technologies and standards function in difficult environments: lacking end-end Internet connectivity, experiencing overload or other resource constraints, and operating in three dimensional space. Through this lens, we demonstrate how to optimize networks to serve marginalized communities outside of first world urban cities and make our networks resilient to natural and political crisis that threaten communication.

Table Of Contents

List of Figures

List of Tables

List of Abbreviations

BCCH Broadcast Control CHannel 54, 56

CCCH Common Control CHannel 54–56, 58

CoLT Cell on Light Truck 231, 232

COW Cell on Wheel 231, 232

CROW Cell Repeaters on Wheel 231, 232

DCCH Dedicated Control CHannel 54

DRL U.S. Department of State's Bureau of Democracy, Human Rights and Labor v, 173, 225

DTN delay-tolerant network 2, 8, 18, 20, 30, 67, 73, 112, 233, 236

ECEF earth centered earth fixed 116

eNodeB LTE base station 10, 26, 27, 30, 35, 48, 52–54, 58, 60–66

FAA United States Federal Aviation Administration 38, 71, 97, 98, 109

FCC United States Federal Communications Commission 6, 22, 24, 34, 52, 231

FUSN flying ubiquitous sensor network 5, 8, 18, 21, 30, 67, 68, 92, 112

GLM generalized linear model 72, 76–80

GSM Global System for Mobile Communications 25, 27

ICT4D information and communications technologies for development 155

IMSI international mobile subscriber identity 59, 153, 154

IoT Internet of things 2, 8, 9, 12, 18, 20, 30, 32, 66–68, 73, 76, 88, 90–92, 139, 236

LLS linear least squares 117, 118, 121, 125–128, 132, 135–138

LR-WPAN low-rate wireless personal area network 8, 20, 68

LTE long-term evolution 2, 5–7, 10, 17, 21–27, 29–32, 34, 37, 39–41, 44, 45, 47, 49, 51–54, 56, 58, 60, 62, 66, 139, 230, 233

MIB master information block 54

MLE maximum likelihood estimation 29, 118, 121, 126–128, 132–138

MME mobility management entity 62

MNO mobile network operator 3, 5, 7, 10, 23, 25, 27, 30, 31, 34, 51, 52, 65, 153, 154, 228, 230–232, 236

MNU Mongolia National University 180

MU-MIMO multi-user multiple-input and multiple-output 19, 20

NAS non-access stratum 54

NGO non-governmental organization 146, 148, 149, 181

NSF United States National Science Foundation v, 51

OSN online social network 13, 14, 142–145, 147, 148, 150, 151, 156, 159–163, 175, 176, 180–182, 184, 187, 189, 191–200, 202, 206, 207, 209–218, 223, 224

PHY physical network layer 27

PRR packet reception rate 8, 9, 30, 31, 68, 73, 74, 77, 82, 84–92, 99, 100, 103, 105–110

PWA progressive web app 11

QoE quality of experience 51

QoS quality of service 3, 51, 56

RAB radio access bearers 55

RF radio frequency 23–26, 36, 50

RFID radio-frequency identification 18

RLC radio link control 55, 56

ROC receiver operating characteristic 84, 106

RPi Raspberry Pi 38, 39, 71, 97, 121–123, 137

RRC radio resource control 54–56, 58, 59, 61

RSRP reference signal received power 23, 25, 26, 31

RSRQ reference signal received quality 25, 26

RSSI received signal strength indicator 8, 9, 11, 21, 28–32, 68, 72, 73, 75–82, 86–92, 98, 99, 101–103, 113–115, 117, 122, 125, 128, 129, 139

RTL-SDR RTL2832U chipset software defined radio 24, 35, 37–40, 42, 48–51

SA spectrum analyzer 35, 37, 42, 48

SDR software defined radio 23–25, 27, 29, 31, 34, 36, 37, 51, 52, 54, 55, 57, 233, 237

SI system information 54, 55

SIB system information block 54–56, 61, 64, 65

SIM subscriber identity module 153

SINR signal to interference plus noise ratio 25, 26

SNR signal to noise ratio 25

SRB signalling radio bearer 54, 55

TDOA time difference of arrival 29

TOA time of arrival 29

TVWS TV white space 18, 25, 32, 230, 232, 233

UAS unmanned aircraft system xii, xvi, 2, 3, 5, 6, 8–12, 15, 18–21, 23–26, 28–32, 34–42, 44, 47–51, 66–71, 73, 75, 76, 80–82, 88–91, 93, 94, 97–102, 107–110, 112–119, 121–126, 128–132, 137–139, 232, 233, 236, 237

UE cellular user equipment xvi, 7, 10, 12, 19, 24–28, 31, 34–36, 38–44, 46, 48, 49, 53–59, 61, 63, 64, 66, 111, 113

UN United Nations 142, 171, 228

USRP universal software radio peripheral 35, 37, 40, 42, 48, 57

VBA voxel-based approximation xviii, 11, 12, 30–32, 113–115, 118, 119, 123, 125, 126, 128–130, 132–139

VPN virtual private network 153, 154, 157–159, 193

WCDMA Wideband Code Division Multiple Access 56

WLS weighted least squares 117, 121

WSN wireless sensor network 2, 7–9, 12, 18, 20, 21, 28, 30, 66, 67, 92, 114, 139, 236

ZINB zero-inflated negative binomial regression 8, 9, 31, 74, 77, 84–87, 93, 100, 105, 106

ZWD Zambia watchdog 194

Chapter 1

Introduction

As of 2019, roughly 46% of the world's population lacks Internet access [1]. The division between those with the ability to access the Internet and those without is frequently referred to as the "Digital Divide" [2, 3, 4]. Those without Internet access predominately live in rural regions, especially in the developing world. The division in access is not a distinct boundary between groups. While some go without Internet access due to lack of infrastructure, others lack the financial means to pay for connectivity [5]. In other cases, those with access may have service so poor that they find it difficult to use the Internet for anything of value to their lives [6]. Two decades after the formulation of the digital divide as a key social challenge for computer networks, the growth in Internet penetration is stagnating [7, 8] and existing tools to assess network access are insufficient.

Wireless networks, especially in rural regions, are frequently used for "Last Mile Network Access" [9, 10]. Wireless deployments are typically cheaper to deploy and maintain then cable counterparts particularly over long distances and in difficult terrain. In these cases fixed wired infrastructure is brought to the edge of a community, and wireless technologies are used to disseminate access across individual households. In some cases even terminating portions of this backhaul use point to point wireless links [11, 12].

Cellular broadband technologies, such as Long-term evolution (LTE) as well as older generation 3G and 2G networks, are frequently used to connect the edges of access networks to the Internet. The assessment of such networks, both in terms of coverage and performance, remains a significant challenge. Governments and corporations frequently use generous propagation models to assess coverage, often overstating availability and not accounting for quality and usability [8, 13, 14].

Inaccuracies in network assessment exist beyond public facing networks. In environmental and agricultural domains, network connected Internet of things (IoT) sensors are increasingly deployed to enable precise measurement and modeling of microclimates [15, 16, 17, 18]. The data from these sensors contribute to improving land use and boosting crop yields, as well as preparing for and responding to natural disasters. As in the case of public Internet use, wireless sensor networks (WSNs) deployed in rural regions frequently encounter a deficit of backhaul connectivity. When end-end Internet connectivity becomes impossible or impractical, Delay-tolerant network (DTN) approaches can provide access [19]. One promising proposal deploys UASs to access remote WSNs, collect data, and deliver that data back to Internet connected gateways [20, 21, 22]. As of 2020, adaptations of this approach are in their infancy. While simulations of wireless networks show capacity for DTNs of sensors utilizing UASs for wide scale data delivery [23], as we explore in chapters 5 and 6, the reality is harsher than simulations project. As in the case of cellular networks, real-world networks encounter a variety of constraints. To turn theory into reality, new network assessment methods are needed.

Even if Internet connectivity is typically available in a region, networks sometimes encounter disruptions. Natural disasters can damage otherwise well provisioned networks when most needed. In addition to directly threatening human lives, natural disasters, such as hurricanes, cripple communication infrastructure. During the 2017 Atlantic hurricane season, three successive hurricanes inflicted significant communication disruptions, despite

prior warning and preparations. After Hurricane Harvey, 4.1% of cell sites across the coasts of Texas and Louisiana were inoperable [24]. Hurricane Irma disrupted 27% of towers in affected areas of Florida [25] and downed 55% of Puerto Rico's cellular towers [26]. Hurricane Maria disabled 95% of cellular sites in Puerto Rico and 76% in the Virgin Islands [27]. As a result, affected citizens were unable to request help or rescue in the face of rising floodwaters.

The impact of these disasters sometimes last months or years. Infrastructure damage to power and communication networks in Puerto Rico took more than 11 months to restore [28] and caused a temporary migration of half a million residents [29]. Unfortunately, such disasters seem to be increasing in frequency and intensity [30, 31].

As evidenced in the case of the 2017 hurricanes, natural disasters drastically change network topography. Previously well provisioned networks become fragmented or overloaded and quality of service (QoS) diminishes. Rapid assessment of the changing network conditions is essential to disaster response and recovery. Mobile network operators (MNOs) already use emerging technologies, such as UASs, to assess physical infrastructure damage after a disaster. The work in this dissertation further expands capabilities, by proposing systems of automated network quality assessment and systems for locating disconnected wireless devices.

Natural disasters are not the only threat to access. For the last decade, civil freedoms have declined across the world [32], especially journalistic freedoms, freedom of speech and expression, and access to information. This trend likewise applies to Internet freedoms. The last decade experienced a sharp rise in censorship, Internet surveillance, targeted disinformation, and cyberbullying [33]. Political crises, such as the refugee dispute in Bangladesh [34] or the dispute in Kashmir [35], entirely disabled local Internet use for days or months. In other cases, governments limit Internet access for particular types of content [36, 37, 38, 39, 40]. Policies and threats also lead to self censorship of content,

as we cover in chapter 9. As in the case of environmental crises, political crises deprive users of access to critical information and affect their ability to communicate.

1.1 Thesis Statement

This dissertation shows that:

> In challenged environments, digital communication infrastructure may be difficult or even impossible to access. This is especially true in rural and developing regions, as well as in any region during a time of political or environmental crisis. We advance the state of the art in wireless networking and security to design networks and applications that rapidly assess changing networking conditions to restore communication and provide local situational awareness.

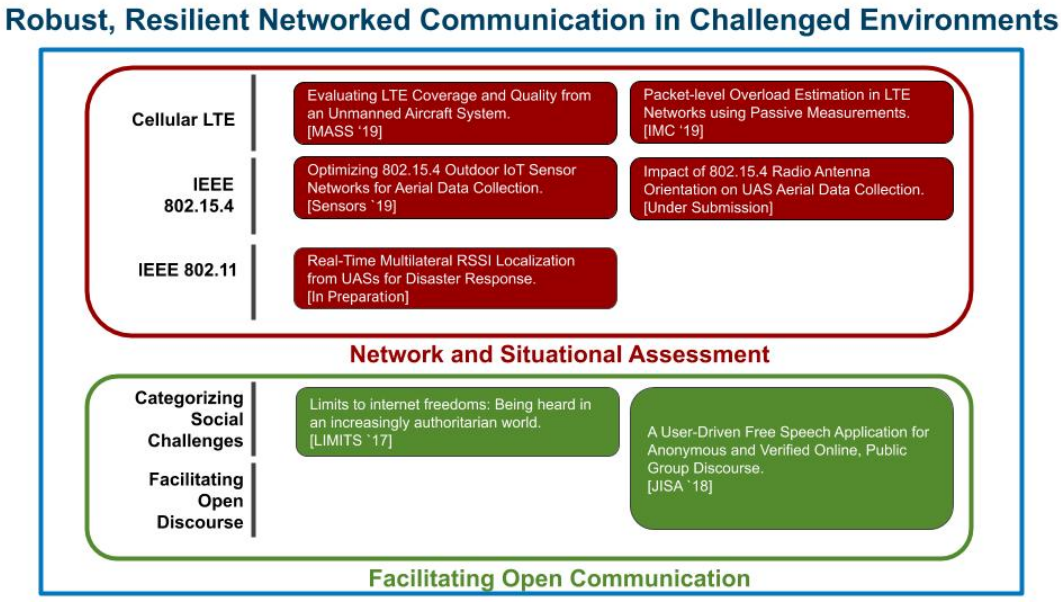


Figure 1.1: Thesis Outline.

1.2 Dissertation Outline and Contributions

The structure of this dissertation is summarized in Figure 1.1 and is organized as follows. The dissertation is divided into two parts. The first section explores new strategies for rapid network planning and providing situational awareness. First, we develop new methods for evaluating LTE cellular connectivity in rural areas and overloaded conditions. Second, we assess IEEE 802.15.4 flying ubiquitous sensor networks (FUSNs) in order to facilitate the use of aerial data collection practices for difficult to access networks, such as after a disaster. Third, we present a novel wireless localization algorithm for UASs using IEEE 802.11. The second part of the dissertation investigates challenges in online communication during a political crisis. First, we categorize social challenges to free speech. Then we present a novel tool that resolves some of these challenges.

1.2.1 Network and Situational Assessment

The first part of the dissertation deals with assessment of wireless networks and provision of situational awareness. We investigate multiple wireless standards, including LTE, IEEE 802.15.4, and IEEE 802.11 (WiFi). This section introduces new systems and algorithms for evaluating network quality, detecting overloading, and locating devices. Our work extends the abilities of network planners and evaluative bodies to assess challenged networks, including rural networks and networks before and after a disaster.

Cellular LTE

Chapters 3 and 4 investigate passive methods of evaluating LTE cellular networks without MNO co-operation. As discussed earlier in this chapter, accurate assessment of cellular networks remains a critical challenge to bridging the digital divide. MNOs lack incentive to construct and share accurate coverage maps or load levels. This work

provides outside parties, such as affected communities and network researchers, the tools to passively evaluate network performance.

Evaluating LTE Coverage and Quality from an UAS: Chapter 3 presents a novel aerial system for evaluating LTE coverage and quality from an UAS. This work focuses on rapid evaluation of large rural areas. Despite widespread LTE adoption and dependence, rural areas lag behind in availability and quality of coverage. While the United States Federal Communications Commission (FCC), which regulates mobile broadband in the United States, reports increases in LTE availability, the most recent FCC Broadband Report was criticized for overstating coverage [41, 42]. Physical assessments of cellular availability and quality are essential for evaluate and predict actual user experiences. However, measurement campaigns can be resource, time, and labor intensive; more scalable measurement strategies are urgently needed.

In this work, we first present several measurement solutions to capture LTE signal strength measurements, and we compare their accuracy. Our findings reveal that simple, lightweight spectrum sensing devices have comparable accuracy to expensive solutions and can estimate quality within one gradation of accuracy when compared to user equipment. We then show that these devices can be mounted on UAS to more rapidly and easily measure coverage across wider geographic regions. Our results show that the low-cost aerial measurement techniques have 72% accuracy relative to the ground readings of user equipment, and fall within one quality gradation 98% of the time. This work was published in the proceedings of the 16th IEEE International Conference on Mobile Ad-Hoc and Smart Systems (MASS) [43].

Packet-level Overload Estimation in LTE Networks: Chapter 4 examines existing cellular networks under load and presents a novel method of packet-level overload estimation in LTE networks using passive measurements. Over 87% of US mobile wire-

less subscriptions are currently held by LTE-capable devices [44]. However, prior work demonstrates that connectivity may not equate to usable service. Even in well-provisioned urban networks, unusually high usage (such as during public events or during a political or environmental crisis) can lead to overload that makes the LTE service difficult, if not impossible to use, even if the user is solidly within the coverage area. A typical approach to detect and quantify overload on LTE networks secures the cooperation of the MNO for access to internal metrics. An alternative approach deploys multiple mobile devices with active subscriptions to each MNO.

As both approaches are resource and time intensive, in this work, we propose a novel method to estimate overload in LTE networks using only passive measurements and without requiring MNO cooperation. We use this method to analyze packet-level traces for three commercial LTE MNOs, T-Mobile, Verizon and AT&T, from several locations during both typical levels of usage and during public events yielding large, dense crowds. This study presents the first look at overload estimation through the analysis of unencrypted broadcast messages. We show that an upsurge in broadcast reject and cell barring messages can accurately detect an increase in network overload. This work was published in the proceedings of the 2019 Internet Measurement Conference (IMC) [45].

IEEE 802.15.4

The need for Internet access extends beyond UEs. As discussed earlier in this chapter, WSNs are increasingly used in environmental monitoring and precision agriculture to monitor and model environmental changes. These models can be critical in predicting, assessing and recovering from disasters. The potential of these systems stems from the aggregation and analysis of multitudes of sensors. When missing an Internet backhaul, as is often the case in a rural area or after a disaster, these systems need an alternative method of access.

The IEEE 802.15.4 standard, once envisioned for low-rate wireless personal area networks (LR-WPANs), is the basis of a number of common standards, such as ZigBee and WirelessHART and SNAP, for home and outdoor IoT WSNs [46]. These standards take advantage of 802.15.4's low power usage and flexible topology. Consumer 802.15.4 radios are widely available in compact form factors, making them ideal for application in environmental and agricultural sensor networks. As we cover in chapters 2, 5 and 6, UASs show promise in FUSNs employing DTN approaches for aerial content delivery. Unfortunately, as we will see in chapter 2, three dimensional communication introduces new challenges, such as high mobility, non-uniform signal radiation patterns and extreme ranges. While ample ground based measurement studies of 802.15.4 exist, rigorous evaluation of 802.15.4 in three dimensional aerial communication remains inadequate.

Optimizing 802.15.4 Outdoor Sensor Networks for Aerial Data Collection: Chapter 5 investigates 802.15.4 network optimization for aerial data collection. Our work is, to our knowledge, the first performance measurement study of 802.15.4 ground-air data collection performance from an UAS. We analyze experimental measurements from an outdoor aerial testbed, examining how factors, such as antenna orientation, altitude, antenna placement, and obstruction affect signal strength and packet reception rate. This work uses Digi WRL-15126 XBee3 transceiver using PCB antennae. These radios are popular due to their compact form factor.

In our analysis, we model and predict the quality of services for aerial data collection, based on these network configuration variables and contrast that with the received signal strength indicator (RSSI) - a commonly used signal strength metric. We find that network configuration plays a significant role in network quality, which RSSI, a mediator variable, struggles to account for in the presence of high packet loss. Instead we propose that packet reception rate (PRR) is a more appropriate network metric and that Zero-inflated

negative binomial regression (ZINB) effectively models PRR. This work provides a first look at optimizing outdoor 802.15.4 networks for aerial data collection. A portion of this work was published in the proceedings of the 5th ACM Workshop on Micro Aerial Vehicle Networks, Systems, and Applications (DroNet) [47] and the full work in the MDPI Journal Sensors [48].

Impact of 802.15.4 Radio Antenna Orientation on Aerial Data Collection: While the work in chapter 5 investigated antenna orientation as a potential variable for RSSI and PRR, we found that the coiled PCB antenna dampens effects, such as toroidal radiation and signal polarization, exhibited by straight wire dipole antennas. In chapter 6, we expand the work of chapter 5 and provide the first study of the difference in performance between straight wire and PCB antenna for 802.15.4 outdoor wireless networks.

We compare external antenna configurations to the commonly used embedded coiled antenna modules. We investigate the impact of antenna type and orientation on PRR and RSSI, again modeling our data using ZINB. For each hardware configuration and orientation, we identify the optimal altitude to fly a UAS. Combined, these two works serve as a comprehensive first look at how hypothetical aerial data collection strategies for 802.15.4 networks would fare in real-world conditions. Our results show that choosing antenna configuration (including type and orientation) for an IoT network comes with trade-offs and an informed choice must consider the intended UAS collection flight plan.

IEEE 802.11

In addition to network quality, the physical locations of devices can be critical for providing situational awareness and network planning. For an outdoor WSN for example, the location of sensors provides valuable input for domain modeling and network

performance. A further (and particularly critical) case of device localization arises during a natural disasters.

During a disaster, first responders often locate survivors as a part of search and rescue operations [49, 50]. Normally UEs, such as smartphones, self report their location via GPS or other self-localization schemes. However, when disasters disrupt communication infrastructure, UEs are no longer connected to a network. To restore connectivity, some propose using LTE equipped UASs [51] to request location. While technologically feasible, policies and user behaviors make this option difficult because LTE frequencies are licensed and tightly regulated. Even in emergencies, actors other than licensed MNOs cannot legally broadcast on these frequencies. Additionally, the physical infrastructure, LTE base stations (eNodeBs), are expensive and power hungry and mounting them on UASs results in brief flight times. The IEEE 802.11 (WiFi) standard provides a lower cost, lower power alternative, but at much shorter ranges.

Another challenge to localizing devices lies in the process of querying the device for the location. There is no universal system for querying devices for their location. Some solutions exist, such as the Facebook disaster check-in program [52]. However, these solutions are limited in scope and require active user participation. In the event that partial connectivity is restored, without returning end-to-end network access, such systems would not function.

An alternative solution involves captive portals, that force user devices to a web page when connectivity is first established. Captive portals have already been successfully used as communication platforms during emergencies [53, 54]. However, traditional captive portals assume continuous connectivity to the wireless AP, which maintains state for the duration of the session. This assumption does not hold for a captive portal hosted by a moving drone, particularly in a disaster scenario with impacted connectivity. Instead, we propose a captive portal that loads on a user device when connectivity with the

drone is established, and which does not require Internet connectivity or connectivity to a separate AP. To this end we developed a progressive web app (PWA), EmerGence, designed to leverage opportunistic Internet connectivity to relay messages from individuals trapped in disaster-hit areas. EmerGence works by bypassing the crippled communication infrastructure to connect disaster victims with needed resources (both human and material) through a combination of a new web development concept called the PWA and served by an UAS. This work was presented at the International Workshop on Mobile Computing Systems and Applications (HotMobile) [55]. This work in turn led to coauthored work on information retrieval from traditional sources, such as social media which is published in the proceedings of The Web Conference (WWW) [56].

Real-Time Multilateral RSSI Localization from UASs for Disaster Response:

In a disaster, users may be unwilling or unable to engage with a system actively. In chapter 7, we propose a system for passive real-time multilateral RSSI localization from UASs using 802.11. We explore how an 802.11-equipped drone can be used to locate an uncooperative device in three dimensional space, based on the RSSI. Even when not connected to an access point, 802.11 devices broadcast packets called *probe requests* that assist in connectivity [57]. These requests can be passively captured and used for localizing devices.

We provide a novel algorithm using VBA for real-time UAS based multilateral localization. Using an aerial testbed, we perform extensive evaluations and compare our solution with other standard approaches. We demonstrate that, in contrast to other approaches, VBA uses fixed computation time and memory as the number of data points increases, while still providing comparable error. Based on our results, we discuss how a continuously updated geographic representation can be used for flight planning and rescue operations post-natural disaster.

In addition to UEs, this algorithm could be applied to IoT devices and WSNs. IoT sensors frequently lack GPS or other localization technologies and tracking their location post deployment can be troublesome [58, 59]. VBA may provide a low cost method of rapidly mapping deployed IoT sensors from a UAS.

1.2.2 Facilitating Open Communication

As described earlier in this chapter, the digital divide goes beyond questions about infrastructure. Even if infrastructure is present, capable of satisfying performance needs, and undamaged by natural disaster, political crises can make the Internet unusable for large demographics. Part II of this dissertation examines the issues surrounding access availability that limits usability. As we will see in this section, while the Internet is a critical tool for communication and knowledge acquisition in societies across the globe, its use has become a battlefield for governments, corporations, and individuals to censor speech and access to information. In order to understand the obstacles to free speech on-line, we partner with social science and film and media studies to categorize the social challenges through user surveys of vulnerable communities and ethnographic interviews. Based on this research we build and evaluate a novel tool for facilitating discourse.

Categorizing Social Challenges

When we began research into facilitating open communication on-line, we eschewed preconceived notions of technological solutions we thought would be useful to target communities. Instead, we conducted interviews on the needs of these communities and determined that our initial concepts did not align with the actual needs of vulnerable populations. Therefore, before developing any tools, we first categorize the unmet social challenges. To conduct these interviews, we sought out a diverse set of participants

with independent and sometimes conflicting agendas. To understand the barriers they encounter, we investigated the competing motivations, the adversaries, and the tools those adversaries use to silence speech and block access.

Limits to Internet Freedoms: Chapter 9 presents research into the use of social media for free speech based on ethnographic research conducted by our team in Lusaka, Zambia; Ulaanbaatar, Mongolia; and Istanbul, Turkey from 2014-2016. We reached out to diverse sets of communities to investigate Internet freedoms and, in particular, their relation to online social networks (OSNs). We use this research as the basis of discussion into the limits, actors, and concerns in this space. We formally interviewed 110 people and had informal conversations with dozens more individuals. While our work provides only a small window into the broad set of limits that individuals encounter in on-line access and speech, the diverse perspectives, cultures, and struggles serve as a platform for understanding the limits to Internet freedoms in a global context. We conclude with a discussion of how design and development choices for technology can affect marginalized communities, as well as the ethical and technical considerations for developing tools and applications that support Internet freedoms.

This work is a novel partnership between social and computing disciplines. The work came at a particularly important time when Turkey experienced a steep change in freedoms, including a massive protest in Gezi Park [60], an attempted coup [61], and subsequent purging of academics [62]. This work included interviews with marginalized groups and individuals directly targeted based on Internet activities. This work was published in the proceedings of the 2017 Workshop on Computing Within Limits (LIMITS) [63].

Facilitating Discourse

As can perhaps be expected, from our research in chapter 9, we found users often turn to technological solutions to combat this threat to civil liberties. For example, popular anonymity tools such as Tor [64] provide network level anonymity, while person to person messaging tools such as Signal [65] and WhatsApp [66] allow private communication using end-end encryption. Nevertheless, safe public communication remains a challenge. Individuals conversing publicly on OSNs open themselves up to legal and physical dangers, encouraging self censorship and stifling discourse. Reputation and trust are likewise eroded. In an era of "fake news", users struggle to identify what OSN accounts and posted content can be trusted [67, 68]. Adversaries to open discourse deploy armies of operatives masquerading as legitimate users to sow division [69, 70]. Even trusted news outlets using OSNs can be hacked to spread mis-information [71, 72, 73]. From this work we discovered that public group discourse while simultaneously maintaining anonymity and preserving reputation on OSNs was a critical unmet need of Internet users.

A User-Driven Application for Anonymous and Verified Online, Public Group Discourse: Chapter 10 is a culmination of our efforts in freedom of speech on-line. In this work, we explore the design process of SecurePost, a novel tool allowing verified group anonymity to those communicating publicly on OSNs. We ground our work in the survey-based research and ethnographic interviews of communities vulnerable to censorship. We explore needs and requirements of users, such as modes of censorship, resistance to network disruption, and appropriate platform consideration. We outline our technological solution and expand on how on-the-ground research of user communities guides technological requirements. While our work focuses on specific communities, not populations of countries as a whole, we believe examining the needs of communities particularly vulnerable to censorship provides a lens through which to understand some

challenges to overcoming censorship more broadly. Portions of this work was published in the proceedings of the 7th USENIX Workshop on Free and Open Communications on the Internet (FOCI) [74] and the full work was published in the Journal of Internet Services and Applications [75].

1.3 Discussion

Computer science research strives to design technology that tackles existing social challenges and extrapolates into the future to develop solutions for emerging problems. This dissertation investigates critical contemporary challenges for the use of wireless networks. Rise in inequality with bidirectional links to the digital divide, natural disasters fueled by anthropocentric climate change, and political crises - from human rights abuses, war, government corruption, and religious conflict - are among the core problems of the next century [76, 77].

This dissertation examines new systems for responding to current and emerging needs for wireless networks. This work looks across the wireless ecosystem of widely deployed standards. We develop new tools to improve network assessment and to provide robust and reliable network communication. By incorporating new technological breakthroughs, such as the wide commercial success of UASs, we introduce novel methods and systems for existing wireless standards for these challenged networks.

We assess how existing technologies and standards function in difficult environments, such as those lacking end-end Internet connectivity, overloaded, resource constrained, and operating in three dimensional space. Through this lens, we can understand how to optimize networks to serve marginalized communities, outside of first world urban cities, and make our networks resilient to natural and political crisis that threaten communication.

Part I

Network and Situational Assessment

Chapter 2

Wireless Networks

Part I of this dissertation deals with the assessment of wireless networks and provision of situational awareness in order to address concerns raised in chapter 1. We investigate multiple wireless standards, including LTE, IEEE 802.15.4, and IEEE 802.11 (WiFi). We introduce new systems and algorithms for evaluating network quality, detecting overload, and locating devices. Our work extends the abilities of network planners and evaluative bodies to assess challenged networks, including rural networks and networks before and after a disaster.

This chapter proceeds as follows. First, we provide an overview of emerging trends in network assessment, including outstanding research problems. Second, we provide an outline of our work in this field. Third, we conclude by highlighting key contributions to the academic community and the broader impacts of our work.

2.1 Background

2.1.1 Unmanned Aircraft Systems (UASs)

Due to the recent commercial success of multi-copters, UASs, with capacities previously unavailable in fixed wing systems, are now widely accessible. Multi-copters maneuver in 3 dimensions, hover in-place, and effortlessly take-off and land. New applications include fields like precision agriculture [78, 79, 80, 15, 81], automated ground surveying [82], structural assessment [83, 84, 85], environmental research [86, 87, 88, 89], and disaster management [90, 91, 92, 93, 94]. Many applications rely on high-bandwidth low latency network traffic, such as white light and infrared imagery.

In addition to communicating with users, UASs can communicate with each other or bridge existing networks by supplementing cellular connectivity [95, 96] or interfacing with one another to provide wider coverage [97, 98]. UASs can integrate into FUSNs, providing DTN delivery to WSNs of IoT and radio-frequency identification (RFID) devices [99, 23, 100, 101]. New wireless protocols, like TV white space (TVWS), allow for novel UAS communication strategies, as well as UAS based spectrum mapping and policing [102]. When mobility is unnecessary, UASs can perch for rapidly deployable infrastructure [103, 104, 105]. In addition to data transfer, UASs have wireless capabilities suitable for locating the UAS when GPS is not present (like in an indoor environment) or for locating other devices [106, 57].

When networks fail to provide end-end Internet connectivity, either through lack of infrastructure or due to damage, UASs can service disconnected networks. UASs can be deployed as aerial network relay nodes [107, 21, 20, 108] or as data mules [22, 109]. In the event of a disruption to an existing network, such as damage to relay towers during a storm, UASs can mend network fragmentation [110, 111, 112], acting as temporary relays or DTN nodes. UASs can likewise supplement existing communication infrastructure, in

vehicular networks [113] for example, as well as in rural applications for environmental monitoring [114, 115] and precision agriculture [116, 117]. To adequately utilize UASs for these types of applications, network planners must understand how best to optimize connectivity so that data is successfully delivered.

2.1.2 IEEE 802.11 (WiFi)

The IEEE 802.11 standard, also known as WiFi, is, perhaps, one of the most pervasive communication technologies in use today, with over thirteen billion estimated WiFi enabled devices in use [118]. WiFi enabled devices range from stationary computers, mobile UEs, including smartphones, to new applications, such as UASs. As 802.11 was originaly designed with two dimensional ground-based communication in mind, consumer wireless network hardware and the 802.11 family of standards are not optimized for highly mobile 3D settings. This can pose problems for the use of these protocols on UASs.

Consumer devices typically employ dipole antennas that radiate in a toroidal pattern (doughnut-like shape perpendicular to the antenna). This signal is uniformly emitted at the same orientation, resulting in polarization [119]. This creates a low signal zone above the antenna tip. On the ground, particularly in indoor spaces, obstruction from walls and ceilings cause signal reflection. This bounce leads to multi-path signal propagation where portions of the signal reflect off surfaces and arrive at the receiver at different orientations [120]. While introducing challenges, such as multi-path fading, this effect reduces the relative significance of polarization and toroidal radiation for indoor networks. Recent 802.11 standards (*ac* and *ax*) use multi-path signal propagation to improve bandwidth via the use of multi-user multiple-input and multiple-output (MU-MIMO) [121].

In contrast, outdoor three dimensional networks typically operate with comparatively little obstruction. In 3D space, the toroidal radiation and signal polarization of dipole

antennas used in consumer electronics negatively impact reception quality for aerial networks [122, 123, 124, 125]. Additionally, when used with UASs, the rate adaptation algorithms of 802.11 do not cope with unusually high mobility [126, 127]. Similarly, technologies that boost throughput via spacial diversity, such as MU-MIMO, do not translate well in outdoor spaces with few sources of reflection [128]. As a result, measurements collected from aerial devices are not directly comparable to terrestrial measurements [129] and additional factors, including UAS altitude, affect performance [130].

2.1.3 IEEE 802.15.4

IoT devices and sensors proliferate across a wide domain of applications, ranging from home automation to environmental monitoring. As of 2018, the number of active IoT devices exceeded seven billion, with the largest growth expected for 802.15.4 LR-WPAN devices [131]. The 802.15.4 standard, once envisioned for wearable networks, is implemented by a number of common standards, such as ZigBee and WirelessHART and SNAP, and has become a common communication standard for IoT devices [46]. Standards designed for complex computing devices do not fit the power and network topology for many IoT applications, including WSNs. For example, the 802.11 standard consumes excessive power and is ill-suited for periodic, low bandwidth communication style typical of WSNs. In contrast, 802.15.4 is optimized for low power, low data rate node-to-node connectivity inside a local sensor network [132]. However, remote WSNs still require an Internet gateway for broader Internet access. In areas without Internet access, due to lack of infrastructure or due to infrastructure damaged in events, such as natural disasters, alternate Internet access technologies are necessary.

UASs are a promising technology for DTN data collection from outdoor WSNs [133]. Environmental and agricultural applications sometimes lack access to existing Internet

backhauls for data delivery because low population densities in rural areas provide little economic incentive for cellular providers to serve these remote areas. As these types of applications may be in difficult or inaccessible terrain spanning large geographic areas, manual "sneaker-net" data collection can be dangerous and labor intensive [134].

Work in 802.15.4 largely focuses on two-dimensional topography common to terrestrial networks. For example, [135] examined 802.15.4 signal propagation, while [136] examined person-to-person communication over 802.15.4 and found mobility challenging in an 802.15.4 network. Additionally, [137] examined enhancements to two dimensional movement and stationary 802.15.4 roadside sensors. [138] evaluated 802.15.4 indoor and outdoor performance in terms of error rate and RSSI. While measurement of 802.15.4 performance in three-dimensional space is limited, from the similarity of the physical layer between 802.15.4 and 802.11 radios, we reasonably expect 802.15.4 to evince the same affects of toroidal radiation of the omni-directional dipole antennae of IoT devices in a three-dimensional environment.

Experimental research on 802.15.4 for UASs is sparse. [139] found that 802.15.4 devices are sensitive to antenna orientation; however, their measurements were collected on sensors within 3m of each other and therefore do not scale well to FUSN scenarios measuring distances in the hundreds of meters [133]. On the other hand, [99] investigates the performance of FUSN in remote locations. In addition, [23] provides a simulation of a WSN with hundreds of nodes, but real-world topography and obstruction was not considered in their methods.

2.1.4 Cellular LTE Networks

With 3 billion users and growing, LTE is the leading mobile network technology worldwide [140]. However, economic incentives often drive LTE and other broadband

technology expansion, concentrating deployment in populated urban areas. Economically marginalized and sparsely populated rural areas remain underserved [141]. In the United States, for example, rural tribal regions suffer from the poorest LTE coverage [142]. Even when cellular providers claim coverage, poor signal quality can limit achievable download data rates far below the mobile broadband threshold, defined by the FCC as a median speed of 10 Mbps [142].

For underserved regions in the United States, the FCC has instituted incentive programs to offset provider infrastructure deployment costs [143, 144]. These programs first determine the bounds of existing coverage and identify coverage deficiencies by semi-annually collecting network connectivity reports from commercial network operators. Every operator that owns cellular network facilities in the United States participates in data collection by submitting a Form 477 [145]. The reported coverage area consists of geo-polygons using operator-defined methodology. Based on this data, the FCC allocates subsidies to incentivize commercial coverage in underserved regions and verifies compliance.

The FCC publicly releases annual Broadband Deployment Reports (e.g. [142]), as well as shapefiles for each operator indicating geographic coverage areas [145]. However, researchers challenge their accuracy [146, 14, 147]; for example, [148, 149] examined a public dataset of speed tests collected by the Measurement Lab and demonstrated that broadband access in Pennsylvania is much lower than claimed in the report. This inaccurate over-reporting can be attributed to the proprietary and often generous propagation models used by network operators [13]. To validate the true state of mobile broadband access, we need publicly controlled methods for measuring coverage areas and signal quality, particularly in typically underserved regions.

To audit provider-reported coverage claims, third parties undertake independent measurement efforts. While the concept of "coverage" remains imprecise [150], network

parameters, such as reference signal received power (RSRP), are typically used to estimate the extent of network availability. Popular public crowdsourcing platforms, such as CellMapper [151] and OpenSignal [152], collect measurements from network users to calculate cellular coverage and signal quality. Data from crowdsourced efforts provide information over time and for a wide range of devices, but these data cluster around major transportation arteries, potentially omitting communities outside of these areas. An alternative collection strategy employs specialized equipment with dedicated users. For example, wardriving typically involves physically navigating difficult terrain in remote areas to record on-the-ground measurements [153]. This method enables greater control over the measurement process and the geographic scope but scales poorly due to considerable time investment and labor costs.

Because existing strategies suffer from the aforementioned drawbacks, alternative solutions are needed for evaluating LTE coverage and signal quality. An ideal strategy should enable quick assessment using measurements of the radio frequency (RF) landscape throughout large areas (on the order of square miles), even if hard-to-access. In addition, new strategies should provide scalability with respect to equipment and human resources. Based on these criteria, off-the-shelf software defined radios (SDRs) offer viable solutions. However, because SDRs can cost anywhere between tens of dollars to a few thousand dollars, it is important to study the relationship between precision of readings and the cost of the equipment, to determine whether more affordable equipment will suffice.

To reduce the human effort of LTE coverage analysis, UASs carry payloads while maintaining appreciable flight times. The availability of low cost, programmable, highly agile UASs has spurred interest in employing aerial RF sensing for cellular coverage mapping [154, 155]. UASs enable coverage for large geographic areas, which may be costly, difficult, or impossible to cover on foot or in land vehicles. MNOs already employ UASs for visually inspecting equipment after natural disasters [156]. Extending UAS

capabilities to include signal measurements is an active area of interest for a variety of wireless applications [47, 48]. These extensions could further enable uses for scalable rural cellular coverage mapping as well as post-disaster recovery efforts. Compact SDRs with high sensitivity, such as the RTL2832U chipset software defined radio (RTL-SDR), prove increasingly useful for LTE applications [157, 158, 159]. However, the high altitude of UAS flight, relative to the ground, poses challenges to the efficacy of these approaches. As antennas on LTE towers are provisioned for ground transmission, the RF radiation pattern picked up at high altitudes may not reflect signal quality on the ground [160].

Even in well provisioned regions, abnormal events overload LTE. Sudden escalation in traffic demands from UEs occur during large gatherings (e.g., street festivals, protests). Similarly, after a disaster, damaged infrastructure and atypical volume of utilization overwhelm previously well-provisioned network. Prior work demonstrates that, even in areas cellular providers claim are well-covered, persistent over-usage due to insufficient capacity can exist [161]. For example, in 2017, Hurricane Maria disabled 95% of cellular sites in Puerto Rico [27]. As a result, affected citizens were unable to request help or rescue in the face of rising floodwaters. In such disaster scenarios, even when cellular towers remain functional, the call volume overloads capacity, causing base stations to reject calls [162, 163, 164]. Those who do get through to emergency call centers likely experience long wait times to speak with a 911 operator [165]. Unfortunately, cellular providers have incentive to state that damaged cellular services have been returned to an operational state. Indeed, after Hurricane Maria, statuspr.org soon reported that over 90% of cell towers were again operational; however, anecdotal evidence indicated such statistics were grossly over-stated.

To remedy this disparity between reported coverage and actual usability, individual users, watchdog groups, and government agencies need tools to verify whether a network is adequately serving customers. After a disaster, the FCC typically receives outage

reports from, for instance, telecoms [166], but the actual usability, due in part to overload, on active towers is difficult to assess without access to the internal network. Ideally, public entities should be able to assess the overload and operational status/usability for a particular base station. Further, they should be able to accomplish this without relying on the cooperation of a MNO.

RF Spectrum Sensing: We can use RF spectrum sensing to passively estimate the coverage of LTE networks using SDRs. [167], for example, showed how low cost SDRs can rapidly evaluate occupancy on TVWS frequencies. Previous studies involving wide-scale cellular sensing include analysis of Global System for Mobile Communications (GSM) pollution [168] and propagation model verification for LTE signals [169].

Considerable prior work focuses on identifying the application of UASs for cellular networks. [154] studies the variation in LTE signal strength and signal to interference plus noise ratio (SINR) for both an underserved rural area and an urban center. A UAS connected to an existing LTE network monitored the LTE signals in a range of different altitudes [155]. The authors found that at 60m to 100m above the ground, LTE coverage probability climbs to 90% and the received power gains $\sim$18 dB with respect to the ground level. This work, however, did not compare measurements taken from the aerial platform to measurements taken from a ground-level UE.

[160] sheds light on the applicability and performance of mobile network connectivity to low altitude UASs by analyzing downlink channel indicators, such as RSRP. [170] examined the variation of RSRP, reference signal received quality (RSRQ), and signal to noise ratio (SNR) throughout a drive-by style campaign in an urban university campus using a passive monitoring device. But little prior work has explored the effective measurement of received signal strength using SDRs accompanied by an adequate validation from UE readings.

Estimation of RSRP plays a vital role in many control plane operations, including inter- and intra- eNodeB handovers as well as several diagnostic methods in LTE networks [171, 172, 173, 174, 175]. For instance, [176] evaluates the performance of RSRP handovers in LTE. The authors observe that a handover margin of 2dB to 6dB (RSRP) leads to an optimal number of handovers without sacrificing much of uplink SINR (for a specific range of user velocity). [177, 175] study the effect of RSRP measurement bandwidth on the accuracy of handovers. From a telecom provider's perspective, this suggests a need for up-to-date, accurate RSRP space-maps for improving service quality.

Several prior works examine the relationships between RSRP, RSRQ and SINR [178, 179, 180, 181], but little work explores the correlation between passive monitoring of LTE channels and ground UE readings. In [182], researchers examine the viability of deploying LTE connectivity using UASs in a rural area. Their results indicate that the coverage outage level increases from 4.2% at an altitude of 1.5 m to 51.7% at 120 m under full load conditions. A related study analyzed a set of live network measurements conducted with an LTE scanner attached to an airborne UAS [183]. The findings suggest improved radio clearance as the UAS increases altitude. The increase in the average number of detected cells, as altitude increases, corroborates these findings.

LTE Message Analysis: In addition to LTE coverage assessment via RF sensing, the analysis of LTE messaging traffic provides insight into the operations of the network. This type of analysis may be difficult. For example, messages transmitted after the LTE connection establishment stage are invisible to a passive device. As a result, little prior work leverages passive measurements to detect overload.

Previous work led to the development of several network analysis tools. `xgoldmon` [184], for instance, can monitor control plane messages over 2G/3G but not LTE. `SCAT` [185] detects problems in cellular networks but is limited to active monitoring on Qualcomm and

Samsung basebands. `QXDM` [186] diagnoses network statistics but is limited to Qualcomm baseband and requires a paid license. While [187, 188, 189] offer very similar feature sets to the tools discussed above, they are not tailored to work with SDRs for passive monitoring. [190] and [161], for example, provides a system for message analysis limited to GSM networks, which relies on UE connected to the network. Given the variety of MNOs simultaneously operating in an area, an approach using passive measurement devices, such as SDRs, without a need for network access is preferable.

Several prior works investigated various congestion control algorithms in LTE networks [191, 192, 193, 194], but little work explores overload detection without involving an active monitoring aspect. [195] used machine learning models to predict network congestion. However, their approach requires considerable historical data and is not suitable for urban sectors where eNodeBs are upgraded regularly to cater to increasing user bases, nor can it be used to assess current overload levels. [196] introduced LoadSense, which offers a measure of cellular load using channel sensing at the physical network layer (PHY). Similarly, [197] allows a client to efficiently monitor the LTE basestation's PHY resource allocation to estimate available bandwidth. [198] propose Cellular Link Aware Web loading (CLAW), which boosts mobile Web loading using a PHY informed transport protocol. Although the aforementioned tools can estimate whether the radio resources are fully allocated, they do not explicitly reveal whether the network is overloaded.

2.1.5 Outdoor Wireless Localization

When LTE networks overload or go offline, other networks, such as fixed broadband, may be similarly affected. For example, during Hurricane Harvey, the Analysis of Network Traffic (ANT) lab at USC/ISI, testing Internet reachability for certain IP blocks, found home network outages as high as 40% in certain areas [199]. Those needing emergency

assistance or rescue had no digital communication channels to fall back on. Lack of technology-assisted communication forces rescuers to rely on sight and sound to locate survivors. These challenges motivate research into new technologies for locating survivors and establishing alternate emergency communication channels.

As with network assessment, UASs are a promising technology used in establishing communication and providing situational awareness in an area immediately post-disaster. First responders adopt UASs for awareness and localization using white-light and thermal cameras [49, 200, 201, 202]. Using these systems, first responders locate and respond to people in distress. In January 2018, for the first time, a piloted UAS performed a real remote water-rescue, using cameras to locate swimmers and actuators to deploy an inflatable raft [50].

The problem of localizing an object using wireless signals is well-researched in the literature [203, 57, 204], particularly in WSNs where sensors self-locate based on reception of beacons from a series of anchor nodes with known locations [205, 206, 207, 208]. In these situations, anchor nodes are fixed, and the sensor uses received beacons to localize its position. In contrast, our work explores the reverse situation, where UEs transmit, and the mobile aerial device performs the localization. Because we target a disaster scenario with a mobile UAS, there are time and computational constraints. For instance, the UAS has payload weight limitations as well as limited battery lifetime. Further, given the urgency of fast and accurate localization in these conditions, our localization algorithm must meet the computational time constraints.

In the current literature, UAS based localization is typically performed to identify a transmitter in 2D space. However, because the height at which a device is located (for example, on the roof of a building) could be used by first responders in their rescue efforts, our work explores full 3D localization. We utilize the RSSI from a transmitting device to approximate distance to a receiver (UAS). Alternative time-based range approximations,

such as time of arrival (TOA) and time difference of arrival (TDOA) [209, 210], are not applicable. Time-based methods require either synchronization between transmitter and receiver or multiple receivers. As we do not assume cooperation from the transmitter, we have neither, making RSSI an appropriate choice.

There has been limited work in using UAS for 802.11 RSSI-based localization. [211] successfully used RSSI to determine a 2D range for 802.11 devices within a 15 meter average square from a quad-copter using maximum likelihood estimation (MLE) performed on a dedicated server (off-the drone). [212] used a fixed-wing UAS to locate device within 50m after 9 minutes and 18m after 15 minutes, relying only on probe requests and using a Bayesian Optimization for 2D localization. [57] uses a quad-copter for 2D localization using a random forest to determine which of six zones in a 120x80 meter search areas a device is located in.

2.2 Part Outline

Our research explores open questions in the field of wireless networks for challenged contexts. Through the utilization of advances in UASs and SDRs, we explore new methods of communication and network assessment. We back our methods and systems with experimental measurements in realistic conditions.

We begin part I by evaluating LTE networks. To tackle the need for low cost, rapid cellular network assessment particularly in rural areas and/or after a disaster, we explore how UASs equipped with low cost SDRs can be utilized for aerial assessment of LTE deployments in chapter 3. This research was conducted in partnership with multiple Native American tribal communities in California and New Mexico. We compare aerial approaches to on-the-ground measurements in order to validate our system with the two dimensional terrestrial user experiences.

While chapter 3 examines LTE coverage, in chapter 4 we investigate network behavior under significant load. As we discussed in the previous two chapters, LTE networks sometimes become overloaded during natural disasters and political crises. Measurements of network load are typically only available to MNOs. Outside agencies, such as federal regulators, first responders, and local administrators have few resources to evaluate the MNO reports. In chapter 4, we utilize passive measurements of control traffic by eNodeBs for evaluating LTE network load.

In addition to connecting humans, wireless technologies, such as IEEE 802.15.4, connect large IoT sensor networks. In remote regions, these WSNs may lack end-to-end Internet connectivity. While FUSN, using UASs for DTN data delivery, offer a promising solution, the effect of three dimensional communication on the IEEE 802.15.4 family of standards remains under evaluated.

Our research fills this gap by evaluating 802.15.4 2.4GHz network performance with measurements from an outdoor wireless aerial testbed. In chapter 5, we analyze experimental measurements from 802.15.4 transceivers utilizing PCB coiled antennae, commonly used in sensor networks. In chapter 6, we expand this research by investigating the difference between internal coiled antenna and straight-wire external antenna modules. For both works, we examine how factors, such as antenna orientation, altitude, antenna placement, and obstruction affect RSSI and PRR. This work is critical to the effective planning of FUSN deployments.

We finish this section with a look at IEEE 802.11. Unlike LTE and 802.15.4, this standard has a wide breadth of research for both terrestrial and aerial networks. In chapter 7, we explore how disconnected wireless devices can be used for disaster communication. We use 802.11 transmissions to locate users, an application highly valuable to search and rescue efforts. In contrast to other localization schemes, this research utilizes full 3D localization using VBA. VBA is a novel algorithm for passive localization, which

is computationally fast and can run on UASs. Past work investigates computationally expensive algorithms that do not scale well with increases in data size. We demonstrate VBA runs in constant time relative to the size of the data. We verify our work in multiple scenarios, varying factors like elevation, search pattern, number of packets per second transmitted, as well as effect of obstruction, and we compare our algorithm to other approaches.

2.3 Key Contributions

This research contributes to the understanding of three highly utilized wireless standards. It provides methods, systems, and algorithms, backed by experimental evaluation.

- For LTE coverage analysis, we show that low-cost SDRs equipped on UASs can detect LTE availability to produce a coverage map, using RSRP measurements comparable to ground-truth measurements from UEs. To the best of our knowledge, our study is the first to capture high precision RSRP measurements with an UAS using low-cost off-the-shelf SDRs.

- For LTE overload detection, we propose a novel method for capturing control plane connection establishment messages via SDRs. Our approach is portable, scalable, independent of any proprietary platform (e.g., Qualcomm, Samsung, etc.), and compatible with any LTE MNO.

- For 802.15.4 networks, we find that network configuration plays a significant role in network quality, which RSSI, a mediator variable, struggles to account for in the presence of high packet loss. Instead, we propose that PRR is a more appropriate network metric. We show that for this application, ZINB effectively models PRR.

Our work is, to our knowledge, the first performance measurement study of 802.15.4 2.4GHz ground-air network from a UAS.

- For 802.11 networks, we deploy an UAS to locate an uncooperative device in three dimensional space using RSSI. We provide a novel algorithm using voxel-based approximation (VBA) for real-time on-the-drone multilateral localization. We demonstrate that, in contrast to other approaches, VBA uses fixed computation time and memory as the number of data points increases, while still providing comparable error. Based on our results, we discuss how a continuously updated geographic representation can be used for flight planning and rescue operations post-natural disaster, and how this algorithm can be extended for other applications, such as IoT sensor network mapping.

2.4 Broader Impacts

In addition to peer reviewed publications and presentations to academics, this work was impactful to the larger community.

- We have partnered with multiple Native American communities to assist in the mapping, evaluation, and deployment of LTE and TVWS technologies. We partnered with Southern California Tribal Digital Village [213], New Mexico Santa Clara Pueblo, and New Mexico Ohkay Owingeh tribe in order to develop technology in line with the needs of these communities.

- We have provided outreach and training at Santa Clara Pueblo, Ohkay Owingeh, and Northern New Mexico Community College, in addition to countless one-on-one discussions of our work with partners and community members.

- Aspects of our work were integrated into classes at the UCSB School for Scientific Thought [214] as part of a five week high school enrichment program, as well as the Curie-osity Project, which introduces girls in grades 4-6 to women scientists and engineers.

- This work was instrumental in mentoring a new wave of scientists, including Ryan Allen who graduated with a Masters degree based on this project and Max Ginier, an undergraduate researcher. Students helped develop methodology, run experiments, perform analysis, and summarize their work in academic publications.

- This work has already been well-received by industry including Google, Cisco, and WeFi.

Chapter 3

Evaluating LTE Coverage and Quality from an Unmanned Aircraft System

As we have discussed in chapters 1 and 2, there is a need for accurate public assessment of LTE network coverage. While MNOs submit estimated coverage maps to regulators, like the FCC, companies may either accidentally or maliciously overestimate coverage. For rural and tribal areas, existing methods of network assessment can be difficult in time and manpower due to physical area and difficulty of terrain. Similarly assessing coverage after a disaster is valuable but may be difficult or dangerous for ground personnel.

New technologies like SDRs and UASs are promising tools for automating LTE coverage evaluation. However, the highy mobile and three dimensional nature of UASs results in signal characteristics that do not fully align with terrestrial counterparts. This discrepancy between aerial and terrestrial systems necessitates extensive evaluation that a proposed sensing system aligns with the on-the-ground experience of UEs.

In this chapter, we assess the accuracy of a low-cost, small form-factor RTL-SDR for sensing ENodeB signal strength over a wide area through integration with an off-the-shelf quad-copter UAS. We first compare reading accuracy of this airborne sensor with commonly used hardware for ground-based wardriving approaches (i.e a Spectrum analyzer (SA) and a Universal software radio peripheral (USRP)). Because no existing studies systematically examine the effect of altitude on signal strength measurements, we fly the UAS at varying altitudes across multiple locations and examine how aerial signal sensing can be aligned to ground-level measures. Because minimal previous research compares observed signal strength between measurements collected by UEs (i.e smartphones, tablets, and hotspots) and UASs, we deployed four cellular devices on the ground, each collecting measurements from different cellular networks, and compared these measurements over the same geographic area to those collected by the RTL-SDR on the UAS. We look to the UE measurements as "ground truth" because the UE readings capture examples of the actual coverage and performance a user, in the given location, would experience with UE.

Our findings reveal that the simple RTL-SDR has comparable accuracy to expensive solutions and can estimate quality within one gradation of accuracy when compared to user equipment. Further, we show that these devices can be mounted on a UAS to more rapidly measure coverage across wider geographic regions. Our results show that the low-cost aerial measurement techniques have 72% accuracy relative to the ground readings of user equipment, and fall within one quality gradation 98% of the time. Our findings, taken together, offer a detailed look at the efficacy of low-cost, public controlled, aerial coverage and quality sensing.

(a) UE (b) Ground Measurement (c) UAS (d) Stationary box.

Figure 3.1: LTE Sensing Equipment.

3.1 System Overview and Methodology

We collected ground and air measurements in two regions in Rio Arriba county, New Mexico over a period of five days, beginning May 28, 2019. For each region, we obtained permission to drive through residential areas, as well as to fly a UAS with a SDR.

In this section, we describe the six unique RF sensing methodologies employed in our analysis. Like all wardriving studies, our work is necessarily limited in scale. However, these ground measurements are uniquely useful for contrasting the efficacy of each measurement technique. In all cases, our methodology is generalizable. Figure 3.1 shows images of many of our sensing set-ups.

3.1.1 Method 1: Ground-Driven UE Sensing

In our wardriving campaign, we record signal strength readings from four Motorola G7 Power (XT1955-5) phones, each running Android Pie (9.0.0). We collect measurements using the Network Monitor application [215]. An external GlobalSat BU-353-S4 GPS connected to an Ubuntu Lenovo ThinkPad laptop gathered geolocation measurements, which we matched to the appropriate ground measurement by timestamp. We outfitted each phone with a SIM card from one of the four top cellular providers in the region: Verizon, T-Mobile, AT&T, and Sprint. The phones recorded signal strength every 10 seconds while we drove at speeds less than 10 miles per hour through the areas of study.

3.1.2 Method 2: Ground-Driven Spectrum Analyzer

We gathered measurements on LTE channel center frequencies with a high-precision Keysight N9340b SA using a ham radio antenna capable of sensing signals up to 3 GHz. The SA was transported inside the measurement vehicle while the antenna was magnetically mounted to the roof.

3.1.3 Method 3: Ground-Driven USRP

We collected center frequency readings with a Ettus Research USRP B200, a versatile SDR widely used for LTE and TV frequency experimentation and sensing. The USRP measured the same set of LTE channel center frequencies as the SA through a ham radio antenna placed beside the identical SA antenna.

3.1.4 Method 4: Ground-Driven RTL-SDR

We also collected center frequency readings with a NooElec RTL-SDR RTL2832U and Elonics E4000 Tuner, an inexpensive software defined radio operating in the 55MHz-1100MHz and 1500MHz-2300MHz ranges. The RTL-SDR measured LTE channel center frequencies through a ham radio antenna placed beside the identical SA and USRP antennas.

Unlike the USRP B200, which when inside a transportable case is bulkier and more expensive, the RTL-SDR is low-cost and protected in a smaller form-factor. This specific model of RTL-SDR covers most LTE frequencies and is simple to equip onto a UAS or deploy at a stationary site for long-term monitoring. This ground-transported RTL-SDR serves as a comparison point for the UAS and longitudinal sensing experiments described subsequent subsections.

3.1.5 Methods 5: Aerial Sensing Platform

Our UAS consisted of a DJI Matrice 100 quad-copter, as shown in Figure 3.1c. The UAS collected signal strength readings via a NooElec RTL-SDR (the same model as used for ground measurements) connected to a Raspberry Pi (RPi) 2—Model B on-board computer via USB. The location of the UAS was recorded from the Matrice 100 on-board GPS, sampling at a rate of 50 Hz and using a UART connection to the Raspberry Pi (RPi).

Horizontal Coverage Mapping

In one set of experiments, we flew the UAS manually at varying speed and elevation (in order to clear obstacles and keep the UAS in line of sight) to map coverage. We attempted to cover the same areas as covered by the UE and ground measurements. For each geographic area, UAS measurements occurred on the same day as the other data collection, but sometimes several hours apart.

Vertical Experiments

To investigate the impact of elevation on signal strength measurements, we performed four sets of vertical-only flights. Each set of flights was conducted in a different geographic region of our measurement area. During each vertical flight, the UAS was raised by 10ft increments approximately every 15 seconds to 100ft. It was then raised in 20ft increments approximately every 15 seconds to 400ft (the maximum United States Federal Aviation Administration (FAA) non-exempt altitude limit).

3.1.6 Method 6: Stationary Box

Because continuous monitoring in an area can be costly in terms of equipment and manpower, coverage mapping is typically completed via sampling over a short time-frame. For example, in our ground sensing driving campaign, we take all samples over a maximum of one hour for each unique location. As part of our study, we seek to verify that this one-shot sensing method is appropriate for estimating long-term spectrum availability.

We therefore measure spectrum occupation over time in a single location to monitor changes. We enclosed a NooElec RTL-SDR (the same model as is utilized by the ground measurements and UAS) run by a Raspberry Pi (RPi) 3 B+ in a weather-proof case with the stock antenna on top the case, shown in Figure 3.1d. Over two days, the RTL-SDR continuously iterated through a pre-programmed list of all 20 known LTE frequencies for the four network providers in the area and recorded signal strength readings for each frequency every three seconds.

This method monitors the stability of the RTL-SDR measurements over time and can indicate the appropriate flight time necessary to generate a consistent measure of signal quality in an area. While this data is not generalizable geographically, it provides insight into the precision of signal strength reading from an RTL-SDR .

3.1.7 LTE Channel Selection by Provider

Before data collection, we compiled a list of LTE cellular frequencies in use by the top providers in the area. This was needed for every sensing method other than the UEs , which pulled the active frequencies automatically for their respective provider. We compiled this list using two complementary processes. First, on each UE we ran CellMapper [151], an Android application that allowed us to query the active frequencies detected by the device for the corresponding LTE provider. We supplemented this list

Table 3.1: Number of overlapping geographic bins by signal collection method.

	UE	**UAS**	**Spectrum Analyzer**	**USRP**	**RTL-SDR**
UE	1152	-	-	-	-
UAS	305	812	-	-	-
Spectrum Analyzer	131	53	1199	-	-
USRP	131	53	1199	1199	-
RTL-SDR	131	53	1199	1199	1199

with a scan using a spare Ettus Research USRP B200, equipped with a wideband LTE dipole antenna [216], connected to a Lenovo ThinkPad laptop running srsLTE[217]. Using srsLTE we performed a scan of all possible LTE frequencies operated in the United States and appended to our list any frequencies not previously discovered. Since UEs choose the strongest frequency to communicate with a nearby base station, this allowed us to locate other frequencies available from nearby cells which the UEs would not use at our test sites but could jump to intermittently. As we moved between regions, we added all newly detected frequencies to the list scanned by all sensors.

The resulting list contained 22 frequencies in operation in the area, served by the four providers. Because the NooElec RTL-SDR is limited to a frequency of 2300 MHz, two of the detected frequencies (2628.8 MHz and 2648.6 MHz) were outside the range frequencies we could sense on the Ground RTL-SDR, Stationary Box, or UAS and are dropped from our analysis.

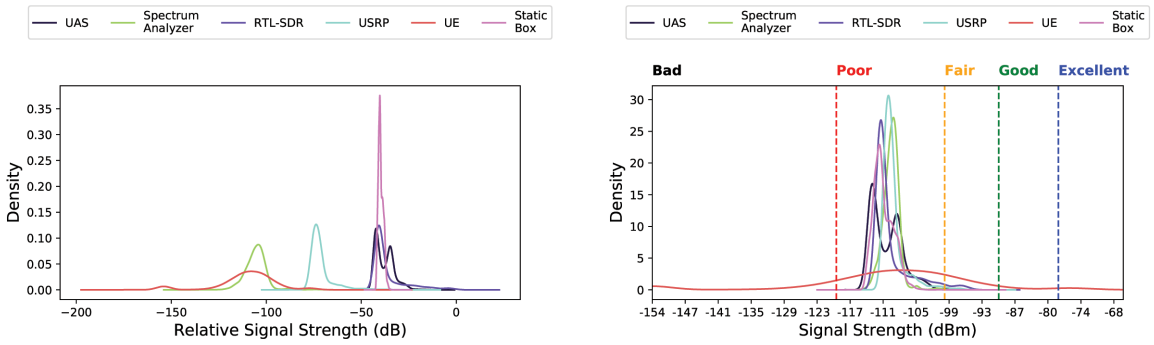
3.2 Analysis

3.2.1 Accuracy of Data Collection Methods

Preparing data for geographical analysis

Because multiple devices and personnel participated in data collection, the data was not sampled at the same exact timestamp or precise GPS location for all methods. For example, both the ground sensors and the UAS passed over the same residential area but may not have covered the same 1 meter GPS coordinate due to road availability. To accurately compare data collected by different methods, we first aggregate data into geographical bins of three decimal places of GPS accuracy, approximately 110 square meters in area. Then, for each method and for each set of readings on different LTE frequencies, we take the mean across all the signal strength values that fall into that geographic bin.

The data collected by the UEs included only the network provider (AT&T, Sprint, T-Mobile, and Verizon), and not the frequency on which the UE was operating. To compare this to the other data collection methods, which report frequency instead of network, we first map each frequency to the corresponding network provider using the frequency list we describe in Section 3.1.7. For each network provider and geographic bin, we then select the frequency with the strongest signal strength and set that as the signal strength for the provider in that bin. This method resulted in 2,637 unique 110m^2 geographic bins. Not every area was sampled by every method. The resulting overlap between methods and geographic area is summarized in Table 3.1.



(a) Original data.

(b) Transformed distributions.

Figure 3.2: Kernel density estimation of distributions by signal collection method.

Transforming Raw Signal Readings

The UEs , spectrum analyzer, RTL-SDR modules (including both the one mounted on the UAS and in the stationary monitoring box), and the USRP all report signal strengths on varying scales. We show the original distribution of the relative signal strengths for each collection method in Figure 3.2a. As we can see, while the distributions have similar normal-like peaks, the offsets and width of the distributions do not match.

While the spectrum analyzer outputs dBm, the other devices report relative signal strengths. As we are interested in the experience of the end user, before comparing data collection methods, we first need to transform the raw relative signal strength readings to match the UEs .

To do so, we first perform a min/max normalization on each method, as shown in Equation 3.1, where $\vec{O}$ is the original data, $m \in \{SA,\ USRP,\ RTL\text{-}SDR,\ UAS\}$ is the method and $\vec{N}$ is the normalized data.

$$\vec{N}_m = \frac{\vec{O}_m - \min(\vec{O}_m)}{\max(\vec{O}_m) - \min(\vec{O}_m)} \tag{3.1}$$

Next we offset and scale the other methods to align them with the signals received by the UE. To do this we randomly selected 50% of our data as a training set. On this training set for each method we find an offset x_m and scaling factor a_m to minimize Equation 3.2, where n_m and n_{ue} are measures taken from the same geographic bin and cellular network provider. If a method does not have a matching UE measurement, it is omitted from the sum.

$$\min_{x_m, a_m} \left(\sum_{n_m \in \vec{N}_m} [a_m(n_m + x_m) - n_{ue}]^2 \right) \tag{3.2}$$

Finally we scale and offset *all of our data* for every method other than the UE by x_m and a_m, and scale back to the readings of the UE as expressed in Equation 3.3, where $\vec{T}_m$ is the resulting transformed data for each method.

$$\vec{T}_m = \left[a_m \left(\vec{N}_m + x_m \right) \right] \left(\max(\vec{N_{ue}}) - \min(\vec{N_{ue}}) + \min(\vec{N_{ue}}) \right) \tag{3.3}$$

By transforming the data we can now compare signal strength readings to one another and to the UE. The resulting transformed distributions are shown in Figure 3.2b. For the rest of our analysis, we use this transformation when we report signal strength values in dBm.

Signal Quality	**Range**	**Color**
Bad	<120 dBm	Black
Poor	-120 to -111 dBm	Red
Fair	-111 to -105 dBm	Orange
Good	-105 to -90 dBm	Green
Excellent	>90 dBm	Blue

Table 3.2: Categorization of signal strength into signal quality bins.

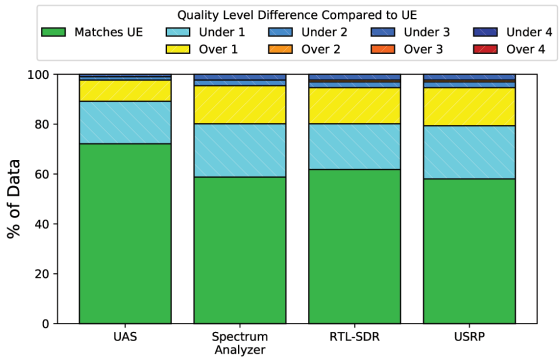


Figure 3.3: Accuracy of signal collection methods as compared to the UE.

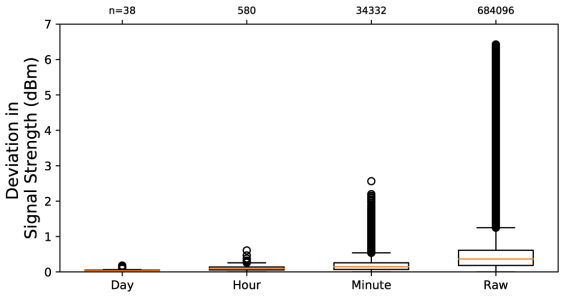


Figure 3.4: Distribution of deviation from mean signal strength of all frequencies.

Estimating Signal Strength

We computed the Pearson correlation on each method pair and found only a weak linear relationship between the collection methods and the readings from the UEs , even after transforming the data. As signal strength can vary, even between different UE device makes and models, we categorize the level of signal quality rather than predicting the exact signal strength a UE would receive in an area by dividing the signal strength levels into five groups, based on criteria in Table 3.2. While there is no standard for defining what LTE signal strength corresponds to what quality, we model our criteria after those suggested by SignalBooster [218].

Based on this categorization we compare how each method sorted the signal strength readings across the geographical bins, using the UE as ground truth. We summarize our results in Figure 3.3. The UAS was most closely aligned with the UE, matching values exactly for 72% of the geographic bins. When allowing for one signal quality of discrepancy (for instance, a method stating a signal was Fair when the UE labeled it as Poor) all methods had over a 95% accuracy, with the UAS again leading with a 98% accuracy. A notable result was that error was skewed towards under-predicting the received signal strength. Accounting for this bias when estimating UE reception would improve accuracy further.

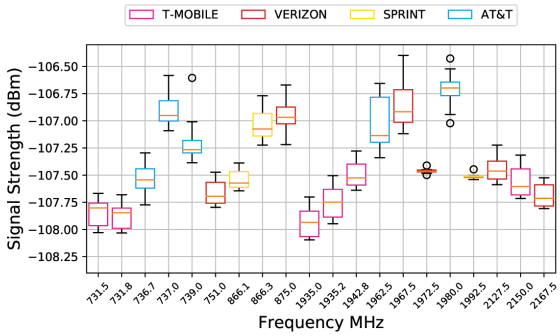


Figure 3.5: Distribution of signal strength by frequency over period of observation, averaged across one hour windows.

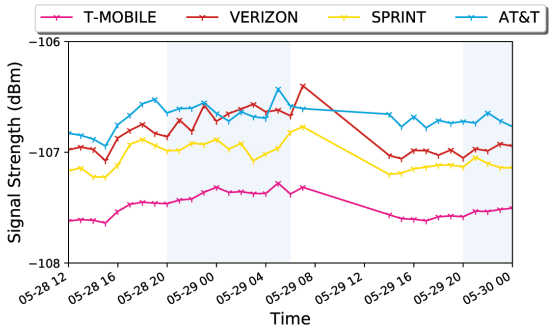


Figure 3.6: Signal strength over time by provider gathered by stationary box. Night hours are shaded in blue.

3.2.2 Longitudinal Analysis

From the stationary radio (introduced in Section 3.1.6) we received 684,096 readings over a period of two days. To measure the relative stability of signal strength readings, for each reading we calculated the deviation from the mean of the corresponding frequency. To determine the stability across different time scales, we re-sampled the data over multiple time scales (1 minute, 1 hour, 1 day), averaged the intermediate readings, and re-computed the deviation of each sample. The resulting time-series is shown in Figure 3.4.

As expected, raw readings (with a sampling frequency of 3 Hz) fluctuated considerably, with a total range of 80dBm and the majority of fluctuations < 7 dBm from the mean. When comparing between minutes the majority of the reading were < 3 dbm from the mean. When comparing hour to hour, the majority of signals deviate < 1 dBm from the mean. Comparing between two days of data, across all frequencies, the majority of signals did not deviate.

We analyzed the distribution of hour to hour signal strengths across the 20 monitored LTE frequencies, as shown in Figure 3.5. The majority of readings across the two day time span fell within 1 dBm of each other. The largest change of signal strengths between

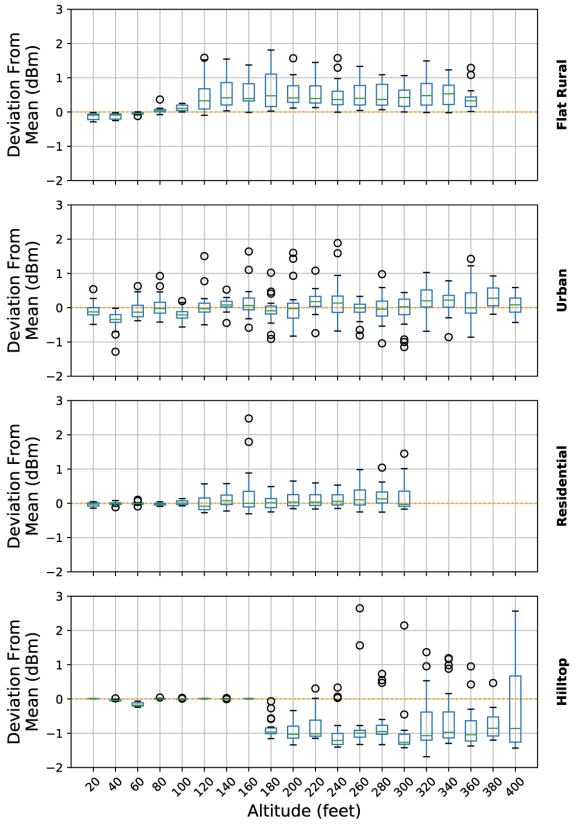


Figure 3.7: Deviation of signal strength from mean.

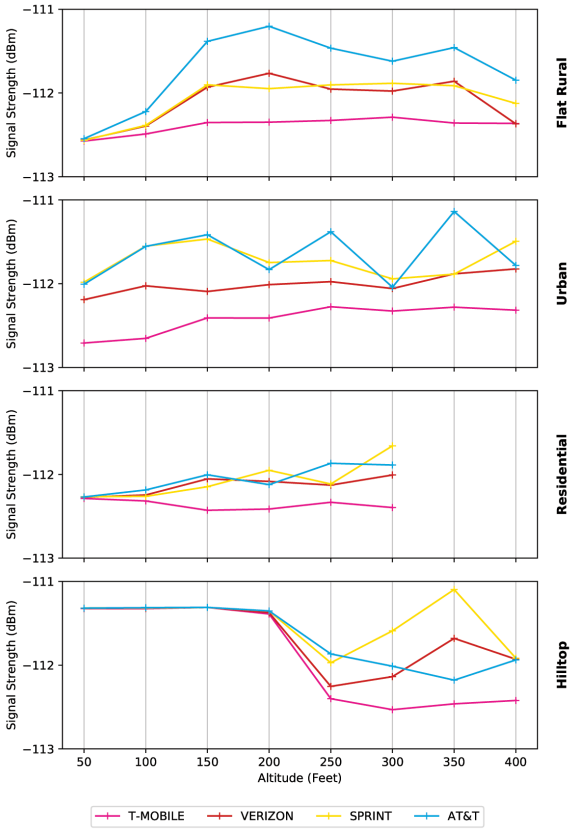


Figure 3.8: Signal strength change by altitude and network.

two hours was observed on 739.0 MHz (utilized by AT&T) and 1967.5 MHz (utilized by Verizon) which exhibited 7dBm changes.

The end user is most impacted by the signal strength of the frequency chosen by the UE. We also examine the hour by hour change of network signal strength. For each operator, for every hour time window, we choose the frequency with the maximum average signal strength. We present the results in Figure 3.6. While there is a slight improvement in signal strength during night time hours, for each network the total hour to hour fluctuations in signal strength are minimal.

3.2.3 Impact of Altitude

To analyze the impact of sampling altitude on signal strength we executed multiple vertical flights in four different locations, as described in Section 3.1.5. In our analysis, we keep the four locations separated and look at how signal strength at each LTE frequency varies with altitude from the ground. To compare between frequencies, we calculate the deviation of each signal strength measure from the mean of that frequency at each location. We then group altitudes into 20 foot bins and examine the distributions of altitudes across those bins at each of the four locations. We present the results in Figure 3.7.

Our results show that signal strength variation can be quite dependent on location. The first and third locations, *Flat Rural* and *Residential* respectively, were located level with a wide area of flat terrain surrounded by low hills. At these locations, the vertical UAS flight showed an overall increase in signal strength as altitude increased across LTE frequencies. This might suggest that in low lying terrain, away from strong cellular readings, coverage mapping may be sensitive to flight altitude.

The second location, *Urban*, was located in a more urban area with better cellular coverage. In this area, altitude did not alter the signal strength of frequencies sensed by the UAS. The fourth location, the *Hilltop*, was located high on a hill approximately 400ft above the *Residential* location. At this location, altitudes over 160ft showed a drop in signal strength across most of the monitored LTE frequencies. One possible explanation is that the aerial vehicle may have difficulty detecting coverage at altitudes significantly higher than the provisioned coverage area.

In addition to examining frequency fluctuations, we examine the received signal strength by cellular network provider. In Figure 3.8 we show the mean change in signal strength across all frequencies with altitude for each network and location. The change is signal is network dependant, and moreover the difference between networks depends on

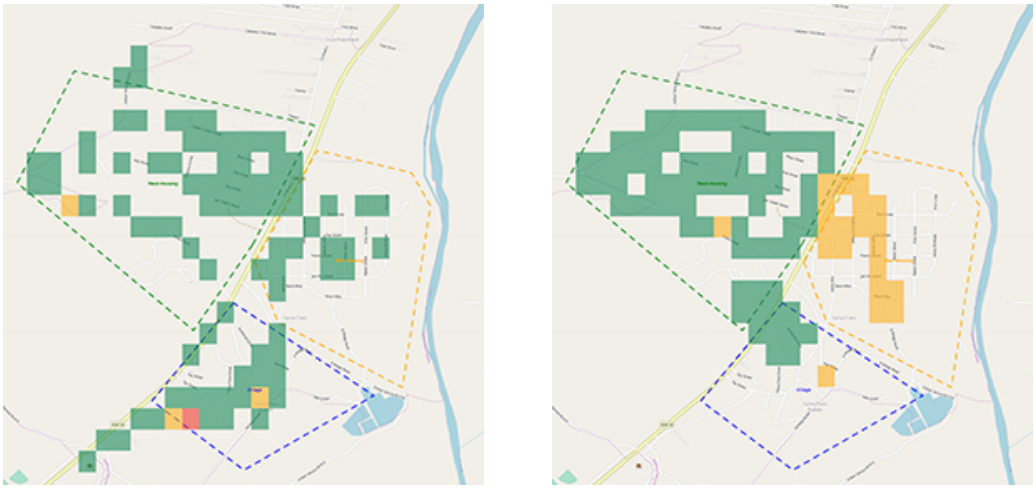
location. A probable explanation for the observed difference is that the ENodeBs serving these networks are in disparate geographic locations, with different signal propagation patterns.

3.3 Discussion

While we observe clear relationships between sensing methods, the relative signal strength values output by the devices are weakly linearly correlated, particularly to the UEs, even after transforming the data to a common reference frame. We believe the problem stems from difficulty in aligning the various methodologies for comparison. Because we were unable to capture the frequencies on which the UEs operated, we compared the frequencies with the highest signal strength for a given method. This may not always match the actual frequency used by the UE. Additionally, the wardriving readings from the RTL-SDR, USRP, and SA are more difficult to collect due to the labor involved. As a result, there are fewer points of geographic overlap than for the UE and UAS measurements.

By categorizing individual signal strengths by quality, mirroring the "bars" of signal strength that a user's device might report, we were able to accurately match these categorical measurements across measurement methods. As the most versatile collection method, the UAS predicted quality within one gradation over 98% of the time. This aerial signal sensing method demonstrates promise as an effective system for wide-scale cellular coverage mapping.

Based on our experimental data, we generated a coverage map for each method and provider. Figure 3.9 shows a portion of the map for Verizon. The readings from the UE are shown in Figure 3.9a and those taken from the UAS on the same day are shown in Figure 3.9b.



(a) Map from UE readings. (b) Map from UAS readings.

Figure 3.9: Example of generated LTE signal coverage map. Colors and values correspond to Table 3.2, with high RSRP in green and low in red.

Next, we evaluated how the design of our aerial data collection impacted the accuracy and precision of UAS signal quality assessment. We considered: 1) how a sample taken at a particular time compares to the overall LTE channel quality during a 24 hour period; and 2) how different UAS altitudes impact this characterization.

Sensing over Time:: Consider the longitudinal analysis of the stationary sensing equipment from Section 3.2.2. For most frequencies, a single flight *is* sufficient; most readings fell within 3 dBm of each other. However, certain frequencies may be less stable. For example, we found that two channels within the two-day deployment showed values varying by up to 7 dBm, which is wide enough to bump a reading by two signal quality levels (e.g. Good to Fair, or even Poor).

As observed in Figure 3.5, RTL-SDR signal strength measurements fluctuate between readings. The average of multiple readings provides a more stable description of signal quality in a geographic region. Because the UAS would need to take dozens of samples

from a geographic region, the flight pattern and maximum flight speed would vary with the desired granularity of the measurements. The UAS must fly at low speeds to achieve high granularity (e.g. building level accuracy) and higher speeds to achieve greater coverage but lower granularity (e.g. neighborhood level). Alternatively, the UAS can fly at higher speeds following a flight plan that conducts repeated measurements at the same location (e.g. through flight loops).

RF signals sometimes fluctuate based on moisture and other weather conditions. In this study, we did not capture the sensitivity of RTL-SDR signal strength measurements to weather fluctuations and seasonal changes. The stationary sensing equipment was deployed during clear days with no rain, a daily temperature high of ≈ 72 and a low of ≈ 52. In future work, it would be informative to examine how signal readings from an RTL-SDR vary over the course of much longer time spans. Such an assessment could reveal whether the RTL-SDR equipped UAS requires calibration depending on current weather. This type of study would also help with understanding how cellular network quality measurements from our system may fluctuate during or after a natural disaster.

Choice of Altitude: To measure how altitude affects signal quality, we look to the analysis in Section 3.2.3. The interaction between altitude and signal strength reception by the UAS is complex. The local geographic topography seems to be the dominant factor in received signal strength. When flying in low valleys, an increase in altitude corresponded to an increase in mean received signal strength. Yet, when ascending from a hilltop above the residential area, signal strength declined. Because the orientation of cellular tower antennas by network providers is provisioned to optimize coverage at elevations of residences and businesses [160], aerial collection at altitudes (in our case approximately five hundred feet above the residential elevation) may see degradation in signal strength.

The effect of altitude on signal quality has implications for evaluating LTE coverage and availability for the occupants of high rise buildings. In dense city centers, we could use aerial systems to map signal strength in three dimensions. Such a measure of signal quality across floors in skyscrapers would not be accounted for by conventional measurement methodologies.

3.4 Conclusion

In this chapter, we have shown that a UAS-mounted RTL-SDR is capable of providing a granular reflection of LTE signal strength. Our low-cost solution enables accurate coverage mapping and quality assessment in regions typically neglected by other forms of assessment. Moreover, our system achieves this without requiring expensive specialized equipment, extensive time commitments, or significant manpower. We hope that our work will pave the way for future solutions that more accurately represent cellular coverage, particularly in those regions that are likely under-served.

So far our evaluation has dealt with LTE coverage. However signal strength alone does not guarantee a high QoS or quality of experience (QoE). In times of crisis, even previously well provisioned LTE networks may become overloaded, degrading performance. In the next chapter, we explore how SDRs can be used to passively detect overloading, without MNO cooperation.

Acknowledgment

This work was done in collaboration with Vivek Adarsh, Udit Paul, Esther Showalter, Ellen Zegura, Morgan Vigil-Hayes and Elizabeth Belding and the incredible support of Jerrold Baca. This work was funded in part through NSF Smart & Connected Communities award NSF-1831698.

Chapter 4

Packet-level Overload Estimation in LTE Networks using Passive Measurements

In the previous chapter, we explored how SDRs can be used by regulators and communities to independently assess LTE coverage. However, as we concluded there is a difference between coverage and usability. This is especially critical during and after a disaster. In the United States for example, after a disaster, the FCC typically receives outage reports from MNOs, for instance [166], but the actual usability, due in part to overload, on active towers is difficult to assess without access to the internal network. Ideally, public entities should be able to assess the overload and operational status/usability for a particular base station. Further, they should be able to accomplish this without relying on the cooperation of a cellular provider.

To address this critical need, in this chapter, we propose a novel solution to infer overload in LTE networks based on messages broadcast by the eNodeB. Through the analysis of multiple message types, we draw clear comparisons between instances of high

network utilization and typical operating conditions for several eNodeBs. Our results indicate that eNodeBs demonstrate measurable performance differences indicative of overload conditions.

Importantly, our solution works without the cooperation of the cellular provider. Using low-cost, readily usable off-the-shelf equipment, we demonstrate that unencrypted broadcast messages sent by the eNodeB [219] on the broadcast channel can be passively collected and analyzed to estimate local overload, and hence usability.

We quantify our results by computing two normalized metrics, which are proportional to the number of connection reject messages and cell barring signals (`cellBarred`), respectively (cell barring signals prohibit UEs from camping on a particular cell). In addition, we evaluate the back-off timer (`waitTime`) encapsulated in each reject message. Note that in LTE, a connection reject message does not contain a rejection case. Consequently, we must use higher `waitTime` values, coupled with high rates of connection request denials, to reveal possible overload.

To test the operation of our system, we perform multiple measurement campaigns: two at events with unusually large crowd gatherings, and two at those same locations but during times of typical usage. Through these measurement campaigns, we collect and analyze over 3.2 million LTE frames. Our analysis indicates that overload on an eNodeB can be identified through an increase in reject messages and mean back-off time. Moreover, these events are often accompanied by a significant increase in cell barring signals. We show that overloaded cell towers frequently deny larger percentages of connection requests and issue higher `waitTime` as compared to typical utilization periods. Further, we observe an unusual number of barring signals prohibiting UEs from camping on their desired eNodeBs.

4.1 Background

In our work we examine cellular transmissions using SDRs. While most of the transmissions on LTE are encrypted between the eNodeB (LTE base station) and UE (such as a cellphone) [220], connection establishment messages are sent in the clear. We use these messages in order to determine overload, as described in the following sections.

4.1.1 Radio resource control (RRC)

The Radio resource control (RRC) protocol [221, 222] supports the transfer of common Non-access stratum (NAS) [223] information (which is applicable to all UEs) as well as dedicated NAS information (which is applicable only to a specific UE). Directed RRC messages (unicast to a single UE) are transferred across Signalling radio bearers (SRBs), which are mapped onto logical channels [224, 225] – either the Common Control CHannel (CCCH) during connection establishment or a Dedicated Control CHannel (DCCH) if the UE is in an active connection state. Similarly, System Information (SI) messages are mapped to the Broadcast Control CHannel (BCCH). Since messages on DCCH are on a private channel, they cannot be decoded by passive monitoring devices.

Common Control CHannel (CCCH): This channel is used to deliver control information in both uplink and downlink directions when there is no confirmed association between a UE and the eNodeB – i.e. during connection establishment. Messages on this channel are transmitted in the clear, and can be passively decoded. We leverage this knowledge to analyze signalling messages and estimate the overload level in an eNodeB.

Broadcast Control CHannel (BCCH): This is a downlink channel that is used to broadcast System information (SI). It consists of the Master information block (MIB) and a number of System information blocks (SIBs). The MIB and SIBs are broadcast through

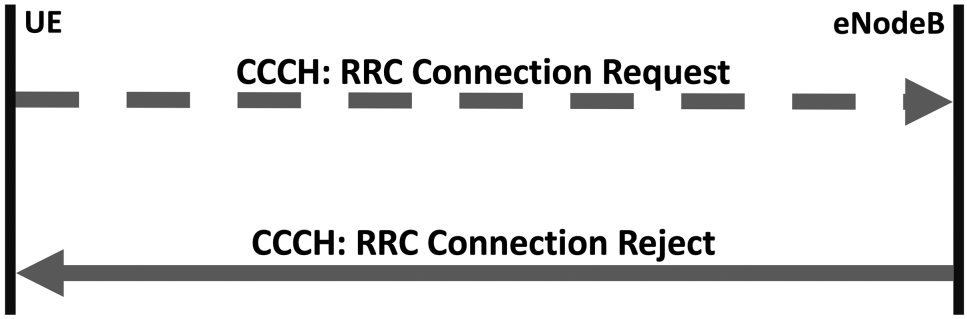


Figure 4.1: Flow diagram for LTE connection reject messaging.

RRC messages. SIB1 is carried by `SystemInformationBlockType1` message. Though there are other SI messages, we focus on SIB1 for the purpose of this study. SIB1 contains the cell barring (`cellBarred`) status, which indicates whether or not a UE may choose the cell. When `cellBarred` status is indicated, the UE is not permitted to select/reselect this cell, not even for emergency calls [226]. In that case, the UE may connect to another cell.

4.1.2 Signalling radio bearer (SRB)

A SRB [227] carries CCCH signalling data. An SRB is used during connection establishment to establish the Radio access bearerss (RABs) and to deliver signalling while on the connection (for instance, to perform a handover, reconfiguration or release). There are three types of SRBs. SRB0 uses the CCCH channel with *transparent mode Radio link control (RLC)* while SRB1 and SRB2 use the dedicated channel with *acknowledged mode RLC*. In other words, SRB0 can be decoded by non-network equipment such as a SDR in the vicinity, while SRB1 and SRB2 cannot. Table 4.1 shows various signalling messages SRB0 carries.

For our study, we focus on `RRCConnectionReject` messages (solid arrow in Figure 4.1) with corresponding `waitTime` (back-off time, before a UE can again initiate a connection)

Table 4.1: Summary of Signalling Radio Bearer 0 (SRB0)

Channel Type	**RLC Mode**
CCCH	Transparent (*Decodable from passive capture*)
Direction	**RRC Message**
Downlink	RRC Connection Setup RRC Connection Reject
Uplink	RRC Connection Request

values, `Connectio- nRequest` messages, and `cellBarred` signals (BCCH). We formulate two normalized metrics based on the percentage of reject messages per request sent and the ratio of `cellBarred` signals to the number of SIB1 messages transmitted over thirty-second time bins.

4.1.3 Managing Overload

Overload management is invoked in order to unburden a cell to an acceptable level when overload is detected, for instance if the cell load remains above a threshold for some continuous period. An alternative strategy, such as that used by Wideband Code Division Multiple Access (WCDMA), is to lower the bit rates of connected users until the load returns to an acceptable level [228]. However, in a pure packet-based system such as LTE, the user bit rate is maintained at the MAC scheduler [229], which already provides a soft degradation of user throughput as the system load increases. Thus, if overload is detected in a cell the system must remove a subset of the connected bearers until the load is reduced to an acceptable level. Admission Control [230] is used to restrict the number of UEs given access to the system, in order to provide acceptable QoS to admitted users.

Listing 4.1: Decoded DL - CCCH message showing RRCConectionReject.


```
"user_dlt": "DLT: 147, Payload: lte-rrc.dl.ccch
 (LTE Radio Resource Control (RRC) protocol)",
"lte-rrc.DL_CCCH_Message_element": {
 "per.choice_index": "0",
 "lte-rrc.message": "0",
 "lte-rrc.message_tree": {
  "per.choice_index": "2",
  "lte-rrc.c1": "2",
  "lte-rrc.c1_tree": {
   "lte-rrc.rrcConnectionReject_element" : {
    "per.choice_index": "0",
    "lte-rrc.criticalExtensions": "0",
    "lte-rrc.criticalExtensions_tree": {
     "per.choice_index": "0",
     "lte-rrc.c1": "0",
     "lte-rrc.c1_tree": {
      "lte-rrc.rrcConnectionReject_r8_element": {
       "per.optional_field_bit": "1",
       "lte-rrc.waitTime": "6"
       "lte-rrc.nonCriticalExtension_element": {
        "per.optional_field_bit": "1",
        "per.optional_field_bit": "1",
        "per.octet_string_length": "2048",
        "lte-rrc.lateNonCriticalExtension":
         "34:07:79:f0:2c:e7:90:00:28:07:63:48:31:b7:90:00:
          38:07:04:f0:22:67:81:08:30:87:9e:40:3f:37:60:70:
          20:27:82:00:21:17:4c:88:36:47:80:00:20:07:15:00:
          2a:97:90:00:28:17:95:30:2a:97:99:30:2c:87:82:00:
          21:07:4c:f0:36:77:85:b0:22:d7:82:30:21:07:82:40:
          21:27:9f:80:2f:d7:68:18:33:f7:84:00:32:07:23:80:
          21:d7:76:f0:2b:77:91:40:28:a7:81:00:30:97:42:00:
          21:17:88:70:24:27:96:00:2b:07:48:00:24:17:66:00:
          23:d7:93:c0:29:f7:94:00:3a:07:50:f0:38:77:68:80:
```


4.2 Implementation

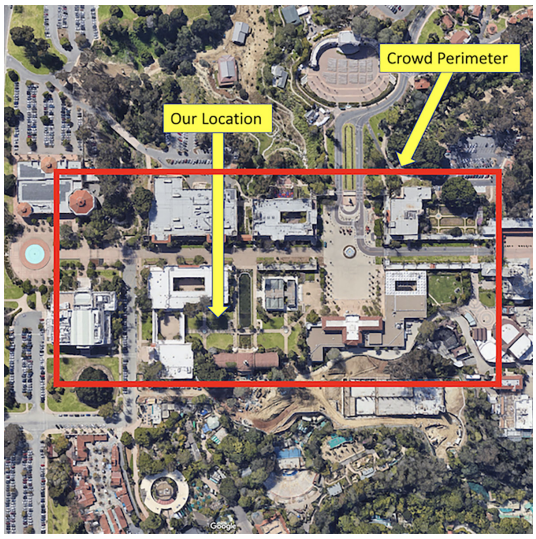
4.2.1 Experimental Setup

In our experimental setup, our receiver is comprised of an Ettus Research USRP B210 [231] SDRs attached to a MPantenna SUPER-M ULTRA Mobile Antenna with a frequency range from 25MHz to 6GHz [232]. The USRP is connected to a Lenovo ThinkPad W550s laptop for data collection and post-processing. We use the `srsUE` mode in the open-source srsLTE software suite [217] to locate available cells in the vicinity by scanning all frequency bands. On the day of the event, we capture broadcast messages in the form of binary I/Q samples using srsLTE's UE `usrp_capture` utility.

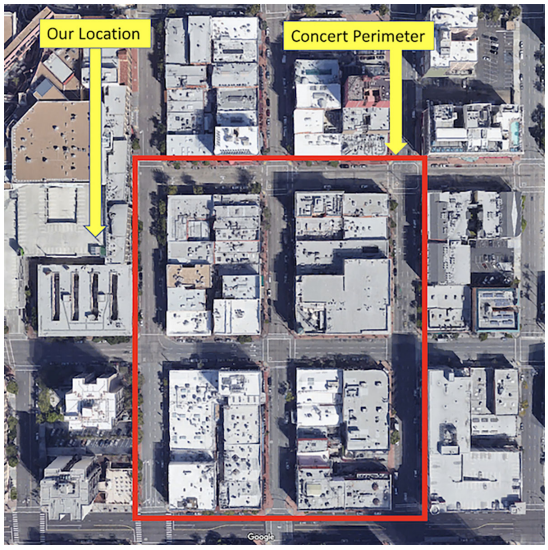
4.2.2 LTE Packet Decoding

We start with converting binary I/Q samples to hexdumps. To investigate the extent of overload on eNodeBs, we then transform the hexdump into network traces using Wireshark's *text2pcap* command [233]. Next, we use *lte_rrc* lua dissectors to decode LTE RRC messages using *tshark* [234]. We employ $lte-rrc.dl.ccch$ and $lte-rrc.ul.ccch$ protocols to decode RRC messages on the downlink and uplink common control channel, respectively. Lastly, we use the $lte-rrc.bcch.dl.sch$ protocol to decode downlink messages on the broadcast control channel.

Listing 4.1 shows a snapshot of the decoded RRC message on the downlink CCCH. We can see the *RRCConnectionReject* message tree along with additional options sent by the eNodeB during the RRC connection establishment phase. Embedded in this message is the *waitTime* parameter. While reject messages provide an indication of overload, we can use the value of the *waitTime* metric as a measure of the severity of overload. The value of *waitTime* is an integer in the range of 1–16, which defines how many seconds the UE should wait after reception of the *RRCConnectionReject* until a new connection can be attempted. According to 3GPP TS 23.401 [235], when rejecting an RRC connection request, the eNodeB indicates to the UE an appropriate timer value that limits further requests, relative to the severity of overload; the more the overload, the greater the *waitTime*. Upon receiving the *RRCConnectionReject* message, the UE starts timer T302 [221], with the timer value set to *waitTime*. The UE is not allowed to send another *RRCConnectionRequest* for mobile originating calls, signalling, terminating access or circuit-switched fallback (CSFB) [236, 237] on the same cell until the expiry of T302. Note that in LTE, the *RRCConnectionReject* message does not contain a *RejectionCause*, therefore *waitTime*, in conjunction with reject messages, is a crucial parameter in assessing the level of overload.



(a) SPD dataset. Balboa Park, San Diego, CA



(b) CSR dataset. Downtown San Diego, CA

Figure 4.2: Google aerial map of experimental datasets.

4.2.3 Datasets

To test our proposed solution, we identify times and locations in which we anticipate cellular overload, capture traces, and then compare network performance in those traces with baselines captured in the same location during normal operating conditions (when no network overload is likely to occur). We select spaces that are anticipated to have large gatherings but that are unlikely to be provisioned for large crowds (i.e. city streets as opposed to stadiums which typically have sufficient network capacity to handle crowds).

Our hypothesis is that during large crowds we will observe higher numbers of `RRCConnectionReject` messages than in times of regular operation. Overall, our dataset consists of over 3.2 million frames, with data collection that lasts for a cumulative duration of about 5.2 hours. While it is not possible to compute the exact number of UEs in the vicinity due to the lack of International mobile subscriber identity (IMSI) number in broadcast messages for security reasons, measuring the number of temporary unique UE IDs (`uniqueUeID`) in RRC Connection Requests allows us to estimate the number of active UEs present nearby.

St. Patrick's Day (SPD)

We collect cellular traces during the 2019 St. Patrick's Day parade adjacent to Balboa Park in San Diego, CA [238]. The parade was held on Saturday March 16^{th}, beginning at 10:00AM and ending around noon, while the public fair lasted through 3:30PM. We physically positioned our data collection devices within the crowd to better assess the eNodeBs serving this particular region as shown in (Figure 4.2(a)). The total duration of data collection is about 76 minutes, which resulted in over 1.1 million LTE frames. We observe 27,349 `uniqueUeID`s.

St. Patrick's Day Baseline (SPD_base)

As a point of comparison for the SPD dataset, we again gather LTE traces from the same location, from 8pm to 9pm on Tuesday March 26^{th}. Collection in the evening on a weekday helped us to avoid unexpected large gatherings in the many venues of the park, while still capturing activity of local nightlife. Compared to the *SPD* dataset we expect this dataset to exhibit low levels of overload, acting as a baseline for the location. Indeed, we see about 6,992 `uniqueUeID`s. We collect a little over 275K frames in 65 minutes.

ShamROCK Concert (CSR)

We collect traces from the ShamROCK concert in the downtown area of San Diego [239] on March 16^{th}. The event started at 7:00 PM and lasted until midnight. We collect 113 minutes (˜1.7 million frames) of traces during this time perioe. This event/location combination (as shown in Figure 4.2(b))100 was selected because we anticipated that the amount of cellular traffic during the event would well-exceed the typical traffic load. Because this location (city streets) does not typically have large crowds, we expect there to be network overload during a large event. This dataset contains 42,433 `uniqueUeID`s.

ShamROCK Concert Baseline (CSR_base)

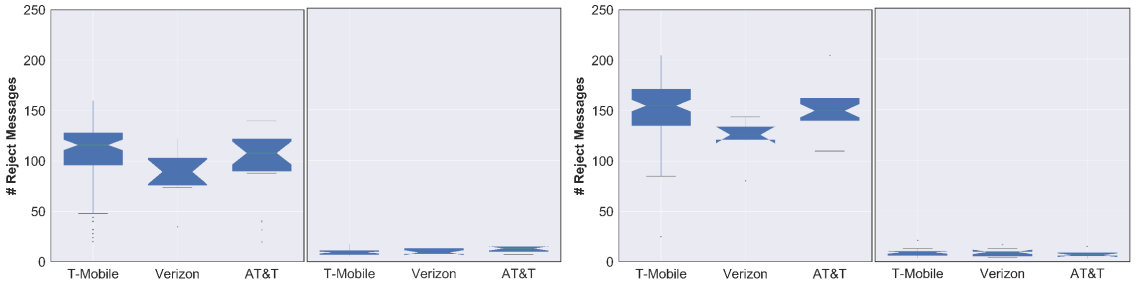
As a baseline to the *CSR* dataset, we capture additional traces (˜135K frames) in the same location from 9:30pm to 10:30pm on March 26^{th}, when the number of pedestrians and amount of vehicular traffic was more representative of normal operating hours. We detect only 3,338 `uniqueUeID`s during this data capture.

4.3 Evaluation

We analyze five RRC elements: (a) `RRCConnectionReject`, (b) `wait- Time`, (c) `RRCConnectionRequest`, (d) `cellBarred` signal and (e) number of SIB1s transmitted (#SIB1). Collectively, we refer to this data as "RRC metrics". We plot the values of these RRC metrics over thirty-second bins. We found that thirty-second bins were appropriate for our analysis because smaller time bins had little to no relative variation between the samples; however, we missed out on important data points when we used sixty-second or larger bins. In our evaluation, we observe that the rate of transmitted `RRCConnectionReject` messages is considerably higher in SPD and CSR than their respective baselines, in accordance with our initial hypothesis. Further, we discover an increase in *cellBarred* signals and $waitTime$ values in overloaded datasets (i.e., SPD and CPR).

4.3.1 Rejects

According to [227], an eNodeB may send an `RRCConnectionReject` in response to the UE's `RRCConnectionRequest` for exactly one of the following three reasons: (i) the eNodeB is overloaded (e.g., severe increase in requesting UEs that the eNodeB cannot accommodate); (ii) the necessary radio resources for the connection setup cannot be

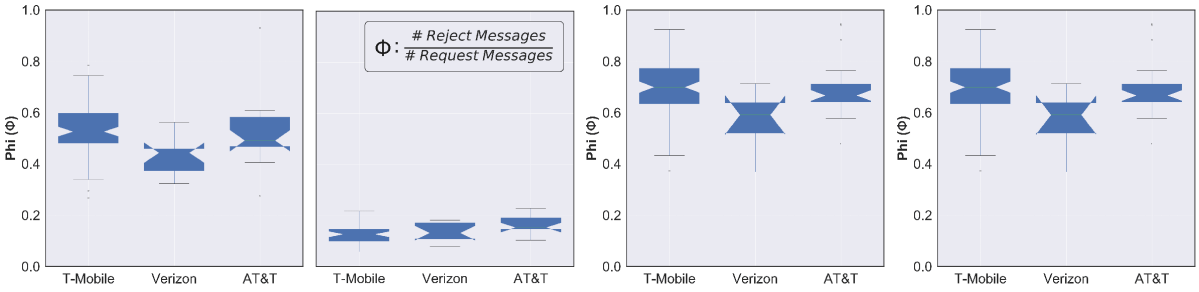


(a) SPD (b) SPD_base (c) CSR (d) CSR_base

Figure 4.3: Number of `RRCConnectionReject` messages transmitted in thirty-second bins.

provided (for instance, damaged equipment on eNodeB that results in limited access to the core network); or (iii) the Mobility management entity (MME) is overloaded. The MME is the key control-node for the LTE access network, which serves several eNodeBs. It is in charge of all the control plane functions related to subscriber and session management. Once the MME detects overload, it transmits an `overload start` message to the affected eNodeBs, signalling them to reject connection request messages that are for non-emergency and non-high priority mobile originated services.

Analysis of the reject messages sent over a fixed time interval can quantify the level of overload in the network. Figure 4.3 illustrates the average number of reject messages transmitted in thirty-second bins. As predicted, we notice significantly more reject messages in the SPD and CSR datasets. Figure 4.3(a) indicates that there are, on average, eight times more reject messages during the SPD parade compared to the SPD baseline (Figure 4.3(b)). Similarly, we observe a fifteen-fold increase in reject messages in Figure 4.3(c) as compared to Figure 4.3(d). This significant increase in reject messages is a clear indication of an increase in cellular network utilization.



(a) SPD (b) SPD_base (c) CSR (d) CSR_base

Figure 4.4: Phi (Φ) measure in thirty-second bins.

4.3.2 Phi (Φ) Measure

To better understand how overload levels vary, we examine a normalized second-order metric. We define the Phi (Φ) measure as the ratio of the number of `RRCConnectionReject` messages to the number of `RRCConnectionRequest` messages. Once again, we choose a bin size of 30 seconds. The Phi measure provides an indication of the severity of overload, as it reflects the percentage of new users who were unable to connect to the network. In future studies, we wish to examine the temporal variation in Phi (or the number of new users that were rejected) in order to quantify the maximum acceptable load threshold in eNodeBs.

As expected, there is a considerable difference between overloaded datasets (i.e., SPG and CSR) and their respective baselines. Figure 4.4(a) shows that Phi is as much as three times that in Figure 4.4(b). This difference is even more pronounced in Figure 4.4(c), where Phi is more than seven times that in Figure 4.4(d). This trend is similar to what we observed in Section 4.3.1. It is also indicative of the relationship between the number of UEs (# `uniqueUeIDs`) to the tendency towards network overload, as is expected.

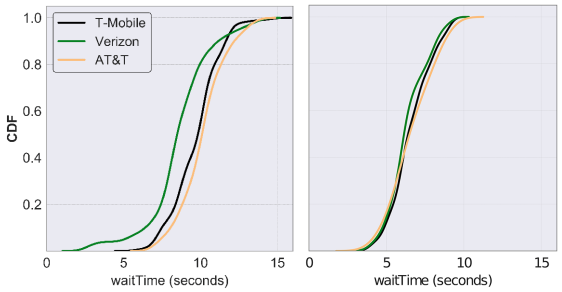
4.3.3 Average `waitTime`

When we compare the average `waitTime` across datasets in Figure 4.5, we observe that SPD and CSR have longer `waitTime`s than their baselines. We also see that AT&T performs worse in SPD, closely followed by T-Mobile. In CSR, T-Mobile appears to perform slightly worse than AT&T. Verizon, however, shows lower `waitTime` in all of the datasets. Note that the sample sizes of these distributions are proportional to the number of reject messages, as shown in Figure 4.3. Nevertheless, all of the telecom providers transmit longer `waitTime`s during increases in traffic demand.

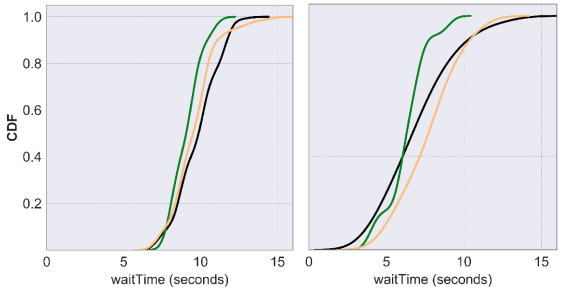
Longer `waitTime` in SPD and CSR is perhaps explained by the high proportion of UEs (# `uniqueUeID`s) in the given location. If the magnitude of UEs is great enough to result in overload, eNodeBs start to curtail overload conditions by engaging proprietary mitigation schemes, one of which is transmitting longer `waitTime`. The overall result is a confirmation of our hypothesis that these messages and parameter values can be used to infer overload. The comparison is noteworthy as it supports our earlier results where we compute `RRCConnectionReject` messages. Average `waitTime` serves as an additional indicator of overloaded eNodeBs.

4.3.4 Omega (Ω) Measure

In addition to the reject messages and their corresponding `waitTime`, `cellBarred` status is a crucial parameter that can indicate overload in an eNodeB. The `cellBarred` status transmitted within a SIB1 message indicates that the UE is not allowed to camp on a particular cell. We suspect that during overload conditions, cells can initiate load balancing by systematically preventing UEs from anchoring on them. In order to evaluate our theory, we analyze `cellBarred` messages to compare the percentage of these messages in our datasets.

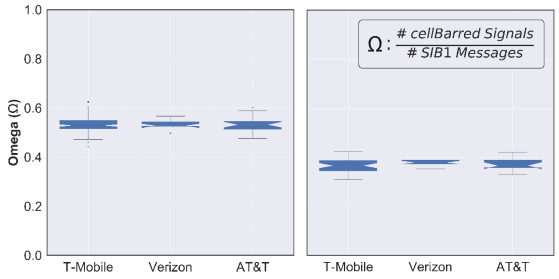


(a) SPD (b) SPD_base

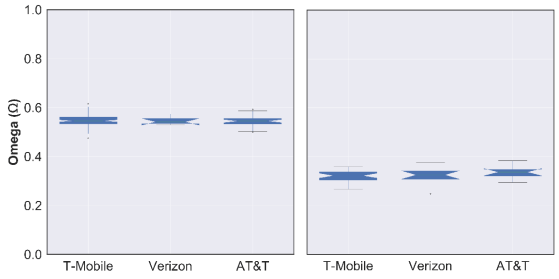


(c) CSR (d) CSR_base

Figure 4.5: Distribution of average `waitTime`.



(a) SPD (b) SPD_base



(c) CSR (d) CSR_base

Figure 4.6: Omega (Ω) measure in thirty-second bins.

The Omega (Ω) metric allows us to measure the ratio of $cellBar-red$ signals transmitted to the number of SIB1 frames received, in thirty-second bins. We use this second-order metric to establish a correlation between Omega and overload. Figure 4.6 depicts the variation in Omega across all datasets. We observe an increase of 20% in SPD and CSR datasets over their respective baselines. This indicates a relationship between cell barring signals and overload, confirming our hypothesis. However, it is interesting to observe that each of the MNOs (i.e., T-Mobile, Verizon and AT&T) have comparable Omega values in SPD and CSR, even though they exhibit noticeably different trends in Figures 4.3 and 4.4. That is because the inherent load-handling capacity of eNodeBs, as well as the density of users served, apparently differs. This suggests that overloaded eNodeBs operating in disparate network conditions prefer to consistently reject incoming connection requests rather than broadcast unavailability (through cell barring messages), regardless of their proprietary overload mitigation schemes.

4.4 Conclusion

In this chapter, we proposed a novel method to assess overload in nearby LTE eNodeBs, utilizing off-the-shelf hardware and without requiring cooperation of the cellular provider. Our analysis offers convincing evidence that messages broadcast by the eNodeB can be used to detect cellular overload using passive monitoring. In future work we will explore how passive overload inference can be leveraged in a system for automated overload mapping using ground-based data collection and UASs, independent of collaboration from a cellular provider. Such tools can be leveraged by regulators and policy makers and allow targeted deployment of alternative communication channels.

Lack of coverage, damaged infrastructure, and overloading are not just a problem for UEs. In rural regions, and in crisis conditions WSNs used for situational awareness and modeling are similarly effected. In the next chapter we shift our focus from communicating with human held devices, to IoT devices and sensors.

Acknowledgment

This work was done in collaboration with Vivek Adarsh, Ellen Zegura and Elizabeth Belding.

Chapter 5

Optimizing 802.15.4 Outdoor IoT Sensor Networks for Aerial Data Collection

The value of Internet connectivity is not limited to human held devices. As we have discussed previously, IoT is a booming industry, with 802.15.4 connected devices showing a large growth. Connected sensors monitoring environment and infrastructure can be invaluable to providing modeling and situational awareness before, during and after disasters. While 802.15.4 can provide local connectivity, in most cases the data is most valuable when it makes it way to the Internet. A promising solution is the development of FUSNs, which utilize UASs to incorporate DTN approaches to WSN data delivery.

Unlike the 802.11 standard, the 802.15.4 radio standard emphasizes energy performance over data rates. 802.15.4, as well as derivative standards, such as Zigbee [18], WirelessHART [46], and Thread [240], are widely used for applications requiring low-rate, low-power communication, such as those utilizing IoT sensors in rural outdoor WSNs. The interaction between these transmission standards and aerial systems is not

well understood, but is essential to the feasibility of 802.15.4 FUSN for a variety of applications, including disaster management, environmental monitoring and precision agriculture [241, 16, 17, 242].

In this chapter, we examine the IEEE 802.15.4 LR-WPAN standard [243] and its use in aerial vehicles. We identify elements critical to successful data collection from an 802.15.4 2.4 GHz network using a moving quad-copter. Our work provides three key contributions: (1) We conduct performance measurements of the RSSI and PRR by evaluating the impact of variables, such as altitude, displacement, antenna orientation, obstruction, transmission rate, and transmitter elevation. (2) We use the results to model the independent impact that each variable has on RSSI and PRR. (3) We show how network configuration can be used to model and predict network performance.

Our work suggests that optimal network performance in an outdoor rural 802.15.4 2.4 GHz network can be achieved by flying the UAS at altitudes of approximately 45 m with the receiver mounted parallel to the ground. The orientation of the transmitter does not have a significant impact on reception. Further, we find that RSSI is a poor indicator of overall network performance.

5.1 Methods

Our experiment comprised approximately nine hours of experimental flight data collected from an outdoor aerial testbed, near the University of California, Santa Barbara in March 2019. We deployed four identical IoT transmitters utilizing the 802.15.4 standard, broadcasting packets at 500 ms intervals at three different locations. To collect the data, we flew a UAS at varying altitudes and distances. The experimental setup is shown in Figure 5.1. We conducted three repeated runs, on multiple days, at each location. Each experiment comprised approximately one hour of flight, divided across thirteen altitudes.

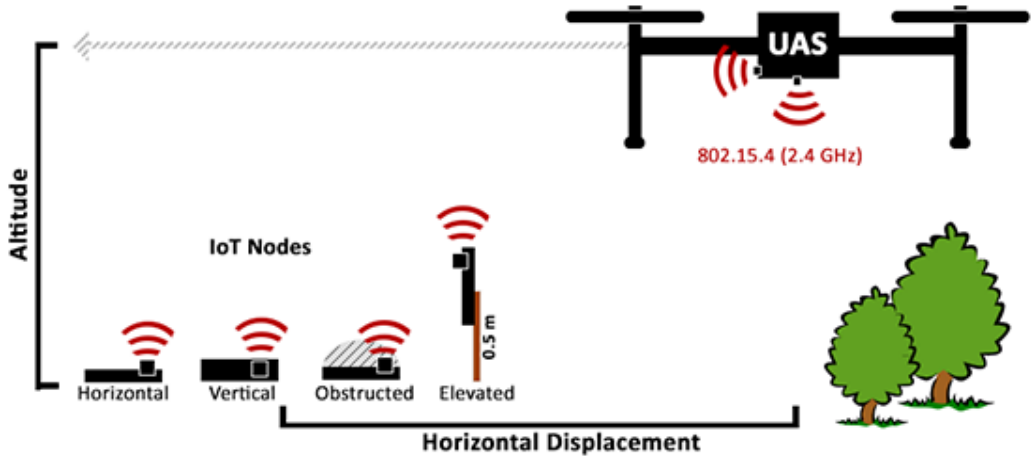


Figure 5.1: 802.15.4 aerial testbed system overview.

5.1.1 Equipment

Our equipment consisted of four transmitters on the ground, each transmitting packets received by a single UAS. To communicate, they utilized a total of six Digi WRL-15126 XBee3 RF 2.4 GHz transceivers using PCB antennae implementing the 802.15.4 standard. The specifications for these transceivers advertise an outdoor range of 1200 m at a power of 8 dBm and a receiver sensitivity of −103 dBm [244]. We utilized the XBee3 2.4 GHz model (as opposed to 900 MHz or 868 MHz models) to make our work comparable to previous research, including work that studies 802.11 at 2.4 GHz [138, 135, 136, 139].

Transmitters

For transmitters, we used four XBee3 radios mounted on SparkFun XBee Explorer boards controlled by a SparkFun Teensy LC. The transmitters were powered by external USB battery packs from varying vendors via a USB-to-Serial converter on the Teensy LCs. The transmitters were configured to broadcast 23 byte packets, every 500 ms. The packets consisted of a randomly generated floating point number to simulate sensor data, as well as unique device and packet identifiers.

(a) Horizontal. (b) Vertical. (c) Elevated 0.5 m. (d) Obstructed.

Figure 5.2: Equipment deployed in the aerial testbed.

Before each experiment, the transmitters were randomly distributed in a line, approximately 11 m apart. Placement was chosen so that there was no obstruction within the 15 cm vicinity of the transmitter. A linear configuration was chosen to ensure the UAS (flown in parallel over the transmitters) covered approximately equal horizontal displacement ranges for all transmitters. The latitude and longitude of the transmitter were recorded manually from repeated readings using a smartphone GPS. Each of the four transmitters was deployed a different configuration, as shown in Figure 5.2, in order to evaluate the effect of the antenna orientation, elevation, and obstruction:

- **Horizontal**: Laid flat, parallel to ground.
- **Vertical**: Placed on edge on ground.
- **Elevated**: Mounted to pole 0.5 m above ground.
- **Obstructed**: Laid flat under 1 quart of debris.

Figure 5.3: DJI Matrice 100.

Unmanned Aircraft System (UAS)

The aerial data collector was a DJI Matrice 100 quad-copter, show in Figure 5.3. Two XBee radios, set only to receive packets, were mounted to the bottom of the UAS with the *horizontal receiver* parallel to the ground and the *vertical receiver* perpendicular to the ground. The Matrice 100 communicates with a remote control at 5.725–5.825 GHz, which is outside the frequency range of the 2.4 GHz XBee nodes. The UAS was flown with no on-board camera. A Raspberry Pi (RPi) 2—Model B served as an onboard computer. The two XBees forwarded packets to the RPi via USB connections. The location of the UAS was recorded from the Matrice 100 onboard GPS, sampling at a rate of 50 Hz and using a UART connection to the RPi.

We flew the UAS over the transmitters in a straight line, at an average speed of 2.2 m/s. The exact flight path and speed varied due to manual execution under varying wind condition on multiple days. Each flight consisted of 13 altitudes (in relation to ground level at the lowest transmitter) of 9m, 12m, 15m, 18m, 21m, 24m, 27m, 30m, 46m, 61m, 76m, 91m, and 122m. Altitudes were originally chosen in feet and converted to meters for this publication, as the FAA regulates altitudes in feet. The maximum altitude of 122m corresponds to the FAA max altitude limit of 400 feet. Each flight varied in total maximum horizontal displacement, but exceeded a minimum horizontal radius of 250m from the closest transmitter to the UAS .

5.1.2 Location

The experiments were conducted at Coal Oil Point UC Reserve, a coastal grassland reserve near the university. The area is relatively flat with minor obstruction due to tall grass and bushes. In the grassland, we experimented in three locations with varying topography:

- **Road**: Transmitters were deployed along a 200 m section of a flat dirt road. The area had the lowest level of natural obstruction among the three experimental sites.
- **Grassy**: Transmitters were deployed in a field with tall grass and nearby bushes, ≈1 m tall, but with the immediate 15 cm around each transmitter unobstructed.
- **Hills**: Transmitters were deployed on the uneven terrain of hills with a shallow trench, (≈0.5 m deep), cut out by erosion. Tall grass and a denser concentration of bushes were prevalent, but the immediate 15 cm surrounding the transmitter was unobstructed.

5.1.3 Modeling RSSI

In our analysis, we analyze the distributions of RSSI by experimental variables that we hypothesize will impact RSSI (displacement, altitude, location, and transmitter/receiver configuration). To evaluate the independent influence of the variables on RSSI, we use a Generalized linear model (GLM) with robust variance estimation. We set RSSI as the outcome variable and natural log of displacement, altitude (13 groups), and configuration group (8 groups) as fixed covariates while also controlling for location (3 groups). We present the results in Section 5.2.1 and the resulting variable estimates in Table 5.1.

5.1.4 Measuring Packet Reception Rate

Physical layer metrics, such as RSSI, are used to passively infer the expected performance of a network. As outdoor IoT applications are delay tolerant in nature not all metrics of always-connected wired and wireless network performance apply. Time-based metrics utilizing *latency* and *jitter* are inappropriate for this application, as the UAS might be acting as a DTN node. While *throughput* is a common metric for 802.11 networks, it may be inappropriate for 802.15.4, since typical IoT applications do not saturate a network's bandwidth. Instead, they optimize for low-power consumption, especially in the context of outdoor sensor networks that may lack access to the power grid. We therefore examine the packet reception rate, which is the number of packets received divided by the number sent. Because each packet loss is wasted energy, PRR is a more appropriate network performance metric for this type of application.

$$\text{PRR} = \frac{\text{\# of packets received}}{\text{time in sector} * \text{transmission rate}} \tag{5.1}$$

We group the experimental data by horizontal displacement into concentric circular sectors radiating out from each transmitter (10 m wide for heat maps, and 25 m wide for models), keeping other experimental variables (altitude, receiver/transmitter configuration, and location) separated. To determine the sector into which a packet from a particular transmitter falls, we compare the UASs high frequency (50 Hz) on-board GPS with the manually recorded transmitter location. To estimate the number of packets sent by a transmitter, we calculate the product of the pre-programmed transmission rate and the time-in-sector by the UAS . We drop groupings where fewer than 5 packets were sent.

5.1.5 Predicting Packet Reception Rate

When planning a deployment, the ability to estimate the expected PRR for a network configuration and drone flight path is critical to ensuring that aerial data collection is successful. To investigate this possibility, we model the expected mean PRR based on our experimental variables.

We evaluated Poisson regression, Zero Inflated Poisson, Negative Binomial, and ZINB as possible models. By running the Vuong's closeness test, we found that ZINB was the best fitting model, as it accounts for the over-dispersion due to high number of PRRs at zero from locations and altitudes that never receive a packet. We therefore modeled both the chances that a packet is received at all, and the estimated number of packets received. We assessed the goodness of fit for our ZINB model by Scaled Pearson Chi-Square criteria, which was close to one and by the Full Log Likelihood criteria. We also compared the observed relative frequencies of the various counts to the maximum likelihood estimates of their respective probabilities. We found that our model was a good fit for the observed data.

The input data to the model were grouped by displacement bins, as described in Section 5.1.4. Before analysis, we randomly divided these data into three three sets: 50% were allocated as a *training set*, 40% were allocated as a validation set, and the remaining (10%) were the *test set*. Roughly 59% of the observations received at least one packet. We used a ZINB model with the number of received packets as the outcome. We set displacement as a categorical variable with seven groups (<50 m, (50 m, 75 m], (75 m, 100 m], (100 m, 125 m], (125 m, 150 m], (150 m, 175 m], and > 175 m), altitude as a categorical variable with 11 groups (9 m, 12 m, 15 m, 18 m, 21 m, 24 m, 27 m, 30 m, 46 m, 61 m, and 76 m), and transmitter–receiver configuration with eight groups (Vertical and Horizontal Receivers paired with Horizontal, Elevated, Obstructed and

Vertical Transmitters) as fixed covariates for both parts of our model. We left out location as it was not a significant variable in this model. We used a natural logarithm of sent packets as an offset in the NB part of the model and control for the number of of sent packets in the ZI part of the model. As before, we dropped bins where fewer than five packets were sent.

5.2 Results

Our results comprise nine hours of collected data, totaling 121,503 received packets, and include only the experimental portion of each flight; landing, takeoffs, and transitions between experiments are omitted. The dataset is available online [245].

5.2.1 RSSI

Past measurement studies of UAS-ground communication have focused on RSSI as a key indicator of performance [126, 122, 123]. RSSI is a common signal strength indicator, often the only signal metric reported by radio modules. Past work has shown that reported RSSI is proportional to the actual received signal strength [246]. We therefore begin our analysis by examining how RSSI is affected by varying experimental variables. Given the literature, we expect:

- **Altitude/Displacement**: Receivers should have the best reception in proximity to a transmitter.
- **Location**: Obstacles introduce interference, so the unobscured *road* should have the best signal.

- **Transmitter/Receiver Configurations**: Horizontal transmitters and receivers should have the best reception. Obstructions should decrease reception, while elevating equipment should increase it.

As described in Section 5.1.3, we model RSSI using a GLM . This allows us to examine the independent influence of our experimental variables on RSSI. We summarize the model estimates in the left part of Table 5.1, predicted values of RSSI with confidence intervals from the GLM . For a baseline, we set what we consider reasonable variable choices for a realistic IoT UAS collected deployment: an elevated transmitter in a location with low obstruction, flying the UAS 101–125 m away, at an altitude of 45 m above ground level, using a horizontal receiver. In our analysis, we contrast other variable choices to this baseline.

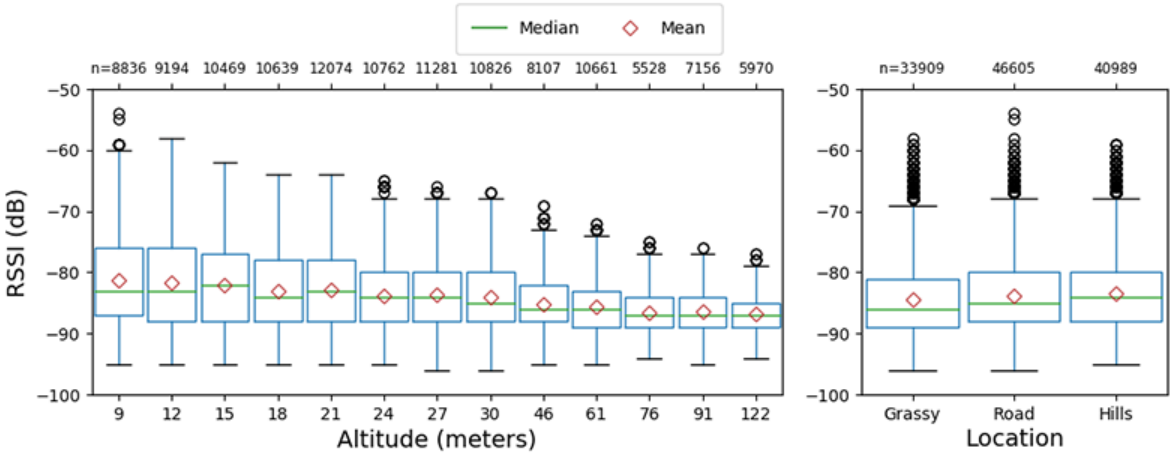


Figure 5.4: Observed RSSI distribution. Grouped by altitude (left) and location (right).

Table 5.1: Estimates by parameter for GLM and ZINB models.

	Parameter	RSSI GLM Model			PRR ZINB Model		
		Estimated Mean RSSI	Mean Confidence Interval		Estimated Mean PRR	Mean Confidence Interval	
Baseline	Rec. Horiz. Trans. Elev. Alt. 150 ft D. 101–125 m Loc. Road	−87.2114	−87.1066	−87.3162	0.204	0.1777	0.2341
Rec. H.	Trans. Horiz.	−86.5816	−86.4676	−86.6956	0.1528	0.1319	0.1769
	Trans. Obs.	−86.8098	−86.6867	−86.933	0.1369	0.1175	0.1595
	Trans. Vert.	−86.5092	−86.3997	−86.6187	0.1865	0.1625	0.214
Rec. V.	Trans. Horiz.	−89.2424	−89.1179	−89.3668	0.1218	0.1046	0.1418
	Trans. Elev.	−89.6226	−89.503	−89.7422	0.125	0.1082	0.1445
	Trans. Obs.	−90.155	−90.0124	−90.2977	0.0877	0.0746	0.1032
	Trans. Vert.	−89.2804	−89.1584	−89.4023	0.1203	0.1038	0.1395
Altitude	9 m	−84.1888	−84.0457	−84.3319	0.1859	0.1626	0.2126
	12 m	−84.2479	−84.107	−84.3889	0.2025	0.1779	0.2304
	15 m	−84.5907	−84.4674	−84.714	0.2147	0.1884	0.2446
	18 m	−85.2649	−85.1463	−85.3836	0.2117	0.1859	0.2411
	21 m	−85.1696	−85.0613	−85.2778	0.2025	0.1781	0.2302
	24 m	−85.8722	−85.7614	−85.983	0.2234	0.1971	0.2533
	27 m	−85.6163	−85.5125	−85.7201	0.21	0.1854	0.2378
	30 m	−85.7763	−85.6692	−85.8834	0.213	0.1873	0.2422
	61 m	−86.8529	−86.7577	−86.9481	0.1897	0.1676	0.2147
	76 m	−88.1146	−88.006	−88.2232	0.1563	0.1354	0.1805
	91 m	−87.4706	−87.3676	−87.5736	-	-	-
	122 m	−87.7332	−87.6294	−87.837	-	-	-
Displacement	$<$ 50 m	−82.1024	−81.9858	−82.2189	0.5518	0.4923	0.6185
	50–75 m	−85.3911	−85.2856	−85.4966	0.3116	0.2746	0.3535
	76–100 m	−86.4356	−86.331	−86.5402	0.258	0.2276	0.2925
	126–150 m	−87.8248	−87.7193	−87.9303	0.1984	0.1715	0.2295
	151–175 m	−88.3346	−88.2282	−88.441	0.1535	0.1302	0.1809
	$>$ 175 m	−88.9355	−88.8277	−89.0434	0.1271	0.1036	0.1559
Loc.	Grassy	−88.2025	−88.0827	−88.3222	-	-	-
	Hills	−87.2797	−87.1742	−87.3853	-	-	-

This table shows the mean estimates for possible experimental variable choices for the two models, using the top row's variable choices as a baseline. Each subsequent row reflects the effect of changing that one variable while keeping all other variables to their baseline settings (Horizontal Receiver, Elevated Transmitter, Altitude of 46 m, Displacement of 101–125 m, at the Road location). Altitudes $>$ 76 m and locations are not in the PRR ZINB model. All p-values for reported data are $<$0.0001.

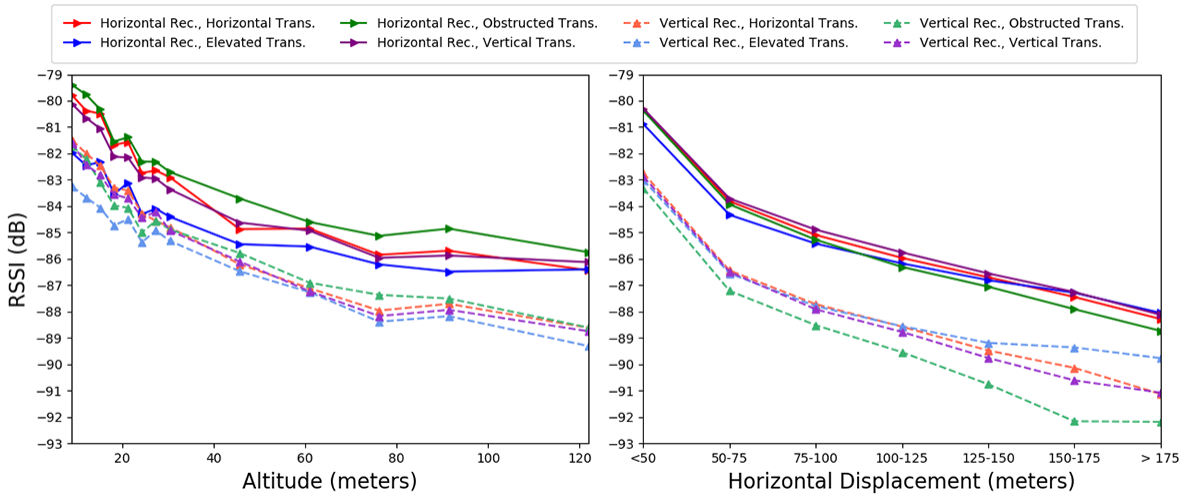


Figure 5.5: Predicted mean RSSI from GLM. Grouped by altitude (left) and horizontal displacement (right).

Altitude

We present the distribution of the reported RSSI by altitude from ground level at the location of take-off in Figure 5.4 (left). As only the radio module reports RSSI, the data represent only packets successfully captured and decoded by the radio module. We observe that higher altitudes have fewer high RSSI values. However, the mean and median values decrease only slightly as altitude increases.

We examine the predicted effect of altitude on RSSI from the GLM averaging across displacements and locations, as presented in the Figure 5.5 (left). As expected, we find that, for all receiver/transmitter configurations, RSSI decreases with altitude; however, this mean decrease is small (<1 dB per 3 m) compared to the wide fluctuations in individual observed RSSI measurements. In Table 5.1, altitude changes the RSSI by a maximum of 4 dB and the trend is not monotonic.

Location

We present the distribution of RSSI for the three locations in Figure 5.4 (right). The mean and median RSSI values remain similar across the three locations. However, counter to expectations, the most obstructed site (Hills) displays the highest median RSSI, and the area of least obstruction (Road) yields the greatest variance in RSSI values, although, once again, the difference in dB is small. When examining the predicted values from GLM, we observe a different pattern with the Road displaying the optimal RSSI but with 1dB of difference.

Horizontal Displacement

Averaging across altitudes and locations, we examine the predicted impact of displacement on RSSI from the GLM . As shown in Figure 5.5 (right), we observe that greater horizontal displacement reduces RSSI slightly (approximately 2 dB per 30 m) for all receiver/transmitter configurations. When comparing displacements to the baseline in Table 5.1, we observe a 1 dB monotonic drop per 25 m displacement. Once again, we note that the mean decrease is small relative to the total observed fluctuations in individual RSSI measurements.

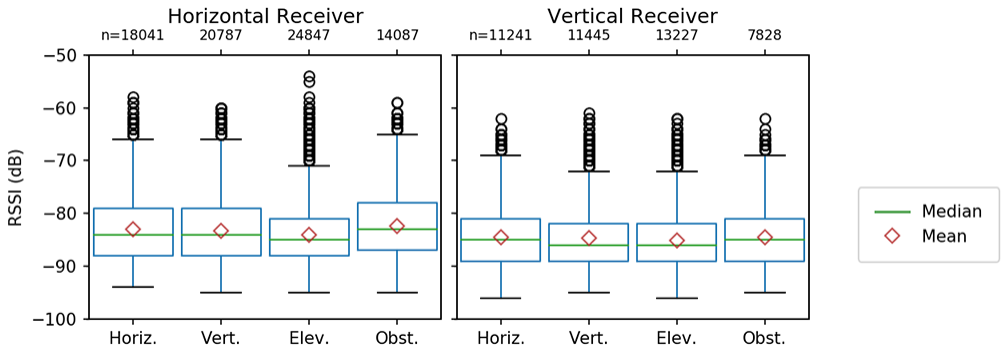


Figure 5.6: Observed RSSI distributions by configuration.

Transmitter/Receiver Configurations

We examine the distributions for data grouped by configuration of the receiver and the transmitter. In Figure 5.6, the left part of the graph represents packets collected by the *horizontal receiver* (mounted on the UAS parallel to the ground), while the right shows data collected by the *vertical receiver* (mounted perpendicular to the ground). For each receiver, the data are further categorized by transmitter configuration on the ground (Horizontal, Vertical, Elevated, and Obstructed).

For the observed distributions, *transmitter* and *receiver* orientation minimally affect mean RSSI. However, *receiver* orientation has a slightly more pronounced impact. All four transmitters had lower observed mean RSSI for the *vertical receiver* than the *horizontal receiver*. These observations are consistent with the predicted values from the GLM . With regards to the effect of *horizontal displacement* on receiver/transmitter configuration in Figure 5.5 (right), *receiver* orientation slightly impacts RSSI, while *transmitter* orientation minimally impacts RSSI. Mirroring this result in Table 5.1, receiver orientation changes RSSI by approximately 3 dB, while the transmitter orientation makes little difference.

When examining the impact of *altitude* on receiver/transmitter configuration performance, we identified a more complicated interaction for the mean RSSI predicted by the GLM in Figure 5.5. While the horizontal *receiver* still performs better overall, the *transmitter* behaviors are contrary to expectations. The obstructed transmitter paradoxically outperforms the elevated transmitter for all altitudes.

The similarity in RSSI between transmitter configurations is contrary to expectations, especially comparing the elevated transmitter, with superior line of sight, to the obstructed transmitter, which is buried in debris. For the observed distributions in Figure 5.6, all transmitters broadcast at the same rate (one packet per 0.5 s), yet the *elevated* transmitter delivers a greater number of packets than the *obstructed* one (in the horizontal case, 10 k

more). Therefore, the obstructed transmitter is in fact delivering fewer packets despite showing an overall better RSSI.

Deviations from Expectations

Overall, the observed RSSI does not match the expectations stated at the start of this section. When examining mean RSSI alone, one might conclude that the experimental variables minimally impact network performance. Further, the slight improvement in RSSI due to obstructing the transmitter proves yet more confusing.

A high RSSI reflects successfully received packets, while lost packets are unaccounted for in the data. Their RSSI is never reported to the receiver module. Therefore, the mean of the received RSSI remains relatively consistent despite changes to experimental variables, especially when compared to the fluctuation in RSSI for fixed experimental variables, due to outside factors, such as external RF interference and minor changes to reception geometry due to small variations in flight paths.

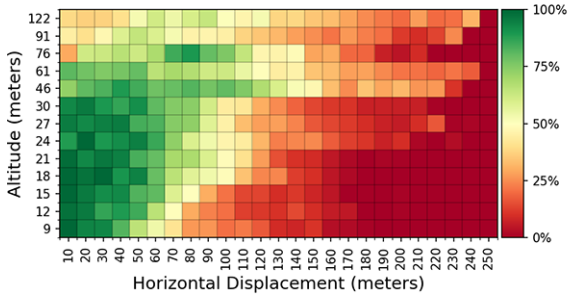
Therefore, for applications involving UAS data collection on 802.15.4, RSSI alone does not provide a complete picture of network performance for our data. As noted above, while our analysis of RSSI showed little mean fluctuation between experimental configurations, our total number of received packets indicates significant differences in network performance not accounted for by RSSI. We therefore must investigate further to adequately assess the performance of our network during aerial data collection.

5.2.2 PRR by Altitude and Displacement

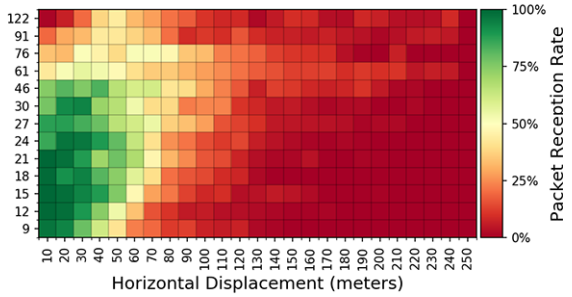
We examine the underlying network performance via the observed PRR. As introduced in Section 5.1.4, we grouped the measurements into 10m concentric circular sectors radiating from each transmitter, resulting in 29,161 grouped measurements. The observed PRR (averaged across runs and locations) is presented in Figure 5.7 with each heat map representing a transmitter/receiver configuration.

As expected, close proximity to the transmitter, in terms of both altitude and horizontal displacement, leads to a higher rate of received packets. A greater horizontal displacement from the transmitter has greater packet loss. When examining PRR, most configurations demonstrate a sharp drop in RSSI for distances over 150 m, while the mean RSSI displays a small decrease. While lower altitudes produce a better PRR at lower horizontal displacements, altitudes between 46 and 76 m show improved performance at greater displacements. Thus, a UAS need not fly low to the ground, where it is more likely to hit obstacles, in order to optimize data collection.

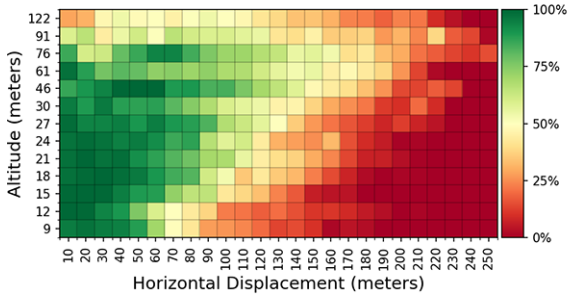
Likewise, *receiver* orientation noticeably impacts loss. For all transmitters, the *vertical receiver*'s PRR is the worst at higher altitudes and displacements. According to Figures 5.7e and 5.7f, the PRR is best for lower altitudes, where elevating the transmitter 0.5 m off the ground overcomes ground obstacles (such as tall grass) to establish line-of-sight with the UAS . Furthermore, Figures 5.7g and 5.7h show that, while the *obstructed transmitter* maintains a high PRR when the UAS is in proximity, the maximum horizontal displacement at which the UAS has good reception is lower than other configurations, especially at lower altitudes. This behavior is masked when examining RSSI.



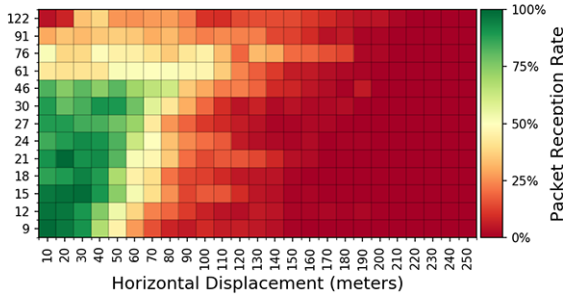
(a) Horizontal Trans. to Horizontal Rec.



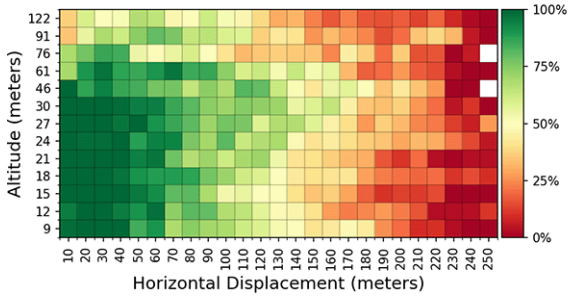
(b) Horizontal Trans. to Vertical Rec.



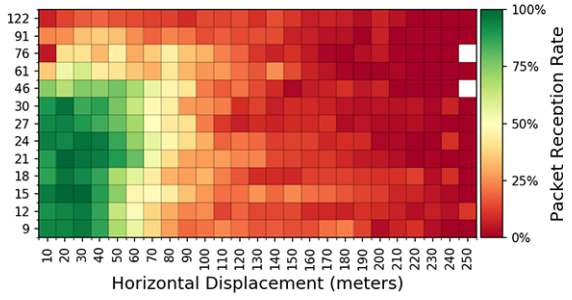
(c) Vertical Trans. to Horizontal Rec.



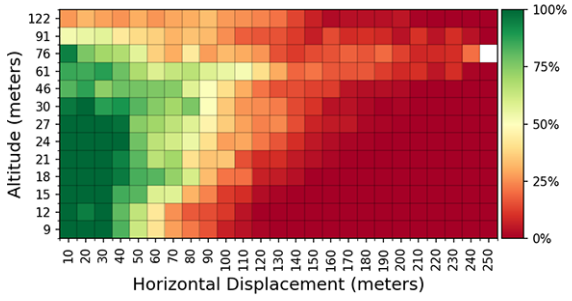
(d) Vertical Trans. to Vertical Rec.



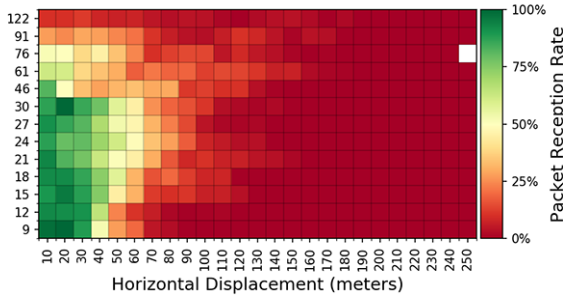
(e) Elevated Trans. to Horizontal Rec.



(f) Elevated Trans. to Vertical Rec.



(g) Obstructed Trans. to Vertical Rec.



(h) Obstructed Trans. to Vertical Rec.

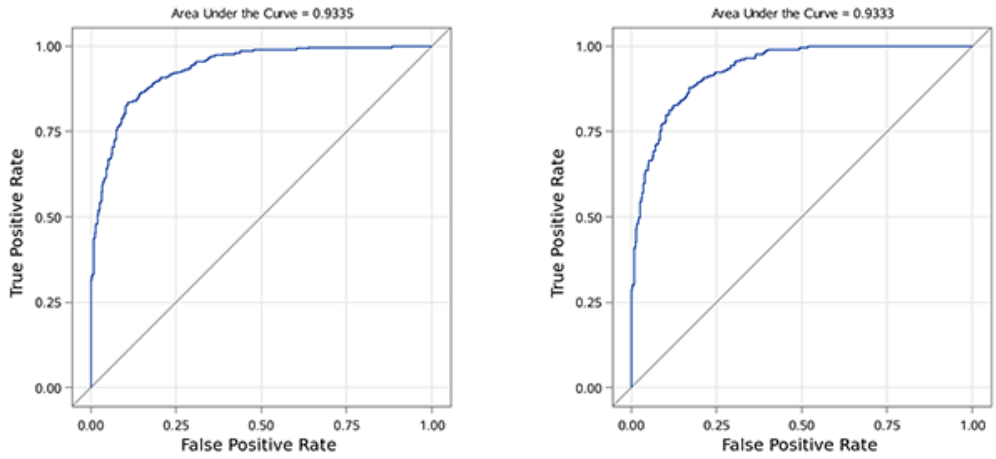
Figure 5.7: Observed packet reception rates grouped by altitude and 10 m displacements from transmitter. Cells with fewer than five sent packets are left blank.

5.2.3 Optimizing PRR

We model how experimental variables impact PRR using a ZINB model, described in Section 5.1.5. The input data consisted of 7,906 observations (divided by displacement bins of 25 m), and divided into a training set (3,910 observations), validation set (3,166 observations), and test set (830 observations). The model's estimated PRRs are presented on the right side of Table 5.1.

Using this model, we can predict (for a given altitude, horizontal displacement, and receiver/transmitter configuration) the expected mean PRR. This serves to help identify the conditions under which a packet of data can be delivered to the drone, as might be required, for example, when planning a deployment.

We can convert the ZINB model, which predicts PRR, into a binary classifier, predicting when at least one packet will be received, by specifying a threshold below which we expect the packet to be lost. This can be done in two ways. In Figure 5.8, we present a Receiver operating characteristic (ROC) graph from the *test set* displaying the possible true and false positive rates based on threshold choices. Figure 5.8a (left) employs the ZI part of the model to calculate the probability that zero packets are received and shows the corresponding ROC. This method does not incorporate the number of sent packets, and could be sensitive to transmission rates. Figure 5.8b (right) first calculates predicted PRR, incorporating number of sent packets in the modeling, and then shows the resulting ROC. For example, picking a PRR threshold of 40% for the *test set*, the resulting binary classifier correctly identifies 86.9% of the observations with a specificity of 86.2% and recall of 87.5%.



(a) Probability of Zero Packets. (b) Packet Reception Rate.

Figure 5.8: ROC curves for ZINB model estimates of test set.

Altitude

Based on the ZINB model of our training set, we examine the effect of altitude on the PRR. Figure 5.9 (left) shows the impact of altitude by group (averaging across displacements). For most transmitter/receiver configurations, PRR peaks at 46 m altitude. As observed in the Figure 5.7 heat maps, the impact of altitude seems tied to displacement. When fixing a displacement, as we do in Table 5.1, the altitude minimally effects PRR with single percent fluctuations until altitude exceeds 61 m.

Horizontal Displacement

We similarly use the ZINB model of our training set to examine the effect of horizontal displacement on the PRR. In Figure 5.9 (right), we show the impact of displacement by group (averaging across all altitudes). Here, the effect is more linear, with displacement causing a monotonic drop in PRR until no packet is received. When comparing to our baseline, in Table 5.1, we observe a nearly 5% drop in PRR per 25 m, after 50 meters.

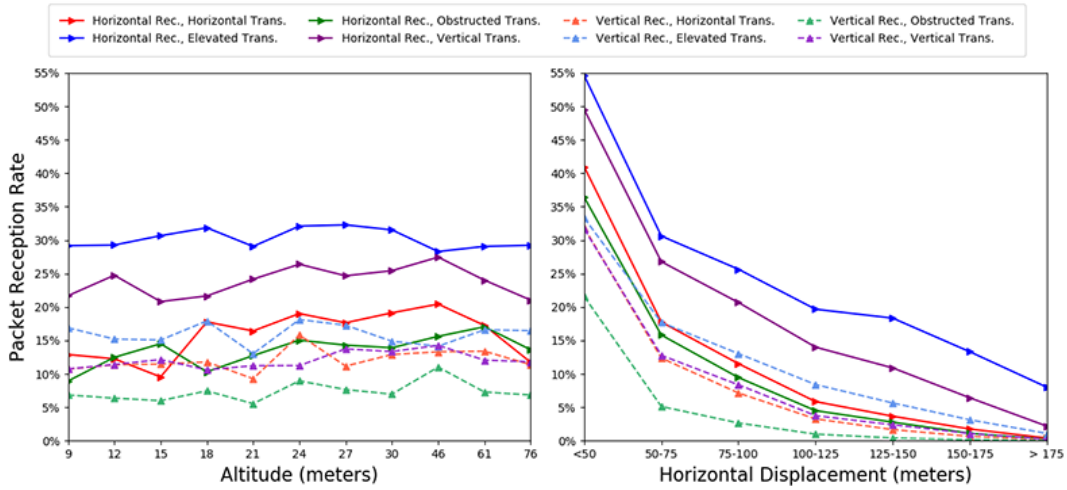


Figure 5.9: Predicted mean PRR from ZINB model for each receiver/transmitter configuration grouped by altitude (left) and horizontal displacement (right).

Transmitter/Receiver Configuration

As expected, the elevated transmitter has a significantly better PRR across all altitudes and displacements, while the obstructed transmitter has the worst. The horizontal and vertical transmitters behave similarly. The horizontal receiver shows better PRR for all groups. When compared to our baseline in Table 5.1, a horizontal receiver gives a 3%–8% increase in PRR. We similarly observe that obstruction causes a 2%–4% decrease in PRR compared to a horizontal transmitter, while elevating the transmitter gives a 5% boost but only in the case of a horizontal receiver.

Predicting PRR from RSSI

Predicting PRR from RSSI can be difficult because RSSI remains unreported when no packets are received (i.e., the PRR is zero). The ZINB model comprises two parts, a logistic regression model predicting the probability any packets are received and a negative binomial model predicting a PRR given that a packet is likely to be received.

To examine whether RSSI, acting as a mediator variable, can be used to model and

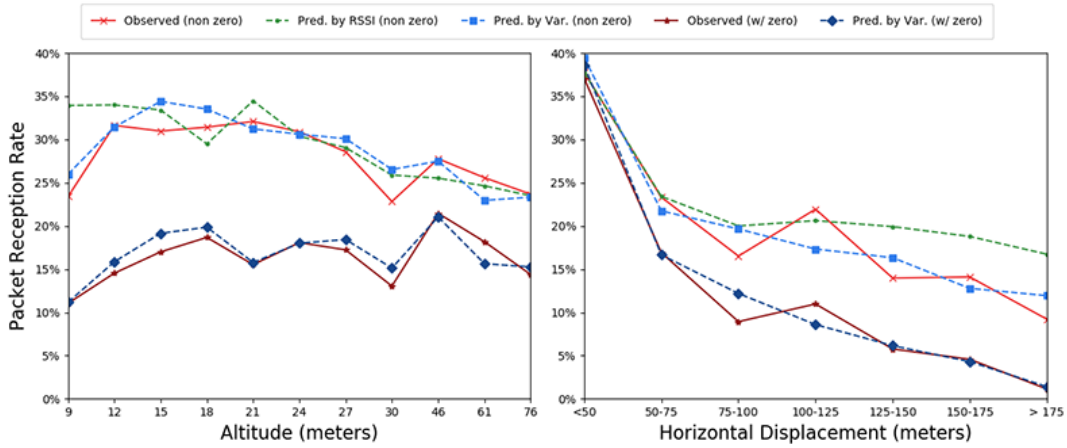


Figure 5.10: Accuracy of predicted mean PRR grouped by altitude (left) and horizontal displacement (right).

predict PRR, we used RSSI to train a negative binomial model on our *training set*. To properly compare it, we also reran the negative binomial portion of the model based on experimental variables for our *training set*, dropping PRRs of zero. We evaluate our model on the *test set*, without PRRs of zero. Figure 5.10 shows the results of the two models. We compare the predictions of those models to the observed PRR in the *test set* (including and not including PRRs of zero), as well as the result of the full ZINB model, which predicts using experimental variables.

When we ignore PRRs of zero and average across other variables, RSSI closely predicts PRR. However, the model that is based on experimental variables, rather than RSSI, is still a better fit to the observed data. While inferring PRR, in order to gauge how many re-transmissions may be required for successful data delivery, may be useful, doing so fails to predict whether packets will be received in the first place. When we compare the RSSI prediction to the observed readings that include PRRs of zero, RSSI over-predicts the real PRR by more than double. In contrast, the ZINB model, trained to predict total packet loss, is a good fit.

5.3 Discussion

Our measurement study provides insight into optimizing an IoT deployment for aerial data collection by a UAS . While existing literature focuses primarily on mean RSSI as a principle indicator for an aerial network's performance, our analysis reveals that RSSI inadequately reflects the network behavior. Because radio modules report RSSI only when a packet is successfully delivered, RSSI fails to capture steep drops in PRR. Instead, in open spaces without strong sources of interference, the geography of a network, such as distance from node and mounting of transmitters, can be used to estimate the expected performance. In this section, we discuss how, based on our results, one can optimize such an 802.15.4 network for aerial data collection.

5.3.1 Effective Reception Range

During initial planning of this measurement campaign, we expected to observe connectivity at distances far greater than those measured. Digi advertises an effective total distance of 1200 m for the XBee3, with the caveat that "[a]ctual range will vary based on transmitting power, orientation of transmitter and receiver, height of transmitting antenna, height of receiving antenna, weather conditions, interference sources in the area, and terrain between receiver and transmitter" [244]. However, in near optimal real-world conditions, we only successfully captured packets at a maximum total distance of 297 m in our experiments. When accounting for the UASs altitude, this corresponds to a 278 m maximum horizontal displacement from the transmitter, but the UAS was highly unlikely to receive a packet at this displacement.

5.3.2 Optimal Altitude

Although low altitudes generally improve PRR and RSSI, higher altitudes provide better connectivity at greater horizontal displacements. We theorize that higher altitudes provide a steeper angle between the UAS and transmitter, which reduces signal blockage from obstacles, such as trees and bushes. An altitude of 46 m provides the best overall reception. This is fortunate as a high altitude allows a UAS to clear most trees and ground obstacles.

5.3.3 Optimal Antenna Orientation

Previous work investigates how radiation patterns cause antennae orientations to affect signal quality. In our work, the antenna orientation of the *transmitter* did not significantly impact performance, while the orientation of the *receiver* had a far greater impact, with the vertically mounted receiver performing substantially worse than the horizontally mounted one.

In our case, this may indicate that signal radiation is a smaller factor than minor fluctuations in line-of-sight. Unlike past experiments that employed a 1.1 inch straight wire as an antenna for a 802.15.4 2.4 GHz CC2420 transmitter [139], our setup employed a coiled embedded antenna directly on the comparatively smaller XBee. While Lymberopoulos et. al. [139] studied distances of <10 m, our work includes signals at distances of >250 m where the topography and antennae geometry might reduce impact of orientation.

The better performance of the vertically mounted receiver could be due to the superior line of sight to the horizontally mounted receiver. The *horizontal receiver* was parallel to ground at all times, while the *vertical receiver*'s own body may have blocked the signal from unfavorable angles.

In further consideration of transmitter orientation, aerial networks might serve as auxiliary modes of connection to on-the-ground infrastructure. For example, cluster heads may use the same 802.15.4 radio to communicate with other nodes and the UAS. In such a case, tailoring ground transmitters to communicate with one another without taking communication with the UAS into account would be optimal.

5.3.4 Elevating Transmitters

In IoT deployments where elevating the transmitter is possible, it is advisable to do so. While the *elevated transmitter* had a very similar RSSI performance to the *horizontal transmitter* on the ground, the PRR of the elevated transmitter was much improved, especially for lower UAS altitudes. The transmitter cleared much of the ground-level obstruction and attained better line-of-sight to the UAS when elevated even half a meter above the ground. The optimal height for a transmitter may depend on the specific local geography.

5.3.5 Obstruction

Sensor nodes are frequently deployed in the field with little protection from extreme weather. Nodes and antennas can be affected by various obstructions, including dirt from rain or wind. While obstruction minimally affects reported RSSI, it significantly impacts PRR. Buried sensors communicate at a high loss rate. However, most packets are lost at greater distances, wasting transmission power. If the IoT device relies on solar power, which may likewise become obstructed, this exacerbates the issue.

5.3.6 Transmission Rate Selection

IoT deployments typically minimize the transmission rate in order to save power. However, we found that achieving this may be difficult when performing aerial data collection with a UAS, since the flight time on consumer multi-copters as of spring 2019 is approximately 20 min (1200 s) per battery. As flight speed correlates with successful data capture at a particular data transmission rate, multi-copter battery capacity constrains viable transmission rates. Fixed-wing aircraft are similarly constrained as they have a minimum flight speed of >10 m/s.

We flew the quad-copter at an average speed of 2.2 m/s (5 mph). On a single battery, at this speed, the UAS could cover 2.6 km. This is already a relatively small coverage area when accounting for a round trip flight—approaching the lowest feasible flight speed for UAS-based data collection.

At our low flight speed, both 500 ms and 1 s inter-packet transmission rates produced similar RSSI and PRR values. While we attempted slower transmission rates, we received too few packets for a meaningful analysis. At one packet per 15 s, we received an average of only 258 packets per transmitter–receiver pair across all locations, altitudes, and repeated trials. For one packet per minute, the average per transmitter-receiver pair was only 127 packets. Given the experimental results, determining the max flight speed that guaranteed delivery at low transmission rates is difficult, as the data is sparse.

This is a potential complication for aerial collection, as high transmission rates would correspond to increased power consumption. Strategies such as pre-scheduling collection windows or a signaling mechanism, during which the transmitters increase rate, could mitigate this issue.

5.4 Conclusion

In this study, we reviewed the effects of altitude, antenna orientation, obstruction, antenna elevation, and transmission rate on RSSI and PRR of a 802.15.4 2.4 GHz IoT network. We found that RSSI is a weak indicator of network performance for aerial data collection of an IoT network, as it poorly reflects high levels of packet loss. When examining reception rate, our experimental variables had far greater variability than RSSI showed.

An 802.15.4 network optimized for ground based communication may not be suitable for aerial data collection without modification and consideration of a variety of factors, such as device placement, altitude of collection, transmission rate, and resilience to obstruction. Our work provides a model for making informed choices about these factors.

As we discussed in this chapter, for this series of measurements we used Digi WRL-15126 XBee3 with PCB antennae. While this is a compact and often used form factor. While our work provides valuable insight for a large class of deployments, this work does not examine straight wire dipole antennas that are also sometimes utilized by WSNs. Past work suggests that these straight wire antennae may be tied to higher impact of the impact of antenna orientation [139]. In the next chapter, we study how antenna type impacts the performance of FUSNs.

Acknowledgment

This work was done in collaboration Ryan Allen, Irina Artamonova and Elizabeth Belding. Thanks to Coal Oil Point UC Reserve for allowing us to utilize your space for our aerial testbed.

Chapter 6

Impact of 802.15.4 Radio Antenna Orientation on UAS Aerial Data Collection

In chapter 5 we discovered that embedded coiled antenna orientation of the transmitters and the receivers had little impact on signal strength. However, we theorized that external antenna modules may be more sensitive to orientation and therefore might display behavior observed in 802.11 works, such as toroidal radiation.

In this chapter, we compare external antenna configurations to the commonly used embedded coiled antenna modules. We study the effects of toroidal radiation and antenna polarization on signal strength. We model our data using a ZINB model. For each hardware configuration and orientation we identify the optimal altitude to fly an UAS. Our results show that choosing antenna configuration (including type and orientation) for an IoT network depends on the intended UAS collection flight plan. For example, in our study, we find that UAS flights with a horizontal displacement to transmitters less than 150 meters optimize when using vertically oriented transmitters with internal

antennae and a UAS flight altitude of 150-250ft. On the other hand, UAS flights with a horizontal displacement to transmitters exceeding 150 meters optimize when using vertical transmitters with external antennae and are not sensitive to flight altitude.

6.1 Methods

The results of this paper are based on experimental data collected using an outdoor aerial testbed at Coal Oil Point Reserve near the University of California, Santa Barbara in March, September, and October 2019. We deployed 802.15.4 transmitters broadcasting packets at 500ms intervals to a mobile UAS. We varied radio hardware, antenna orientation, altitude, distance, and amount of obstruction. We performed multiple repeated measures, varying sensor placement and UAS flight path.

6.1.1 Equipment

The experiments used a single UAS and two sets of 802.15.4 Digi 2.4GHz Xbee3 radios, a set of transceivers with integrated antennae, and a set of transceivers with external antennae. The specifications for both sets of transceivers advertise an outdoor range of 1200m at a power of 8dBm and a receiver sensitivity of -103dBm [244]. To evaluate the choice of antenna type, we first conducted experiments utilizing radio modules with integrated antennae, then swapped for external antenna modules described in this section.

Integrated Antenna Modules: The first set of six transceivers comprised Digi WRL-15126 XBee3 using PCB antennae; we refer to these as *integrated antenna* modules. These radios are popular due to their compact form factor. Because the antennae are integrated into the circuit board, changing antenna orientation also rotates the entire radio module. For transceivers with integrated antennae, we define *horizontal orientation* as one where

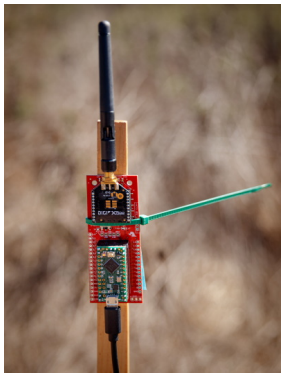
(a) Horizontal Internal.

(b) Vertical Internal.

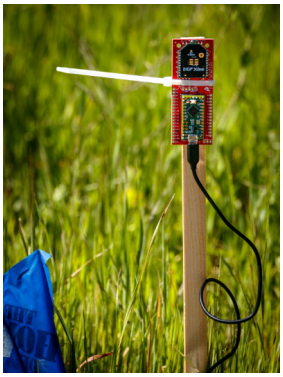
(c) Obstructed External.

(d) Horizontal External.

(e) Vertical External.



(f) Elevated External.



(g) Elevated Internal.

(h) UAS: DJI Matrice 100.

Figure 6.1: Equipment deployed in our aerial testbed.

the radio module circuit board is parallel to the ground, as shown in Figure 6.1a. We define *vertical orientation* as one where the circuit board is perpendicular to the ground, as shown in Figure 6.1b.

External Antenna Modules: The second set of six transceivers were Digi WRL-15130 XBee3 with an external WRL-00145 RP-SMA 2.2dBi Duck Antenna [247]; we refer to these as *external antenna* modules. The external antennae allow control of antenna orientation independent of the radio module. For transceivers with external antennae, we define *horizontal orientation* as one where the antenna is parallel to the ground, as shown in Figure 6.1d. We define *vertical orientation* as one where the antenna is perpendicular to the ground, as shown in Figure 6.1e.

Transmitters: For each set of Xbee3 radio modules, four of the six were designated as transmitters. These were mounted on SparkFun XBee Explorer boards controlled by a SparkFun Teensy LC, powered by external USB battery packs from varying vendors via a USB-to-Serial converter on the Teensy LCs. For both antenna variants, the transmitters were programmed to broadcast 23 byte packets every 500ms. The payload of the packets consisted of a randomly generated floating point number (simulating numerical data of a potential attached sensor), as well as device and packet identifiers.

At the start of each experiment, the transmitters were randomly placed approximately 10 to 15 meters apart, avoiding obstruction within the 15cm vicinity of each transmitter. The latitude and longitude of the transmitter were recorded manually via a GPS.

Each of the four transmitters was, as shown in Figure 6.1, deployed in one of four unique configurations: horizontal, vertical, elevated, and obstructed. The *horizontal transmitter* was placed on the ground with its antenna in the horizontal orientation (as previously defined). Similarly, the *vertical transmitter* was placed on the ground with its antenna in the vertical orientation. The *elevated transmitter* was mounted to a pole one

half meter above the ground with its antenna in the vertical orientation. The *obstructed transmitter* was laid flat on the ground with its antenna, in a horizontal orientation, covered with one quart of debris consisting of dirt and wood chips.

Unmanned Aircraft System: Packets broadcast by the transmitters were collected using an unmanned aircraft system. For the UAS, we utilized a DJI Matrice 100 quadcopter, as shown in Figure 6.1h. The Matrice 100 communicates with a remote control at 5.725 - 5.825 GHz, which is outside the frequency range of the 2.4GHz XBee nodes. The UAS was flown manually with no attached camera. A Raspberry Pi (RPi) 2 - Model B served as an on-board computer. The location of the UAS was recorded from the Matrice 100 on-board GPS, sampling at a rate of 50Hz and using a UART connection to the RPi.

When evaluating antenna type we used a matching pair of receivers on the UAS. So when evaluating internal antennae, four of the six XBee3 radios with internal antennae were used as ground based transmitters and two of the six were mounted to the bottom of the UAS. Likewise for the external antennae tests. These modules acted exclusively as receivers set to capture only. We oriented the antennae of the receiver in two configurations. The *horizontal receiver* had its antenna in a horizontal orientation, while the *vertical receiver* had its antenna in a vertical orientation. The two XBees forwarded packets to the RPi via a USB connection.

We flew the UAS over the transmitters at an average speed of four meters per second. The exact flight path and speed varied due to manual execution under varying wind conditions. Because the FAA, which governs the airspace over our testbed, regulates altitude in feet, we represent altitude (relative to ground level at the start of flight) in feet, while keeping displacement in meters. Flights for the integrated antennae were at altitudes of 50ft, 100ft, 200ft, 300ft, and 400ft at horizontal displacement of up to 250-325 meters from the closest transmitter to the UAS. When conducting flights for the external

antenna modules, we saw an improved reception range for certain configurations and altered our flight plan to include altitudes of 150ft and 350ft and horizontal displacement of up to 650 meters. As the FAA limits max altitude to 400 feet, we restricted our maximum experimental flight altitude to match (the EU similarly limits flight to 120m $\approx$ 394ft).

6.1.2 Experimental Area

The experiments took place outdoors at Coal Oil Point UC Reserve, a coastal grassland near the university. The experiments were conducted in a relatively flat area with some ground level obstruction due to tall grass and shrubs. The transmitters were placed so that the 15cm around each transmitter was clear of any obstruction, in areas with no tall shrubs in the 2m vicinity. For UAS horizontal displacement of over 350 meters, tall clusters of trees lined the sides of the flight path, however the UAS kept a line of sight corridor to the deployment area.

6.1.3 Measurements

Received Signal Strength Indication: RSSI is a common indicator of signal strength; in fact it is often the only signal quality metric reported by radio hardware. As a result, prior research of air to ground networks has relied on RSSI as a key indicator of network performance [126, 122, 123].

However, as our past work has shown [47, 48], RSSI may not be the ideal indicator for 802.15.4. RSSI is typically calculated by the receiver from successfully received packets. When conditions are poor and RSSI is low, the packets may not be received by the receiver and, as a result, their RSSI may not be recorded. Therefore, in past outdoor aerial 802.15.4 measurements, we observed that the mean of the received RSSI

remains relatively consistent despite changes to experimental variables, such as adding obstruction. In contrast, with fixed transmission frequency, significant shifts to the total number of received packets suggest significant differences in network performance between configuration scenarios not accounted for by RSSI.

In this work, while we provide an overview of RSSI for comparability to past work, we focus on packet reception as a more definitive network quality metric for the outdoor aerial 802.15.4 data collection use case.

Packet Reception Rate: Unlike other applications, where throughput, latency, and jitter are the principle metrics of performance, IoT applications are often delay tolerant and do not saturate network bandwidth. Instead, IoT networks try to minimize power consumption, especially outdoors where there may not be access to grid power. Similarly a UAS has a limited battery, and hence limited flight time. Therefore to maximize performance of an aerial data connection, we seek to minimize the number of failed transmissions. Accordingly, we measure the PRR, which is the number of packets received divided by the calculated number sent, as the principle metric of 802.15.4 IoT network performance:

$$\text{PRR} = \frac{\text{number of packets received}}{\text{time of UAS in sector} * \text{transmission rate}}$$

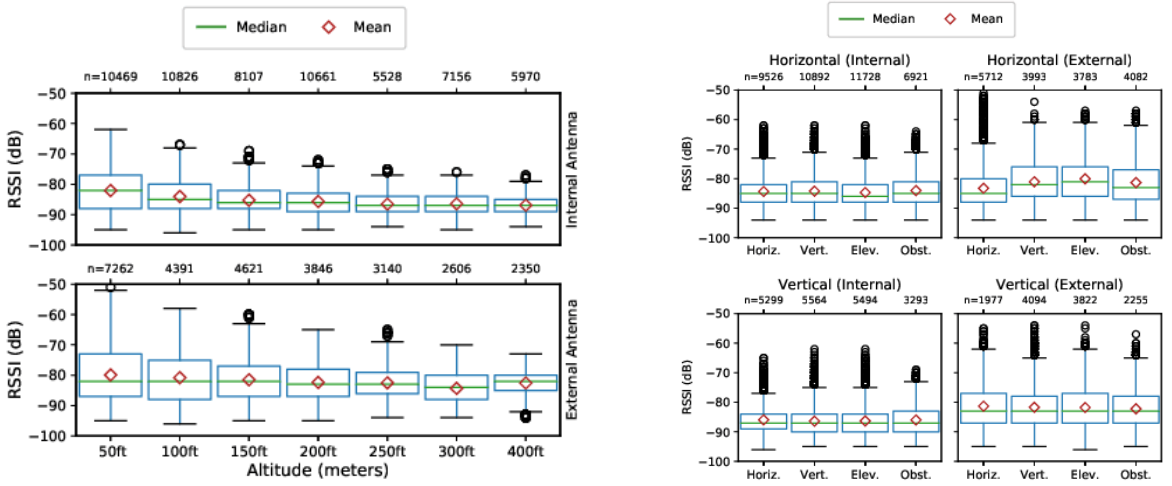
Because PRR only makes sense over an aggregate of readings, we group the experimental data by horizontal displacement from the corresponding transmitter into concentric circular sectors, 25m wide, radiating out from each transmitter, keeping other experimental variables separate. To determine the sector into which a packet from a particular transmitter falls, we compare the UAS's on-board GPS with the manually

recorded transmitter location. We estimate the number of packets sent by a transmitter by taking the product of the transmission rate and the time-in-sector occupied by the UAS. We drop measurement windows where fewer than five packets were sent.

6.1.4 Modeling Packet Reception Rate

To study the effect of each variable on PRR, we model the expected mean PRR using ZINB. From past work [48], we found that ZINB was the best fitting model, as it accounts for the over-dispersion and high number of PRRs at zero from locations and altitudes that never receive a packet. We therefore model both the chances that a packet is received at all and the estimated number of packets received. As the internal and external data sets had some experimental differences (the internal antenna had fewer displacement bins, and one fewer altitude), we constructed two independent ZINB models for each antenna type. We assessed the goodness of fit for our ZINB models by Scaled Pearson Chi-Square criteria, which were close to one and by the Full Log Likelihood criteria. We also compared the observed relative frequencies of the various counts to the maximum likelihood estimates of their respective probabilities. We found that our models were a good fit for the observed data.

To prepare for modeling we aggregated our readings into sectors, as discussed in the last section. We separated our data into internal and external antennae. We then randomly divided each group data into two sets: 60% was designated as a *training set*, while the remaining 40% was designated as a *test set*. We used a ZINB model with the number of received packets as the outcome. For the model of external antennae we set displacement as a categorical variable with 25 groups (25m to 625m), altitude as a categorical variable with 8 groups (50ft to 400ft). For the model of internal antennae we grouped displacement as 11 groups (25m to 275m) and altitude with 7 groups (omitting



(a) Grouped by altitude and antenna.

(b) Grouped by antenna and configuration.

Figure 6.2: Distributions of RSSI from aerial measurements.

350ft). For both models the transmitter-receiver configuration was 8 groups *(Vertical & Horizontal Receivers paired with Horizontal, Elevated, Obstructed & Vertical Transmitters)* as fixed covariates for both parts of our model. We used a natural logarithm of sent packets as an offset in the NB part of the model and control for the number of sent packets in the ZI part of the model.

6.2 Evaluation

6.2.1 Received Signal Strength Indicator (RSSI)

For each received packet we logged the RSSI reported by the receiver modules on the UAS. In particular, we examined how RSSI changed based on antenna type (internal vs. external), antenna orientation of the transmitter and receiver, amount of obstruction, and altitude.

Altitude: To examine the impact of altitude, we group the data by antenna type and altitude. We present the distribution by group as a box plot in Figure 6.2a. These distributions include data from all horizontal displacements and transmitter configurations. The top plot shows the distribution of observed measurements from the set of transmitters with internal antennae, and the bottom shows those with external antenna.

The median and mean RSSI for both internal and external antenna remained close to one another with fluctuations within $\pm 2dB$ for each altitude. The minimum RSSI values remained likewise fixed. However the interquartile ranges shrank at higher altitudes, as the UAS increased in total vertical distance from the transmitters. Notably the external antennae' maximum RSSI is $\approx 10dB$ better than the internal antennae, across all altitudes.

Transmitter Configuration: To examine the impact of antenna orientation and obstruction, we group the data by antenna type, orientation, and obstruction. We present the resulting box plot in Figure 6.2b. The two left plots show results from internal antenna modules for both the transmitter and receiver, while the two plots on the right show those for external modules. The top two plots correspond to horizontal antenna orientations for the receivers, while the bottom two plots show vertical receiver orientation. Each of the four subplots show the four possible transmitter configurations: horizontal antenna orientation, vertical orientation, elevated transmitter with a horizontal orientation, and obstructed transmitter with horizontal orientation.

We can see that the mean and median RSSI across all conditions remain within $\pm 3dB$ of each other, with similar interquartile ranges. Overall, the external antennae perform better than the internal antennae, the majority of the packets are received with higher RSSI.

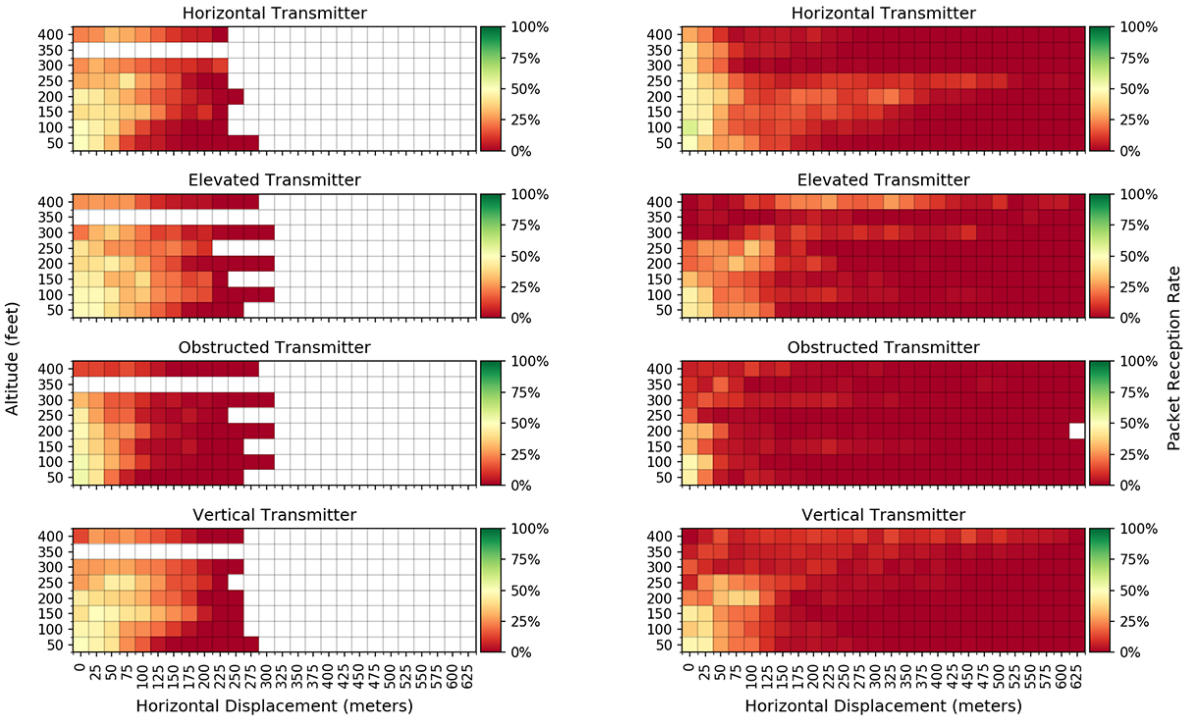
The internal antennae display unusual behavior in the performance of the elevated and obstructed transmitters, showing that, contrary to expectation, the obstructed transmitter performs better than the elevated transmitter. When controlling for the number of packets sent, we found that while the obstructed transmitter RSSI is not significantly different, the obstructed transmitter successfully delivered less than half the number of packets. This strongly suggests that RSSI is not a reliable indicator of network performance for this application.

6.2.2 Packet Reception Rate (PRR)

Because RSSI does not provide a comprehensive look at network performance, we focus the majority of our analysis on packet loss by examining PRR. As explained in Section 6.1.3, we group our observations into 25m sectors. In our analysis we omit observations where fewer than five packets were sent, resulting in 12,891 total observations (8,591 observations for the external antenna set and 4,300 observations in the internal antenna set). Roughly 36% of the groups received at least one packet.

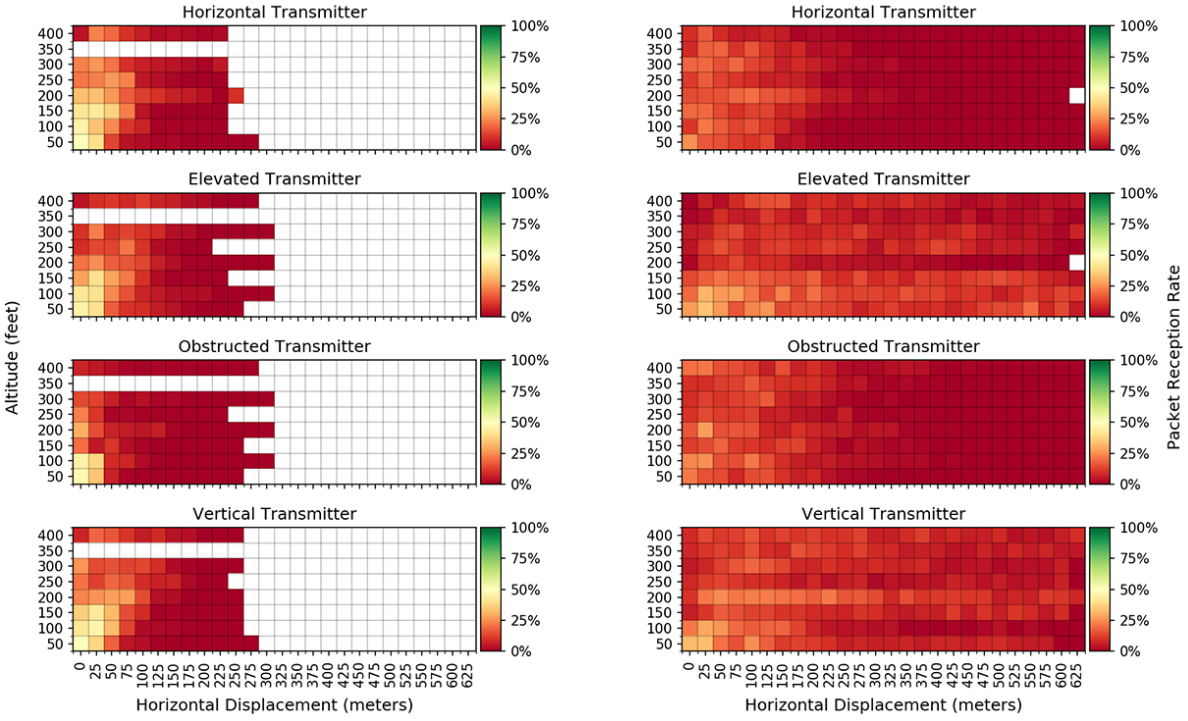
A heatmap of the observed PRR grouped by antenna type, antenna orientation, altitude, and horizontal displacement, averaged across multiple runs, is shown in Figure 6.3. Results for the internal antenna modules are shown in the left (Figures 6.3a and 6.3c), while those of the external modules are shown in the right (Figures 6.3b and 6.3d). Each figure is broken down by transmitter configuration with the color of each square representing the average PRR of a horizontal displacement sector at a particular altitude. White squares indicate fewer than five packets were sent at those variable conditions and so are omitted from the heatmap.

The internal antenna flights have fewer (displacement, altitude) pairs filled than the external antennae. As we performed those experiments first, we tailored the flight plan



(a) Horizontal receiver (internal antenna).

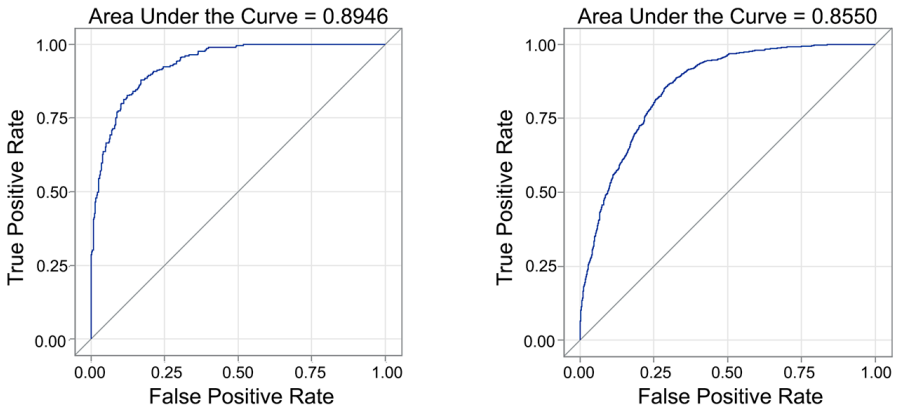
(b) Horizontal receiver (external antenna).



(c) Vertical receiver (internal antenna).

(d) Vertical receiver (external antenna).

Figure 6.3: Observed packet reception rates grouped by altitude and 25 m displacements from transmitter. Cells with fewer than five sent packets are left blank.

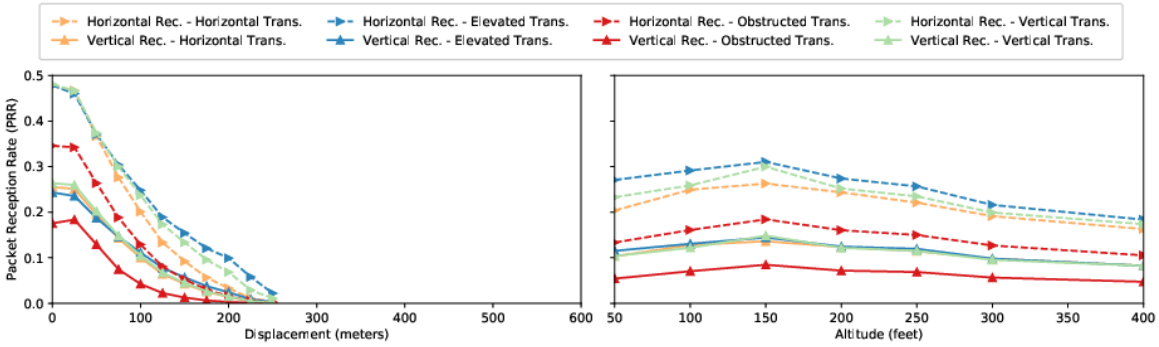


(a) Internal antenna model. (b) External antenna model.

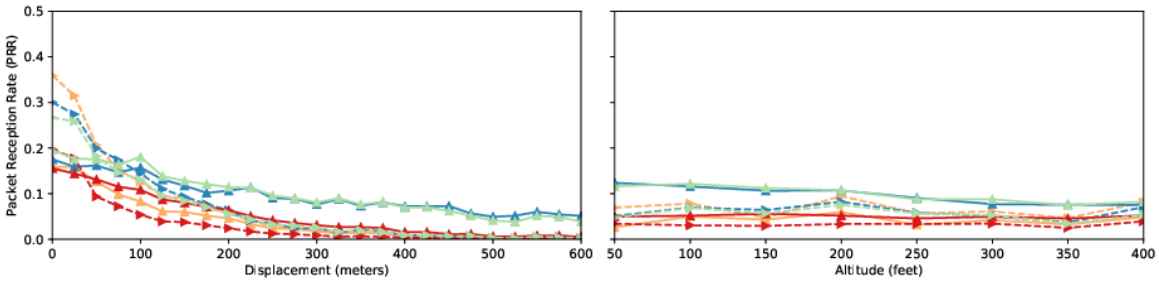
Figure 6.4: ROC curves for ZINB models evaluated on test sets.

to preliminary results on that hardware type. As we found near total loss at distances greater than 250 meters, we limited our flights around that range. In contrast, the external antennae showed greater reception range, so we tripled our maximum horizontal displacement (further displacements was limited by restrictions on our airspace) and added an additional altitude measure of 350ft. Due to the significantly increased flight times from the increased horizontal range, we restricted altitude to 50, 100, 150, 200, 250, 300, 350, and 400 feet.

Modeling PRR: From the training set, we constructed a pair of ZINB models of PRR (one for internal and one for external antennae), as described in Section 6.1.4. Using these models, we can predict (based on input variables of hardware type, antenna orientations of transmitter and receiver, horizontal displacement, and altitude) the expected mean PRR. We utilize these model to study the effects of each variable on PRR. Network planners could likewise use this type of model, perhaps expanding the training set to fit more geographic topographies, to plan out sensor network equipment deployment and aerial data collection.



(a) Internal antenna.



(b) External antenna.

Figure 6.5: Observed PRR grouped by altitude and displacement, varying antenna type and configuration.

We can turn the ZINB model, which predicts a rate, into a binary classifier, predicting when at least one packet will be received, by specifying a threshold below which we expect the packet to be lost. We verify the resulting classifier based on the test set. We present a ROC curve of the ZINB for both models executed on the *test set* in Figure 6.4. This displays the possible true and false positive rates based on threshold choices. For example, given an external antenna and a PRR threshold of 30%, the resulting binary classifier correctly identifies 77% of the observations with a recall of 78% and specificity of 76%.

A key focus of this study was to examine the difference between the internal coiled circuit board embedded antennae that we previously evaluated [47, 48] and the external straight wire antennae. To make this comparison, we examine the differences between the PRR of the four receivers (external horizontal, external vertical, internal horizontal, and

internal vertical) across UAS altitudes and horizontal displacements. Due to limits in manpower, this work does not examine mixing transmitter and receiver antenna types (for example internal transmitter with external receiver). For this study, we assume that antenna types are homogeneous between transmitter and receiver, we will examine heterogeneous antennae types in future work.

Transmitter Configuration: First, we compared the PRR based on transmitter configuration, examining antenna type, obstruction, altitude, and orientation. In Figure 6.5, we present plots from the model for the PRR grouped by altitude and displacement, separated by internal and external antennae.

The difference between configurations of transmitters with internal antennae was less stark than the external antenna models. For the internal antennae, as expected, the vertical elevated antenna with the horizontal receiver showed the highest PRR rate, followed by the vertical ground level antenna with the horizontal receiver, then followed by the vertical elevated with a vertical receiver. The worst PRR was, again as expected, the horizontal obstructed transmitter with the vertical and horizontal receivers. This order held across altitudes and displacements.

For transmitters with external antenna modules, there was a higher variance in PRR by configuration. The vertical transmitters (both elevated and not) paired with the vertical receiver had a PRR nearly double that of the other antennae configurations. Unlike the internal antennae, for the external set, altitude did not make a difference on PRR. For the external set, the vertical transmitters outperformed the horizontal transmitters, particularly when paired with vertical receivers.

Interestingly, the obstructed horizontal transmitter had a higher PRR than the horizontal unobstructed transmitter when collected by the vertical receiver. The obstructed horizontal transmitter had the worst PRR when collected by the horizontal receiver.

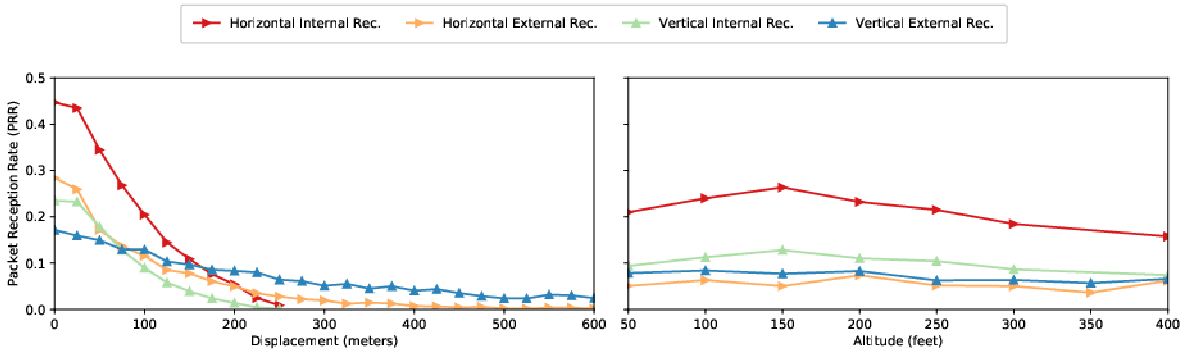


Figure 6.6: PRR grouped by receiver type.

Receiver Configuration: Next, we compare the PRR by receiver across all transmitter configurations of the matching antenna type. We present plots from the model for the PRR grouped by altitude and displacement in Figure 6.6. Overall, the receivers with internal antennae exhibited higher PRRs, regardless of orientation.

When examining the effect of receiver orientation, we found that for the internal antenna hardware, the horizontal receiver produced a better PRR across all altitudes and displacements than the vertical receiver. The behavior of the internal antenna modules is attributable to the mounting of the receivers. While the horizontal receiver was mounted on the bottom of the UAS, with clear line of sight to the ground at all times, the vertical receiver was mounted on the side of the UAS undercarriage with the mounting potentially interfering with line of sight at some angles.

The effect of receiver orientation on the external antenna modules is more complex. At smaller displacements ($< 100m$), the receiver with horizontal antenna orientation exhibited a higher PRR, while at greater displacements the vertical antenna had higher PRR.

Horizontal Displacement: As expected, PRR drops off as horizontal displacement between the UAS and transmitter increases, influenced by the inverse square law of signal strength decay. The internal antennae sets have an overall higher PRR at close displacements, but as displacement increases to 250m, PRR falls to near zero. In contrast the external antennae, while overall yielding a lower PRR, still perform well at extreme distances of $> 600m$.

Altitude: For internal antenna modules, as seen in Figure 6.5a (right), collecting at altitudes of around 150ft maximizes PRR (the exact optimal altitude depends on antenna configuration). In contrast, for the external antennae shown in Figure 6.5b (right), there is no optimal altitude. For external antennae higher altitude corresponds to a higher PRR up to the FAA limit of 400ft (likely there is an optimal altitude past this limit).

The toroidal radiation pattern of the external dipole antenna introduces additional considerations when selecting an altitude for UAS data collection. Dipole antennae radiate outwards, perpendicular to their orientation, with a cone of low signal strength at the tip of the antenna (the exact radiation pattern can be found in [247]). While overall lesser horizontal displacements produce a higher PRR, for *vertical transmitters* there is a dip in PRR at displacements $< 100m$ for altitudes $> 200ft$. This is most clearly seen in Figures 6.3b and 6.3d. In contrast, external horizontal transmitters perform better across all altitudes at smaller displacements.

6.3 Conclusion

When evaluating the efficacy of an outdoor aerial assisted data collection strategy for a sensor network, network administrators need real-world models of optimum flight altitude, expected reception rates, and maximum effective horizontal displacement for reliable data reception. Our work provides a foundation for understanding 802.15.4

2.4GHz outdoor performance for three dimensional network communication using physical experimentation.

The widely used XBee3 module evaluated in this work has an advertised effective operating range of 1200m [244], but, as our evaluation shows, variables such as altitude of UAS, antenna type, antenna orientation (of transmitter and receiver), evaluation, and obstruction can dramatically limit the maximum horizontal displacement with a usable PRR. Moreover there is not a one size fits all configuration.

For an outdoor grassland setting, results show that if the UAS flight plan is expected to come close to the transmitters ($< 150m$), then an altitude of 150-250ft with internal antennae consisting of an elevated vertical transmitter and a horizontal receiver produce the best PRR. For effective data collection at greater displacements ($> 150m$), external antennae consisting of a vertical transmitter and a vertical receiver are optimal regardless of flight altitude.

In our experiment packets were broadcast by the transmitters at a rate of 500ms, which is highly energy intensive for deployments that might need to operate on battery power for weeks or months. Unfortunately, experimentation on lowering transmission rate proved challenging due to the power requirements of the UAS. As each UAS battery provides just over 20 minutes of useful flight, experimentation on slow transmission rates makes collection of a meaningful amount of data challenging. Initial results showed that the UAS would have to slow flight speed substantially for slower transmission rates. To address this issue there is active research on using low power radios for "waking" 802.15.4 radios for transmission [248]. While our dataset will likely not generalize to all terrain and geographies (e.g. an urban sensor deployment), we believe our method of modeling and insights into three dimensional performance of 2.4GHz 802.15.4 under a variety of antenna configurations is highly transferable to future work in 802.15.4 analysis and to real-world sensor network planning.

So far in this dissertation, we have focused on evaluating and restoring connectivity to networks. In the next chapter we explore how wireless transmission capability of UEs can be valuable, even in the absence of a functioning network.

Acknowledgment

This work was done in collaboration Maxton Ginier, Ryan Allen, Irina Artamonova and Elizabeth Belding. Thanks to Coal Oil Point UC Reserve for allowing us to utilize your space for our aerial testbed.

Chapter 7

Real-Time Multilateral RSSI Localization from UASs for Disaster Response

In chapters 5 and 6 we evaluated the feasibility of 802.15.4 DTN data delivery from FUSN. Unlike 802.15.4, the IEEE 802.11 family of standards has relatively short communication ranges, typically under 100 meters. Further as we discuss in chapters 1 and I, due to the wide range of device types and security concerns consumer mobile devices do not have a standardized way to share data such as location. In this chapter we explore how the passive monitoring of 802.11 transmissions, even after network failure can be used to aid in disaster response applications.

While white-light or thermal imaging techniques may be viable in the case of a disaster, poor visibility, visual obstruction, and thermal variations can impede a visual-only system. When other communication channels are unavailable, UAS based radio localization can augment visual systems by providing localization of mobile devices in the possession of affected individuals. WiFi is ubiquitous in consumer devices, and the MAC address in

802.11 packet headers facilitates distinguishing the source of a transmission.

WiFi broadcasts, such as probe requests sent by user devices to locate an access point when no existing connection is present [57], can be used to passively locate UEs carried by users during a disaster. Disaster victims might be unable to actively engage with an emergency system due to physical impairment or time constraint. In addition, victims may not have prepared for an emergency situation ahead of time (i.e. by pre-loading emergency apps). We, therefore, do not assume cooperation from mobile devices. This precludes devices from self-reporting location obtained via GPS or using custom communication protocols.

In such cases, rescuers need a way to quickly and passively locate people on the ground. Our work addresses this problem through voxel-based approximation (VBA), a novel algorithm for UAS-based multilateral localization. A key difference between our work and prior localization algorithms is that most prior work focuses on the user device performing the localization through passive reception of beacons from multiple stationary transmitters. In contrast, VBA operates on a *single mobile UAS*. It enables the UAS to localize a transmitting device on-the-ground, through the aggregation and analysis of RSSI data points collected as the UAS moves over the affected terrain. With this information, emergency personnel can locate and assist an individual in a situation where standard communication avenues (cell towers, broadband Internet) have become inoperable due to power outages, damage, or capacity overload.

In this chapter, we describe the operation of VBA and evaluate its efficacy in locating WiFi-enabled devices through physical experimentation using an in-flight UAS. Through multiple scenarios, we evaluate VBA and compare it to other approaches, such as linear least squares and maximum likelihood estimation. Our results show that VBA has comparable error rates to established algorithms, while maintaining a constant run time and memory requirements for each new update. This is key to the deployability of any

localization algorithm, as UASs are constrained in power, due to their use of batteries; and on-board computational and storage resources, due to maximum weight requirements.

Our specific contributions are as follows:

- We present a novel *mobile-UAS-based localization algorithm*, where a mobile in-air vehicle localizes a stationary transmitting device in 3D space. Our approach reverses the standard localization scenario, and breaks new ground in successful device location.
- Unlike past algorithms for multilateral localization, VBA computation time and storage for each measurement update is constant, making it ideal for real-time localization and location updates that leverage the entire received data set.
- We validate and evaluate our work through extensive physical in-the-air experimentation and explore the efficacy of 802.11 2.4GHz localization in the context of a disaster response. We demonstrate that VBA outperforms alternative approaches in resource requirements while maintaining comparable localization error, even in a "mock disaster" scenario.

7.1 3D Aerial Localization

As the UAS moves across a search path, it collects broadcast 802.11 packets sent by the device. Even when not connected to an access point, 802.11 device broadcasts packets called probe requests that assist in connectivity. A UAS can monitor broadcasts and correlate them with its own position coordinates at time of reception via an on-board GPS. It filters packets based on their MAC address and extracts the RSSI values for a transmitter. Unlike WSN approaches, where nodes, numbering in the tens to hundreds, broadcast periodically (generating time series of readings), our approach can receive

thousands to tens of thousands of measurements in the case of an active 802.11 stream. When operating in a disaster area the UAS can filter the devices by MAC address and investigate each location to see if someone needs help.

As the UAS flies, it continually updates its estimate of the transmitter's location, based on newly arriving broadcasts. The UAS could use this data for real-time flight-planning, converging on the geographic location of the transmitter, and facilitating immediate response by emergency personnel. Methods relying on post hoc analysis are less suited for emergency scenario applications because they cannot react as quickly.

As we explore in Section 7.1.3, common methods of multilateral localization grow in computation time and storage space with the size of the data. Because of our time and storage constraints, these solutions are far from ideal. To solve this problem, we present VBA, a novel algorithm for generating a real-time heat-map matrix for approximate multilateral localization. This method updates a representation of a physical space in the form of a matrix of location probabilities. At any point, the UAS can query the matrix for the most probable location(s) of the transmitter. We explain the process of localization and our contribution in further detail in this section.

As stated above, 802.11 broadcasts are sorted by MAC source address, and only broadcasts from a single source are used for that device's location calculation. For the remainder of this chapter, we focus on the task of locating a single transmitter. For multiple sources the same approach can be re-run for each MAC address.

7.1.1 Computing Distance from RSSI

As the UAS travels, it captures 802.11 broadcast packets and extracts the RSSI, as reported by the wireless adapter. From the RSSI, we can compute the approximate relative distance between the transmitter and the UAS, using an appropriate path loss model

[249]. For the purpose of our experiments, which are primarily in an open space with little obstruction or interference, we use an extension of the Friis transmission equation [250, 249] (as shown in equation (7.1) where d is the distance to the transmitter, d_0 is a reference distance, α is the signal decay constant and f corresponds to the frequency on which the UAS is listening - in this case, that of a target 802.11 channel).

$$PathLoss = 10\alpha \log_{10}(\frac{d}{d_0}) + 20\log_{10}(f) + 20\log_{10}(\frac{4\pi}{c}) \tag{7.1}$$

Assuming free space transmission (setting α to 2), equation (7.1) can be simplified as (7.2), where A is some calibration constant for a wireless adapter, that we can compute through experimentation.

$$d = 10^{\frac{RSSI - 20\log_{10}(f) - A}{20}} \tag{7.2}$$

7.1.2 Localizing the Transmitter

As we aim to locate an object using a single UAS, we will assume a slow moving transmitter relative to the agile UAS. This would be consistent with natural disaster scenarios, such as flooding or mudslides where survivors may be confined to high-ground with limited mobility.

As the UAS moves in three-dimensional space, it uses GPS to note its own coordinates. For each received 802.11 packet, it collects the time, latitude, longitude, altitude, MAC address of the transmitter, and the calculated distance to the transmitter using (7.2). As GPS coordinates are spherical, we can convert them into a Cartesian earth centered earth fixed (ECEF) representation to effectively work with them.

Given a set of four measurements, we, in an ideal setting, employ quadrilateral localization to locate an object in three-dimensional space using a system of equations of

the form $(x_m - x)^2 + (y_m - y)^2 + (z_m - z)^2 = d_m^2$, where (x_m, y_m, z_m, d_m) corresponds to the m$^{\text{th}}$ measurement by the UAS, and (x, y, z) to the position of the transmitter. Note that using only three measurements would give two possible solutions, hence two possible locations. A fourth measurement refines this to one location.

In practice, wireless transmissions are not ideal. The RSSI measurement of any given packet can be prone to error, due, for example, to interference, multi-path, or angle of antenna [122, 251, 252]. To refine the location, the UAS can continue to collect measurements. Each measurement can then be used to recompute the predicted location of the transmitter through multilateral localization. This gives a system of equations (7.3) that grows as more measurements are collected.

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 = d_1^2 \\ (x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 = d_2^2 \\ \vdots \\ (x_m - x)^2 + (y_m - y)^2 + (z_m - z)^2 = d_m^2 \end{cases} \tag{7.3}$$

We can convert this into linear matrix form by subtracting each m$^{\text{th}}$ equation from the m-1$^{\text{th}}$ equation, as is done in the two dimensional case in [253]. This gives m-1 system of linear equations of form (7.4), which can be rewritten as $\boldsymbol{Ax} = \boldsymbol{b}$ as shown in (7.5).

7.1.3 Multilateral Localization

$\boldsymbol{A}$ is a matrix $\boldsymbol{m}$ x 3 where $\boldsymbol{m}$ can number in the millions of measurements, each with a certain error. So the system of equations (7.5) is overdetermined and, moreover, likely inconsistent. Therefore, the goal of our localization is, given $\boldsymbol{m}$ measurements, to find the optimal location $\boldsymbol{x}$ that best fits the available data. Common ways of performing this optimization are either to approximate using linear least squares (LLS) [254], weighted

Table 7.1: System of equations for 3D multilateral localization

$$\begin{cases} 2x(x_2 - x_1) + 2y(y_2 - y_1) + 2z(z_2 - z_1) + x_1^2 - x_2^2 + y_1^2 - y_2^2 + z_1^2 - z_2^2 = d_1^2 - d_2^2 \\ 2x(x_3 - x_2) + 2y(y_3 - y_2) + 2z(z_3 - z_2) + x_2^2 - x_3^2 + y_2^2 - y_3^2 + z_2^2 - z_3^2 = d_2^2 - d_3^2 \\ \vdots \\ 2x(x_m - x_{m-1}) + 2y(y_m - y_{m-1}) + 2z(z_m - z_{m-1}) + x_{m-1}^2 - x_m^2 + y_{m-1}^2 - y_m^2 + z_{m-1}^2 - z_m^2 = d_{m-1}^2 - d_m^2 \end{cases} \tag{7.4}$$

$$\begin{bmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) & 2(z_2 - z_1) \\ 2(x_3 - x_2) & 2(y_3 - y_2) & 2(z_3 - z_2) \\ \vdots & \vdots & \vdots \\ 2(x_m - x_{m-1}) & 2(y_m - y_{m-1}) & 2(z_m - z_{m-1}) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1^2 - d_2^2 - x_1^2 + x_2^2 - y_1^2 + y_2^2 - z_1^2 + z_2^2 \\ d_2^2 - d_3^2 - x_2^2 + x_3^2 - y_2^2 + y_3^2 - z_2^2 + z_3^2 \\ \vdots \\ d_{m-1}^2 - d_m^2 - x_{m-1}^2 + x_m^2 - y_{m-1}^2 + y_m^2 - z_{m-1}^2 + z_m^2 \end{bmatrix} \tag{7.5}$$

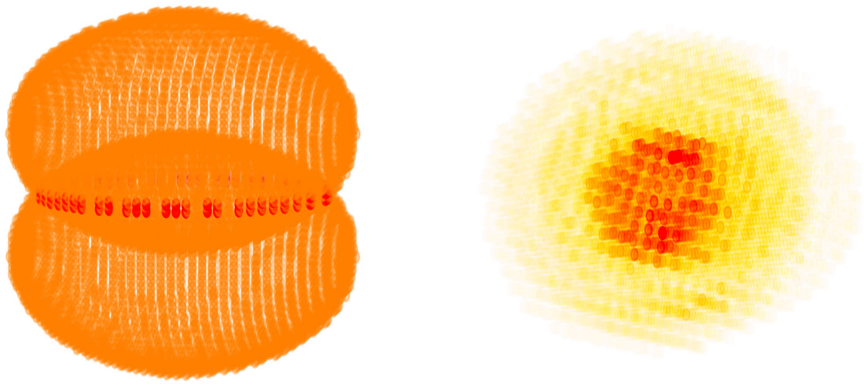
least squares (WLS) [255, 256], or use MLE to approximate the coordinates [205, 257].

As the UAS continues to collect measurements, an ideal system would use each new measurement to refine the localization without recomputing for the entire dataset. This would enable real-time applications, such as dynamically adjusting a flight-plan based on most accurate data. While these established approaches work well for a given data set, they are not well-suited for continuously reanalyzing data at each collection step. Suppose, each time we collected data, we re-ran our algorithm. For $\boldsymbol{m}$ measurement both LLS and MLE take $\boldsymbol{O(m)}$ time for a single optimization run. Re-running LLS or MLE each time a new packet is collected would yield $\boldsymbol{O(m^2)}$ time.

In a disaster scenario, there may not be a reliable backbone link to utilize off-UAS computational resources, forcing us to use on-board computing power limited by weight and energy. Further, the time sensitive nature of this task encourages utilizing real-time analysis. Algorithms where the computational load increases, based on the size of the current data, are not well suited for this task.

7.1.4 Voxel-based approximation (VBA)

To meet resource constraints, we developed Voxel-based approximation (VBA) - an algorithm for approximating transmitter location. We divide the geographic area of the



(a) The circle formed by the intersection of spheres has the highest probability in this scenario.

(b) Example probability heat-map for VBA search.

Figure 7.1: Example probability heat-map for VBA search. Darker reds represent higher probability.

search operation into cubes represented by voxels. Voxels are like pixels and contain values at three dimensional coordinates that are recorded by their position in a 3-dimensional array. Each voxel represents a cube of physical space and stores the likelihood that the transmitter is in that cube. How much physical space a voxel represents can scale based on the size of the total physical 3D volume of the UAS operation, and the available computing and memory of the device running the localization. For example, given a $\mathbf{500x500x500}$ $\boldsymbol{meters^3}$ search area, we can set a maximum 3D matrix size of $\mathbf{100x100x100}$ which stores $\mathbf{100^3}$ voxels, each corresponding to a $\mathbf{5}\boldsymbol{m^3}$ cube of physical space.

For each measurement, the predicted location of the transmitter can be described as positioned somewhere on the surface of a sphere, centered at the UAS with a radius equal to the computed relative distance. When two spheres intersect, as shown in Figure 7.1a, a point in the area of intersection of their surfaces is the most likely location of the transmitter. Multiple measurements can further refine the probability matrix for the position of the transmitter, as seen in Figure 7.1b, which is produced from experimental

data from one of our experiment trials. In the figure, the darker areas represent regions of higher probability for the transmitter location.

Each time a measurement is received, we identify the voxels of the surface of a sphere $\boldsymbol{S}$ of radius $\boldsymbol{d}$, and increase the value of those voxels by a scalar $\boldsymbol{\beta}$. Because transmitters may be moving, old information in the matrix needs to expire. For this we introduce a decay constant $\boldsymbol{\lambda}$ in range $\boldsymbol{(0, 1]}$, which gradually eliminates stale values in the matrix. For cases where there is little mobility, we can set this to $\mathbf{1}$. Each update takes the form (7.6). This can be done in $\boldsymbol{O(g)}$ time, where $\boldsymbol{g}$ is the number of cells in $\boldsymbol{M}$.

$$\boldsymbol{M = \lambda M + \beta S(x, y, z, d)} \tag{7.6}$$

In a disaster, transmitters are most likely positioned on the ground, inside buildings, or on roof tops. We might, therefore, choose to assign greater weights to these guesses, by specifying a topography $\boldsymbol{T}$ for a physical region, as opposed to, for instance, guesses suggesting that a transmitter is floating in the air. We can represent this by taking the Hadamard product of $\boldsymbol{S}$ and $\boldsymbol{T}$:

$$\boldsymbol{M = \lambda M + \beta S(x, y, z, d) \circ T} \tag{7.7}$$

If we are entirely disinterested in points deviating from our topography (floating transmitters, transmitters under the ground, etc.), we can assign the corresponding weights in $\boldsymbol{T}$ to zero. This could be used to speed up computation.

At any point, the system utilizing this localization algorithm can generate a heat-map of the probability that a transmitter is at a voxel by normalizing the matrix. Again, this can be done in $\boldsymbol{O(g)}$.

A notable property of this algorithm is that each update cost is determined by the size of the matrix and not the input data, and that the storage size is constant. Each new

measurement updates the matrix without re-computing using past measurements. For $\boldsymbol{m}$ measurements, our approach takes $\boldsymbol{O(m \cdot g)}$ time and $\boldsymbol{O(g)}$ space, where $\boldsymbol{g}$ is specified at initialization. In contrast, traditional approaches, such as LLS, WLS, and MLE take $\boldsymbol{O(m^2)}$ time and $\boldsymbol{O(m)}$ space. Further, unlike traditional approaches, computational and memory requirements are scalable at the cost of precision. We can select a matrix size $\boldsymbol{g}$ meeting our hardware constraints. As computational capabilities of UASs grow, the algorithm can scale to take advantage of these resources.

7.2 Methods

To understand both the efficacy of our approach and real-world constraints of localization from an aerial system, we conducted extensive empirical measurements to evaluate our system in a semi-controlled environment. The goal was to understand the impact algorithmic design choices (such as the size of the approximation matrix, and the use of pre-defined topography) have on accuracy over time, and compare our approach to two other commonly used approaches: Linear least squares (LLS) Maximum likelihood estimation (MLE).

7.2.1 Equipment

In our experiments, our UAS was a DJI Matrice 100 [258] that communicates with a remote control at 5.725 - 5.825 GHz (outside our monitoring frequency of 2.4GHz). The UAS was equipped with a pair of Alfa network adapters running the Atheros AR9271 chipset located at opposite sides of the UAS, as shown in Figure 7.2. Both adapters ran in monitor mode, listening to channel 6. In a real use case, we could utilize more adapters, each with a dedicated channel or implement a round robin scanning technique to scan all 802.11 channels. We used Raspberry Pi (RPi) 2 Model B for on-board computation.

Figure 7.2: UAS used in VBA localization evaluation.

We ran Tshark [234] to log, for each received packet, time of reception, RSSI, MAC address, and channel. Using a UART connection, we interfaced with the quad-copter to access its on-board telemetry, including GPS. Data was collected at 50HZ and synchronized with the Tshark data.

The transmitter (the device the UAS was trying to locate) was a Lenovo X230 Laptop with an Intel Centrino Advanced-N 6205 wireless card. We determined the ground truth of the transmitter location via averaging GPS points from a GlobalSat BU-353-S4 GPS attached via USB. This location was later used in our analysis to compute error rate. The Lenovo laptop was connected to a Raspberry Pi (RPi) 3, acting as an 802.11n Access Point broadcasting on channel 6. In our initial test cases, the Lenovo transmitted a constant stream of UDP packets using iperf to the RPi. We vary and reduce the rate of this transmission based on experimental scenarios as specified in section 7.2.2.

We used a laptop instead of a mobile device to facilitate our experiment. A laptop is easier to control without additional programming, and is more predictable in terms of energy optimization and program execution. We also use a constant stream of UDP packets, which would be unlikely in a real disaster scenario. Instead we would expect more infrequent probe requests. We look at the effect using fewer packets has on accuracy

in our analysis. The primary goal of this experimentation is to evaluate our approach in a more controlled setting, we leave closer simulation to disaster conditions as future work.

7.2.2 Experimentation Scenarios

We explored four test scenarios. For each scenario the receiver (the UAS) listened to the packets sent between the transmitter (Lenovo) and the access point (RPi 3). Each scenario was conducted in three different locations. Each location was a grassy area with minimal change in elevation, minimal obstructions from trees and bushes, and minimal interference from other 802.11 devices (there was a nearby 802.11 wireless mesh in the vicinity). At each location we performed two trials per scenario and configuration (elevation, presence of obstruction), yielding a total of 48 trials.

We varied the starting point of the UAS in each trial; we only relocated the transmitter when relocating to a new location. We manually piloted the UAS in all trials, which, when coupled with wind, resulted in some variation in velocity. As a result, the total time each trial took to complete varied as well. Overall, each trial took between 5 to 10 minutes, depending on the specific scenario.

Scenario 1 - Vertical Fly-over: In this scenario, we fly over the transmitter in a straight line, at altitudes (above ground level) of 20, 30, 40, and 50 ft, as shown in Figure 7.3. We maintain a 10MBps UDP stream between the transmitter and access point with a 1k buffer size using iperf. This trial examines both the accuracy of VBA, as well as the effect of the UASs altitude on localization accuracy.

Scenario 2 - Lane Search: In this scenario, the UAS flies in a lane search-pattern at a constant altitude of 25 ft above ground level. An example GPS trace from one of the test runs is shown in Figure 7.4. The left of the figure shows the trace projected onto the horizontal plane, while the right the full 3D space. We again maintain a 10MBps UDP

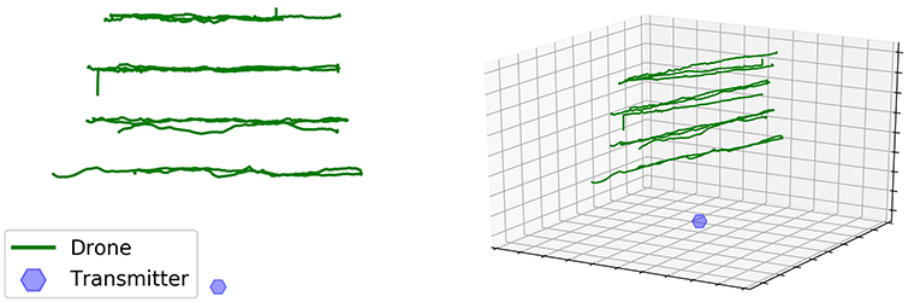


Figure 7.3: Flyover UAS path of VBA experimental GPS traces.
YZ plane (left) and 3D (right).

stream. This evaluates localization in a pre-defined search pattern, where the UAS does not change behavior based on estimated location.

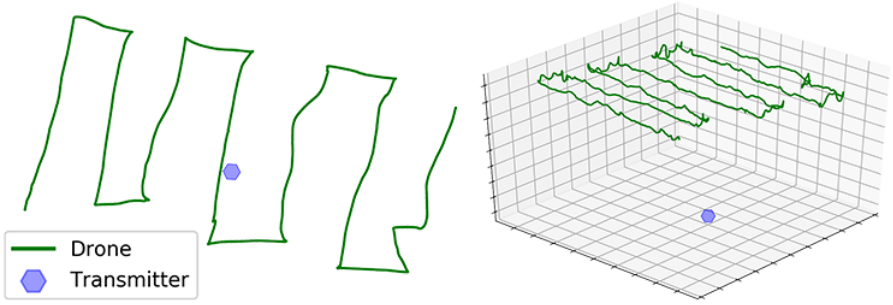


Figure 7.4: Lane search UAS path of VBA experimental GPS traces.
XY plane (left) and 3D (right).

Scenario 3 - Inward Spiral: In Scenario 3, we perform an inward spiral search pattern. An example GPS trace from one of the runs is shown in Figure 7.5. For this scenario we fly the UAS at an altitude of 25ft above ground level starting at the outer edge and slowly spiraling in to the transmitter. As the UAS nears the transmitter, it drops in elevation until it is hovering a meter above it, at which point the trial ends. This trial mimics the situation in which the UAS was able to use some combination of localization methods to hone in on an object and examines the change in accuracy as the UAS flies in the circular pattern around the object. In this scenario, we again manually pilot the UAS and maintain a 10MBps UDP stream.

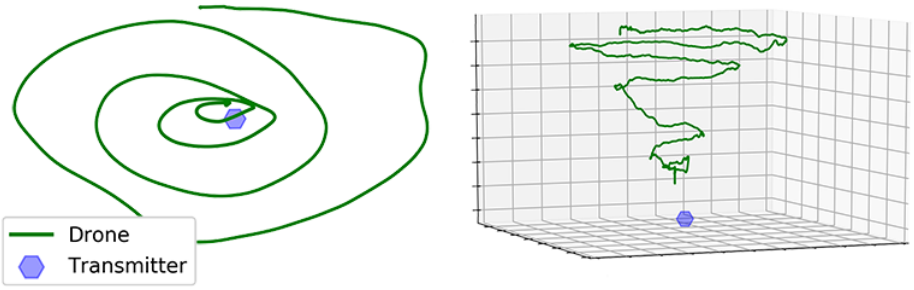


Figure 7.5: Inward spiral UAS path of VBA experimental GPS traces.
XY plane (left) and 3D (right).

Scenario 4 - Mock Disaster: Scenarios 1 through 3 enable us to evaluate the basic operation of VBA without confounding factors, such as signal obstructions. However, in order to more closely model a disaster scenario, we add some realism. In this scenario, the transmitter sends only 1 packet per second to the access point (in a real situation, this would be analogous to a periodic probe request sent by a mobile device), resulting in fewer total RSSI measurements with longer inter-packet durations. We again perform an inward spiral search pattern, as in scenario 3.

We perform two variations of this scenario. In Scenario 4A, the transmitter is placed in an unobstructed location on the ground. In Scenario 4B, we close the laptop, place a book on it, and a box over it, adding objects around it to obstruct its transmissions. This mimics the more realistic situation where the device might not have line of sight connectivity to the UAS (i.e., if it was placed in a bag or pocket, or buried under light debris).

7.2.3 Algorithm Implementations

We implement all our algorithms in Python 3.6 due to its versatility and ease of cross platform compatibility. We endeavored to provide an equal comparison between the algorithms. While it might be possible to further optimize the performance of LLS, we

base our work on generalized representations from literature.

For our evaluation, we record all data from the UAS and then run it through each algorithms off-line. This allows us to compare multiple approaches using the same data. In a real deployment, the UAS would run the algorithm in real-time without necessitating postponed analysis.

Voxel-based approximation (VBA) Implementation

While the majority of VBA (7.7) is written in Python, the most computationally expensive portions of our algorithm (the embedding of spheres into the voxel matrix and associated memory operations) are written as a C++ extension module to Python. This puts it on equal footing with LLS and MLE, for which we use pre-written optimized modules.

As in all of our experiments the transmitter is not moving. We set the decay constant $\boldsymbol{\lambda} = \mathbf{1}$. We will evaluate the effect of motion on our localization algorithm in future work. In all scenarios, we defined the physical area of operations (for the VBA algorithm) as the 150m horizontal radius around the transmitter with a possible elevation range of -50m to 100m above and below the transmitter. From experimental data we found that this range covered the maximum flight area of the UAS (100x100x15 meters) for all experiments with room for predictions.

To test the scalability of our algorithm in terms of error and computation time, we evaluate multiple matrix sizes as summarized in Table 7.2. Each matrix size corresponds to the same physical 300x300x150 meter space around the transmitter, but changes the ratio of the physical volume that each voxel in the matrix represents.

For each scenario and matrix size, we run an analysis without specifying topography (allowing any point, even those hovering in the air as valid guesses), and we re-run the analysis specifying a horizontally-flat topography $\boldsymbol{T}$ at the elevation of the transmitter.

Table 7.2: Evaluated VBA matrix sizes.

Name	Size $x \cdot y \cdot z$	Ratio
G10	$\mathbf{10 \cdot 10 \cdot 5}$	**1** voxel : $\mathbf{30\ m^3}$
G50	$\mathbf{50 \cdot 50 \cdot 25}$	**1** voxel : $\mathbf{6\ m^3}$
G300	$\mathbf{300 \cdot 300 \cdot 150}$	**1** voxel : $\mathbf{1\ m^3}$

This limits guesses to be within $\pm\mathbf{1}$ voxel elevation to the transmitter, but does not restrict the horizontal search space.

Linear least squares (LLS) Implementation

We compare our algorithm to the unweighted Linear Least Squares of equation (7.5), as implemented in the Python SciPy module version 1.0 [259], which uses the optimized gelsd LAPACK driver. During experimentation, we also compared to the slightly computationally faster QR decomposition of the same linear system using the Numpy QR decomposition and the SciPy triangulation solver. QR produces nearly identical results to LLS, so we omit it from error rate comparison graphs, though the results are computed at a lower computation cost. We explore this further in Section 7.3.6.

Maximum likelihood estimation (MLE) Implementation

The approach of Maximum Likelihood Estimation for localization sets up a likelihood function based on a model using the measured data and tries to find the parameters that maximize that likelihood function. We assume that our variables are independent and follow a Gaussian normal distribution. Hence, we must find the $\boldsymbol{x, y, z}$ that maximizes the likelihood function (7.8) given $\boldsymbol{m}$ measurements, with each measurement comprised

of coordinates of the UAS (x_i, y_i, z_i) and a computed distance d_i from (7.2).

$$L(x, y, z) = \prod_{i=1}^{m} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{\frac{-\left(d_i - \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}\right)^2}{2\sigma^2}} \tag{7.8}$$

$$LL(x, y, z) = \sum_{i=1}^{m} \left(d_i - \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}\right)^2 \tag{7.9}$$

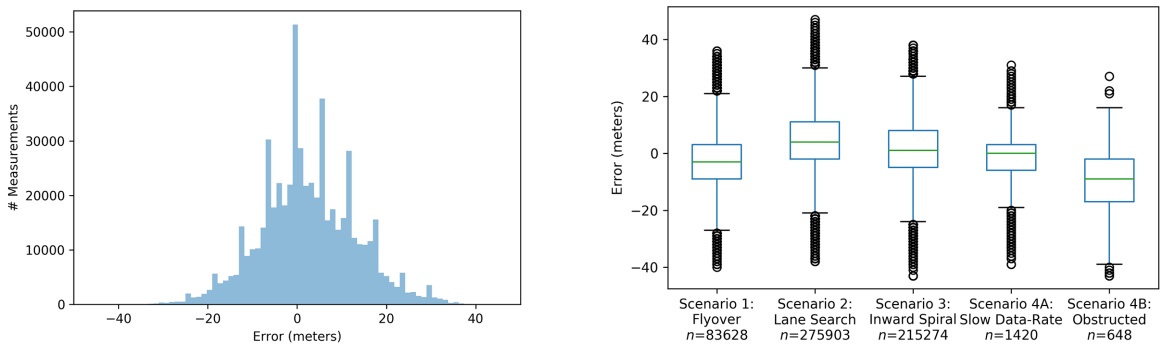
We determine the parameters x, y, z that minimize equation (7.9) by running the SciPy minimize method, which uses the L-BFGS-B algorithm [260] for bounded minimization (with bounds matching the 300x300x150m geographic area as configured for VBA).

7.3 Evaluation

As described in Section 7.2.2, we run 48 flight experiments in four different scenarios, localizing the transmitter using VBA, LLS, and MLE. Due to time constraints and the computational cost of LLS and MLE (whose run-time time grows with the size of the data), we re-sampled the data at 100ms by taking the mean of the time window. This allows a direct comparison of VBA to these other algorithms. We evaluate the performance of VBA compared to other algorithms in terms of error rate and computation time under multiple scenarios and configurations of parameters with the goal of understanding the efficacy of VBA and the constraints of WiFi based localization.

7.3.1 RSSI to Distance Conversion

As outlined in Section 7.1.1, the first step of localization is to convert RSSI measurements to distances using equation (7.2). To do so we used a calibration constant $A = 30$, which we computed by minimizing the mean error from initial calibration experiments



(a) Histogram of errors across all scenarios.

(b) Boxplots grouped by scenario.

Figure 7.6: Distribution of distance errors from VBA experiments.

when prototyping the VBA algorithm.

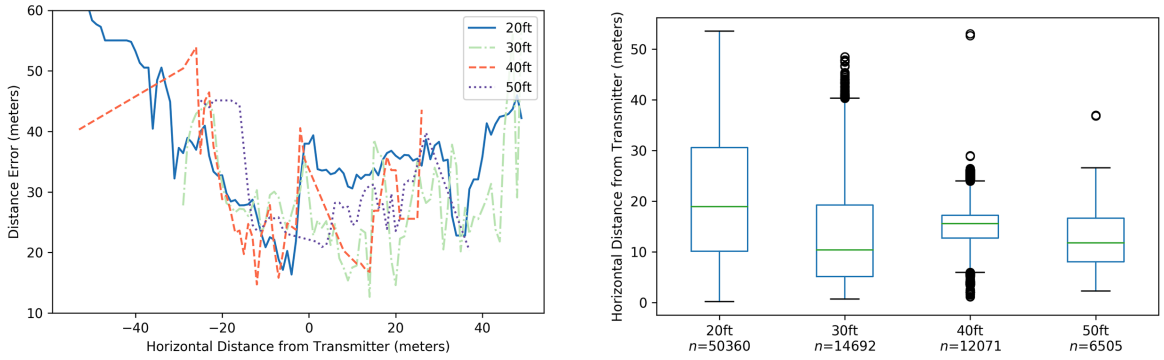
We compared the error (in meters) using the difference between the known GPS coordinates of the UAS and transmitter as ground truth. The distribution of errors across all scenarios and trials is summarized in Figure 7.6a. Our wireless adapters only return RSSI in integer quantities, which leads to gaps in computed distances. This, in turn, may contribute to the blocky distribution.

Figure 7.6b further delineates the distributions by test scenario. As expected, the conversion process produces a Gaussian-like distribution of errors. The mean of the errors shift based on the scenario setup.

7.3.2 Scenario 1: Effect of Altitude

For Scenario 1, we average trials across all three locations, differentiating by altitude. Because the trials take different lengths of time to complete due to variation in speed and starting location, we align the trials based on the horizontal displacement from the transmitter. As the UAS passes the transmitter we assign negative displacement values.

We compute error as the difference between two points in 3D space: the actual location



(a) Error rate of VBA G50.

(b) Effective range by altitude.

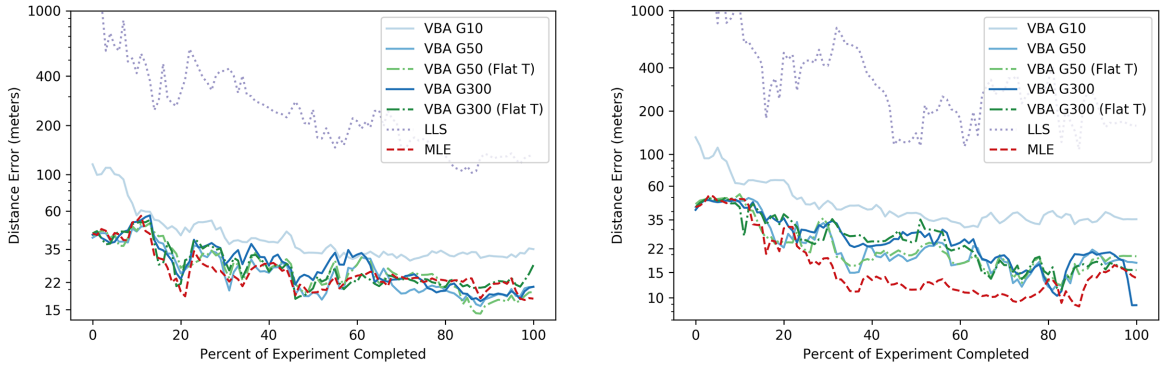
Figure 7.7: Scenario 1: Altitude Variation

of transmitter (as determined by GPS) and the predicted value of transmitter. As VBA can return multiple predictions of equally likely candidates, we take the root mean square of the distances of predictions to the ground truth. We group our data by altitude for each algorithm and configuration.

We present one algorithm as an example (VBA with a matrix size G50 as defined in Table 7.2) in Figure 7.7a. As is the case with VBA G50, when examining other algorithms, the effect of altitude on error rate is not quite clear. This is could be for a number of reasons. One factor could be the 1-dimensional flight pattern. By flying in a straight line, there is insufficient orthogonality in the data to effectively predict a point in the full three dimensional space.

Increases in altitude result in changes in overall distance from the transmitter. As the UAS increases in altitude above the transmitter, the maximum horizontal displacement from the transmitter at which it is still able to capture packets shrinks. This can be seen in the sharp jumps and missing data in the figure at greater horizontal distances.

In Figure 7.7b, we examine the distribution of distances at which we receive packets in Scenario 1. While at 25 foot altitude, packets as far as 54 meters away horizontally



(a) Scenario 2 (Lane Search). (b) Scenario 3 (Inward Spiral).

Figure 7.8: Localization error rate scenarios 2 and 3.

are received; as the altitude increases, the horizontal range where packets are received shrinks.

Because a UAS has to fly high enough to avoid hitting obstacles, we want to pick the highest possible altitude that still has good reception. For our other scenarios, we chose an altitude of 25ft, which should safely clear trees and shrubs in the surrounding area. However in an urban environment this altitude would not be enough to clear all obstacles, such as buildings. This may mean that a UAS operating with this system would likely need to be programmed with an obstacle avoidance algorithm, such as [261].

We initially experimented with altitudes $>$ 50 feet, but the low number of received packets per trial (zero to a dozen) forced us to limit our UAS height to 50ft. For localization at higher altitudes, it may be necessary to switch to a different type of signal with longer propagation characteristics, such as cellular.

7.3.3 Scenario 2: Lane Search

For scenario 2, we again average across all trials and locations. The horizontal displacement is more variable for this scenario. The UAS periodically moves closer and

then farther away from the transmitter. Hence instead of displacement, we align our trials based on the completion percentage of the individual trials as defined by the reception of first and last packet. We show the error by algorithm in Figure 7.8a. Note this figure uses a logarithmic scale for the Y-axis.

As the scenario progresses, accuracy slowly improves for all algorithms. All forms of VBA outperform LLS, and perform as well as or better than MLE. The smallest size matrix, G10, has a higher average rate of error than larger matrix sizes. Despite having higher granularity, G300 does not have lower error. This may be because a larger matrix size is more sensitive to error and spheres that would intersect on a coarser voxel mapping can miss each other when using finer granularity. This sets G50 as the most suitable size for this scenario.

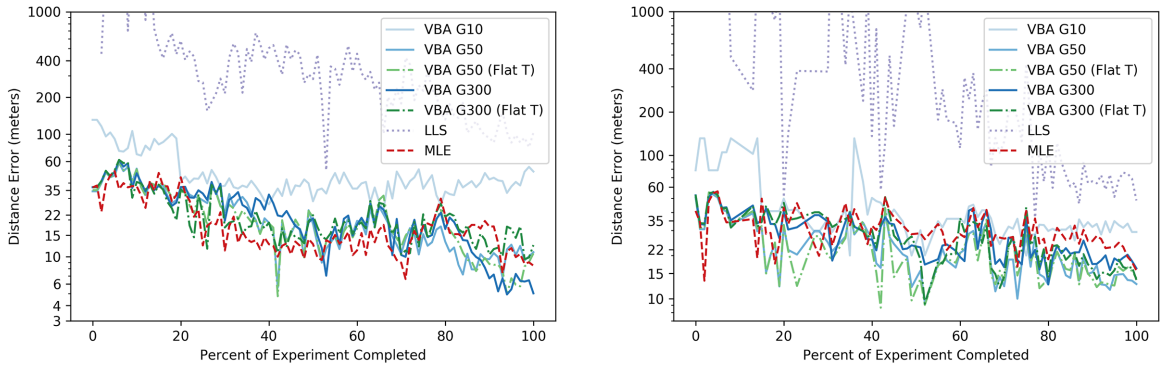
Specifying a topography does not negatively effect error. For most of the scenarios, VBA G50 (Flat T) and VBA G300 (Flat T) do just as well or better than their unbounded counterparts.

7.3.4 Scenario 3: Inward Spiral

In scenario 3, as shown in Figure 7.8b, the results follow a pattern similar to the lane search. This scenario is analogous to a situation where the UAS uses some mixture of localization techniques to identify the general area of the transmitter and then dynamically updates its course to circle in on the object.

The middle matrix size of G50 again yields the best VBA performance. VBA does best overall, with a lower error compared to the lane search. LLS has higher error and more inconsistency. MLE has lower error across the trial duration compared to the other algorithms, but as we explore in Section 7.3.6, this comes at much higher computational cost. If computational power is limited, VBA is the better choice in this scenario.

7.3.5 Scenario 4: Mock Disaster



(a) Scenario 4A (Unobstructed).

(b) Scenario 4B (Obstructed).

Figure 7.9: Scenario 4 error rates, aligned by percent of experiment completed.

Scenario 4 has fewer data points to use in the localization algorithm due to the lower packet transmission rate of the transmitter. This is analogous to a more realistic scenario where, instead of a continuous stream of packets, a device periodically sends an 802.11 probe request, as the 802.11 standard dictates when the interface is on but the device is not connected to an access point. The lower data packet rate produces sparser data, and as a result the error graphs are more variable.

In the unobstructed case (scenario 4A), as shown in Figure 7.9a, VBA performs similar to the inward spiral search. Despite receiving fewer measurements, VBA has a similar error pattern to scenario 3, decreasing as more measurements are collected (albeit slower than in the case of the higher transmission rate). MLE, however, performed worse compared to the previous scenario.

The algorithms perform similarly in the obstructed case (scenario 4B), as shown in Figure 7.9b. There are higher variations in error, as even fewer packets were received than in scenario 4A. As a result, there is overall a slightly higher error rate. However, the results are very promising, as despite receiving fewer packets and the presence of obstructions, the error rate remains comparable to scenario 3. Notably, unlike scenario 3

where MLE attained the lowest error rate, multiple configurations of VBA matched the performance of MLE in scenario 4A, and others outperformed MLE in scenario 4B.

The unobstructed case, performed similarly with fewer variations in error and overall slightly lower error rate. The results are very promising, as despite receiving fewer packets and the presence of obstructions, the error rate remains comparable to scenario 3. Notably, unlike scenario 3 where MLE attained the lowest error rate, multiple configurations of VBA matched the performance of MLE in scenario 4A, and others outperformed MLE in scenario 4B.

While the algorithms operated with lower measurement count without a significant increase in error in scenarios 4A and 4B, the horizontal displacement from the transmitter at which packets were received varied , as shown in Figure 7.10. The lower packet sending rate and physical obstruction combine to narrow the effective transmission range. This issue is irrespective of the algorithm used for localization, but a limitation of packet transmission and reception using WiFi. In future work, we plan to look at alternatives, such as cellular.

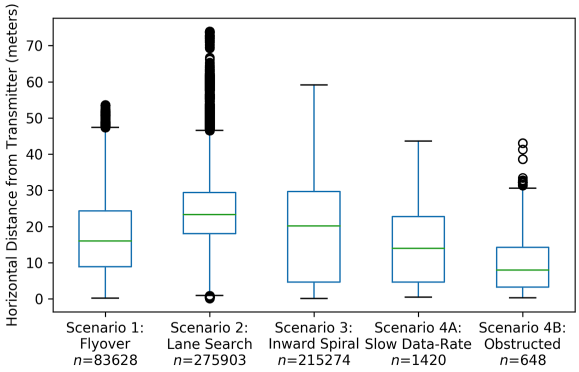


Figure 7.10: Horizontal reception range by scenario.

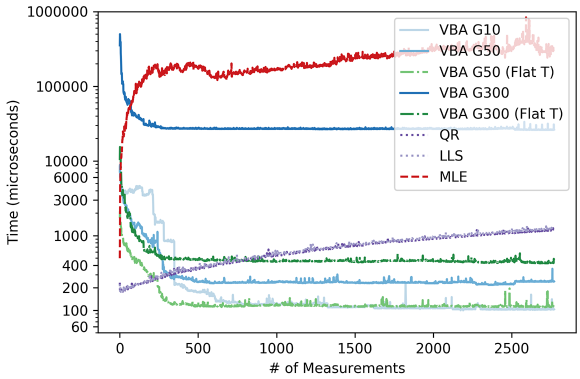


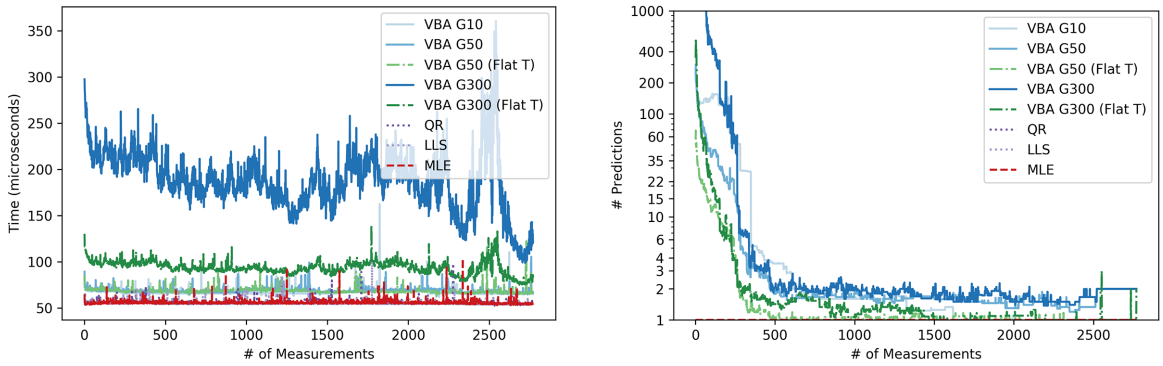
Figure 7.11: Total algorithm run time per measurement.

7.3.6 Computation Time

We evaluate the computation time of VBA and compare it to LLS, QR decomposition, and MLE. For each algorithm, each prediction is divided into two steps: an *UPDATE* step, which adds new measurements to the model; and an *APPROXIMATION* step, which uses the model to generate a best prediction of transmitter location. We record the duration of each step and the sum is the total computation time to calculate a prediction for the transmitter's location. We average across all scenarios and trials based on the number of measurements collected.

The total compute time (in microseconds) for each new prediction is shown in Figure 7.11. As expected, while the VBA algorithm maintains a relatively constant cost, LLS compute time increases as more measurements are added. QR decomposition is an alternative to LLS, producing identical results (in most cases) but with slightly lower computation time for any given count. However, it too follows the same trend of increasing computation time as the size of data grows. MLE has a much higher computational cost. Further, prediction time also increases as more measurements are collected.

When we examine the performance of VBA by matrix size, the results are as expected. Smaller matrix size leads to faster computation. VBA G10 is faster than VBA G50, which



(a) Update time per measurement.

(b) Total number of predictions by algorithm.

Figure 7.12: Closer look at algorithm performance.

is, in turn, faster than VBA G300. As shown in the previous section, the average error rate of the larger matrix G300 is no better than G50, but comes at a high computation premium, which makes it unsuitable for this situation.

When specifying a topography, we see a substantial performance improvement, in terms of computation time. When we specify a flat topography, we can eliminate most of the 3D matrix from the search (the voxels correspond to empty space). As we see in the figure, this means we can have comparable error with substantially faster run time. Our future work includes examination of the trade-offs of using more complex topographies. For the first 500 measurements, VBA has a spike in computation time. If we break down the components of computation time we can identify the source of the spike. When we examine the *UPDATE* step of our algorithms, as we show in Figure 7.12a, we see a relatively constant computation time. Note, unlike past figures, this figure is not using a logarithmic scale. In the case of VBA, the *UPDATE* is the process of computing a new sphere and embedding it into the matrix. In the case of LLS, it is adding a row to the system of equations (7.5). For MLE it is expanding a list of known measurements.

We see that a large part of the work for VBA is done in the update portion of the

algorithm as VBA computes a sphere and updates the corresponding grid values, which is computationally expensive. On the other hand, LLS, QR, and MLE have minimal work to perform the update. The somewhat downward trend of VBA is likely due to sphere size, which is determined by distance. Smaller distances generate smaller spheres, which are faster to compute. Because scenarios 3 and 4 end with the UAS in proximity of the transmitter and have a longer run time (and as a result more measurements), they likely account for the dip. In the figure we see that no large initial spike exists in the update portion of VBA, leading us to conclude that the initial spike is part of the *APPROXIMATION* step.

To account for the initial spike we looked at the difference in the output of VBA and the other algorithms. While the other algorithms produce a single prediction, VBA generates a probability matrix for the entire physical space. When we query VBA in the *APPROXIMATION* step, we are asking for a list of coordinates with the highest possible probability and there could be more than one candidate. The number of guesses by algorithm is shown in Figure 7.12b. While LLS and QR stay constant at 1, VBA starts with thousands of possible guesses, that are then refined to a handful of possibilities. This most likely accounts for the initial spike in total prediction time, which could potentially be optimized in future work.

We also ran VBA G50 with no re-sampling on the RPi, with an *UPDATE* and *APPROXIMATION* computed for each measurement received. We found on average that VBA G50 took 25 seconds with a maximum time of 250 seconds. Overall, in our experimentation, each trial took between 5 to 10 minutes, depending on the specific scenario. On average running VBA takes 20% of the total experiment time and meets the constraints of a real-time system, even when running on low performance hardware.

7.4 Conclusion

The rise in availability and sophistication of UASs has led to exciting new possibilities for disaster response. While many existing commercial applications focus on vision-based solutions, wireless sensing shows promise as a useful tool for locating affected individuals.

To this end, we proposed a Voxel-based approximation (VBA) algorithm for real-time multilateral localization. Unlike other commonly used algorithms, VBA computation time does not increase as measurements are added to the location model. Through extensive testbed evaluation, we demonstrated that localization errors are low in all cases when compared to unweighted Linear Least Squares and were comparable to Maximum Likelihood Estimation. Further, VBA outperformed LLS (including the QR decomposition variant) and MLE in computation time. While LLS, QR, and MLE have increasing costs per measurement, VBA maintains a bounded computation time proportional to the size of the matrix. Similarly, while the data storage requirements for LLS and MLE grow with the data, VBA requires a fixed amount of memory based on matrix size and independent of the size of data, which can be preallocated at initialization. These characteristics make VBA well-suited for use on resource-constrained devices such as UASs, and in time-constrained scenarios such as search and rescue.

We found that specifying a topography did not have a significant impact on error rate but nonetheless saved computation time. In future work, we would like to examine the impact more complex topographies (for example topographies including buildings) have on performance and error. Even with infrequent transmission rates and obstruction, VBA maintained error rates less than 35 meters, and in the most difficult case, Scenario 4B, outperformed our implementation of MLE.

While LLS and MLE provide a single location point, VBA maintains a probability matrix for the entire physical space and produces multiple predictions. This feature is

likely to be useful for path planning of UASs. A UAS could get a relative search area from the first received packet. As the UAS closes in on the search area and more packets are received, the number of probable locations to search narrows. The UAS can investigate multiple points in order of probability.

While VBA preformed well under the imposed constraints, the use of 802.11 2.4GHz and RSSI are challenging for disaster response applications. Even in open space experiments, with minimal reflection and radio interference, RSSI distance conversion (equation (7.2)) has a wide distribution of error (across our experiments a standard deviation of 10 meters). Additionally, commercial wireless adapters provide only integer level RSSI values, which leads to jumps in possible range estimates.

The use of 802.11 is further complicated by the effective range of consumer 2.4 GHz WiFi devices. Even at low altitudes of 20ft, the effective horizontal range of 802.11 (from our experiments) is approximately 50 meters, with most measurements received within only 30 meters of the transmitter. Altitudes above 50 feet receive few to no packets. Given a typical building is about 13 feet per floor [262], at 50ft a UAS could not clear four story buildings. Thus the current version of VBA is well-suited to suburban residential areas, but would need further development for metropolitan areas with taller buildings. This suggests a need for obstacle avoidance, and also limits the field of view for applications that also rely on a white light or thermal camera. Further, a 50 meter effective horizontal range shows that our 10-30 meter error from our experiments does not substantially shrink the search space.

As VBA is not tied to the underlying range estimation algorithm, VBA could be adapted to other technologies such as 802.15.4 and LTE. While the ranging computation would need to be modified, the constant time performance of VBA would enable real-time localization on resource constrained hardware such as UASs. Applications could include search and rescue, as well as network mapping, for example of IoT WSNs.

So far in this part of the dissertation, we have explored scenarios for evaluating and improving the physical infrastructure of wireless networks. Unfortunately, as we will discuss in part II of this dissertation, infrastructure is only one possible hindrance to communication. In the remaining portion of the dissertation we will cover how social and political factors may prohibit or limit communication on computer networks, and new ways of protecting the ability to freely communicate online.

Acknowledgment

This work was done in collaboration with Elizabeth Belding. Thanks to Coal Oil Point UC Reserve for allowing us to utilize your space for our aerial testbed.

Part II

Facilitating Open Communication

Chapter 8

Free Speech Online

As discussed in chapter 1, the Internet is a critical tool for communication across the globe. As of 2019, 4.1 billion people, roughly 54 percent of the world's population, are connected to the Internet [1]. A substantial part of Internet activity comprises users who learn, play, converse, and access content for work, entertainment, and intellectual growth. OSNs, such as Facebook and Twitter, have in particular become major hubs of communication. In June of 2017, Facebook surpassed 2 billion active users [263]. Individuals, companies, and bots generate a massive amount of content. Every second, users worldwide post an average of 6,000 tweets on Twitter [264] and "like" an average of 9,200 items a second on Facebook [263]. When asked what aspects of the Internet users value, the ability to freely speak, share, and trust were among the top concerns for users [265].

The Internet is the dominant tool for people to access and share information. As a result, the Internet is central to free speech and dissemination in today's world. Recognizing the importance of online discourse, the United Nations (UN) considers that the protection of human rights, especially free speech, should fully extend to the Internet [266, 267]. Despite the UN's assessment, governments, corporations, and individuals often restrict what users can say online and punish those with dissenting views. The user stated

values of free speech, sharing ideas, and trust are under threat [7]. Around the world, Internet freedoms are restricted [268, 269, 270, 40, 36, 39]. Even liberal democracies, typical advocates of free speech, increasingly restrict content [33, 271, 272, 273, 274]. While the United States acted in defense of these freedoms world wide in 2010 [275], the current United States administration campaigned, among other things, on "closing that Internet up in some way" [272] and has eliminated net neutrality [276]. Even as the Internet continues to expand, digital freedoms are a resource under threat. Globally, suppression of free speech and press freedoms is on the rise [277, 32], accompanying a rise in authoritarianism [278, 279], which threatens the foundations of functioning democracies.

Users often turn to technological solutions to combat such threats to civil liberties. For example, popular anonymity tools such as Tor [64] provide network level anonymity, while person-to-person messaging tools, such as Signal [65] and WhatsApp [66], allow private communication using end-end encryption. Nevertheless, safe public communication remains a challenge. Individuals conversing publicly on OSNs open themselves up to legal and physical dangers, encouraging self censorship and stifling discourse.

Reputation and trust are likewise eroded. In an era of "fake news", users struggle to identify which OSN accounts and posted content can be trusted [280, 281, 67, 68]. Adversaries to open discourse deploy armies of operatives masquerading as legitimate users to sow division [69, 70]. Even trusted news outlets using OSNs can be hacked to spread misinformation [71, 72, 73].

Existing counter-censorship tools inadequately address censorship of OSNs. Groups seeking to broadcast their ideas to the widest possible audience prefer OSNs as the platform for communication due to their pervasiveness. Although these groups seeks to maintain credibility, they may wish to retain anonymity out of fear of reprisal. However, these two requirements, credibility and anonymity, are difficult to address simultaneously.

8.1 Outline

Our research aims to understand core issues around freedom of speech online for communities that are particularly vulnerable to censorship, such as journalists, political activists, and minorities and to develop technological solutions. Over four years we visited communities in Zambia, Turkey, and Mongolia to investigate issues of free speech on the Internet and in particular OSNs.

As summarized in the introduction, our work focused on two major issues: 1) categorizing social challenges and 2) facilitating open discourse. Chapter 9 characterizes social challenges to free speech. It is based on work published in ACM LIMITS '17 [63]. We identified common actors and methods of censorship as well as some technological needs unmet by existing tools. We found that public group discourse, while maintaining anonymity and preserving reputation on OSNs, was one such unmet need.

In Chapter 10, we expand our characterization of social challenges and show how close partnership with stakeholders can be used to develop tools for facilitating open discourse online. Developed around the requirements we identified, in this chapter we present SecurePost, a novel tool for verified group-anonymity on OSNs. This tool was developed through research and partnerships with affected individuals as a means of balancing personal anonymity with group credibility. We present the research undertaken for developing SecurePost, its technical contributions and operation, and an evaluation of our work. This chapter includes work published in USENIX Workshop on Free and Open Communications on the Internet [74] and the Journal of Internet Services and Applications [75].

8.2 Key Contributions

This work has progressed both the understanding of global issues surrounding freedom of speech and provided new systems for protecting those freedoms.

- Over the course of our project, we conducted 109 interviews and surveyed 526 individuals in Lusaka, Zambia; Istanbul, Turkey; and Ulaanbaatar, Mongolia. Partnering with interdisciplinary experts in computer science, communication, and media studies, we used this data to characterize social challenges to free speech online, particularly as it relates to OSNs.

- We documented that preserving reputation while maintaining anonymity for group discussion on OSNs is a need that existing technologies do not adequately address.

- We developed a novel tool called SecurePost, which allows individuals to share a single group identity while retaining individual anonymity on OSNs. SecurePost comprises three coordinated modules: an Android application that allows group members to post content and manage membership; a proxy server that relays posts to social networks; and a browser extension that allows members of the public to verify those posts. Together the modules provide group-anonymity coupled with an ability to verifying the integrity and authenticity of posts.

- This work evaluated the use of time-synchronized hash chains for group account sharing and provided an evaluation of why such a scheme is unsuitable for group verification.

- We introduced an image based cryptographic signature for use on social media that used compression resistant encoding. This process is applicable to platforms that limit character count.

- Finally, this work outlined practices for working with stakeholders in order to provide culturally appropriate technological solutions using iterative design practices.

8.3 Broader Impacts

This work has already made an impact on the broader global community.

- The application has had over 400 installs and users have formed 68 groups in multiple countries, including the USA, Mongolia, Turkey, Zambia, as well as other countries where censorship is a problem (59 countries in total).

- Our team conducted broader Internet and security trainings with community partners where we trained over 300 individuals, including in Beirut, Lebanon; Lusaka, Zambia; and Ulaanbaatar, Mongolia. For the trainings, we collaborated with vulnerable organizations and individuals, such as LGBTQ centers, HIV prevention centers, non-governmental organizations (NGOs), journalists, political activists, minority groups, human rights groups, and university members.

- Aside from the academic venues previously stated, we presented our work at a number of public venues, including the Internet Freedom Festival 2016 in Valencia, Spain; After Tahrir: Egyptian Revolutionary Experiences and Future Visions 2016 in Santa Barbara, USA; and RightsCon '16 in San Francisco, USA.

Chapter 9

Limits to Internet Freedoms: Being Heard in an Increasingly Authoritarian World

This chapter explores the growing limits to free speech based on research conducted in Lusaka, Zambia; Ulaanbaatar, Mongolia; and Istanbul, Turkey from 2014-2016. As part of our research we reached out to diverse sets of communities to investigate Internet Freedoms and in particular their relation to the use of OSNs. We use this research as the basis of discussion into the limits, actors, and concerns in this space. Over the course of our research, we formally interviewed 110 people and had informal conversations with dozens more individuals. While our work provides only a small window into the broad set of limits that individuals encounter in on-line access and speech, the diverse perspectives, cultures, and struggles serve as a platform of understanding the limits to Internet freedoms in a global context.

9.1 Limits to Speech and Access

During our research, we sought out a diverse set of individuals, with independent and sometimes conflicting agendas. To understand the barriers they encounter, it is helpful to explore the competing motivations, the adversaries, and the tools they use to silence speech and block access. We provide an overview existing tools, areas of growth, and a discussion about some of the ethical considerations when designing free speech technology.

9.1.1 Seeking the Voices

Before understanding the limits on digital free speech and access, we identify the groups facing these barriers. From our research across the three countries, we interviewed a multitude of groups that struggled to access and post content on-line. These groups include: political activists, the press, minority groups, watchdogs and NGOs, and unaffiliated citizens.

Political Activists: Political activists are the most common targets of censorship. Ruling politicians silence and discredit political rivals both physically and digitally. In Turkey and Zambia the ruling parties exercise legal suppression of dissenting opinion, shut down websites and arrest opposition leaders. Voices speaking out against the current government are prime targets for censorship.

The Press: Journalists shared similar accounts. In Turkey, news organizations, like Zaman [282], are physically raided and journalists are arrested for publishing content that defies the government. In Zambia, radio stations and newspapers are likewise raided and, in multiple reported incidents, shut down for printing, streaming, and publishing physical and digital content. When press organizations are shut down, some reporters continue to work as citizen journalists, publishing news on blogs and OSNs. Many face

arrests, lawsuits, and censorship of their content. Even single tweets on topics such as governmental corruption, lead to arrest of journalists [283].

Minority Groups : People face censorship for reasons other than speaking out against the government. Those investigating minority issues are especially vulnerable. Journalists reporting on Kurdish treatment in Turkey face arrests, confiscation of their devices, and bullying. LGBTQ activists in Mongolia struggle with language censure that prohibits posting impolite words, including sexual terminology, even when using medically appropriate language [284]. When soliciting information about safe sex, their material is labeled pornographic in nature and prohibited. In Zambia, we interviewed an HIV health center that faced issues of getting past the stigma of the disease. Minority groups often look to technology to overcome societal barriers and engage open discussion.

Watchdogs and NGOs : Watchdogs and NGOs also conveyed difficulty in reporting factual information. In Mongolia, groups are sued for libel when reporting on corporate environmental damage, and free press watchdogs face opposition when reporting on government crackdown on media. In Zambia, NGOs overseeing water projects are opposed by people unwilling to report corruption due to pressure from corporations and local governments. As these groups rely on accurate information to function, censorship and external interference inhibits their success.

Unaffiliated Citizens: Unaffiliated citizens are also not exempt. In Turkey, we interviewed a gay man who was arrested and fined over a tweet [285]. Due to laws against insulting the government of Turkey, a single tweet is enough to warrant arrest. This discourages people from speaking out in the first place. Even if individuals do not find themselves in violation of the law, they can become collateral in large-scale censorship efforts. In times of social conflict, governments, like Turkey, shut down access to websites

for all citizens [286]. Aside from government pressure, people living in Mongolia, Zambia, and Turkey looking for information such as LGBT issues face on-line bullying and social stigma.

9.1.2 Assessing the Adversaries

There are groups whose goals motivate them to restrict Internet freedoms. Agents imposing these limits are adversaries of free speech and access. They include: government, corporations, and communities.

The Government: The most dominant adversary is usually the government. Governments all over the world litigate and enforce censorship of content [36, 37, 38, 39, 40]. Governments may do so to proscribe social norms, to stifle minority opinions, to ensure "safety", or out of political self-interest - suppressing news that would make the government look bad. Of the three countries in which we conducted research, the government of Turkey most directly imposed limits on free speech. Many times in the past several years, the Turkish government used technology to censor voices and cut off access to OSNs including Twitter, Facebook, WhatsApp, and YouTube [286]. Turkey also aggressively enforced laws by policing content posted on-line, tapping phone conversations, and arresting political dissidents. In Mongolia, our interviews identified governmental focus on filtering speech, and banning and blocking sites based on content. In Zambia, our interviews suggested a government that acted through arrests and lawsuits to silence opposition.

Corporations: Corporations are another critical adversary to on-line speech and access. They bring lawsuits against journalists and individuals under libel laws, using these suits as a deterrent from reporting on issues of corruption and environmental damage. In particular, OSNs play a large role in imposing limits on speech. By tracking users and

gathering personal information, large OSNs like Facebook and Twitter provide tools for others to reveal identities of users. Reporting tools that can be used to flag posts as improper can also be used by other adversaries to silence speech. Additionally, by isolating users in content bubbles of like-minded users and suggested posts, users are shielded from dissenting ideas and opinions. Even if speech makes it onto OSNs, the echo chamber effect [287] can prevent it from ever being viewed or heard.

Communities: The last and often most influential adversary limiting free speech is a person's community. Individuals that post views on controversial issues can be targeted by cyber-bullies. Journalists reporting on sensitive topics, such as on Kurdish issues in Turkey, face constant barrage of hateful posts. In Zambia, it is difficult to voice an opinion in an on-line forum. A user's ethnicity, gender, and past posting record stereotypes the user. Resulting responses from the on-line community frequently target the physical characteristics or past political affiliation, over the content of discussion. Even in the confines of one's household, people encounter limits to their on-line freedoms. During a security training for a gender-based violence center in Mongolia, we heard accounts of how husbands and partners break into email and OSNs accounts of their wives. The goal is to monitor communication and content access, and the result can be domestic violence.

Even when adversaries do not specifically target an individual they can force self-censorship. The same adversary that limits on-line communication can restrict physical media and create a conversational stigma. With an adversary in every corner, Internet freedoms are severely limited. When people are afraid to ask questions, or do an Internet search for fear of reprisal, they are cut off from resources that could improve their physical and mental well-being.

Figure 9.1: Political Cartoon by Kiss Brian Abraham commenting on the use of technology for freedom of expression in Zambia.

9.1.3 Techniques to Limit Freedoms

Adversaries place limits on Internet freedoms through legal action, technology, threat and violence, and control of infrastructure. These techniques allow adversaries to censor content, track users, and log communications.

Legal Action: Governments pass laws criminalizing discussion of certain topics. Even if no further action is taken, those laws serve as a deterrent for voicing opposing view points. Some laws, such as those banning insult of government officials in Turkey, are far reaching and suffice as cause for prosecuting individuals perceived as political threats. For anything ranging from public criticism to satirical tweets, celebrities, newspapers, activists, and unaffiliated citizens face criminal charges [288, 289].

Sometimes these lawsuits border on the absurd, such as when Dr. Bilgin Ciftci, a Turkish physician, went to court over a meme he created, comparing Erdogan, president of Turkey, to Gollum, a fictional character from Tolkien's *Lord of the Rings* [290]. The doctor lost his job and is facing 2 years in prison. The outcome of his court case hinges on testimony from a panel of experts expounding on the moral character of Gollum in order to determine whether the meme was insulting the public official. Such wide enforcement makes even mundane opposition to the government dangerous for an individual.

In countries with strong libel laws, like Mongolia, politicians and corporations sue against unflattering reporting by alleging wrongful defamation of character. Using expensive legal teams, these libel plaintiffs are able to silence opposing viewpoints, intimidating those who may not have the monetary resources from fighting a legal challenge. Threats of libel suits also act as a deterrent.

Unlike voiced speech, which needs to be recorded to be saved, on-line content is tracked and archived. Even passing thoughts, formulated into late night tweets, that are later deleted, become a matter of record and can be called into evidence at a later date. When every word written can be used against them at a later date, individuals self-censor themselves when posting on-line.

Technology: In addition to retroactive enforcement of laws, adversaries use technologies for proactive censorship of content. Governments, such as Turkey, can enact broad DNS and IP bans that block entire sections of the web, targeting news and dissenting opinions [268]. As regimes become more restrictive they may block specific types of network streams, such as Virtual private networks (VPNs) and Tor connections as was seen in Turkey last December [291]. Other governments, like those in China, employ comprehensive filtering of websites by topics and keywords [292]. Mongolia takes a more direct approach by mandating website hosts to install a program to filter content, including comments for slander and rude language, based on an extensive banned word list [293]. This makes access to local content on topics, such as sex, difficult.

When access is allowed, governments actively work to identify users. Internet service providers and MNOs are often forced to register IPs and subscriber identity modules (SIMs) to real names. This allows arrests and intimidation, even on un-named accounts on-line. Some tracking is harder to detect. For example, some governments and individuals have deployed IMSI catchers, which are fake cellular towers that intercept calls and texts.

IMSI catchers can log communication and register a phone's presence at a location, such as a protest [294]. While, in some cases, it is possible to bypass the tracking and censorship technologies with the use of proxy servers and VPNs, this brings other limitations that will be discussed later in the chapter.

Infrastructure: Some limits to on-line speech manifest in a block to on-line access itself. Areas that are rural, underdeveloped, or war-torn, may lack the infrastructure to access the Internet in a meaningful way. This lack of infrastructure can be a byproduct of economic disincentives, difficulty due to physical obstacles, such as terrain, weather, and distance, or in some cases deliberate neglect. In Mongolia, towns we visited on the railway lack Internet access due to the tough terrain, expensive upkeep due to weather, inaccessibility, and lack of economic prospects for MNOs. The lack of incentive is typical of rural communities, including those in Zambia and other parts of the world. As mentioned, sometimes infrastructure neglect is deliberate as it is a way of suppressing a particular community. In the case of the Za'atari refugee camp, Internet access was deliberately not provided in order to discourage refugees from encroaching on the labor market of Amman through on-line work [295].

Even when existing infrastructure is present, access to it can be rescinded. Governments may block Internet and mobile access for a region in response to events, such as protests. During such times, all citizens, not just members of the protest, lose access to news, communication, as well as access to digital financial transactions. This has been the case, among others, in Turkey, Egypt, and Syria [269, 296, 297]. Even when there is no deliberate block, protests or natural disasters can overload mobile networks and disrupt Internet access intermittently [298].

When infrastructure is present and functional, the cost, speed, and quality of Internet connectivity may restrict usage to a particular socio-economic class. Additionally, upload

and download bandwidth and costs are not always symmetrical for users. Internet service providers regularly provide plans that allow downloads at a disproportionately faster rate than uploads. While users may have the capacity to consume content, their ability to voice their own ideas and culture might be limited due to upload caps. Projects that claim to provide free access, such as Facebook's Internet.org [299], may limit which websites are freely available to subscribers. Limiting access to infrastructure can be profitable to companies aiming to control consumer choice but detrimental to user freedoms.

While infrastructure limits to Internet access is comprehensibly studied by information and communications technologies for development (ICT4D) literature, it is important to emphasize how lack of Internet access can be used to suppress the voice of a particular group or minority on-line. Connecting communities to the Internet amplifies their voice both globally and domestically. Hindering connectivity intentionally or through neglect censors a community and muffles their voice. The drawing shown in figure 9.1, by Kiss Brian Abraham, reminds viewers that while freedom of expression through technology is an active part of Zambian cities, those in rural Zambia lack the infrastructure to participate.

Threats and Violence: The enforcement of laws sometimes results in physical altercations. Before going through due process in court, police may make a show of violence when apprehending suspects. A manager was beaten by police at Komboni Radio in Zambia, which offers both radio broadcast in Lusaka and on-line streaming, and the radio station was temporarily shut down [300]. Government shows of force during arrest act as a deterrent for others thinking of speaking out. Even when no probable cause exists for arrest, police and government agents may use force to intimidate journalists or, in some cases, seize belongings. The search and seizure of devices is a major barrier to journalists reporting in areas with no Internet access [301]. Reporters who rely on their phones and

laptops to store and ferry footage from conflict zones and are especially vulnerable to seizure as a means of censorship.

Aside from government agents, violence or the threat of violence is a powerful demotivator on-line. Individuals who post on-line expose themselves to cyber-bullying. Bullies attack users personally using identifying information to tailor attacks. In an attack known as doxing, bullies find and release personal contact information, such as phone, email, or address, thereby inviting escalation against an individual [302, 303]. When personal information is known, attacks can escalate from threats to acts of violence. Associates and family members are likewise potential targets. Not only is this a technique for silencing the target, but acts it as a deterrent for others.

When on-line users post or search for content, they open themselves to targeted acts of hate. For example, searching or posting on LGBT topics can lead others to label individuals as non-heteronormative. These labels can impact job availability, interpersonal behavior, and trigger threats both from society at-large and at home [304]. At the gender-based violence center in Mongolia, we heard accounts that husbands assault their wives based on search history or OSNs posts. When inquiry leads to such grave consequences, individuals are unlikely to take the risk and engage on-line.

There are many techniques that adversaries use to limit Internet freedoms, as we explored in this section. The barriers go beyond technological, extending into the legal, social, and economic. Aspects of these techniques can be countered by tools that, among others, circumvent censorship, anonymize users, and obfuscate communication.

9.2 Limitations of Existing Tools

While aspects of the techniques to limit speech can be circumvented by large libraries of existing tools, these tools have limitations. In our research we sought to assess the successes and shortcomings of these tools to understand the capabilities individuals have to overcome limits. We found existing tools are often a poor fit for marginalized communities and fail to overcome limits in effective communication. To tackle imposed or naturally occurring barriers to on-line speech, users need tools from multiple technological facets.

9.2.1 Censorship Circumvention

As explored earlier, one of the direct ways that adversaries limit access to particular content is through the censure of websites. To overcome these blocks, users use circumvention tools. A common tactic is tunneling content through unblocked devices, such as proxy servers, that sit outside the control of a censoring adversary. Users funnel their normally blocked request via this proxy. The proxy relays requests and mirrors the responses from the desired website. While this technique is popular in regions where governments or corporations block content, it comes with some security drawbacks. Proxies are able to read requests made by the user and modify the results. Free proxy services allow user access in exchange for injecting advertisements into web pages. Users lose the ability to trust responses as the third party advertisers can modify web pages. The proxies are able to intercept user requests and can monitor any unencrypted user materials, such as passwords. Un-encrypted requests and responses can still be intercepted by Internet service providers as well as intermediate network routers. Adversaries monitoring the network can link these activities to a particular IP address.

Similar to proxies, VPNs are used to relay network traffic from a device through an intermediary. This can allow users to access censored content. While proxy servers

are typically used on an application basis, VPNs can be used to project network traffic through a remote server. Unlike proxies, the requests sent through VPNs are encrypted along with all other traffic while en-route to the VPN server. Once at the VPN server, unencrypted requests and responses can still be read and modified. Due to the added computation cost of encrypting and decrypting communication, VPNs are rarely free, and the encryption adds a processing cost to the user's device, which narrows the accessibility to certain socio-economic classes.

Many of the individuals we encountered in our research had heard of proxies and VPNs. In Turkey, where active IP and DNS filtering is common, many of the users, even those less technically proficient, had used proxies or VPNs as a tool for bypassing censorship blocks. Few, however, knew about the benefit to anonymity these approaches provided.

9.2.2 Anonymity

In addition to blocking content, adversaries track users for legal prosecution or as an intimidation tactic. Tracking identities in turn promotes self-censorship by the user. By providing a means of censorship-circumvention, proxies allow possible anonymity between requester and the desired website. However, if requests and responses are not encrypted, outside parties, such as governments and corporations, can still track users. Worse, proxies themselves may keep logs of interactions and share them with adversaries, either willingly or through subpoena. Proxies may keep lists of requested IPs, linked to users, which can serve as a hit-list for an adversary.

VPNs are only marginally better than proxies for anonymity. While all traffic to VPN servers is encrypted, the IPs of both the requesting device and VPN are visible on the network. If both the requester and the VPN is within a part of the network

monitored by an adversary, traffic analysis can link the IPs to the final destination. While VPNs are employed by businesses, the act of using a VPN can still raise suspicion by an adversary policing censorship circumvention. If the VPN server is compromised or legally vulnerable to subpoena by an adversary, it may still be possible to get full access records. Corporations, such as Netflix [305], limiting certain groups of users from accessing content can also block VPN use.

Another popular approach to proxies, which provides better anonymity, is Tor [306]. Tor is a network of proxies that relay encrypted data. Anyone can volunteer to become a relay of this network by running freely available software. Users connect to the network and funnel TCP streams through an entry node in the Tor network. The stream is relayed across multiple Tor relays. For a given stream, each Tor relay only knows the IPs of its two neighbors. Intermediate relays do not know the IP of the original requester nor of the destination. The packets in the stream are encrypted multiple times, like layers in an onion, with ephemeral keys of the intermediate nodes [64]. This type of system makes traffic analysis, linking the requester and intended destination, difficult. However, if enough Tor nodes are compromised, then traffic analysis is possible [307]. Even if adversaries are unable to de-anonymize traffic, downloading or using Tor is visible on the network and can flag an individual as a person of interest. In our research, we found that many technically proficient users knew about Tor, but that new users found the concept confusing and suffered language barriers when attempting to install and use it themselves.

Most OSNs force users to reveal real identity. For example, Facebook, imposes a real name policy as part of the terms of service [308, 309], while Twitter requires a phone number to create an account. Additionally, OSNs log access, including the IP of each request. Corporations can sell this data to other adversaries. Other websites can reuse tracking cookies left by OSNs to identify users. Governments can requisition these user records [310]. Using a service, such as a proxy server, VPN, or Tor can mask the accessing

IP. However, if the account is ever accessed by a device with an IP tied to the user, that single interaction suffices to de-anonymize the entire account. As mentioned, in addition to monitoring censorship infractions, governments and corporations may go after users deemed offensive or dangerous by tying words or site visits to identity.

Revealing personal identities exposes users to threats from governments and corporations as well as bullying and violence from on-line and real world communities, including family members. When users know they are tracked, they self-censor posts and queries, which limits both speech and access to vital information. Even when adversaries lack the ability to identify users on IP, they can de-anonymize users based on the content they post. While using OSNs, users frequently post identifying information. An account using a fake name that posts a personal photo can instantly identify the individual. Less obvious details can still allow adversaries to guess identities, for example naming a school, age, and town of birth might be enough to uniquely identify an individual. As users generate content they expose identifying information. A dedicated attacker can correlate this information and call out the identity of the user. Doxing the user by an adversary exposes them to the threats mentioned in previous sections.

Use of anonymity tools, such as proxies, coupled with meticulous discipline can help protect users from adversaries. Unfortunately, self-censoring all identifying information limits the content an individual is able to post and access. It is difficult to have frank conversations about personal issues, with the worry that every word can be used to reveal identity and expose the user to danger.

9.2.3 Reputation

While anonymity can help individuals access information and post without retribution, anonymous communication has drawbacks. Personal investment brings with it accountability, and while those seeking genuine discourse can use anonymity to be heard, others can use anonymity as a tool to attack. Without the reputation of the individual, anonymous accounts can have difficulty fighting for credibility. This is especially difficult for journalists and media outlets whose credibility is tied to reputation. During interviews in Turkey, we repeatedly heard that journalists were unwilling to use anonymity tools as it would strip them of credibility and prevent them from doing their job.

Nevertheless, over time, even anonymous accounts can earn credibility. Groups that share factual information on anonymous OSNs pages or blogs can build a reputation of credibility, tied to an assumed identity. In Turkey, an anonymous Twitter account, going by Fuat Avni, delivered information ostensibly from within the Turkish government. By repeatedly posting credible information, the account gained millions of followers and became the target of a government investigation [311].

Unfortunately, anonymous groups suffer from a variety of problems. When accounts are blocked or removed, the credibility chain is disrupted. Reestablished groups must provide evidence of continuity or risk forfeiting established reputation. Infiltrators joining the group or seizing the account can tarnish reputation as readers struggle to determine what information is factual and what is planted. Loosely formed groups that span accounts can have unclear affiliations. While the hacker group "Anonymous", for example, has some degree of reputation, almost any anonymous account can claim membership, muddying the message and reputation of the group. Cryptographic signatures can validate assumed identities, but are difficult to use in the OSNs context.

Additionally, there is a difference between anonymity of a group and that of an individual. A media organization may wish to retain its identity and reputation in on-line communication while protecting individuals in that organization from prosecution. The Zambia Watchdog used this approach, combining public and anonymous sources under a single identity to publish critiques of the government and expose corruption [301].

9.2.4 Broad Reach

When individuals and groups manage to make it on-line, their voices are only heard if they are able to reach a breadth of people. There are many tools that enable secure end-to-end encrypted communications for email, messaging, and content sharing. These tools are somewhat effective at disseminating information in a group securely but do little to communicate with broader audiences. Individuals and groups we interviewed were primarily interested in OSNs due to the ability to reach a large audience. Tools with narrow audiences limit viability in many of the use cases. Speaking to an empty room does little to share ideas.

9.2.5 Crowding Out

When a post makes it to OSNs, overcoming the many barriers, it can still be silenced. Governments and corporations increasingly deploy bots, automated programs behaving like users, to crowd out dissenting voices [312, 313]. In comment sections on OSNs and news sites, automated posts can overwhelm real discussion. On sites using ranking algorithms, bots can down-vote posts, forcing them into obscurity. Some governments, such as Russia, go further and employ real people in "troll" farms [314, 70] to control the direction of discussion and suppress opposing viewpoints.

Even mechanisms enacted to protect users are frequently exploited. Reporting functionality, present on most OSNs, allows users to flag posts as harassment or indecent. This is helpful in preventing cyber-bullying. Unfortunately, adversaries use bots or trolls to falsely report posts, generating mass complaints towards a user. Russia has been aggressive in silencing opposing views from popular accounts by falsely flagging content as containing violence or pornography, resulting in temporary and permanent account bans [315]. These attacks exploit automated moderation algorithms of platforms, such as Facebook, to temporarily or permanently ban accounts, thereby silencing dissenting voices.

9.2.6 Technical Literacy and Language

While a wide library of tools, including those discussed, exist to overcome limits to Internet freedoms, there is often a capacities mismatch between the developers and users. One of the most direct issues is language. Many security tools and corresponding instructions are only available in a small set of languages. When discussing security in Turkey, we attempted to introduce users to Tor. We found that Orbot, an Android application for Tor, was not available in Turkish. This was a barrier to usage as all instructions and user interfaces required explanation and translation. No application we examined had a Mongolian translation. While Zambia uses English as its official language, the 73 Zambian native languages were also absent. For a quick overview of language availability for a sampling of tools please refer to Table 9.1. Lack of instructions in a native language limits the ability to understand and use tools effectively.

Security tools are frequently used by those in computing fields who already have some level of technical literacy. Proper use of tools requires an understanding of the threat, purpose of the tool, and its limitations. In our interviews we found variation in

Table 9.1: Number of languages in which tools are available.

Tool	**Languages**
Privacy Badger (Chrome) [316]	10
Confide (iOS and Android) [317]	15
Tor Browser [306]	16
Orbot (Android Tor App) [318]	25
Signal (iOS and Android) [65]	36
HTTPS Everywhere (Chrome) [319]	48

technical expertise. While some were proficient and, in many cases, using tools for on-line interactions, many others were far less technically literate. Many did not understand the mechanisms behind tracking or censorship, when they were vulnerable, or how to protect themselves. Those working in journalism, in highly dangerous conditions may have the interest but lack the resources to get the necessary training to overcome limits. Learning to use tools in a non-native language compounds the issue.

Individuals working with technology are not always literate in the vulnerabilities of their on-line activities. Users often do not worry about security and anonymity until they become targets themselves. When training undergraduates in computer science at Mongolian National University, the group showed little initial interest in learning about security tools. When we showed them a live demo of intercepting complete web pages running over HTTP on an unsecured wireless access point, the level of interest in protecting their identity and communication increased dramatically. Simply making users aware what aspects of their on-line activity is visible and to whom is a powerful first step to raising interest and overcoming future limits.

Even if an ideal tool existed to overcome each technical limitation, language and digital literacy would still hinder adoption. Access to language and technical experience may be tied to particular groups of individuals based on access to education and socio-economic

status. The design and translation of tools can determine who is able to overcome the limits and speak, and who remains silent.

9.3 Discussion

The capacity to speak and be heard is a powerful force with both societal and ethical implications. The decisions behind design, implementation, and deployment of technologies that overcome these limits can have the power to define which groups and ideas promulgate on the Internet. Empowering Internet speech is vital as it shines light on injustices, empowers minorities, breaks cycles of poverty, and assists individuals to succeed. However, the same tools empowering free speech can also be used for hate speech, planning acts of violence, destabilizing governments and societies, or even reinforcing socio-economic divides by favoring particular groups of individuals. The authors of the tools play a crucial role in deciding who these tools empower.

9.3.1 Impact of Design

For a tool to overcome a limit, it has to be used. As discussed in previous sections, even existing tools are not suitable for users who may lack the knowledge, experience, or income to use them. From our research, we observed that proficiency in English and technical literacy tend to favor those who are wealthier and live in large cities.

Language

When developing tools that enable Internet freedoms, the choice of languages to support has consequences. Every country in the world has users that speak major languages such as Mandarin, Spanish, and English, but many countries only have partial adoption [320]. Picking a language can alienate portions of the population for which the

language is non-dominant. Language expectation may bias toward a particular socio-economic class [321]. People who engage in international business or higher education may be more likely to speak a major language. Even without creating new tools, translating existing tools to new languages can reduce the adaptation barrier for currently restricted minorities. Selectively distributing tools can amplify a subset of voices over others. Neglecting to translate a tool that provides freedoms for some, effectively limits freedoms of others.

Technical Literacy

Alongside language is the expectation of technical literacy. Tools that are hard to use and setup, or those with poorly explained limitations can alienate and even endanger groups. While information technology professionals may have the technical understanding to use or learn to use existing tools, the same is not true for users from all domains. From our experiences, journalists and civil rights advocates, especially those who have little funding for I.T. support, face difficulties setting up and using existing tools. Worse still, groups with poor backgrounds in cyber-security may not understand the threat model that a particular tool is designed to counter, leading to a false feeling of security.

Even if a tool is available in a language the user understands, without comprehension of the full security context and without an intuitive design, the user may not be able to use it effectively. Like language, the design and usability of a tool can segregate populations. Ensuring that an application is clear to a novice extends the application's reach and ability to empower. Conversely, ignoring the design and ease of use of a tool can disproportionately favor those with the education and experience to use it, or those with the economic advantage to hire someone who can.

Device and Platform

Choosing a platform or operating system for a security tool limits the user demographic that a tool empowers. Requiring a Twitter account, for example, may alienate users who would otherwise be interested in the security tool, but who lack interest in starting a Twitter account. When applied on a global scale, alienated demographics could comprise the majority of entire countries. In our research we found a high usage of Twitter in Turkey, but when talking to activists in Mongolia and Zambia, we found nearly all favored Facebook.

Likewise, the choice of operating system can segregate populations of users. This is especially true for mobile applications that have experienced rapid growth and change. Selecting iOS over Android can alter the types of groups who are able to use a mobile application. The version of operating system can further subdivide groups. In Istanbul, we found newer Android phones running the latest operating system were quite common; however, when working in Zambia we found phones running operating systems as old as Android 1.6. Android applications not targeting such old versions would not run. Adding backwards compatibility to applications can increase development time, complexity, and complicate usability testing. On the other hand, restricting operating system type or version limits the tools to those who can afford newer devices.

It is important to note that while mobile-broadband usage in the developing world is limited, it is the primary method for Internet access. As of 2016, 41% of the population in the developing world had mobile-broadband subscriptions compared to 8% with fixed-broadband subscriptions [322]. Throughout much of the developing world, mobile devices are the primary means of accessing the Internet. Technologies that are not accessible via mobile, segregate users for whom this is the only method of access.

Ownership of a suitable device, like a smart phone, is still a limit. While most of the people we talked to in the capital cities owned smart phones, in rural communities this is not the case. In Zambia, for example, a 2015 study found only 51% of the population actively used mobile devices and only 13.5% of those devices were smartphones [323]. While it is impossible to tailor a software tool for communities with no hardware, these groups should still factor in ethical considerations. As societies become reliant on technology for protecting freedoms, those without the proper hardware fall further behind.

Connectivity and Power

Lacking access to power and Internet connectivity can be a limit to speech. Between no access and reliable access is a gray zone in which much of the world resides [324]. In tool design, connectivity and power are commonly treated as binary, either present or absent. In reality, Internet access can be unreliable, expensive, or incredibly slow. Power is similarly unreliable. In rural areas, blackouts may be frequent and brown outs, when voltage drops bellow operating norms, may be common. Applications built on the assumption of low latency, high bandwidth, and continuous power may be unusable for these communities. Like other design choices, the network and power requirements of tools selects the demographic that they empower. Developers can overcome some of these restrictions through techniques such as caching data, bundling server requests, and minimizing local computation. Optimization of tools for resource poor environments takes development time and adds complexity. Failure to design and test for situations of limited resources favors those in richer conditions.

9.3.2 Security

While technologies can overcome limits on speech and access, they can present a danger to their users. Even if empowering users is not the priority to tool developers, user safety should be. If a tool is poorly explained, users may not realize that they are not protected against specific threats. For some, speech can put them in danger, leading to incarceration, economic hardship, violence, or even death. While tools typically try to grow a user base, advertising to users without adequately preparing them can do more harm than good. Even experienced users may grow complacent from a feeling of security and make mistakes that expose them to threats.

Like other tools, software focusing on Internet freedoms occasionally have bugs or oversights that create vulnerabilities. For low-risk individuals, a vulnerability may pose little threat. For high-risk individuals, who are under scrutiny by adversaries with high levels of network control, a single vulnerability can suffice to identify users or provide evidence for incarceration. Tool designers are responsible for the integrity of their tools. Like other concerns, keeping tools up-to-date and informing users of potential problems may be harder in particular communities. Users lacking affordable Internet access may not keep their applications updated. Similarly, users who side-load applications due to blocking of larger repositories may never receive application updates. These users might be exposed to vulnerabilities for which their software was never patched. Alternate delivery systems, as well as resource-aware update sizes can help protect these users.

9.3.3 Ethical Concerns

Misuse for Harm

Some worry that agents seeking to do harm will misuse tools intended for Internet freedoms. Encryption tools enabling human rights activists to talk without fear can be used by terrorist groups to coordinate attacks. Tools allowing circumvention of censorship for tasks such as gaining knowledge about safe-sex practices can be used to access bomb-making instructions. Further, free speech entails the possibility of hate speech. Anonymity tools can protect the identity of activists, but also of cyber-bullies. When working on these technologies, there is an ethical concern that in the course of empowering communities, they would cause collateral harm.

One possible justification goes as follows. While marginalized groups are silenced, those seeking to cause harm, like terrorists, have the funding and expertise to build comparable tools for themselves or enlist others to do it for them. Even if researchers did not build these particular tools, bad actors would still have the capabilities to do harm. If developers stopped building encrypted communication applications that keep individuals safe from oppression, terrorists could still build the same type of application for themselves.

Anyone suspicious of this justification might instead suggest that concerns of freedom, especially of vulnerable populations, typically trumps concerns of safety. Fear of wrongdoers intentionally corrupting tools for malice should not come in the way of protecting the oppressed or empowering the marginalized. Designing tools that are resilient to misuse is not always possible. Sometimes it is possible, however, to mitigate the potential harm.

Suppressing Speech of Others

Even when tools make it to intended audiences they can still be abused. When interviewing marginalized groups about the types of capabilities they would like to have on-line, some desired tools to silence or attack those that speak negatively against them. If the point of access and speech is an exchange of ideas, not all communities, even those silenced themselves, are initially interested in allowing others to talk. Developers can be mindful of this ethical concern, and focus on technologies that empower speech without suppressing the speech of others.

Interfering with Other Nations

Another ethical concern is the right to interfere in other societies and cultures. Often technologies are developed in first-world nations, but the technologies can be used anywhere. This may explicitly or implicitly bias development and usage towards groups similar to the developers. To empower speech, developers may target marginalized groups on foreign soil and not have personal stake in the ramification. Sovereign governments, sometimes put there by democratic vote, may actively impose the limits that technology aims to overcome. The counter argument is that free-speech and Internet access are human rights. Most democratic governments, as well as the UN [266, 267], recognize this. Just as we have duties to recognize and prevent other human rights violations, we have an ethical responsibility to support freedom of speech and access across national lines. The marginalized may not have the access or resources to help themselves.

9.4 Conclusions

Indicators point to a world that is becoming increasingly authoritarian. The precious resource of Internet freedoms is actively and intentionally limited by governments, corporations and communities. If, as a society, we place value in the rights of individuals to seek information and share their concerns and experiences, then overcoming those limits is a growing challenge. While technology can help tear down these barriers, it sometimes leads to externalities in the form of undesirable consequences.

When developing technologies supporting Internet freedoms, the design of applications has profound ethical implications. There is a balance between satisfying a human right and exposing others to danger. Empowering the speech of one group could mean suppressing speech of another. Tool developers can mitigate these risks while broadening access.

Developers often build from personal experiences, targeting users of their country and background, but the impact of their decisions often reaches far beyond the confines of their society. Successful tools are not confined to a single country or demographic. The Internet, as well as the ecosystem of tools that use it, is global and pervasive. Factoring in the experiences of users across the world, such as language, technical knowledge, and resource availability, can have profound impacts on peoples lives.

While a large library of security tools exists, there are under-served areas. Problems, such as maintaining reputation while preserving anonymity, the crowding out of voices using bots and trolls, and communicating despite network interruption continue to be areas of growth. Even existing technologies are often limited in their use due to the technical knowledge gap and language requirements associated with using them. As the Internet continues to grow and mature and new applications as well as censorship tools become available, so too will the need for new technologies to counter them. The next chapter explores ways of facilitating open discourse in the presence of adversaries.

Acknowledgment

This work was done in collaboration Lisa Parks and Elizabeth Belding. I would also like to thank all the other people who have helped with this work. Many thanks to members of our team including Danny Iland, Miriam Metzger, Hannah Goodwin, Kristin Hocevar, Lisa Han, Ariel Hasell, and Rahul Mukherjee for conducting interviews and surveys as well as providing analysis. Finally thanks to our many overseas partners for the hospitality, patience, and many hours of work. This work was funded by the DRL.

Chapter 10

A User-Driven Application for Anonymous and Verified Online, Public Group Discourse

Chapter 9 provided characterization for problems that groups and individuals may face when attempting to openly communicate online, including some of the limitations that existing tools exhibit. This chapter extends the characterization offered by ethnographic interviews with survey-based research to identify common unmet needs in Lusaka, Zambia; Ulaanbaatar, Mongolia; and Istanbul, Turkey. The chapter then proceeds to show how we fuse social and technological research to design a novel tool that satisfies these needs. While our work focuses on specific communities, not populations of countries as a whole, we believe examining the needs of communities particularly vulnerable to censorship provides a lens through which to understand some challenges to overcoming censorship more broadly. The ethics and permissibility of censorship is outside the scope of this work.

We begin this chapter by presenting the methodology behind both the social and technical aspects of our work in Section 10.1. In Section 10.2, we present the results of our survey analysis that serves as a basis for understanding people's uses of social media and their censorship concerns. We next explore some of these issues more closely, concentrating on findings from interviews and journalistic accounts in Section 10.3. Based on the social analysis we present our software solution, SecurePost, in Section 10.5. We then evaluate SecurePost using anonymous in-app surveys and usage statistics in Section 10.6. Finally, we summarize and draw implications from our work in Section 10.7.

10.1 Methodology

To investigate the role of the Internet and OSNs in free speech, we sought out communities vulnerable to censorship. Between 2013 and 2016, we visited three regions: Lusaka, Zambia; Istanbul, Turkey; and Ulaanbaatar, Mongolia. These three regions have diverse geopolitical contexts, different levels of socio-economic development and dissimilar cultural and historical backgrounds. In each region we focused on communities particularly vulnerable to censorship. We sought to understand the needs and challenges of these specific communities as a cross-section of global issues surrounding Internet censorship, particularly focusing on free speech on OSNs.

By combining social science and technological research, we developed a software tool that fulfills previously unmet needs to aid at-risk communities. In our work we involved participants in the iterative development of a technological solution that is well-suited to their use-case. Given the diversity of the three selected regions, we believe our solution would be applicable to other communities with similar needs, more broadly bolstering freedom of speech online.

Because there is extensive documentation of censorship in China [325, 326, 327, 328, 329] and the Middle East [330, 331, 332], we chose to focus on countries less studied at the time of data collection: Zambia, Turkey, and Mongolia. This enabled us to examine a broader scope of communities in terms of freedom of speech and censorship. Our goal is to use experiences with these three target communities as a lens to understand global issues surrounding censorship on OSNs and the Internet overall, and to facilitate the development of tools for protecting free speech.

When we first planned our work, Turkey was beginning to increase its censorship efforts, enabling our team to observe responses to increasingly visible and common censorship practices. Like Turkey, Zambia also experienced an increase in government censorship during the course of our study. By contrast, censorship in Mongolia was relatively static and less widespread [333]. Our team identified collaborators in these countries through previously existing contacts.

Our interdisciplinary team consisted of experts from computer science, communication, and media studies, from departments spanning across physical science, social science, and the humanities. Over the course of our project, we conducted 109 interviews and surveyed 526 individuals. We used a combination of ethnographic analysis of the interviews and descriptive statistics on the survey data to understand Internet access patterns, identify barriers to free speech, and assess shortcomings of existing tools that assist groups and individuals in communicating safely. We obtained IRB approval prior to conducting our fieldwork.

We used snowball sampling to recruit respondents for the surveys and in-depth interviews. We calculated descriptive statistics (frequencies) for all variables of interest, including socio-demographics. We compared participants in three countries using chi-square tests for categorical variables and Fisher's exact test for dichotomous variables.

Table 10.1: Demographics of survey respondents from the three sampled countries.

Survey Question	Zambia %	Zambia (#)	Turkey %	Turkey (#)	Mongolia %	Mongolia (#)	All %	All (#)
Gender								
Male	50%	(52)	52%	(81)	40%	(100)	46%	(233)
Female	50%	(51)	48%	(74)	60%	(151)	54%	(276)
Age								
Under 20	21%	(22)	7%	(10)	13%	(27)	13%	(59)
20-39	71%	(74)	78%	(120)	74%	(162)	75%	(356)
40 or more	8%	(8)	15%	(23)	13%	(27)	12%	(58)
Education								
Less than primary school	1%	(1)	0%	(0)	1%	(2)	1%	(3)
Primary School	1%	(1)	1%	(2)	2%	(5)	2%	(8)
Secondary School	37%	(39)	31%	(48)	31%	(77)	32%	(164)
Higher Education / University	61%	(64)	68%	(106)	66%	(163)	65%	(333)
Total # Surveyed		(106)		(166)		(254)		(526)
Dates Surveyed		Dec 7-18 2013		Dec 7-19 2014		Jun 16-Jul 4 2015		

10.1.1 Survey

The survey based research aimed to better understand use of information and communication technology in our three target communities and to gauge the opinions of members of our sample on issues of Internet freedom, censorship, and media trust. When interpreting the data, it is important to note that convenience sampling was used for the respondents, so these data may not accurately represent either the population of the respective countries as a whole or residents of the surveyed cities.

The demographics of survey respondents are provided in detail in Table 10.1. The gender of all survey respondents is roughly equally split between male and female, with slightly more female respondents overall. Across the three countries, we note the high prevalence of responses from individuals with university education and those between the ages of 20-39. This is likely due to our affiliations with universities and organizations, such as LGBTQ centers, with university aged staff. We note that because our focus was on studying communities vulnerable to censorship, the demographics and responses are not necessarily representative of the full populations of these countries.

Zambia

The Zambian survey was distributed in the capital city of Lusaka, using the Open Data Kit survey software on Nexus tablets between December 7 and 18, 2013. Open Data Kit allowed us to securely store encrypted versions of the survey response data without the need for Internet access. The total number of completed surveys was 106.

The convenience sample consisted primarily of individuals who work in media-related fields (e.g., radio stations, newspapers, news websites, blogs, and technology). The research team also recruited respondents from a media institute and a computer lab and technology hub where people can learn computer skills (e.g., game design and coding). The

surveys were also distributed at ConnectForum, a technology conference for women hosted jointly by the government and local companies. The women attending the conference were individuals who either worked at technology companies or aspired to work in the field of Internet communication technologies.

Participants in our sample were highly educated, with 98% of participants completing secondary school or higher, compared to the overall Zambian population where about 26% of females and 44% of males have only completed secondary school [334]. The sample was of a higher income level than average for Zambia and contained fewer unemployed. Students, journalists, and people working in the banking industry and in IT were also overrepresented. However, the sample did reflect the population in terms of age, religion, and ethnicity with the exception that it contained more Europeans. This is likely due to conducting the study in Lusaka, which has a higher number of expatriate workers [335] than in other areas of the country.

Turkey

The Turkish survey was distributed in Istanbul using both a paper survey and an Internet-based interface between December 7 and 19, 2014. Using this combination of in-person (paper and pencil) and mediated (online) administration helped researchers reach the largest possible sample with the resources available to the survey team for this part of the project. The total number of completed surveys was 166.

The sample consisted primarily of young people, including students and people working in media-related fields. The gender of survey respondents was relatively balanced. However, many survey respondents were of a lower income level and significantly higher education level than the general population in Turkey [336]. This is likely because of the number of undergraduate and graduate university students in the sample. Furthermore, the majority of respondents indicated they were not affiliated with the most powerful political party in

Istanbul at the time the survey was administrated (AKP), suggesting that the sample may be more representative of political minority opinions on the topic of Internet freedom than the average Turkish population. The research team relied on academic contacts at a University in Istanbul to recruit respondents.

Mongolia

The Mongolian survey was distributed in Ulaanbaatar, Mongolia's capital, using paper surveys in the Mongolian language between June 16 and July 4, 2015. Surveys were administered primarily by Mongolian undergraduate students who attend Mongolia National University (MNU), a project partner. These students were supervised by MNU professors. The total number of completed surveys was 254.

The sample consisted primarily of young people, including students and others who were relatively highly educated. In 2011 64.2% of the population of Ulaanbaatar was under 35 years old, suggesting that the number of younger respondents to the survey is somewhat representative of demographic trends in Ulaanbaatar [337]. The sample included slightly more female than male respondents. The survey was administered by Mongolian undergraduate students, who had greater access to individuals of similar characteristics, such as education level.

10.1.2 Interviews

Ethnographic interviews provided a qualitative dimension to our research on the underlying issues associated with Internet and OSN free speech and censorship experiences. We engaged in both in-depth interviews and informal conversations with a wide range of individuals and organizations, including journalists, political activists, ethnic minorities, law makers, educators, LGBTQ center employees, gender-based violence center employees,

government watchdogs, and others affected by censorship.

We reviewed contextual data such as political histories and news media reports about online censorship concerns in each country and identified key concerns and individuals. We then consulted with local partners, who were affiliated with universities and NGOs, to develop lists of potential informants. Working with our partners, we reached out to and scheduled interviews with many of these informants. We also used the snowball method to expand our informant lists while in each country.

We conducted more than 35 interviews in each country. Sometimes these interviews were with individuals and sometimes with small focus groups. In total we interviewed more than 150 people across the three countries. Theoretical saturation was reached when we encountered informants identifying the same or similar censorship agents, concerns, sites, and sources. We conducted some interviews in English and worked with translators to conduct others. Since many of our informants have been on the front lines of free speech struggles, we anonymized their identities and securely stored all interview data.

We used grounded theory [338] to analyze our interview data and extrapolated informants' censorship concerns and tactical responses to them based on our close analysis of the transcribed interview data.

10.1.3 Application Development

Based on qualitative and quantitative data, we identified previously unmet challenges to the use of online social networks for these target communities. We then built a software solution, called SecurePost [339], which addresses these challenges by enabling new ways for balancing anonymity and trust in OSNs.

We developed software in parallel to our field visits, and conducted user testing of prototypes while in the field. During each visit partners were able to try our most recent

prototypes and provide feedback. In this manner we received iterative feedback from our partners that helped us refine our application and further understand the requirements of these communities. We evaluated our software using a pair of optional anonymous surveys administered through the SecurePost application.

10.2 Survey and Interview Results

From the survey data we evaluated how vulnerable populations use and access the Internet, and in particular OSNs. We present the survey results and provide discussion supported by the ethnographic interviews. Note that an anonymity agreement in our IRB precludes us from making direct quotes of the interviews. For our analysis, we examined popular OSN platforms and the types of activities users engage in when using social media. We then looked at the difficulties that people experience when accessing the Internet and OSNs, including access disruptions and censorship. Finally, we compared how free users felt when using the Internet and the types of behaviors they engaged in when faced with censorship.

10.2.1 Internet and Online Social Network Usage

As the basis of planning a technical solution, we examined usage modality and preferred platforms. We asked respondents about their Internet and OSN usage, including frequency of usage and their preference of social network. We summarize the key findings in Table 10.2.

Across all three countries, the majority of respondents stated they use the Internet every day or more (with Mongolia to a lesser extent than the other countries). For all respondents, 94% stated they use the Internet at least once a week, indicating a high utilization of the Internet in these communities.

Table 10.2: Internet and Online Social Network Usage

Survey Question	Zambia %	Zambia (#)	Turkey %	Turkey (#)	Mongolia %	Mongolia (#)	All %	All (#)
Use Internet								
Less than once a week (or never)	2%	(2)	1%	(2)	11%	(27)	6%	(31)
Once a week	4%	(4)	1%	(1)	8%	(19)	5%	(24)
2-3 times a week	13%	(14)	5%	(9)	31%	(77)	19%	(100)
Every day or more	81%	(85)	93%	(154)	50%	(125)	70%	(364)
Devices Used to Access Internet								
Can select multiple								
Desktop	26%	(27)	48%	(79)	49%	(121)	44%	(227)
Laptop	80%	(85)	80%	(132)	53%	(131)	67%	(348)
E-reader/Tablet	17%	(18)	31%	(52)	14%	(35)	20%	(105)
Smart Phone	74%	(78)	84%	(140)	72%	(177)	76%	(395)
Basic/Feature Phone	13%	(14)	1%	(1)	14%	(34)	9%	(49)
Social Network Use								
Never	4%	(4)	0%	(0)	0%	(0)	1%	(4)
Less than once a week	1%	(1)	1%	(2)	5%	(11)	3%	(14)
Once a week	4%	(4)	1%	(1)	7%	(17)	4%	(22)
2-3 times a week	9%	(9)	8%	(13)	29%	(70)	18%	(92)
Every day or more	83%	(86)	90%	(150)	60%	(145)	74%	(381)
All Social Networks Used								
Can select multiple								
Facebook	91%	(96)	87%	(145)	97%	(237)	93%	(478)
Twitter	52%	(55)	76%	(126)	34%	(82)	51%	(263)
Google+	52%	(55)	57%	(95)	58%	(143)	57%	(293)
Instagram	3%	(3)	72%	(119)	35%	(86)	40%	(208)
LinkedIn	34%	(36)	31%	(51)	7%	(16)	20%	(103)

Further, the majority of users stated they use OSNs every day or more. Facebook, Twitter, and Google+ were used by the largest number of respondents (it is unclear how many people differentiated Google+ from Google search and other Google services; interviewees did not report significant usage of, or interest in, the Google+ OSN). Other OSNs included YouTube and WhatsApp, and to a lesser frequency Viber, Snapchat, Vimeo, WeChat, Tumbler, Pintrest, and Vine.

The number of users stating they use OSNs daily slightly out-paces the number of Internet users. When looking at individual responses, the same respondents claim higher OSN usage than total Internet usage. This is likely because some users perceive mobile applications and OSNs as something separate from web browsing as a whole.

The results from these communities fit into global trends for Internet and OSN usage. As of 2017, the International Telecommunication Union estimated 3.6 billion people, roughly 47% of the world's population, use the Internet [340]. Young people are at the forefront of adoption with 70% of people between 15 and 24 years old online [341]. Globally, online social networks are some of the most visited websites [342]. As of October 2017, Facebook had over 2 billion active users (1.3 billion daily), while Twitter had over 328 million active users [343]. These statistics highlight the importance of OSNs as communication platforms when confronting censorship.

We asked respondents what device they used to access the Internet. The most used devices were smart phones and then laptops. Globally, as of 2017, it is estimated that 58% of the world's population had a mobile-broadband subscription, while only 13% had fixed broadband subscriptions [341]. Further, the annual growth rate of global mobile broadband subscriptions (20%) is out-pacing that of fixed broadband subscriptions (9%) [341]. These statistics highlight the current and future needs for mobile-device-based solutions in both the cases of our target countries, and in the larger global context.

In order to identify a suitable operating system for development, we asked responders about their device model and operating system. In 2013, survey responses from Zambia indicated that Blackberry was the most popular type of phone (n=33, 34%), followed by Android (n=24, 25%). Nokia phones were next in popularity (n=20, 21%), followed by the iPhone (n=10, 10%). Since then, Android has become dominant in Zambia, accounting for close to 50% of the market share [344]. This is in line with anecdotal evidence from respondents during our return visit in 2016. Responses from the 2015 Mongolia survey indicated that Android was dominant (n=140, 59.8%), followed by iOS (n=76, 28.6%). Relatively few people had other operating systems, including Blackberry (n=10, 4.3%).

We did not collect data on device model and operating systems for participants in Turkey. According to Stat Counter, a website that tracks browser and mobile OS usage worldwide, as of September 2017 Android accounted for 80% of the mobile operating market share in Turkey, followed by iOS with 18%, while all other operating systems accounted for less then 2% of the market share [345].

These trends are reflected globally. As of end of May 2017, Android made up 86% of the global market share [346]. Android's popularity, especially in the developing world, may be due to the availability of cheap phones, from a variety of manufacturers running the Android OS, saturating markets in Africa and the Middle East [347]. Many of these phones run older versions of the Android OS, have limited storage and computational power, and do not receive software upgrades. Nonetheless, when developing tools that work in the developing world, Android applications allow access to a large population.

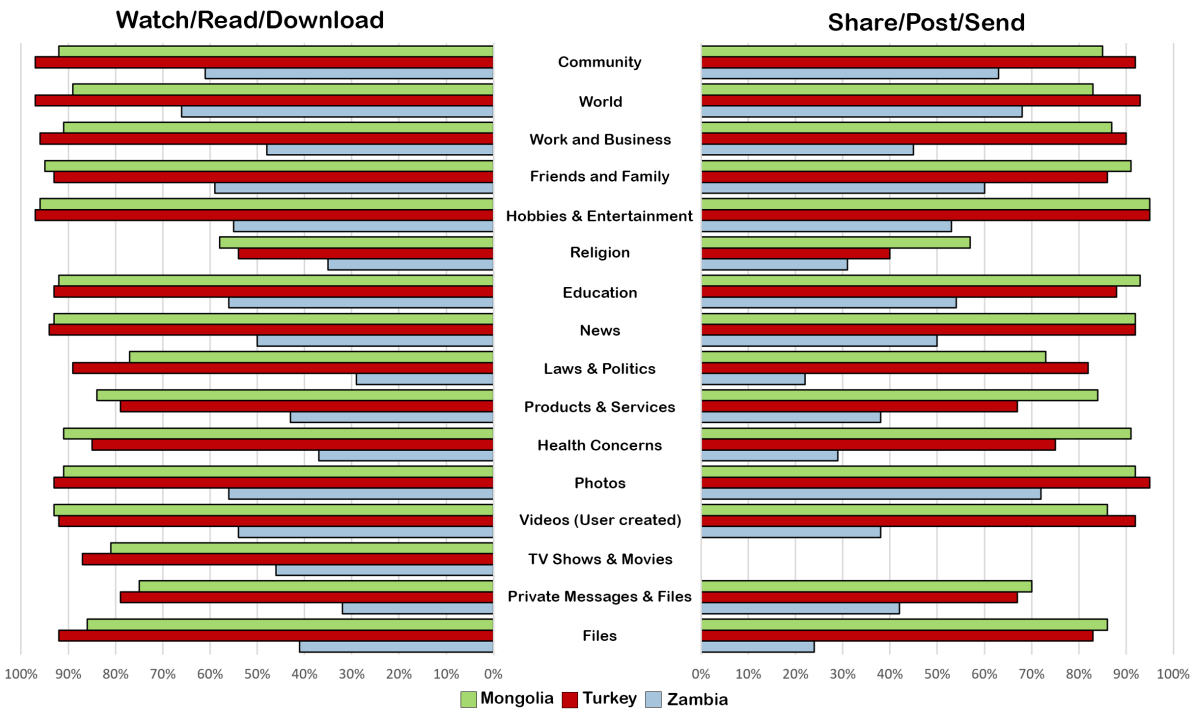


Figure 10.1: Type of users activities on OSNs. Respondents could pick any number of activities. The left part of the chart represents activities that consumed content and on the right activities that generated content.

10.2.2 Activities on Online Social Networks

Respondents engage in a wide variety of activities in online social networks. In particular, respondents most commonly consumed content about their hobbies and entertainment, things happening in their community and the world, as well as information about their friends and family. Respondents most commonly produce content such as photos and information about hobbies and entertainment. A full breakdown is shown in Figure 10.1.

In Turkey and Mongolia, we also asked about the frequency of information consumption and generation. Respondents in Turkey indicated that they most frequently share, post, or send information about things happening in the community and around the world, as well as hobbies, news, photos. Content was less frequently shared or consumed on topics

such as religion, products, and health.

In Mongolia, respondents indicated that they most frequently share and post information about hobbies or entertainment, as well as friends and family, education, photos, news, health, and work. Content was less frequently shared or consumed on topics such as religion, laws and politics, and things happening around the world.

Some activities in which users engage, such as posting about religion, health, and politics, have some relevant topics of discourse that could endanger the user. We explore this in Section 10.3.

10.2.3 Internet Disruption

In the survey, we examined the frequency and perceived reason for disruptions that respondents experienced in attempting to use the Internet and OSNs. A summary of our results is presented in Table 10.3. Across the three countries, 78% of respondents reported at least sometimes experiencing disruption to their activity, most commonly in Zambia (96%).

In Zambia and Mongolia, power outage was reported as a major reason for disruptions. This was less of a problem in Turkey, which has better infrastructure. In all three countries, particularly in Zambia, poor network reliability was a major reason for access disruptions. Government or other censorship which was not frequently cited as a common cause of disruptions in Zambia (7%), was reported more common in Mongolia (19%), and was cited as a major cause of disruptions in Turkey (44%).

Such temporary disruptions emphasize the need for tools that are tolerant to networking delays when fetching and delivering content. Permanent disruptions, like those imposed by government censorship, require circumvention tools to bypass those blocks.

Table 10.3: Disruption of Internet and OSN Usage

Survey Question	Zambia %	Zambia (#)	Turkey %	Turkey (#)	Mongolia %	Mongolia (#)	All %	All (#)
Frequency of Internet Disruption								
Almost Never	4%	(4)	25%	(41)	27%	(68)	22%	(113)
Sometimes	42%	(44)	48%	(79)	61%	(152)	53%	(275)
Often	25%	(26)	19%	(31)	10%	(25)	16%	(82)
Very Often	29%	(30)	8%	(13)	2%	(5)	9%	(48)
Reason for Internet Disruption								
Can select multiple								
Power outage or electrical problems	48%	(51)	15%	(24)	51%	(128)	39%	(203)
Unpaid fees for Internet service	29%	(31)	7%	(11)	32%	(80)	23%	(122)
Poor network connection or service	91%	(96)	68%	(112)	44%	(110)	61%	(318)
Government or other censorship	7%	(7)	44%	(73)	19%	(47)	24%	(127)
Unknown Reason	3%	(3)	15%	(25)	10%	(25)	10%	(53)
Other Reason	5%	(5)	5%	(8)	8%	(20)	6%	(33)

10.2.4 Censorship

In addition to investigating how and for what purposes respondents use OSNs, we investigated how their use is impacted by censorship. We summarize these results in Table 10.4.

Across all three countries, only 11% described feeling "very free" to express themselves on the Internet and OSNs. The majority (53%) felt only "a little free" or "not free at all". In Turkey especially, 37% of responders said they did not at all feel free to express themselves on OSNs.

When asked how often users modify their behavior to protect against government or other monitors seeing things that they post in OSNs, respondents in Mongolia and Zambia demonstrated no statistically significant differences from each other. In those two countries, the majority of users (64% in Zambia, 56% in Mongolia) reported at least sometimes modifying behavior in OSNs (p=0.225). On the other hand, in Turkey, 88% of users reported never or almost never modifying behavior, a statistically significant difference when examining all three countries (p=0.000). This is interesting because out of the three countries, Turkey has a higher incidence of government censorship [333], and - as observed earlier - more users reported feeling not free.

When asked about methods users took in modifying behavior, Mongolia, Zambia, and Turkey had no statistically significant differences for most strategies. About a quarter used secure connections (p=0.554), a few hid their names (8%, p=0.166), and some avoided using the Internet or OSNs entirely (8%, p=0.119). For self-censorship, however, respondents in Turkey again differed in behavior. While the majority of users in Zambia (77%) and Mongolia (67%) self-censored what they posted online (p=0.893), users in Turkey (59%) reported significantly less self-censorship (p=0.0076).

Table 10.4: User Behavior and Perceived Freedom on OSNs

Survey Question	Zambia %	Zambia (#)	Turkey %	Turkey (#)	Mongolia %	Mongolia (#)	All %	All (#)
How 'Free' Users Feel on Internet & OSNs								
Not free at all	32%	(34)	37%	(60)	19%	(47)	27%	(141)
A little free	19%	(20)	21%	(34)	31%	(78)	26%	(132)
Somewhat Free	40%	(42)	30%	(48)	38%	(96)	36%	(186)
Very Free	9%	(9)	12%	(20)	12%	(29)	11%	(58)
How Often Users Modified Behavior								
Never	19%	(18)	71%	(114)	12%	(29)	32%	(161)
Almost never	18%	(17)	17%	(28)	32%	(79)	25%	(124)
Sometimes	47%	(46)	9%	(15)	34%	(83)	29%	(144)
Often	11%	(11)	1%	(2)	11%	(26)	8%	(39)
Very often	5%	(5)	1%	(2)	12%	(30)	7%	(37)
How Users Modified Behavior								
Can select multiple								
Limit/censor what to post	77%	(82)	59%	(98)	67%	(164)	66%	(344)
Use a secure connection	25%	(26)	20%	(33)	24%	(59)	23%	(118)
Do not use a real name	4%	(4)	8%	(14)	10%	(24)	8%	(42)
Avoid Internet or Social Network	3%	(3)	8%	(14)	9%	(22)	8%	(39)

A possible explanation for these behaviors comes from our interviews. In Turkey, many of the journalists and activists we interviewed conveyed a sense of obstinance to censorship efforts. Those we interviewed emphasized the importance of exercising free speech and were willing to continue to do so, even after arrest. And, as noted earlier, a proportionately larger number of our survey respondents in Turkey were members of political opposition groups.

10.3 Categorizing Barriers to Free Speech

As we observed in the previous section, while most respondents do not feel free to express themselves on OSNs, very few (8% across all three countries) reported posting anonymously. Despite this, in the ethnographic interviews, respondents repeatedly brought up issues concerning anonymity, as well as the ways anonymity impacts trust. To understand how a tool could address these concerns, we focused on four central elements: (1) The negative consequences of the use of real-world identity. (2) Difficulties individuals experienced in retaining anonymity. (3) The impact of anonymity on reputation. (4) Problems with retaining reputation and trust in spite of active censorship. (5) Additional methods adversaries use to limit free speech. We use this discussion as the basis of our software based solution.

10.3.1 Using Real-World Identity on Social Media

Authoritarian governments pass strict laws curtailing free speech. In Turkey, for example, following a 2005 restructure of the Turkish penal code, Turkey passed Article 301, which prohibits the "denigration of the Turkish nation," and Article 216, which bans "inflaming hatred and hostility among peoples" [348]. These laws have been used to target journalists, artists, and unaffiliated individuals for criticizing government, policy, and religion in any medium [288, 349]. Zambia likewise recently saw multiple arrests including an opposition leader [350] as well as an engineering student critical of the president [351], on the grounds of defaming Facebook posts. Globally, 27% of all Internet users live in countries where individuals have been arrested for posting, sharing, or liking a post on social media [352].

Libel laws are also a weapon that companies, organizations, and wealthy individuals can use to curtail freedom. These adversaries sue for defamation or insult - wrongfully

in many cases - in response to stories on OSNs and other media [353]. Expensive legal fees and threat of financial ruin silences those who do not have the monetary resources to defend themselves. This creates a chilling effect on free speech [354].

This is especially a problem in Mongolia where libel laws are frequently used as a way to silence the press [355]. In Mongolia, the burden of proof for libel rests with the defendant and libel constitutes a criminal charge. Journalists reporting on topics such as corporate corruption must prove their reporting is true and accurate and evidence not containing original copies and notarized documents may be thrown out as inadmissible [356]. During the Mongolian interviews, self-censorship due to threat of a libel lawsuit was a frequent topic of concern. In 2015, after our visit, multiple individuals were arrested in separate cases over posts on Twitter [357].

Revealing identity and personal details can make an individual a target by members of their physical or online community. Expressing views or interests that go against societal norms can impact job availability and interpersonal relations [304].

From our interviews, we heard how journalists reporting on topics opposing dominant political identities, such as on Kurdish issues in Turkey, face constant barrage of hateful posts. In Zambia, we heard how elements of a user's identity, such as ethnicity, gender, and past posting record, stereotypes the user to the point of overshadowing the discussion of substantive content.

Threats may extend into an individual's home. When interviewing members of a gender-based violence prevention center in Mongolia, we heard stories of threats coming from a person's own family. Family members of some of these women would monitor posts and private messages on social media and punish perceived affronts with physical violence.

10.3.2 Balancing Reputation and Anonymity on OSNs

With so many threats and such serious consequences, many users self-censor their online posts. Yet, as we observed in Section 10.2, some users still remain adamant about speaking their minds, putting them in potential jeopardy. While some post anonymously or use fake aliases, relatively few reported doing so in our surveys. In part, this may be due to the difficulty of being anonymous online. Major social networks, like Facebook, impose a real name policy as part of their terms of service [308, 309]. Other sites, like Twitter, may require identifying information, such as a phone number to create or verify an account.

Even if information is not provided by the user, OSNs still log the IP address of the requests. This information can be subpoenaed by governments [310, 358]. Requiring all users to use anonymity services like VPNs or Tor [64] for each post is unrealistic, since groups can be composed of posters with varying technological expertise, and a single mistake can be costly.

If adversaries lack the ability to identify users based on IP, they can still de-anonymize users based on the content they post. Users of social media regularly post identifying information. An account posting a personal photo can identify an individual. Other information posted by a user can inadvertently allow adversaries to guess identities. For example, revealing details about education, past residences, and events might be enough to uniquely identify an individual. A dedicated attacker could use information exposed by the account over a period of time to establish identity, exposing the user to threat. Adversaries will *dox* users, publishing identifying information, which may escalate digital conflict to physical confrontation [359, 360]. Even if meticulous discipline is followed, self-censoring all identifying information limits an individual's ability to communicate about their personal experience.

10.3.3 Impact of Anonymity on Reputation

Hiding identity likewise has an impact on trust of the posted content. Readers of social media use the reputation of an account to gauge trustworthiness of the content. The reputation of and past experience with an organization, such as a newspaper, leads readers to believe in the content of that account and differentiate it from fake news promulgated by bad actors. The reputation of a source and the identity and history of the author are the first things fact checkers and librarians recommend that readers check when evaluating if a story is fake [361, 362].

In the case of the Zambia watchdog (ZWD) online news service, we interviewed journalists who said they maintained anonymous blogs to avoid government intervention when discussing sensitive topics. Others we talked to in Zambia noted that the lack of real-world identities by ZWD journalists and lack of a physical address for the organization weakened accountability and verification of content. Respondents said this led ZWD to lose credibility among some of its readers [301].

Increasingly, hostile governments and other adversaries use fake accounts to orchestrate attacks. Governments, like Russia, hire paid *trolls* to carry out legitimate sounding discourse online [281, 70, 314, 69]. They automate attacks using *bots* to generate a massive number of messages that flood OSNs [312, 313], and set up anonymous *sybils*, which are fake accounts posing as legitimate users [363], as vessels for these agents. These accounts post spam [364] and fake-news [365] in coordinated attacks to steer conversation and drown out competing ideas. Users attempting genuine discourse, especially those using anonymous accounts that are hard to distinguish from the attackers, can get lost in the noise [280]. Our interview pool included victims of these false reporting attacks, who reported their accounts being banned due to third party reports.

Journalists and public figures whose careers are tied to reputation may find it difficult

to utilize the protections offered by anonymity in light of the need to build and maintain a reputation that garners trust with readers. For example, a lawyer and LGBTQ activist who was politically active against Turkey's president Recep Tayyip Erdoğan was arrested, convicted, and fined over a tweet [285] criticizing the president. Subsequently, he was arrested again for unrelated charges, including membership to HDPIstanbul, a WhatsApp group belonging to the Peoples' Democratic Party [366, 367]. Despite harassment by the Turkish government, he continues to actively use social media and maintains his real name and identity online.

From our interviews, after facing legal jeopardy, some reported adopting anonymity in order to maintain jobs unrelated to their online activities. However, several journalists and activists stated that, despite past and future threats, it is important to publicly stand up using their real names and identities as an act of opposition, personal pride, as well as to garner trust.

10.3.4 Maintaining Reputation After Censorship

An additional dimension to online identity and the decision to practice anonymity is the ability for governments and individuals to censor content on OSNs. Governments may block entire websites [268] or ask OSNs to censor a particular account or post. For example, Twitter maintains an entire system for withholding posts that are censored in a particular country [368].

Aside from legal requests, governments and groups exploit weaknesses in mechanisms built to protect users. Most OSNs have a reporting functionality that users can employ to report cyber-bullying and flag content that may not be appropriate, such as pornography, for a particular communication channel, such as a university's homepage on Facebook. Free speech adversaries exploit these reporting tools to falsely flag dissenting opinions.

Governments and other adversaries deploy bots and trolls to report posts. This usually leads to an automatic account ban, and may take a long time for OSNs to rectify. Russia, for example, deploys trolls to censor popular accounts by reporting content as threatening violence or containing pornographic material [315].

In our interviews, people reported similar instances, where their account would be banned after crowds of users repeatedly mis-flagged content. This led to repeatedly creating new accounts. When legitimate OSN accounts (anonymous or not) are compromised and users switch to new accounts, the trust of the readers may be imperiled. New accounts have to demonstrate continuity with previous ownership, and re-establish readership. This can be even more challenging for anonymous accounts that can't rely on real-world identify to demonstrate continuity. Adversaries can take this opportunity to impersonate accounts to confuse readers and further degrade trust.

10.3.5 Geographical Censorship

In some instances, instead of targeting web sites or individuals, governments or organizations censor entire geographical areas by restricting all Internet access. In some areas access is restricted permanently, such as in the Za'atari refugee camp, where access was not provided in order to discourage refugees from encroaching on the local labor market [295]. Other times access can be cut off in response to events such as protests, which is frequently the case in Turkey - especially in Kurdish regions [269, 296, 297, 369].

From our interviews, citizen journalists in Kurdish areas of Turkey reported encountering these types of tactics [370]. When reporting in areas where Internet is severed, they record content and store it on their device for later publication. However they expressed concerns that sometimes their device will be seized and searched. Similar reports emerged during our interviews of journalists in Zambia as well [301]. Much like the case of access

disruption due to electricity, this scenario is a notable consideration for the design of technology to protect freedom of speech.

10.4 Outlining Requirements for a New Tool

The quantitative survey results complemented with the qualitative interviews, global statistics and journalistic reports, provide an outline to understand and address some of the threats and needs of users when protecting freedom of speech online.

OSNs Highly Utilized: From Section 10.2 we saw that the Internet, and in particular OSNs, are heavily utilized by our partner communities. OSNs are used daily for posting and consuming content on a variety of activities. However, the majority of users report feeling only "a little free" or "not free at all" when using them and, consequently, users modify their behavior. This results in some self-censorship of content, limiting what they discuss on these platforms to reduce the threat of legal and physical harm from governments, corporations, and sometimes even family members.

Smart Phones Are a Primary Mode of Access: In these communities, smart phones are the dominant way for individuals to access the Internet and OSNs, with this trend continuing to increase globally. When designing tools for this type of community, support for smart phones running Android (the most used OS type) gives widest user coverage. For our own technological solution, in the context of limited developer resources, we chose Android phones as the primary platform for developing and deploying our tool. Given the popularity of Facebook and Twitter (which shares similar text based posting behaviors with Facebook), and based on feedback from our local partners, we elected to support these two platforms for our work.

Backwards Compatibility: Given the wide abundance of older devices - especially in Zambia and Mongolia, we sought to develop a tool that was backwards compatible with older operating systems. This puts restrictions on the types of both native and external APIs that developers can utilize during development. Often developers target newer versions due to limitations of older software and hardware. For example, while some of our respondents used devices running Android 2.x, Signal [65] limits its support to Android 4.0 and above.

Network Disruption: Respondents reported disruptions to power, communication, and instances of government restrictions on Internet access to geographic areas. Tools catering to people who may live in areas where network and power are unreliable have to account these limitations. Applications should limit power and data consumption and provide a user experience that does not rely on continuous connectivity. In our tool we implemented local caching of content that was tolerant of network delay.

Confiscation and Search: Due to temporary or permanent lack of Internet connectivity, as we discuss in Section 10.3.5, journalists sometimes ferry information back to Internet connected sites. During this time devices could be seized, which suggests a need for local data encryption.

Anonymity Protection: As discussed in Section 10.3, we observed that when groups use social media to spread news and ideas, there is a tension between anonymity and trust. While other tools, such as WhatsApp[66] and Signal [65], address aspects of individual security and *private* group communications, there are still unmet needs in *public* group communication in the presence of censorship. Users that reveal real-world identities open themselves to both legal and physical threats. However, users face many technical challenges when staying anonymous on OSNs.

Reputation Preservation: Users also find it difficult to retain reputations and trust, especially when posting anonymously. Individual anonymous users may find it difficult to build trust, competing with armies of *bots*, *trolls* and *sybils*. Groups find it difficult to maintain reputation after hacking, infiltration or censorship. There is a need for new tools to address ways of preserving reputation while maintaining anonymity for group discussion on OSNs.

Appropriateness to Intended Users: Lastly, a security tool has to ultimately be usable to the intended population. This extends to both the behaviors of individuals and the local culture where it will be used. A technical solution that works well in one context (for example tested and developed in a western University) might not be suitable for applications in other contexts. In the next section we discuss how our team addressed the aforementioned requirements, including instances where iterative feedback from partner communities steered the technological design of the application to suit local needs.

10.5 SecurePost: Verified Group-Anonymity

To address the constraints outlined in the previous section we developed an application called SecurePost, that allows individuals to share a single group identity while retaining individual anonymity on OSNs. SecurePost is comprised of three coordinated modules: an Android application that allows group members to post content and manage membership; a proxy server that relays posts to social networks; and a browser extension that allows members of the public to verify those posts. Together the modules provide group-anonymity coupled with an ability to verifying the integrity and authenticity of posts.

Using SecurePost, group members can make posts to shared OSN accounts, while masking their individual identity. The connection is routed through the companion proxy server, hiding the IP address from the OSN platform, which may cooperate with a hostile government. The identity of a poster is likewise hidden from other group members, giving plausible deniability for any given post. For members of the public looking at the posts on social media, it appears as if posts from an account have a single author with no way of identifying individual posters or a group's membership roster. In this manner, a group is able to build a social media presence while retaining anonymity of its members.

Because OSN accounts can be seized or infiltrated, SecurePost provides tools to verify that content is genuine and unmodified. SecurePost allows groups to attach cryptographic signatures to every post in order to verify the authenticity and integrity of messages. A companion browser extension allows anyone, even readers not part of the group, to verify posts directly on the social network web page. Using the extension, readers can know if a signed post came from a trusted member of the group and if its text has been modified, even if the OSN account has been compromised.

In the event that a group has been compromised, SecurePost allows dedicated administrators to retain membership control. If administrators see erroneous posts coming from the group, they are able to reset membership, expelling all other members, and invalidating past posts. This allows group leaders to protect the group in the event of infiltration as well as warn readers, protecting the group's reputation.

Together group-anonymity with a layer of verification provides a mechanism for groups to balance personal anonymity with building a trusted group identity.

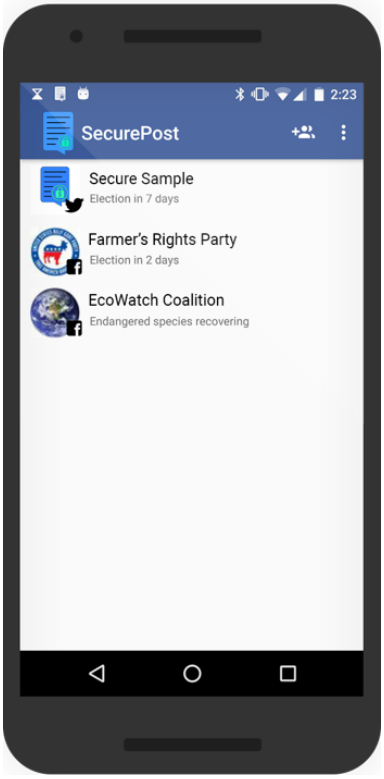


Figure 10.2: Group Overview.

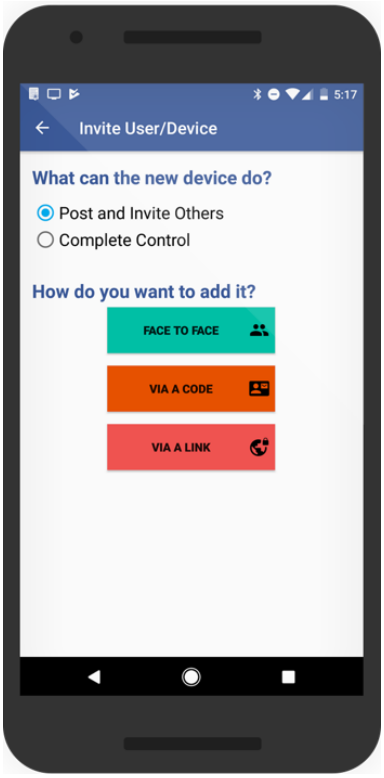


Figure 10.3: Invite Wizard.

10.5.1 The Mobile Application

At the center of SecurePost is the mobile application. Through the application, users are able to form groups, manage membership, as well as make and view posts to social media. As mentioned in the previous sections, due to resource constraints, we developed our application exclusively for Android, as it is the most common mobile operating system. Mindful of the prevalence of cheaper devices in the developing world running older operating system versions, we support Android API level 10. This makes our application backwards compatible with devices running Android 2.3.3 and above. As of October 2017, this accounts for 99.9% of Android devices registered with Google [371].

Forming a SecurePost Group

Each SecurePost group is tied to a corresponding OSN account. Currently SecurePost supports Facebook and Twitter, but its modular design can be extended to include any similar platform. SecurePost allows users to take part in multiple groups, without the need to switch accounts. Figure 10.2 shows the group overview screen of the application, where the user is a member of three groups.

To set up a SecurePost group, users require a corresponding OSN account on the platform of their choice. They can either use an established account (such as a Twitter handle for a newspaper) or set up an anonymous account. The application presents the platform specific API based login web page to the user. After a user logs on, the OSN platforms returns an access token. This is a common design pattern used by other social media applications. The access token is forwarded to the proxy server (discussed in Section 10.5.2) and discarded by the application.

The group creator is granted administrative rights to the group, and can post, invite others, and manage the group. As this initial step requires direct contact with the social media platform, to avoid associating an IP address to the user, we recommend that this step is done through an anonymity service such as Tor [64]. Other than this initial creation step, all application traffic is routed via the proxy server. Notably the user name and password are only needed during this initial group creation process. Inviting new members does not require the sharing of login credentials (a method of sharing social media accounts frequently cited by interviewees).

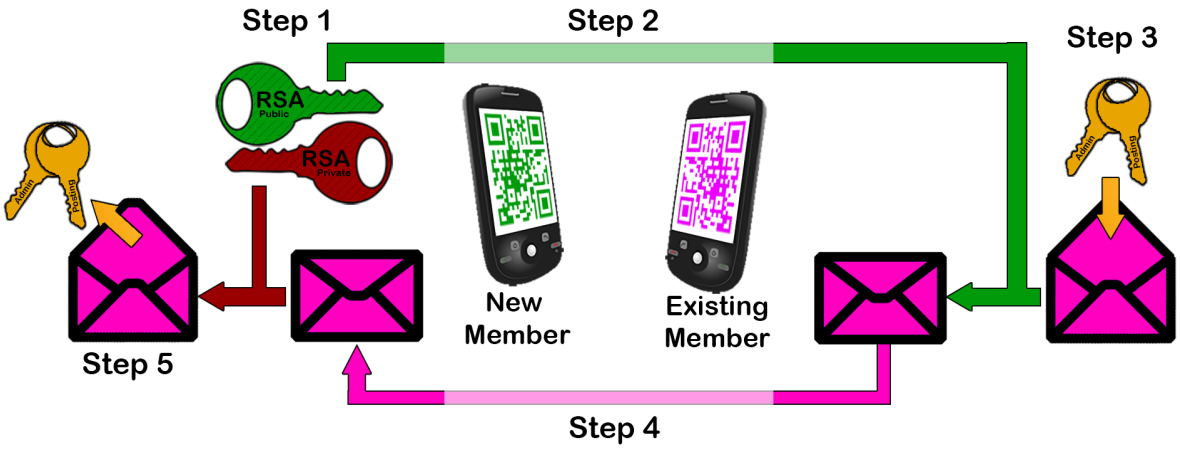


Figure 10.4: Process Flow of Visual Invitation Scheme. Step 1: generate ephemeral asymmetric key pair. Step 2: ephemeral public key sent as QR code. Step 3: ephemeral public key encrypts group credentials. Step 4: send encrypted invite as QR code. Step 5: ephemeral private key decrypts group credentials.

Group Membership

Instead of relying on password sharing, SecurePost uses public key cryptography to authorize users. When creating an account, the app generates two asymmetric 2048 bit RSA key pairs. The app stores the private keys, and transmits the public keys to the proxy server. One key pair grants *posting credentials* while the other pair grants *administrative credentials*.

These two keys signify two classes of users: *administrators* (holding both posting and administrative keys) and *posters* (holding only the posting key). Anyone with a private key to the group can transfer a copy of that key via the application to grant access at the same privilege level or lower. In other words, an administrator can recruit administrators and posters while posters can only recruit other posters.

The complexity of the key exchange is hidden from the user via a graphical user interface (shown in figure 10.3). The application provides a number of methods of recruitment: Face-to-Face credential exchange, a sharable token, and a link.

Face-to-Face Visual Recruitment: When in close geographic proximity, SecurePost offers a secure key exchange via visual scanning of QR codes, as shown in Figure 10.4. The exchange happens in two steps between a **recruit** (someone who wants to join the group) and an **existing group member**:

1. The **Recruit** initiates join process by visually showing an **existing group member** a join request containing a QR code.
   (a) The **recruit** seeking to join the group generates a 2048 bit RSA ephemeral key pair *(EP_PUB & EP_PRV)*.
   (b) The **recruit** shows *EP_PUB* in the form of a QR code *(QR_PUB)* to the **existing group member**.
   (c) The **existing member** scans *QR_PUB* and extracts *EP_PUB*.
2. The **existing group member** responds by sharing encrypted group credentials with the **recruit** in the form of a QR code.
   (a) The existing group member takes the group credentials of a desired privilege level *(CRED)* - i.e. the group's private keys - and encrypts them using *EP_PUB* as *EN_CRED*.
   (b) The **existing member** then displays *EN_CRED* back in the form of a new QR code *(QR_INVITE)* to the **recruit**.
   (c) The **recruit** scans *QR_INVITE* and extracts *EN_CRED*.
   (d) The **recruit** decrypts *EN_CRED* using *EP_PRV*.
   (e) The **recruit** now has *CRED* and is part of the group.
   (f) Both the **recruit** and **existing member** discard the ephemeral keys *EP_PUB* and *EP_PRV* as they are no longer needed.

Once the recruit possesses the private keys, they are a part of the group: they can authenticate with the proxy and are ready to post. By using a two step process, an adversary visually observing the exchange would be unable to decrypt the group credentials without the recruit's ephemeral private key.

Alternative Recruitment: When physical proximity is not possible or unsafe, SecurePost allows alternate recruitment strategies via the use of an invite code or link. These strategies allow users to transmit keys via secure communication channels established outside the application.

For these strategies the existing group members use the graphical interface to initiate recruitment. The application encodes the the corresponding private keys into an invite code or link. The group member can then manually copy this code or use the Android share intent to paste it into a secure application of their choice, for example an end-end encrypted messaging client. The recruit pastes this code into their SecurePost application (or clicks the link) which decodes the private keys, granting group membership. However, if this link or code is intercepted an adversary may be let into the group.

This option gives group members more flexibility but also more responsibility. We added this option in response to user testing, as users wanted a way to asynchronously invite users without physical access. Note that in both recruitment schemes the group members retain custody of the group's private keys. The proxy server only has access to the public keys.

Group Administration

Authentication of group membership is verified by the proxy server using the *administrative* and *posting* public keys. These key are unique to the group and not the user. This approach intentionally omits a user registry. From the perspective of the proxy server and OSN platforms, each group appears as a single entity. The number and identities of group members is only known out of band through social interaction and is not retrievable from any part of the system.

As any group member can invite others, by compartmentalizing recruitment history from other group members, it is possible to hide the full membership roster from any single group member. In this manner, groups can enlist confidential contributers.

This structure imposes limitations on group administration. As there is no user registry or unique identifier, SecurePost lacks the ability to rescind membership to an individual user. Instead, if the group is compromised, an administrator must reset membership entirely. In this process the administrator performing the reset generates new key-pairs and transmits the public keys to the server (much like the initial process of group creation). All prior members are expelled (including other administrators) and have to be re-invited - this time hopefully with a higher level of scrutiny. In addition past posts are invalidated; they continue to remain on social media but are no longer marked as verified by the browser extension.

Social Media Posts

Once users join a group they are able to post directly to the associated OSN account. To members of the public (and other group members) it appears as if all a group's posts come from a single entity.

In addition to providing anonymity, SecurePost also offers verification for posts. Before

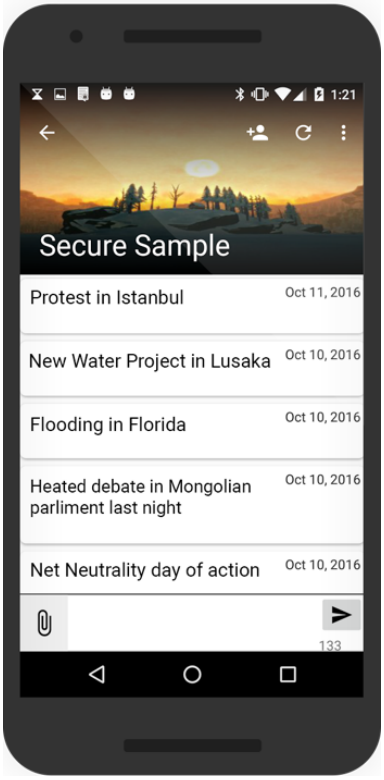


Figure 10.5: View of an Individual Group.

Figure 10.6: Setting a Self-destruct Password.

transmission, each post is signed using the *posting key* of the group using SHA-256. This signature can then be used to verify the integrity and authenticity of the post. The application automatically verifies and displays posts from other group members as shown in Figure 10.5. The general public can verify the posts on OSNs via the use of the browser extension described in Section 10.5.3. To handle situations where connectivity is disrupted - such as power failure or regional government censorship - posts are stored locally and forwarded to the proxy server when Internet connectivity is reestablished.

Several other algorithms have been proposed to provide group anonymity. Ring signature cryptography schemes, as proposed by Rivest, Shamir, and Tauman [372], provide a similar group anonymity guarantee. Signed messages can be authenticated as being authored by the owner of one the public keys, but it is not possible to determine

which one.

However, the ring signature verification process requires the public keys of all group members. This is problematic for both design and implementation reasons in the context of SecurePost. Ring signatures used publicly in this application would leak metadata about the group membership. Specifically, an adversary could always determine the number of members, and could monitor the public key set for changes to determine when members join the group. Additionally, the amount of data that would need to be encoded into the account's profile image for the plugin verification to work would scale linearly with group size, potentially placing a cap on group size.

Group signatures, as proposed by Chaum and van Heyst [373], are similar to ring signatures, except they would allow the group owner to de-anonymize posts by members using the group owner's secret key. This would potentially resolve the problem of 'bad actors' joining the group, by allowing the group owner to identify and expel users who post content that the group owners do not agree with. This would allow the group administrator to identify and expel posters of problematic content from the group without completely purging group membership. However, this kind of cryptosystem reduces the strength of deniability, since there exists a person who can prove that a particular group member wrote the post. This puts the group owner at risk from external actors, who may be motivated to threaten or harm the group owner if they do not de-anonymize a particular post author.

Multimedia Posts: Initially, we only planned to support posting and verification of text. However, in our interviews, respondents stressed the importance of posting images as well as audio and video. Respondents found multimedia content, such as images of police impropriety, is an effective tool for reader motivation. This was particularly important to respondents in areas where Internet connectivity is disrupted and risk of

device confiscation is high. In this case, users wanted to store images in an encrypted manner and queue them for posting when connectivity is available.

As a response we added the ability for users to post images. Unfortunately, as of now, SecurePost does not allow the verification of validity and authenticity for the posted images as OSNs compress and alter images prior to display. Thus the bitwise verification system we use for text post does not translate to multimedia posts. In the current implementation images are marked as unverifiable. Verification of images as well as support for other multimedia content are future work.

Secure Storage

The application works even without continuous Internet access. Previous posts are cached, and new posts are stored locally and transmitted when Internet connectivity is reestablished. As devices can be searched, lost, stolen, or confiscated, all user data are encrypted. In scenarios where a region temporarily or permanently lacks Internet access, users can utilize SecurePost to prepare and ferry posts for delivery when they re-establish connectivity.

Previous posts, group memberships, keys, and other identifying information are stored using SQLCipher [374] (an encrypted database for Android). When the application is first launched, users choose an application-wide password for unlocking the app. This is needed each time the app is started in order to decrypt the database. While running, the application displays a persistent notification reminding the user that the database is unlocked. Dismissing this notification rapidly locks the database and closes the application.

The application password is unique to the device and not shared with the proxy server. In the event of password loss, the data are not recoverable. Users who forget the password are offered the option to wipe the data, allowing them to start over. This does unfortunately erase a user's group memberships as credentials are all locally stored.

Our application is tailored to populations that often use cheaper devices running older operating systems. Older devices lack support for modern features such as full disk encryption or a secure enclave [375]. Our application-level encryption is done independently of any operating system-level encryption that the device may support. It can be used in conjunction with other security methods, and adds a layer of security for older devices. If the phone is confiscated by an adversary, the adversary would need to perform a costly attack to decrypt data, which would still not expose group membership.

Option to Self-Destruct: To handle the case of device confiscation, raised as a concern in the interviews, SecurePost implements an optional "self-destruct" password. This is set up in addition to the regular application password and is shown in Figure 10.6. The self destruct password can be revealed if the user is under duress. Entering the password is visually presented as an incorrect password attempt while in reality it irretrievably wipes the content of the application.

Matching Cryptographic Design With Needs

Our current invitation process is a reflection of feedback from our partners. In the initial designs of SecurePost, we envisioned scenarios where users could form short-term groups at physical gatherings such as protests.

In this initial design users shared a secret pass phrase to join a group. Anyone with the pass phrase could post. Administration was only possible by the owner of the OSN account using the OSN account username and password. When a user joined a group, the app generated a time-synchronized hash chain using the pass phrase as a seed. To post, the group member signed the post using the hash from the current time slot which was verified by the proxy server. Using the browser extension, anyone from the public could verify any past post by using the signature of the most recent one. As time passed

the hash chain shrank until it expired, at which point the group was dissolved.

Initially we were optimistic about this solution as it allowed an easy way for groups to spontaneously form. However through interviews and user testing we were forced to reconsider our approach. Users explained the importance of long-lasting groups that build trust and credibility over time. Despite short lived activities like protests, once users bonded together, they disliked the idea of auto-expiring membership. When recounting Gezi park protests in Istanbul, protesters noted the importance of continuing to grow activist groups after the event.

Sharing a long password also proved a usability challenge. We required users to come up with and share a long password or phrase, independent of the OSN account and unique for each group. Users had a hard time remembering passwords and resorted to using simple easy to guess passwords. The situation was complicated further by having an application password, a self-destruct password and the possibility of adding multiple groups.

Lastly, as mentioned, the hash chain approach used posts from the latest time block to validate posts from previous time blocks. This meant that : (1) the most recent post could not be validated and (2) posts in the current time block could not be validated. In testing, users stressed that the most recent post is often the most important as it is the most timely, and were confused why there was no mechanism for validation.

From this feedback we settled on the key based approach described in detail earlier in this section. The major design change is described here to highlight the importance of understanding the user community in system design. Through the course of this project, our team was able to move from preconceived understanding of the technical needs of users to a tailored approach through social analysis and iterative user testing.

10.5.2 The Proxy Server

The Android application communicates with a companion proxy server. Aside from the initial group creation (that uses the native social media platform authorization API) all content flows through this proxy. The proxy keeps group-level state and masks the individual user's IP address from the OSNs.

As discussed in Section 10.5.1, when groups are first created through the SecurePost application, the OSN platform issues an access token that the app forwards to the proxy. The app also creates and forwards the group's public *posting key* and the public *administrative key*. The proxy stores the OSN access token and uses it to make posts and change banner and profile images. The proxy also stores the public keys and uses them to verify posting and administrative rights. If the group is reset by the administrator through the application, the keys are updated but the OSN access token remains consistent. Notably the proxy does not ever receive the private keys for the group.

Since OSNs log the IP addresses of users, they may be compelled by governments to identify users and produce access logs. The proxy masks individual users' IP addresses. The proxy itself does not log IP addresses or keep any individual user metrics. Because the server may become compromised, it does not store private keys for groups. Adversaries who gain access to the proxy would be able to make posts using the OSN access token but not sign them. Any posts made by adversaries using the OSN access token from the proxy server would lack signatures and show up as invalid. The OSN access token could still be rescinded by a group member with the login information to the OSN account.

The server consists of a standalone Java application running Jetty utilizing a MongoDB No-SQL database for storage. Interaction with the application is implemented through a JSON REST API running over HTTPS. If the proxy needs to be scaled, multiple instances of the proxy can be spun up synchronizing via the MongoDB.

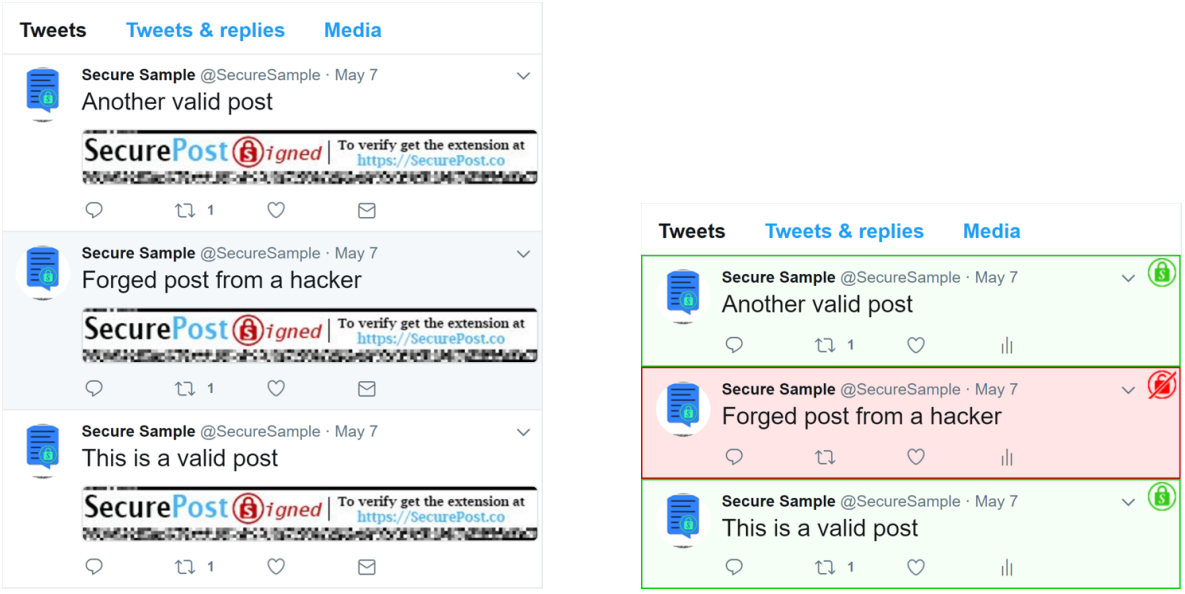
Currently, we run an instance of the proxy on Amazon Web Services. By default, users of the SecurePost application utilize this instance. As users may want to audit the source code and run their own instance, the project is open source and we allow easy configuration of the application to point to a different proxy server instance.

Unlike OSN platforms that have financial incentive to cooperate with adversaries, anyone can run a SecurePost proxy on any machine of their choosing. While some may choose to run it in a data center (possibly in another country outside the jurisdiction of their government) others may choose to do so on anonymous machines. This flexibility allows a particular group to retain control of their security and the point of failure.

If a proxy is blocked, the posts made by the group will still be visible and verifiable for the general population. There would be a brief disruption in posting content for group members, but as the group administrators have full access to the app and proxy, they could migrate to a new IP or new machine to bypass this block. Further, as we elaborate in Section 10.5.3, as long as the verification protocol is not altered, the same browser extension can run irrespective of the back-end proxy with no reconfiguration.

10.5.3 The Browser Extension

SecurePost is designed to be used to disseminate information to the public. Posts made from the application are posted directly to social media. The browser extension allows any member of the public, irrespective of group membership to verify any post made with SecurePost. Currently the browser extension is implemented as a cross-browser extension and is compatible with Chrome, Firefox, and Opera. It runs independent of the proxy server, and uses only the contents of the OSN web page to verify posts.



(a) Unmodified Twitter Feed.

(b) Using the Browser Extension.

Figure 10.7: Example of a Twitter feed using the optional verification feature of SecurePost. In the unmodified feed (left) each post has an associated signature image. Using the browser plugin (right) The signature image is automatically hidden and posts are highlighted to show validity. Mousing over the image in the top right corner of a post shows the user the reason that the post was marked.

Post Signature

The application automatically generates signatures for each post using the private *posting key* for the group. If the SecurePost group enables the option to use verification, the proxy appends an image containing the cryptographic signature when posting to the OSN, as shown by Figure 10.7a. As OSNs compress images, typically as JPGs, we use a compression resistant encoding for the image based signature. The signature is encoded as a series of monochromatic 3x3 pixel squares aligned to the row/column boundaries that match the JPG encoding algorithm. The encoded signature presents to the user similar to TV "snow" at the bottom of an image that informs the reader how to verify the post.

The browser extension processes this banner, decoding the signature back to binary. To verify the signature, the extension requires the *posting public key*. This key is encoded

by the proxy, in the same manner as the signature, into the profile image on Twitter and the cover image on Facebook. This occurs when the group is first created and whenever the images are changed. The public posting key is likewise decoded by the browser and, with the aid of the signature, is used to verify authenticity and integrity of the post.

Notably this technique does not require co-operation with the social media platform or the proxy server. The same extension can verify messages from multiple groups which may be using different proxy servers provided they use the encoding protocol for message signatures. This removes the need for non group members to access the proxy, reducing the risk of traffic analysis by an entity with a sufficient view of the network. Additionally this approach reduces the server load on the proxy as the number of readers requiring verification would likely be magnitudes greater then content creators.

Verifying Posts

When installed, the extension automatically verifies posts when on a supported OSN web page. The extension first determines whether the account uses SecurePost by reading a pixel pattern encoded in the corner of the profile image, and if so goes on to validate the posts in the feed. The extension uses a mix of color and symbols to inform the user of the validity of a post, as shown in Figure 10.7b. It also hides the signature images from the user to improve the user experience.

Posts made directly to social media, without the use of the SecurePost app, do not have access to the *private posting key* and are marked as unverified. Similarly, as discussed earlier, multimedia posts, made through the app are not able to be verified. As the signature is based on the content of the post, copying another post's signature image is insufficient to falsely validate a post. Users are graphically presented with the reason that a post is not valid or not able to be validated by hovering over an icon in the corner of the post.

Figure 10.8: Original text-based signature. Original experimental text-based signature using CJK Unified Ideographs.

If somehow an adversary took full control of the OSN account and set up a new SecurePost group, they could change the public key and make new posts that are verifiable. However all previous posts (that readers trusted) will show up as no longer verified to any member of the public using the browser. There is no way for them to create new verified posts that use the previous signing key. This would be a strong indicator to the public that there was a serious problem and perhaps a change of ownership has taken place.

Design process for image-based signatures

The ability to verify posts is one of the key facets of our work. As our partners wanted both Twitter and Facebook support, we needed a solution that was sensitive to the character limits of these OSNs, which at the time, was 140 characters on Twitter.

In our initial design, we experimented with text-based signatures using 21-bit CJK

Unified Ideographs. This character set has high bit-density per character, which allowed us to maximize bit count while minimizing the number of characters. An example of this approach is shown in Figure 10.8.

While doing user testing and interviews in Mongolia, participants expressed that this approach was unsuitable due to local cultural norms. In Mongolia there are strong tensions with China, and social ramifications for perceived affinity for one of its neighbors. By using characters associated with China or Korea, even if the characters did not correspond to actual or Korean text, users exposed themselves to perceptions of siding with these countries. This was particularly problematic for groups already marginalized.

In response to this social constraint, we moved to an image-based signatures which fits within cultural norms while still satisfying character limits. This change in design highlights the value in understanding the social context for which software is developed.

10.6 Usage and Evaluation

SecurePost is freely available on the Google Play Store. The browser extension, which allows users to verify posts, is available on the Google Chrome Web Store for free. Our app is available in 7 languages: Arabic, English, French, Mongolian, Russian, Spanish, Turkish. All modules, including the Android application, proxy server, and browser extension are open source and available through BitBucket and on our website [339].

10.6.1 Satisfying Design Requirements

As *OSNs are highly utilized* by our partner communities, we focused our work on the two highly utilized social networks: Twitter and Facebook. Since *smart phones are the primary method of Internet access*, our technical solution was designed to work with Android smart phones, while providing *backwards compatibility* for older devices.

In our work we identified public group communication that *protected anonymity* while *preserving reputation* as an unmet need for our partner communities. We address the need for *anonymity* by allowing groups to share a single OSN account while maintaining individual anonymity. By sharing access keys, there is no group roster. IPs are kept hidden from OSNs via the use of a proxy that groups can retain full ownership and control over.

To *maintain reputation* we provide a method of verifying post authenticity with the use of cryptographic signatures. Members of the public can verify that a post came from a group using a companion browser. Posts that are modified are no longer verified, and accounts that are seized or hacked are unable to post verifiable posts. If a group is ever compromised administrators can reset membership and invalidate past posts to signal to readers that there is a problem and they should re-evaluate the truthfulness of current posts. In this manner groups can build reputation, and signal if that reputation might be compromised.

In future work we are interested in exploring ways of migrating groups between OSN accounts using the same signature keys and proxy servers. This would allow groups to switch to a new OSN account in light of censorship while preserving reputation, and could allow verification of identity across services. We are also exploring ideas of tiers of user class so that posts from a designated core group of users can retain verification after group reset.

As participants, such as journalists, expressed a need to report from areas with periodic or permanent *network disruption*, SecurePost maintains a local cache of posts. Users can view past posts without connectivity, and compose posts that are delivered when connectivity is re-established. In this manner users can ferry information from disconnected geographic regions.

As storing content locally leaves users vulnerable to *device confiscation and search*,

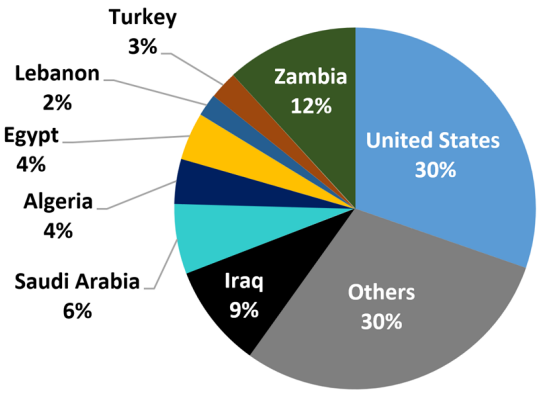


Figure 10.9: SecurePost application installation by country as of October 2017.

SecurePost encrypts all data using an application-wide password. In the event that a user is under duress they can provide a false password that wipes the application data including posting keys.

While designing the application we received iterative feedback from participants that tailored design decisions to be *appropriate for our intended users*. As outlined in the previous section we modified our initial approaches to fit cultural norms (such as switching from CJK text-based signatures) and fit with user social behavior (such as abandoning the initial hash-chain approach).

10.6.2 Usage Survey

As anonymity is at the core of the design philosophy, SecurePost does not collect usage data beyond that which is necessary for functionality of the app and server, as well as basic install statistics collected by the Google Play Store. Based on data from the Google Play Store, as of October 2017, we have had over 400 installs of our application. As we show in Figure 10.9, our users reside not only the countries we were explicitly targeting: USA, Mongolia, Turkey, Zambia, but in other countries where censorship is a problem (59 countries in total). SecurePost users have so far formed 68 groups and made 336 posts.

Table 10.5: Results of initial demographic survey. Collected by when application is first run.

Survey Question	%	(#)
Gender		
Male	62%	(95)
Female	35%	(54)
Other	3%	(4)
Age		
$\leq$ 20 years old	42%	(67)
21-30 years old	49%	(79)
31-40 years old	6%	(9)
41-50 years old	2%	(3)
$\geq$ 51 years old	1%	(1)
Education		
Primary school	1%	(2)
Secondary school / High school	12%	(18)
Higher education or university	87%	(130)
Preferred Language		
Arabic	2%	(1)
English	90%	(43)
Spanish	8%	(4)
Application/OS Language		
Arabic	0.3%	(1)
English	95%	(314)
Spanish	2%	(8)
French	1%	(3)
Russian	0.3%	(1)
Turkish	0.6%	(2)
Total # Surveyed	(329)	

When the application is first run, users are given the option to provide optional demographic information. The data to date are summarized in Table 10.5. As of October 2017, we received 329 responses. From this survey, we found the majority (62%) of users identified as male. Most were 30 years old or younger (91%) and had some higher education (87%). Because we led user studies and presented our application at universities and organizations where college-age students are likely to work, we expected our users to fall into this demographic. In developing countries, younger generations and men are also more likely to use the Internet as a whole [341].

For language, few respondents (15% of total responses) stated a preferred language. Those that did largely preferred English, the dominant language in Zambia and the United States, which makes up a large portion of all installs.

Our survey also automatically registered the language to which the application was currently set. The application language is a match-up between operating system defaults and the seven supported languages of the application. Languages not supported by the application would be reported as the next likely language, usually English. The exact implementation varies by Android version and is detailed in [376]. In this metric, English is again the dominant language (95%), suggesting that our translations are not heavily utilized.

After using the application for three days, users were asked if they would take an optional survey based on their experiences using SecurePost. This follow-up survey is also available through an in-app menu. A subset of the survey questions is summarized in Table 10.6. A total of 22 users have so far completed this survey, which is a small sample. We are working to increase the response rate by re-offering the survey.

Responses to the follow-up survey indicate that the majority of respondents found SecurePost to be at least somewhat useful (73%). When asked to what extent they thought that using SecurePost has improved their confidence in using online social networks, most

Table 10.6: Results of follow-up survey. Collected by application after three days of use.

Survey Question	%	(#)
How useful is SecurePost to user		
Not useful at all	4%	(1)
A little useful	23%	(5)
Somewhat useful	41%	(9)
Very Useful	32%	(7)
SecurePost Improved confidence with OSNs		
Not improved confidence	9%	(2)
Improved confidence a little	41%	(9)
Improved confidence somewhat	45%	(10)
Improved confidence a lot	5%	(1)
Feel Freer on OSNs		
A lot less free	4%	(1)
A little less free	9%	(2)
Equally as free	4%	(1)
A little more free	50%	(11)
A lot more free	23%	(5)
Total # Surveyed	(22)	

said it improved their confidence a little, somewhat, or a lot (91%). When asked to what extent users feel freer in their ability to express themselves on Twitter and/or Facebook without concerns about surveillance or security of the messages sent when using SecurePost, most said they feel freer to some extent (73%).

10.7 Conclusion

Our research seeks to understand and address barriers to free speech on the Internet for vulnerable communities. While there are many significant differences across these diverse communities, we observed particular patterns in free speech challenges among Internet and OSN users across these contexts. Social research revealed user concerns ranging from account disruptions to online credibility issues to equipment seizures. We built a novel tool to address such challenges in partnership with affected communities.

SecurePost allows users greater control of anonymity through group-based communication on OSNs. Through verified group anonymity, users build trust and reputation as a collective, without exposing the identities of individuals. By allowing communities to set up their own instances of the SecurePost proxy server, users do not have to trust OSNs to protect IPs and identities of individual members.

Using SecurePost, an administrator can share OSN accounts without sharing the account passwords and hence maintain control of the account. If the account is seized or hacked, the browser extension can still identify fraudulent or edited posts using the cryptographic signature.

We developed SecurePost to support the most popular OSN platforms (Facebook and Twitter) and device types (Android smart phone), providing compatibility for older devices lacking some of the security features (e.g. encrypted storage) of newer phones. The application also provides a means of storing messages for later delivery to counter

network disruption due to power loss or government censure. Because the project is open source and designed for modularity, other similar platforms and systems can be incorporated in the future.

Like other anonymity applications, individual users can still be revealed by posting personal information in the contents of a message. Given time and access to the file structure of the device, it is also possible for the encrypted storage to be decrypted, for example via a brute force attack. Additionally, adversaries with a sufficient view of the network may still implement de-anonymization through timing analysis. We hope to address these vulnerabilities in future work.

A frequent concern for the development of security applications is the potential misuse by ill-meaning organizations, like terrorist cells. Because the data are all posted publicly, our app does not expand the capabilities of malevolent secret communication. While our tool allows users to remain anonymous, it does not prevent OSNs from censoring accounts or content. We leave the decision of what constitutes a danger to the OSN.

Because all elements of our software are open source, communities do not have to trust us (the developers) or OSNs to protect the identities of individual members. They can audit and improve the code, and can set up their own instances of the SecurePost proxy server that is isolated from developers.

Finally, our work provides insight into community collaborations. By partnering with locals and understanding social context, specific needs, and user behaviors, we were able to come up with a novel method of adding verification to non-cooperative online social networks.

Acknowledgments

This work was done in collaboration with Danny Iland, Miriam Metzger, Lisa Parks and Elizabeth Belding. I would also like to thank all the other people who have helped with this work. Many thanks to members of the social team including Hannah Goodwin, Kristin Hocevar, Lisa Han, Ariel Hasell, and Rahul Mukherjee for conducting interviews and surveys as well as providing analysis. Thanks to Ben Zhao, Divya Sambasivan, and Pritha Narayanappa for helping develop SecurePost. Thanks to Irina Artamonova for helping with the statistical analysis of our survey data. And finally thanks to our many overseas partners for the hospitality, patience, and many hours of work. This work was funded by the DRL.

Chapter 11

Future of Wireless Communication

Computer science researchers strive to design technologies that tackle existing social challenges and that extrapolate into the future to develop solutions for emerging problems. Our work comes at a critical time in our society, when social inequality and natural and political crises are on the rise. The co-occurring major natural and political crises that frequently intersected with our work motivated the research presented thus far. To conclude this dissertation in this chapter, we explore this additional context and motivation, as well as examine emerging communication technologies and how our work integrates with the broader vision of future wireless connectivity.

11.1 A Disruptive Future

11.1.1 Natural Crises

The majority of this research was conducted in California, which has seen a rise in natural crises in the last decade. During our research, California experienced the wettest winter in a century [377], causing landslides and flooding that damaged infrastructure, including major transportation arteries (such as the primary freeway leading to our

(a) Sherpa Fire. (b) Whittier Fire.

Figure 11.1: Photos of recent California wildfires.

University) and the Oroville dam [378]. At the other extreme, dry summers fueled brutal wildfire seasons, including the 2016 Sherpa fire (Figure 11.1a), the 2017 Whittier Fire (Figure 11.1b), the 2017 Tubbs fire in Sonoma (the second most *damaging* wildfire recorded in California history [379]), the 2017 Thomas fire in Ventura and Santa Barbara (the second *largest* fire recorded in the state's history [380]), the 2018 Camp fire (the most *damaging* wildfire recorded in California history [379]) and the 2018 Mendocino Complex Fire (the *largest* fire recorded in the state's history [381]). Note that these record setting events all occurred in the last few years.

Elsewhere in the world, similar trends of record breaking natural disasters resulted in numerous deaths and damage to infrastructure. In 2017, a heavy monsoon season caused flooding in India, Bangladesh, and Nepal, killing over a thousand people and impacting tens of millions [382]. In the Atlantic, three devastating hurricanes (Harvey, Maria, and Irma) impacted countries in and around the Gulf of Mexico, including the Dominican Republic, Cuba, and the United States. In Houston, for example, an estimated 136,000 structures (10 % of total structures) were flooded [383]. From 2019 through 2020, Australia and Brazil have experienced the largest wildfire season on record with millions of hectacres burned [384, 385].

The impact of these disasters lasts from months to years [28] and results in mas-

sive ecological damage, mass migration, and threatens the economic stability of these regions [29, 386, 387]. Natural disasters damage otherwise well provisioned networks when critically needed. In addition to directly threatening human lives, natural disasters cripple communication infrastructure. Fueled by climate change, such disasters are increasing in frequency and intensity [30, 31]. With increases in severity come concomitant increases in the number of affected individuals. For example, 160 countries have more than one-fourth of their total population in areas at relatively high mortality risk from one or more hazards [388].

Wireless networks of the future will need to be resilient to these emerging natural crises and must be designed accordingly. Disaster response and recovery requires rapid assessment of the changing network conditions. The work in this dissertation further expands capabilities by proposing systems of automated network quality assessment and systems for locating disconnected wireless devices. Our work paves the way for passive evaluation of networks without relying on cooperation from MNOs, allowing independent assessment when time and resources are limited.

11.1.2 Political Crises

In addition to an increase natural disasters, civil freedoms are declining globally [32]. Despite the UN's assessment that the protection of free speech should fully extend to the Internet [266, 267], governments, corporations, and individuals often restrict what users can say online and punish those with dissenting views. The last decade experienced a sharp rise in censorship, Internet surveillance, targeted disinformation, and cyberbullying [33]. Adversaries to open discourse deploy armies of operatives masquerading as legitimate users to post "fake news", eroding reputation and trust of content [67, 68, 69, 70, 280, 281].

Our work came at particularly critical time for Turkey. During our research, Turkey

(a) Women's March. (b) March For Science.

Figure 11.2: Photos of recent protests in Santa Barbara, California.

experienced a steep change in freedoms co-occurring with a massive protest in 2013 at Gezi Park [60], and an attempted coup in 2016 [61]. The Turkish government responded by purging academics, blocking access to networks, and deploying a massive censorship campaign [62, 269, 369, 389]. Recently, censorship has been utilized in a number of other countries as response to political conflict [36, 37, 38, 39, 40]. In some cases this took the form of total Internet access disruptions in Bangladesh, Indian province of Kashmir, and Myanmar [34, 35, 390]. In other cases, governments limit Internet access for particular types of content [284].

Our research progressed both the understanding of global issues surrounding freedom of speech and provided new systems for protecting those freedoms. We collaborated with interdisciplinary experts and community partners to characterize social challenges to free speech online across three different regions. To address unmet needs, we introduced a novel system of verified group anonymity, which allows users to build trust and reputation as a collective, without exposing the identities of individuals.

As in the case of a natural disaster [164], large protests, such as the 2017 Women's March (Figure 11.2a) and the March for Science (Figure 11.2b), can overwhelm previously well provisioned cellular networks [391]. Our work provides methods of analyzing network

load without cooperation from MNOs.

11.1.3 The Digital Divide

As of 2019, only 54 percent of the world's population had Internet access [1]. In the United States, the richest country by GDP, one third of the population lacks broadband access. Rural Americans are 12% less likely than their urban counterparts to have access to home broadband [392]. As more of our society becomes reliant on Internet connectivity, those without access can fall behind. For example, in the U.S., 17% of surveyed teens said they often or sometimes were unable to complete a homework assignment because they do not have reliable access to a computer or internet connection [393]. The digital divide can amplify existing societal disparities, such as participation in arts and culture [394], health literacy [395], and even transportation [396]. Rural and tribal communities are particularly likely to lack access [3, 4].

As we have previously discussed, LTE is often used for mobile broadband to connect rural communities and bridge the digital divide. The assessment of such networks, both in terms of coverage and performance, remains a significant challenge due to the overly optimistic reports of availability and failures to account for quality and usability. Our work on passive assessment of coverage and load fills this gap.

Our work engaged communities both locally and globally. We have applied our work for LTE and TVWS assessment to advise network planners of tribal communities in southern California and New Mexico (Figure 11.3a). Further, in conducting research on freedom of speech online, we provided insight into the needs of developing communities in Zambia, Turkey and Mongolia, including the Ger District of Ulaanbaatar shown in Figure 11.3b. Inequality, natural disasters, and political crises are just three of the emerging problems of the next century to which computer networks will need to respond [76, 77].

(a) Santa Clara Pueblo, New Mexico, USA

(b) Ger District, Ulaanbaatar, Mongolia.

Figure 11.3: Photos of rural partner communities.

11.2 Technologies for the Future

Governments and companies are already preparing for these new realities. In the United States, the FCC released the *Wireless Resiliency Cooperative Framework*, which has been signed by all the top MNOs, including Verizon, AT&T, Sprint, and T-Mobile. The framework expresses "voluntary industry commitment to promote resilient wireless communications and situational awareness during disasters" [397]. Technology giants are developing technologies to handle these emergent threats. Our work provides valuable insights into deficiencies of existing technologies, as well as tools and techniques that will remain relevant for the next generation of communication standards.

11.2.1 LTE Networks

In fact, MNOs made notable improvements in disaster preparation in recent years. When a violent weather event is predictable, cellular companies have a variety of tools at their disposal: they deploy and top-off fuel powered backup generators; they place Cell on Wheels (COWs), Cell on Light Trucks (CoLTs), Cell Repeaters on Wheels (CROWs), at strategic locations where outages are anticipated; and they increase staffing at emergency command centers, among others [398]. While these strategies prove helpful for anticipated

natural disasters, unanticipated natural disasters (such as wildfires), disasters where impact is hard to forecast (such as the path of a hurricane), or disasters that become worse than predicted (such as the extensive flooding seen in Hurricane Harvey) can make these solutions unfeasible. For instance, fires and flood waters can prevent the placement of COWs, CoLTs, and other emergency machinery in needed locations. Once placed, they cannot easily be moved or refueled. Our work investigates rapidly evaluating coverage and overload, which can inform the decision making process of where and when to deploy COWs, CoLTs, and CROWs.

11.2.2 Aerial Applications

UASs are increasingly proposed for rapidly extending wireless connectivity [399]. UASs offer a wide variety of connectivity options, limited primarily by payload weight constraints. For instance, as mentioned earlier, MNOs already employ UASs for visually inspecting equipment after natural disasters [156]. The rapidly changing topography of an aerial wireless network, which can provide a backhaul link from the UAS to the Internet, offers both challenges and opportunities for connectivity innovation and optimizing performance [400]. In the simplest case, the UAS could serve as a cellular relay, patching connectivity holes between operational base stations and users [401, 402, 403, 404, 96, 405] or between COWs and users. UASs can even be perched for rapidly deployable infrastructure [103, 104, 105]. Our work paves the way for detecting and automatic aerial patching of coverage gaps and supplementing cellular capacity in the event of overload.

11.2.3 TV White Spaces

When cellular base stations are unavailable, non-operational, or overloaded, TVWS links provide a promising option for backhaul connectivity. TVWS links are characterized

by long transmission ranges and good propagation and penetration properties, making them ideally suited for extending connectivity in difficult-to-reach areas. UASs equipped with lightweight TVWS radios have already been used for spectrum monitoring [102], and cities, such as Oxford, UK are exploring their use for flood monitoring [406]. In cases where TVWS links are impractical or unavailable, DTN literature offers a variety of solutions for delayed communication [407]. Our work has covered a portion of this space via the use of UASs for rapid frequency scans of incumbent transmitters. This work parallels the LTE coverage assessment in chapter 3 and uses low cost SDRs on aerial platform for frequency sensing and coverage mapping. This work has already helped tribal communities plan networks in California and New Mexico.

11.2.4 5G Cellular

At the time of writing this dissertation, 5G cellular networks dominate the discussion on future networks. Fifth generation cellular technology promises a connectivity panacea with a dense grid of intelligent wireless devices that will deliver "multi-Gbps peak rates, ultra-low latency, massive capacity, and more uniform user experience" [408]. According to Verizon, "users will know it as one of the fastest, most robust technologies the world has ever seen. That means quicker downloads, outstanding network reliability and a spectacular impact on how we live, work and play." [409] Despite the intense optimism about these networks, the proposed approaches to 5G networks pose many of the same types of problems as existing cellular technologies. Further, the focus on increasing network throughput leads to complications that can hamper rural roll-out and operations post crises.

As with all prior cellular technologies, build outs will occur first in urban areas. According to [410], urban centers, such as Chicago, Minneapolis, Denver, Atlanta, and

Washington DC already have some areas of 5G coverage. Deployment within a handful of suburban cities, such as Panama City and Sioux Falls, is on the horizon. These selected regions have a common set of characteristics on which 5G technology depends: contemporaneous end-to-end connectivity, economic gain from infrastructure deployment, and high speed backbone links (either pre-existing, or easily deployed).

For rural regions, there are few to no announcements of upcoming coverage. Indeed, rural and tribal regions in the United States currently have the lowest cellular coverage and Internet availability, lagging far behind urban and suburbans centers. While only 10% of Americans lack broadband access, that number increases to 39% in rural areas [411]. According to the 2018 FCC Broadband Report, broadband adoption rates on tribal lands is only 32.6%, and, while LTE deployment in urban areas of the U.S. has seen impressive growth in the last three years, deployment on tribal lands remains flat [142]. This coverage gap is largely due to economics. With their sparser population densities and wider geographic areas, rural cellular and Internet coverage is more expensive to deploy and offers significantly lower return on investment. As a result, we observe a rural/urban digital divide across the U.S., as well as in many other developed regions of the world [4, 3, 412]. This divide is occasionally narrowed through government subsidies, such as the E-Rate program [413, 144, 414, 415, 141], but so far this has resulted in only pockets of coverage in most rural and tribal communities.

To achieve top 5G data rates and ultra-low latencies, 5G radios will need to be densely deployed and have fiber backhaul. While most urban areas are already well-equipped to provide both the fiber backhaul and the infrastructural support needed to physically place 5G radios (i.e. light poles with electrical power supplies), these resources are rarely available in rural areas, and completely unavailable in the density needed to achieve multi-Gbps, ultra-low latency connectivity to rural homes and work places. For example, in the *5G Rural First* project headed by Cisco and the UK Government's Department for

Digital, Culture, Media, and Sport, the pilot network in northern Scotland used a fiber backhaul to an urban center in Glasgow hundreds of miles away [416]. As we discussed in chapter 1, fiber and other wired solutions are a persistent existing challenge for many rural communities, including tribal land, and those in the developing world. A cellular solution that requires fiber connectivity is unlikely to reach these communities in the foreseeable future.

Because 5G radio deployments are anticipated to be ultra-dense in urban areas, it is likely that these networks will operate over mmWave links, which are needed to achieve the ultra-high data rates. These mmWave links only cover a couple hundred meters per base station. However, because of the large geographic spans with lower population densities that need coverage in rural areas, current plans for 5G usage in these spaces tend toward "low-band" communication in the 600-900MHz frequency range, or the "sub-6 mid-bands" in the 3.5-6 GHz range. These bands have better propagation properties, at the expense of lower data rates and greater latencies. Hence, while rural regions may benefit from the architectural features of 5G such as network function virtualization, they are unlikely to receive the massive boosts in data rates and reductions in latencies that urban regions will experience and that are the hallmark of 5G connectivity.

Despite these limitations, there are many types of challenged networks that could vastly benefit from both the 5G architecture and superior broadband capabilities. As an example, rural regions typically lack state-of-the-art hospitals, and often lack medical specialists. As a result, tele-medicine and remote health care can be critical in these areas. Current tele-health applications leverage remote video consultation, and remote surgery via high speed ultra-reliable networks is on the horizon. Residents of rural regions stand to benefit the most from these advanced applications, and need the infrastructure to support them. As a second example, first responders to natural disasters (e.g., wildfires, hurricanes, earthquakes and tornadoes) have critical multi-modal connectivity needs,

ranging from real-time voice to ultra-low latency video, for instance from surveillance UASs [202, 200, 201, 417, 418, 419, 420].

As we have seen repeatedly in this dissertation, natural disasters can damage or down cellular towers, as well as electrical grids [164, 24, 26, 25, 421, 422]. In such cases, rural regions are particularly susceptible, as they tend to have fewer pathways (often only one) between the community and the Internet [144, 414, 423, 424]. Further, as wildfires become increasingly frequent and unpredictable, first responders and fire fighters often must gather in remote, offline locations, yet could still benefit from the same real-time applications that require at least local connectivity (e.g. UAS surveillance video). Ironically, in exactly the moments during and after a disaster, where connectivity is most needed, connectivity is typically unavailable. If a fiber optic cable or related equipment is damaged, in proposed 5G networks, all local access will be severed. In the worst cases, the lack of connectivity can make the difference between life and death [425, 426, 427, 383].

As a result, the roll-out of 5G is unlikely to change the needs or capabilities of the communities in challenged regions in the near future. The research presented in this thesis will continue to be applicable to 5G enabled communities. For example, given the wider range of frequencies utilized by 5G networks with a heterogeneity in range and speed, the need for accurate and rapid coverage mapping grows. In the 5G landscape, the wide discrepancy in frequency characteristics (such as permeability of surfaces) will be even harder to model, making the previous binary coverage MNO self-reporting strategies employed by regulators even more ineffective. An area covered by lower throughput "low-band" frequencies may for example have no coverage of high throughput mmWave. Likewise, 5G's capabilities to enable IoT applications is projected to spur the growth of WSNs. When 5G connectivity is disrupted, DTN data delivery approaches, like those outlined in this dissertation, will remain necessary.

11.3 Conclusion

In this dissertation, we presented research into critical contemporary challenges for the use of wireless networks. Our work covers widely deployed standards across the wireless ecosystem. We develop new tools to improve network assessment and to provide robust and reliable network communication. Our work tackles issues, including the measurement of rural network coverage, assessment of network under crisis conditions, aerial connectivity, disaster applications, and censorship. By incorporating new technological breakthroughs, such as UASs and SDRs, we introduced novel methods and systems for these challenged networks. As we discussed in this chapter, our insights are applicable to many of the emerging wireless communication technologies.

We have evaluated how technologies and standards function in difficult environments, such as those lacking end-end Internet connectivity, overloaded, resource constrained, and operating in three dimensional space. Through this lens, we demonstrated how to optimize networks to serve marginalized communities, outside of first world urban cities, and make our networks more resilient to natural and political crisis that threaten communication.

Bibliography

[1] International Telecommunication Union, "ICT facts and figures 2019." https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf, 2019.

[2] National Telecommunications and Information Administration, "Falling through the net: Defining the digital divide." https://www.ntia.doc.gov/ntiahome/fallingthru.html, July, 1999.

[3] A. Perrin, *Digital gap between rural and nonrual America persists*, *Pew Research* (May, 2019).

[4] D. West, *Rural and urban America divided by broadband access*, *Brookings* (July, 2016).

[5] J. A. Van Dijk, *Digital divide research, achievements and shortcomings*, *Poetics* **34** (2006), no. 4-5 221–235.

[6] P. DiMaggio, E. Hargittai, *et. al.*, *From the 'digital divide'to 'digital inequality': Studying internet use as penetration increases*, *Princeton: Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University* **4** (2001), no. 1 4–2.

[7] Internet Society, "Global internet report 2016." https://future.internetsociety.org/2016/wp-content/uploads/2016/11/ISOC_GIR_2016-v1.pdf, 2016.

[8] International Telecommunication Union, "Last-mile internet connectivity toolkit." https://www.itu.int/en/ITU-D/Technology/Documents/RuralCommunications/20200120%20-%20ITU%20Last-Mile%20Internet%20Connectivity%20Toolkit%20-%20DraftContent.pdf, Jan, 2020.

[9] A. R. Islam, N. Selvadurai, and G. Town, *Wireless broadband technologies for regional and rural australia: A last-mile perspective*, *Telecommunications Journal of Australia* **58** (2008), no. 2/3.

[10] H. Galperin, *Wireless networks and rural development: Opportunities for Latin America*, *Information Technologies & International Development* **2** (2005), no. 3 pp–47.

[11] R. K. Patra, S. Nedevschi, S. Surana, A. Sheth, L. Subramanian, and E. A. Brewer, *WiLDNet: Design and implementation of high performance WiFi based long distance networks.*, in *NSDI*, vol. 1, 2007.

[12] L. Subramanian, S. Surana, R. K. Patra, S. Nedevschi, M. Ho, E. A. Brewer, and A. Sheth, *Rethinking wireless in the developing world*, in *HotNets*, 2006.

[13] Government Accountability Office, *Broadband Internet: FCC's Data Overstate Access on Tribal Lands*, September, 2018.

[14] Rural Wireless Association, "Challenges faced by small wireless providers in measuring LTE coverage." `https://ruralwireless.org/rwa-welcomes-fcc-investigation-into-violation-of-mobility-fund-phase-ii-mapping-rules`, Dec, 2018.

[15] M. H. Anisi, G. Abdul-Salaam, and A. H. Abdullah, *A survey of wireless sensor network approaches and their energy consumption for monitoring farm fields in precision agriculture*, *Precision Agriculture* **16** (2015), no. 2 216–238.

[16] T. Kalaivani, A. Allirani, and P. Priya, *A survey on Zigbee based wireless sensor networks in agriculture*, in *3rd International Conference on Trendz in Information Sciences & Computing (TISC2011)*, IEEE, 2011.

[17] A. Kumar and G. P. Hancke, *Energy efficient environment monitoring system based on the IEEE 802.15.4 standard for low cost requirements*, *IEEE Sensors Journal* **14** (2014), no. 8.

[18] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, and Y. F. Hu, *Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards*, *Computer communications* **30** (2007), no. 7.

[19] F. Z. Benhamida, A. Bouabdellah, and Y. Challal, *Using delay tolerant network for the internet of things: Opportunities and challenges*, in *2017 8th International Conference on Information and Communication Systems (ICICS)*, pp. 252–257, IEEE, 2017.

[20] M. A. Marinho, E. P. De Freitas, J. P. C. L. da Costa, A. L. F. de Almeida, and R. T. de Sousa, *Using cooperative MIMO techniques and UAV relay networks to support connectivity in sparse Wireless Sensor Networks*, in *ComManTel*, IEEE, 2013.

[21] E. P. De Freitas, T. Heimfarth, I. F. Netto, C. E. Lino, C. E. Pereira, A. M. Ferreira, F. R. Wagner, and T. Larsson, *UAV relay network to support WSN connectivity*, in *ICUMT*, IEEE, 2010.

[22] I. Jawhar, N. Mohamed, J. Al-Jaroodi, and S. Zhang, *A framework for using unmanned aerial vehicles for data collection in linear wireless sensor networks*, *Journal of Intelligent & Robotic Systems* **74** (2014), no. 1-2.

[23] J. R. Martinez-de Dios, K. Lferd, A. de San Bernabé, G. Núnez, A. Torres-González, and A. Ollero, *Cooperation between UAS and wireless sensor networks for efficient data collection in large environments*, *Journal of Intelligent & Robotic Systems* **70** (2013).

[24] FCC, "Communications Status Report for Areas Impacted by Tropical Storm Harvey, August 27, 2017." `https://apps.fcc.gov/edocs_public/attachmatch/DOC-346369A1.pdf`, Aug, 2017.

[25] FCC, "Communications Status Report for Areas Impacted by Hurricane Irma - September 11, 2017 ." `https://transition.fcc.gov/Daily_Releases/Daily_Business/2017/db0911/DOC-346655A1.pdf`, September, 2017.

[26] FCC, "Communications Status Report for Areas Impacted by Hurricane Irma - September 7, 2017 ." `https://transition.fcc.gov/Daily_Releases/Daily_Business/2017/db0907/DOC-346607A1.pdf`, September, 2017.

[27] FCC, "Communications Status Report for Areas Impacted by Hurricane Maria - September 21, 2017." `https://transition.fcc.gov/Daily_Releases/Daily_Business/2017/db0921/DOC-346840A1.pdf`, September, 2017.

[28] F. Robles, *Puerto Rico spent 11 months turning the power back on*, *New York Times* (Aug, 2018).

[29] M. Echenique and L. Melgar, "Where Puerto Rico's Residents Migrated Since Maria." `https://www.citylab.com/environment/2018/05/watch-puerto-ricos-hurricane-migration-via-mobile-phone-data/559889/`, May, 2018.

[30] U. Irfan and B. Resnick, *Megadisasters devastated America in 2017. and they are only going to get worse*, *Vox* (Jan, 2018).

[31] NOAA, "Billion-dollar weather and climate disasters." `https://www.ncdc.noaa.gov/billions`, Jan, 2018.

[32] *Freedom in the world 2019*, *Freedom House* (Feb, 2019).

[33] A. Shahbaz and A. Funk, *Freedom on the net 2019: The crisis of social media*, *Freedom House* (2019).

[34] J. Emont, *Bangladesh cuts mobile access to Rohingya refugees*, *Wall Street Journal* (Sep, 2019).

[35] V. Goel, K. D. Singh, and S. Yasir, *India shut down Kashmir's internet access. now, 'we cannot do anything.'*, *New York Times* (Aug, 2019).

[36] S. Kelly, M. Earp, L. Reed, A. Shahbaz, and M. Truong, *Privatizing censorship, eroding privacy*, *Freedom House* (Oct, 2015).

[37] *Ethiopia: Government blocking of websites during protests widespread, systematic and illegal*, *Amnesty International* (Dec, 2016).

[38] C. Arthur, *Egypt blocks social media websites in attempted clampdown on unrest*, *The Guardian* (Jan, 2016).

[39] K. Lim and E. Danubrata, *Singapore seen getting tough on dissent as cartoonist charged*, *Reuters* (Jul, 2013).

[40] T. B. Lee, *Here's how Iran censors the Internet*, *The Washington Post* (Aug, 2013).

[41] A. Izaguirre, *Critics say FCC report overstates broadband availability*, *Associated Press* (May, 2019).

[42] J. Kahan, "It's time for a new approach for mapping broadband data to better serve americans." `https://blogs.microsoft.com/on-the-issues/2019/04/08/its-time-for-a-new-approach-for-mapping-broadband-data-to-better-serve-americans/`, Apr, 2019.

[43] M. Nekrasov, V. Adarsh, U. Paul, E. Showalter, E. Zegura, M. Vigil-Hayes, and E. Belding, *Evaluating LTE coverage and quality from an unmanned aircraft system*, in *Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, IEEE, 2019.

[44] Microsoft Research, "The 2018 Microsoft airband initiative." `https://blogs.microsoft.com/uploads/prod/sites/5/2018/12/MSFT-Airband_InteractivePDF_Final_12.3.18.pdf`, December, 2018.

[45] V. Adarsh, M. Nekrasov, E. Zegura, and E. Belding, *Packet-level overload estimation in lte networks using passive measurements*, in *Proceedings of the Internet Measurement Conference*, pp. 158–164, 2019.

[46] T. Lennvall, S. Svensson, and F. Hekland, *A comparison of WirelessHART and ZigBee for industrial applications*, in *2008 IEEE International Workshop on Factory Communication Systems*, IEEE, 2008.

[47] M. Nekrasov, R. Allen, and E. Belding, *Performance analysis of aerial data collection from outdoor IoT sensor networks using 2.4 GHz 802.15.*, in *Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, pp. 33–38, ACM, 2019.

[48] M. Nekrasov, R. Allen, I. Artamonova, and E. Belding, *Optimizing 802.15.4 outdoor IoT sensor networks for aerial data collection*, *Sensors* **19** (2019), no. 16.

[49] T. OBrien, R. Durscher, and C. Briggert, "The use of remotely piloted aircraft systems (RPAS) by the emergency services." `http://www.eena.org/download.asp?item_id=207`, Nov, 2016.

[50] I. Kwai, *A drone saves two swimmers in Australia*, *New York Times* (Jan, 2018).

[51] X. Lin, V. Yajnanarayana, S. D. Muruganathan, S. Gao, H. Asplund, H.-L. Maattanen, M. Bergstrom, S. Euler, and Y.-P. E. Wang, *The sky is not the limit: LTE for unmanned aerial vehicles*, *IEEE Communications Magazine* **56** (2018), no. 4 204–210.

[52] Facebook, "Crisis response." `https://www.facebook.com/about/crisisresponse/`, 2020.

[53] M. L. Kraushar and R. E. Rosenberg, *A community-led medical response effort in the wake of Hurricane Sandy*, *Disaster medicine and public health preparedness* **9** (2015), no. 4.

[54] K. G. Panda, S. Das, D. Sen, and W. Arif, *Design and deployment of UAV-aided post-disaster emergency network*, *IEEE Access* **7** (2019).

[55] U. Paul, M. Nekrasov, and E. Belding, *Emergence: A delay tolerant web application for disaster relief*, in *Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications*, pp. 167–167, 2019.

[56] U. Paul, A. Ermakov, M. Nekrasov, V. Adarsh, and E. Belding, *Outage: Detecting power and communication outages from social networks*, in *Proceedings of The Web Conference 2020*, 2020.

[57] V. Acuna, A. Kumbhar, E. Vattapparamban, F. Rajabli, and I. Guvenc, *Localization of WiFi devices using probe requests captured at unmanned aerial vehicles*, in *Wireless Communications and Networking Conference (WCNC)*, IEEE, 2017.

[58] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, *Locating the nodes: cooperative localization in wireless sensor networks*, *IEEE Signal processing magazine* **22** (2005), no. 4 54–69.

[59] F. Caron, S. N. Razavi, J. Song, P. Vanheeghe, E. Duflos, C. Caldas, and C. Haas, *Locating sensor nodes on construction projects*, *Autonomous Robots* **22** (2007), no. 3 255–263.

[60] O. Zihnioglu, *The legacy of the Gezi protests in Turkey - after protest: Pathways beyond mass mobilization*, *Carnegie Europe* (Oct, 2019).

[61] *Turkey's coup attempt: What you need to know*, *BBC News* (July, 2016).

[62] H. Pamuk and E. Toksabay, *Purge of academics leaves future of Turkish universities in doubt*, *Reuters* (Mar, 2017).

[63] M. Nekrasov, L. Parks, and E. Belding, *Limits to internet freedoms: Being heard in an increasingly authoritarian world*, in *Proceedings of the Third Workshop on Computing Within Limits*, ACM LIMITS '17, June, 2017.

[64] P. Syverson, R. Dingledine, and N. Mathewson, *Tor: the second generation onion router*, in *Proceedings of the USENIX Conference on Security Symposium. USENIX Association*, SSYM'04, (Berkeley, CA, USA), USENIX Association, 2004.

[65] "Signal." `https://signal.org`.

[66] WhatsApp Inc., "Whatsapp." `https://www.whatsapp.com/`.

[67] B. Carey, *How fiction becomes fact on social media*, *The New York Times* (Oct, 2017).

[68] S. Earle, *Trolls, bots and fake news: The mysterious world of social media manipulation*, *Newsweek* (Oct, 2017).

[69] L. Benedictus, *Invasion of the troll armies: 'Social media where the war goes on'*, *The Guardian* (Nov, 2016).

[70] A. Chen, *The Agency*, *The New York Times* (Jun, 2015).

[71] D. Jackson, *Ap twitter feed hacked; no attack at White House*, *USA Today* (Apr, 2013).

[72] G. Bowden, *BBC Northampton Twitter account issues Donald Trump shot tweet after 'hack'*, *Huffington Post* (Jan, 2017).

[73] M. Ingram, *Twitter hack takes over accounts to spread fake news*, *Fortune* (Jun, 2017).

[74] M. Nekrasov, D. Iland, M. Metzger, B. Zhao, and E. Belding, *SecurePost: Verified Group-Anonymity on Social Media*, in *Proceedings of the 7th USENIX Workshop on Free and Open Communications on the Internet FOCI*, USENIX, Aug, 2017.

[75] M. Nekrasov, D. Iland, M. Metzger, L. Parks, and E. Belding, *A user-driven free speech application for anonymous and verified online, public group discourse*, *Journal of Internet Services and Applications* **9** (2018), no. 1 21.

[76] World Economic Forum, "The global risks report 2020." `http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf`, Jan, 2020.

[77] K. Brown, "5 global issues to watch in 2020." `https://unfoundation.org/blog/post/5-global-issues-to-watch-in-2020/`, Jan, 2020.

[78] J. Torres-Sánchez, F. López-Granados, N. Serrano, O. Arquero, and J. M. Peña, *High-throughput 3-D monitoring of agricultural-tree plantations with unmanned aerial vehicle (UAV) technology*, *PLOS ONE* **10** (June, 2015).

[79] C. Zhang and J. M. Kovacs, *The application of small unmanned aerial systems for precision agriculture: a review*, *Precision agriculture* **13** (2012), no. 6.

[80] J. Valente, D. Sanz, A. Barrientos, J. d. Cerro, Á. Ribeiro, and C. Rossi, *An air-ground wireless sensor network for crop monitoring*, *Sensors* **11** (2011), no. 6.

[81] F. López-Granados, J. Torres-Sánchez, A. Serrano-Pérez, A. I. de Castro, F.-J. Mesas-Carrascosa, and J.-M. Peña, *Early season weed mapping in sunflower using uav technology: variability of herbicide treatment maps against weed thresholds*, *Precision Agriculture* (2015) 1–17.

[82] B. P. Fitzpatrick, *Unmanned Aerial Systems for Surveying and Mapping: Cost Comparison of UAS Versus Traditional Methods of Data Acquisition*. PhD thesis, University of Southern California, 2015.

[83] J. M. Teixeira, R. Ferreira, M. Santos, and V. Teichrieb, *Teleoperation using Google glass and AR, drone for structural inspection*, in *Virtual and Augmented Reality (SVR), 2014 XVI Symposium on*, pp. 28–36, IEEE, 2014.

[84] G. Caroti, I. M.-E. Zaragoza, and A. Piemonte, *Accuracy assessment in structure from motion 3D reconstruction from UAV-born images: The influence of the data processing methods*, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **40** (2015), no. 1.

[85] S.-s. Choi and E.-k. Kim, *Building crack inspection using small UAV*, in *Advanced Communication Technology (ICACT), 2015 17th International Conference on*, IEEE, 2015.

[86] J. W. Durban, M. J. Moore, G. Chiang, L. S. Hickmott, A. Bocconcelli, G. Howes, P. A. Bahamonde, W. L. Perryman, and D. J. LeRoi, *Photogrammetry of blue whales with an unmanned hexacopter*, *Marine Mammal Science* (2016).

[87] C. Suduwella, A. Amarasinghe, L. Niroshan, C. Elvitigala, K. De Zoysa, and C. Keppetiyagama, *Identifying mosquito breeding sites via drone images*, in *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, pp. 27–30, ACM, 2017.

[88] A. Lucieer, D. Turner, D. H. King, and S. A. Robinson, *Using an unmanned aerial vehicle (UAV) to capture micro-topography of antarctic moss beds*, *International Journal of Applied Earth Observation and Geoinformation* **27** (2014) 53–62.

[89] R. Schiffman, *Drones flying high as new tool for field biologists*, *Science* **344** (2014), no. 6183 459–459.

[90] J. Scherer, S. Yahyanejad, S. Hayat, E. Yanmaz, T. Andre, A. Khan, V. Vukadinovic, C. Bettstetter, H. Hellwagner, and B. Rinner, *An autonomous multi-UAV system for search and rescue*, in *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ACM, 2015.

[91] P. Royo Chic, E. Pastor Llorens, M. Solé, J. M. Lema Rosas, J. López Rubio, and C. Barrado Muxí, *UAS architecture for forest fire remote sensing*, in *ISPRS Proceedings 2011*, pp. 1–4, 2011.

[92] M. Quaritsch, K. Kruggl, D. Wischounig-Strucl, S. Bhattacharya, M. Shah, and B. Rinner, *Networked UAVs as aerial sensor network for disaster management applications*, *Elektrotechnik und Informationstechnik* **127** (2010), no. 3.

[93] S. M. George, W. Zhou, H. Chenji, M. Won, Y. O. Lee, A. Pazarloglou, R. Stoleru, and P. Barooah, *Distressnet: a wireless ad hoc and sensor network architecture for situation management in disaster response*, *Communications Magazine, IEEE* **48** (2010), no. 3.

[94] K. Daniel, B. Dusza, A. Lewandowski, and C. Wietfeld, *Airshield: A system-of-systems MUAV remote sensing architecture for disaster response*, in *Systems Conference, 2009 3rd Annual IEEE*, IEEE, 2009.

[95] A. Bujari, C. E. Palazzi, and D. Ronzani, *FANET application scenarios and mobility models*, in *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, ACM, 2017.

[96] O. Andryeyev and A. Mitschele-Thiel, *Increasing the cellular network capacity using self-organized aerial base stations*, in *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, ACM, 2017.

[97] Y. Gu, M. Zhou, S. Fu, and Y. Wan, *Airborne WiFi networks through directional antennae: An experimental study*, in *Wireless Communications and Networking Conference (WCNC)*, IEEE, 2015.

[98] T. Simon and A. Mitschele-Thiel, *Next-hop decision-making in mobility-controlled message ferrying networks*, in *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ACM, 2015.

[99] R. Kirichek and V. Kulik, *Long-range data transmission on flying ubiquitous sensor networks (FUSN) by using LPWAN protocols*, in *Distributed Computer and Communication Networks*, Springer International Publishing, 2016.

[100] J. Ueyama, H. Freitas, B. S. Faiçal, P. Geraldo Filho, P. Fini, G. Pessin, P. H. Gomes, and L. A. Villas, *Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks*, *Communications Magazine, IEEE* **52** (2014), no. 12.

[101] Y. Ma, N. Selby, and F. Adib, *Drone relays for battery-free networks*, in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, SIGCOMM '17, (New York, NY, USA), pp. 335–347, ACM, 2017.

[102] A. Trotta, L. Bedogni, M. Di Felice, L. Bononi, and E. Natalizio, *Enhancing TV white-spaces database with unmanned aerial scanning vehicles (UASVs)*, in *Proceedings of the 2nd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ACM, 2016.

[103] C. E. Doyle, J. J. Bird, T. A. Isom, J. C. Kallman, D. F. Bareiss, D. J. Dunlop, R. J. King, J. J. Abbott, and M. A. Minor, *An avian-inspired passive mechanism for quadrotor perching*, *IEEE/ASME Transactions on Mechatronics* **18** (2013), no. 2.

[104] M. A. Estrada, E. W. Hawkes, D. L. Christensen, and M. R. Cutkosky, *Perching and vertical climbing: Design of a multimodal robot*, in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE, 2014.

[105] M. T. Pope, C. W. Kimes, H. Jiang, E. W. Hawkes, M. A. Estrada, C. F. Kerst, W. R. T. Roderick, A. K. Han, D. L. Christensen, and M. R. Cutkosky, *A multimodal robot for perching and climbing on vertical outdoor surfaces*, *Trans. Rob.* **33** (Feb., 2017).

[106] A. Rubina, O. Artemenko, O. Andryeyev, and A. Mitschele-Thiel, *A novel hybrid path planning algorithm for localization in wireless networks*, in *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, ACM, 2017.

[107] N. H. Motlagh, T. Taleb, and O. Arouk, *Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives*, *Internet of Things* **3** (2016), no. 6.

[108] J. Ueyama, H. Freitas, B. S. Faiçal, P. Geraldo Filho, P. Fini, G. Pessin, P. H. Gomes, and L. A. Villas, *Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks*, *IEEE Communications Magazine* **52** (2014), no. 12.

[109] D. Palma, A. Zolich, Y. Jiang, and T. A. Johansen, *Unmanned aerial vehicles as data mules: An experimental assessment*, *IEEE Access* **5** (2017).

[110] M. Erdelj and E. Natalizio, *UAV-assisted disaster management: Applications and open issues*, in *ICNC*, IEEE, 2016.

[111] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, *Help from the sky: Leveraging UAVs for disaster management*, *Pervasive Computing* **16** (2017), no. 1.

[112] S. M. Adams and C. J. Friedland, *A survey of unmanned aerial vehicle (UAV) usage for imagery collection in disaster research and management*, in *Workshop on Remote Sensing for Disaster Response*, vol. 8, 2011.

[113] B. Galkin, J. Kibilda, and L. A. DaSilva, *Coverage analysis for low-altitude UAV networks in urban environments*, in *GLOBECOM*, IEEE, 2017.

[114] J. C. Hodgson, S. M. Baylis, R. Mott, A. Herrod, and R. H. Clarke, *Precision wildlife monitoring using unmanned aerial vehicles*, *Scientific reports* **6** (2016).

[115] M. Rossi, D. Brunelli, A. Adami, L. Lorenzelli, F. Menna, and F. Remondino, *Gas-drone: Portable gas sensing system on UAVs for gas leakage localization*, in *SENSORS*, IEEE, 2014.

[116] C. Gevaert, J. Tang, F. García-Haro, J. Suomalainen, and L. Kooistra, *Combining hyperspectral UAV and multispectral Formosat-2 imagery for precision agriculture applications*, in *WHISPERS*, IEEE, 2014.

[117] V. Puri, A. Nayyar, and L. Raja, *Agriculture drones: A modern breakthrough in precision agriculture*, *Journal of Statistics and Management Systems* **20** (2017), no. 4.

[118] Wi-Fi Alliance, "Wi-Fi Alliance celebrates 20 years of Wi-Fi." `https://www.wi-fi.org/news-events/newsroom/wi-fi-alliance-celebrates-20-years-of-wi-fi`, June, 2019.

[119] Z. N. Chen, X. Qing, T. S. P. See, and W. K. Toh, *Antennas for WiFi connectivity*, *Proceedings of the IEEE* **100** (2012), no. 7.

[120] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, *802.11 with multiple antennas for dummies*, *ACM SIGCOMM Computer Communication Review* **40** (2010), no. 1 19–25.

[121] O. Bejarano, E. W. Knightly, and M. Park, *IEEE 802.11 ac: from channelization to multi-user MIMO*, *IEEE Communications Magazine* **51** (2013), no. 10 84–90.

[122] C.-M. Cheng, P.-H. Hsiao, H. Kung, and D. Vlah, *Performance measurement of 802.11a wireless links from UAV to ground nodes with various antenna orientations*, in *Computer Communications and Networks, 2006. ICCCN 2006. Proceedings. 15th International Conference on*, IEEE, 2006.

[123] E. Yanmaz, R. Kuschnig, and C. Bettstetter, *Channel measurements over 802.11a-based UAV-to-ground links*, in *GLOBECOM*, IEEE, 2011.

[124] E. Yanmaz, R. Kuschnig, and C. Bettstetter, *Achieving air-ground communications in 802.11 networks with three-dimensional aerial mobility*, in *INFOCOM, 2013 Proceedings IEEE*, IEEE, 2013.

[125] J. Allred, A. B. Hasan, S. Panichsakul, W. Pisano, P. Gray, J. Huang, R. Han, D. Lawrence, and K. Mohseni, *Sensorflock: an airborne wireless sensor network of micro-air vehicles*, in *Proceedings of the 5th international conference on Embedded networked sensor systems*, ACM, 2007.

[126] M. Asadpour, D. Giustiniano, and K. A. Hummel, *From ground to aerial communication: dissecting WLAN 802.11n for the drones*, in *Proceedings of the 8th ACM international workshop on Wireless network testbeds, experimental evaluation & characterization*, ACM, 2013.

[127] C. Lima, E. Silva, and P. Velloso, *Performance evaluation of 802.11 IoT devices for data collection in the forest with drones*, in *GLOBECOM*, IEEE, 2018.

[128] S. Hayat, E. Yanmaz, and C. Bettstetter, *Experimental analysis of multipoint-to-point UAV communications with IEEE 802.11 n and 802.11 ac*, in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2015 IEEE 26th Annual International Symposium on*, pp. 1991–1996, IEEE, 2015.

[129] B. Van den Bergh, T. Vermeulen, and S. Pollin, *Analysis of harmful interference to and from aerial IEEE 802.11 systems*, in *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ACM, 2015.

[130] S. Munari, C. E. Palazzi, G. Quadrio, and D. Ronzani, *Network traffic analysis of a small quadcopter*, in *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, DroNet '17, (New York, NY, USA), ACM, 2017.

[131] K. L. Lueth, "State of the IoT 2018: Number of IoT devices now at 7 billion." `https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/`, Aug, 2018.

[132] I. Howitt and J. A. Gutierrez, *IEEE 802.15.4 low rate - wireless personal area network coexistence issues*, in *IEEE Wireless Communications and Networking (WCNC)*, vol. 3, Mar, 2003.

[133] Y. Qin, D. Boyle, and E. Yeatman, *Efficient and reliable aerial communication with wireless sensors*, *IEEE Internet of Things Journal* **6** (Oct, 2019).

[134] B. DeRenzi, Y. Anokwa, T. Parikh, and G. Borriello, *Reliable data collection in highly disconnected environments using mobile phones*, in *Proceedings of the 2007 workshop on Networked systems for developing regions*, ACM, 2007.

[135] S. Hara, D. Zhao, K. Yanagihara, J. Taketsugu, K. Fukui, S. Fukunaga, and K.-i. Kitayama, *Propagation characteristics of IEEE 802.15.4 radio signal and their application for location estimation*, in *Vehicular Technology Conference*, vol. 1, IEEE, 2005.

[136] E. Miluzzo, X. Zheng, K. Fodor, and A. T. Campbell, *Radio characterization of 802.15.4 and its impact on the design of mobile sensor networks*, in *EWSN*, Springer, 2008.

[137] M. Khanafer, M. Kandil, R. Al-Baghdadi, A. Al-Ajmi, and H. T. Mouftah, *Enhancements to IEEE 802.15.4 MAC protocol to support vehicle-to-roadside communications in VANETs*, in *IEEE International Conference on Communications (ICC)*, May, 2019.

[138] M. Petrova, J. Riihijarvi, P. Mahonen, and S. Labella, *Performance study of IEEE 802.15.4 using measurements and simulations*, in *WCNC*, vol. 1, IEEE, 2006.

[139] D. Lymberopoulos, Q. Lindsey, and A. Savvides, *An empirical characterization of radio signal strength variability in 3-D IEEE 802.15.4 networks using monopole antennas*, in *EWSN*, Springer, 2006.

[140] GSMA, "The mobile economy report." `https://www.gsma.com/mobileeconomy/wp-content/uploads/2018/05/The-Mobile-Economy-2018.pdf`, 2018.

[141] K. Salemink, D. Strijker, and G. Bosworth, *Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas*, *Journal of Rural Studies* **54** (2017).

[142] FCC, "Broadband Deployment Report." `https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2018-broadband-deployment-report`, February, 2018.

[143] FCC, "Connect America Fund (CAF)." `https://www.fcc.gov/general/connect-america-fund-caf`, February, 2017.

[144] J. E. Prieger, "Mobile data roaming and incentives for investment in rural broadband infrastructure." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3391478, 2017.

[145] FCC, "FCC Mobile Broadband Dataset." https://www.fcc.gov/form-477-mobile-voice-and-broadband-coverage-areas, 2019.

[146] J. Kahan, "It's time for a new approach for mapping broadband data to better serve Americans." https://blogs.microsoft.com/on-the-issues/2019/04/08/its-time-for-a-new-approach-for-mapping-broadband-data-to-better-serve-americans/, Apr, 2019.

[147] Rural Wireless Association., "RWA Calls for FCC Investigation of T-Mobile Coverage Data." https://ruralwireless.org/rwa-calls-for-fcc-investigation-of-t-mobile-coverage-data, 2018.

[148] S. Meinrath, H. Bonestroo, G. Bullen, A. Jansen, S. Mansour, C. Mitchell, C. Ritzo, and N. Thieme, *Broadband availability and access in rural Pennsylvania*, .

[149] Measurement Lab, "Open Internet Measurement." https://www.measurementlab.net.

[150] FCC, "Understanding wireless telephone coverage." https://fcc.gov/consumers/guides/understanding-wireless-telephone-coverage-areas.

[151] CellMapper. https://www.cellmapper.net.

[152] OpenSignal. https://www.opensignal.com/.

[153] C. Hurley, R. Rogers, F. Thornton, and B. Baker, *Wardriving and Wireless Penetration Testing*. Syngress, 2007.

[154] M. C. Batistatos, G. E. Athanasiadou, D. A. Zarbouti, G. V. Tsoulos, and N. C. Sagias, *LTE ground-to-air measurements for UAV-assisted cellular networks*, in *12th European Conference on Antennas and Propagation (EuCAP)*, IET, 2018.

[155] G. E. Athanasiadou, M. C. Batistatos, D. A. Zarbouti, and G. V. Tsoulos, *LTE ground-to-air field measurements in the context of flying relays*, *Wireless Communications* **26** (2019), no. 1.

[156] C. Desmond, *Verizon uses drones during disasters like Hurricane Harvey*, *Robotics Tomorrow* (Nov, 2017).

[157] Osmocom, "RTL-SDR." https://osmocom.org/projects/rtl-sdr/wiki/Rtl-sdr.

[158] F. Minucci, S. Rajendran, B. V. d. Bergh, S. Pollin, D. Giustiniano, H. Cordobés, R. C.-P. Fuchs, V. Lenders, *et. al.*, *Electrosense-spectrum sensing with increased frequency range*, in *Proceedings of the 2018 International Conference on Embedded Wireless Systems and Networks*, Junction Publishing, 2018.

[159] J. A. del Peral-Rosado, J. M. Parro-Jiménez, J. A. López-Salcedo, G. Seco-Granados, P. Crosta, F. Zanier, and M. Crisci, *Comparative results analysis on positioning with real LTE signals and low-cost hardware platforms*, in *7th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)*, IEEE, 2014.

[160] X. Lin, R. Wiren, S. Euler, A. Sadam, H.-L. Maattanen, S. D. Muruganathan, S. Gao, Y.-P. E. Wang, J. Kauppi, Z. Zou, and V. Yajnanarayana, *Mobile networks connected drones: Field trials, simulations, and design insights*, *arXiv* (2018).

[161] P. Schmitt, D. Iland, M. Zheleva, and E. Belding, *HybridCell: Cellular connectivity on the fringes with demand-driven local cells*, in *IEEE INFOCOM*, 2016.

[162] *Houston emergency officials tell 911 callers not to hang up*, *CBS News* (Aug, 2017).

[163] *Earthquake calls flooded 911 dispatch center*, *Napa Valley Register* (September, 2014).

[164] N. Ungerleider, *Why your phone doesn't work during disasters - and how to fix it*, *Fast Company* (Apr, 2013).

[165] C. Shu, *Coast Guard asks people stranded by Harvey to call them instead of posting on social media for help*, *TechCrunch* (Aug, 2017).

[166] FCC, "Communications Status Report for Areas Impacted by Tropical Storm Harvey, September 1, 2017." `http://transition.fcc.gov/Daily_Releases/Daily_Business/2017/db0901/DOC-346475A1.pdf`, September, 2017.

[167] A. Saeed, K. A. Harras, E. Zegura, and M. Ammar, *Local and low-cost white space detection*, in *37th International Conference on Distributed Computing Systems (ICDCS)*, June, 2017.

[168] O. Genc, M. Bayrak, and E. Yaldiz, *Analysis of the effects of GSM bands to the electromagnetic pollution in the RF spectrum*, *Progress In Electromagnetics Research* **101** (2010).

[169] S.-K. Noh and D. Choi, *Propagation model in indoor and outdoor for the LTE communications*, *International Journal of Antennas and Propagation* (2019).

[170] O. Simpson and Y. Sun, *LTE RSRP, RSRQ, RSSNR and local topography profile data for RF propagation planning and network optimization in an urban propagation environment*, *Data in brief* **21** (2018).

[171] C. Lin, K. Sandrasegaran, H. A. M. Ramli, and R. Basukala, *Optimized performance evaluation of LTE hard handover algorithm with average RSRP constraint*, *CoRR* **abs/1105.0234** (2011).

[172] K. Dimou, M. Wang, Y. Yang, M. Kazmi, A. Larmo, J. Pettersson, W. Muller, and Y. Timner, *Handover within 3GPP LTE: Design principles and performance*, in *70th Vehicular Tech. Conf.*, IEEE, 2009.

[173] P. Legg, G. Hui, and J. Johansson, *A simulation study of LTE intra-frequency handover performance*, in *72nd Vehicular Tech. Conf.*, IEEE, 2010.

[174] D. Aziz and R. Sigle, *Improvement of LTE handover performance through interference coordination*, in *69th Vehicular Tech. Conf.*, IEEE, 2009.

[175] J. Kurjenniemi, T. Henttonen, and J. Kaikkonen, *Suitability of RSRQ measurement for quality based inter-frequency handover in LTE*, in *Intern. Symposium on Wireless Communication Sys.*, IEEE, 2008.

[176] M. Anas, F. D. Calabrese, P. E. Mogensen, C. Rosa, and K. I. Pedersen, *Performance evaluation of received signal strength based hard handover for UTRAN LTE*, in *65th Vehicular Tech. Conf.*, IEEE, 2007.

[177] J. Kurjenniemi and T. Henttonen, *Effect of measurement bandwidth to the accuracy of inter-frequency RSRP measurements in LTE*, in *19th International Symposium on Personal, Indoor and Mobile Radio Communications*, IEEE, 2008.

[178] F. Afroz, R. Subramanian, R. Heidary, K. Sandrasegaran, and S. Ahmed, *SINR, RSRP, RSSI and RSRQ measurements in long term evolution networks*, *International Journal of Wireless & Mobile Networks* (2015).

[179] C. Ide, B. Dusza, and C. Wietfeld, *Performance of channel-aware M2M communications based on LTE network measurements*, in *24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, IEEE, 2013.

[180] C. Ide, R. Falkenberg, D. Kaulbars, and C. Wietfeld, *Empirical analysis of the impact of LTE downlink channel indicators on the uplink connectivity*, in *83rd Vehicular Tech. Conf.*, IEEE, 2016.

[181] C. S. Park and S. Park, *Analysis of RSRP measurement accuracy*, *Communications Letters* **20** (2016), no. 3.

[182] H. C. Nguyen, R. Amorim, J. Wigard, I. Z. Kovacs, and P. Mogensen, *Using LTE networks for UAV command and control link: A rural-area coverage analysis*, in *86th Vehicular Tech. Conf.*, IEEE, 2017.

[183] R. Amorim, H. Nguyen, P. Mogensen, I. Z. Kovács, J. Wigard, and T. B. Sørensen, *Radio channel modeling for UAV communication over cellular networks*, *Wireless Communications Letters* **6** (2017), no. 4.

[184] T. Engel, "Xgoldmon." `https://github.com/2b-as/xgoldmon`.

[185] B. Hong, S. Park, H. Kim, D. Kim, H. Hong, H. Choi, J.-P. Seifert, S.-J. Lee, and Y. Kim, *Peeking over the cellular walled gardens-a method for closed network diagnosis*, *IEEE Transactions on Mobile Computing* **17** (2018), no. 10.

[186] Qualcomm, "QXDM." `https://www.qualcomm.com/documents/qxdm-professional-qualcomm-extensible-diagnostic-monitor`, 2012.

[187] D. Spaar, "Tracing LTE on the phone." `http://www.mirider.com/weblog/2013/08/index.html`, 2014.

[188] SRLabs, "SnoopSNitch." `https://opensource.srlabs.de/projects/snoopsnitch`, 2014.

[189] Y. Li, C. Peng, Z. Yuan, J. Li, H. Deng, and T. Wang, *Mobileinsight: Extracting and analyzing cellular network information on smartphones*, in *ACM MobiCom*, 2016.

[190] P. Schmitt, D. Iland, and E. Belding, *SmartCell: Small-scale mobile congestion awareness*, *IEEE Communications Magazine* **54** (2016), no. 7 44–50.

[191] M. Tavana, A. Rahmati, and V. Shah-Mansouri, *Congestion control with adaptive access class barring for LTE M2M overload using kalman filters*, *Computer Networks* **141** (2018) 222–233.

[192] S. Duan, V. Shah-Mansouri, Z. Wang, and V. W. Wong, *D-ACB: Adaptive congestion control algorithm for bursty M2M traffic in LTE networks*, *IEEE Transactions on Vehicular Technology* **65** (2016), no. 12.

[193] A. K. Paul, H. Kawakami, A. Tachibana, and T. Hasegawa, *An AQM based Congestion Control for eNB RLC in 4G/LTE Network*, in *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2016.

[194] R. Kwan, R. Arnott, R. Trivisonno, and M. Kubota, *On pre-emption and congestion control for LTE systems*, in *IEEE 72nd Vehicular Technology Conference-Fall*, 2010.

[195] P. Torres, P. Marques, H. Marques, R. Dionísio, T. Alves, L. Pereira, and J. Ribeiro, *Data analytics for forecasting cell congestion on LTE networks*, in *2017 Network Traffic Measurement and Analysis Conference (TMA)*, IEEE, 2017.

[196] A. Chakraborty, V. Navda, V. N. Padmanabhan, and R. Ramjee, *Coordinating cellular background transfers using loadsense*, in *Proceedings of the 19th annual international conference on Mobile computing & networking*, pp. 63–74, ACM, 2013.

[197] X. Xie, X. Zhang, S. Kumar, and L. E. Li, *pistream: Physical layer informed adaptive video streaming over LTE*, in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pp. 413–425, ACM, 2015.

[198] X. Xie, X. Zhang, and S. Zhu, *Accelerating mobile web loading using cellular link information*, in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, ACM, 2017.

[199] USC - ANT Lab, "Evaluation of Hurricane Harvey's Effects on the Internet's Edge." `https://ant.isi.edu/outage/ani/harvey/index.html`, Sep, 2017.

[200] P. Rudol and P. Doherty, *Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery*, in *Aerospace Conference, 2008 IEEE*, IEEE, 2008.

[201] A. Birk, B. Wiggerich, H. Bülow, M. Pfingsthorn, and S. Schwertfeger, *Safety, security, and rescue missions with an unmanned aerial vehicle (UAV)*, *Journal of Intelligent & Robotic Systems* **64** (2011), no. 1.

[202] S. Waharte and N. Trigoni, *Supporting search and rescue operations with UAVs*, in *Emerging Security Technologies (EST), 2010 International Conference on*, IEEE, 2010.

[203] F. Gustafsson and F. Gunnarsson, *Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements*, *IEEE Signal processing magazine* **22** (2005), no. 4.

[204] Y. Zhou and F. Wong, *Relative localization for small wireless sensor networks*, in *Ad Hoc Networks*. Springer, 2017.

[205] N. Patwari, R. J. O'Dea, and Y. Wang, *Relative location in wireless networks*, in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 2, IEEE, 2001.

[206] J. Xu, J. He, Y. Zhang, F. Xu, and F. Cai, *A distance-based maximum likelihood estimation method for sensor localization in wireless sensor networks*, *International Journal of Distributed Sensor Networks* **12** (2016), no. 4.

[207] R. Qiang, W. Wang, H. Wang, P. He, and W. Huang, *3D maximum likelihood estimation positioning algorithm based on RSSI ranging*, in *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Mar, 2017.

[208] X. Li, Z. D. Deng, Y. Sun, J. J. Martinez, T. Fu, G. A. McMichael, and T. J. Carlson, *A 3D approximate maximum likelihood solver for localization of fish implanted with acoustic transmitters*, *Scientific reports* **4** (2014).

[209] G. Shen, R. Zetik, and R. Thoma, *Performance comparison of TOA and TDOA based location estimation algorithms in los environment*, in *Positioning, Navigation and Communication, 2008. WPNC 2008. 5th Workshop on*, IEEE, 2008.

[210] X. Ma, Q.-Y. Huang, and X.-M. Shu, *A new localization algorithm of mobile phone for outdoor emergency rescue*, in *Intelligent System Design and Engineering Applications (ISDEA), 2013 Third International Conference on*, IEEE, 2013.

[211] J. Sundqvist, J. Ekskog, B. J. Dil, F. Gustafsson, J. Tordenlid, and M. Petterstedt, *Feasibility study on smartphone localization using mobile anchors in search and rescue operations*, in *Information Fusion (FUSION), 2016 19th International Conference on*, IEEE, 2016.

[212] M. Carpin, S. Rosati, M. E. Khan, and B. Rimoldi, *UAVs using bayesian optimization to locate WiFi devices*, *arXiv preprint arXiv:1510.03592* (2015).

[213] Southern California Tribal Chairmen's Association, "Southern California Tribal Digital Village." `https://sctca.net/southern-california-tribal-digital-village/`.

[214] School for Scientific Thought , "School for scientific thought - courses." `https://sst-csep.cnsi.ucsb.edu/content/courses`, Jan, 2020.

[215] B. Lubek, "Network Monitor." `https://github.com/caarmen/network-monitor`.

[216] Taoglas, *Apex III Wideband 4G LTE Dipole Terminal Antenna*, 2019.

[217] srsLTE. `https://github.com/srsLTE/srsLTE`.

[218] SignalBooster, "What is strong signal in dBm for 4G?." `https://www.signalbooster.com/blogs/news/differences-between-3g-1x-vs-4g-lte-signal-strength-in-dbm`, Apr, 2016.

[219] 3GPP TR 36.802, *Evolved Universal Terrestrial Radio Access (E-UTRA); NB-IOT; Technical Report for BS and UE radio transmission and reception*, February, 2016.

[220] 3GPP TS 36.508, *Evolved universal terrestrial radio access (E-UTRA) and evolved packet core (EPC); common test environments for user equipment (UE) conformance testing*, Apr, 2017.

[221] 3GPP TS 36.331, *Evolved universal terrestrial radio access (E-UTRA); radio resource control (RRC); protocol specification*, January, 2016.

[222] 3GPP TS 25.331, *Radio resource control (RRC); protocol specification*, October, 2014.

[223] 3GPP TS 24.301, *Non-access-stratum (NAS) protocol for evolved packet system (EPS)*, June, 2011.

[224] 3GPP TS 36.211, *Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation*, June, 2016.

[225] 3GPP TS 36.212, *Evolved universal terrestrial radio access (E-UTRA); multiplexing and channel coding*, April, 2017.

[226] 3GPP TS 36.304, *Evolved universal terrestrial radio access (E-UTRA); user equipment (UE) procedures in idle mode*, January, 2012.

[227] 3GPP TS 36.300, *Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description; stage 2*, January, 201.

[228] S.-H. Oh and Y.-H. Kim, *Policy-based congestion control in WCDMA wireless access networks for end-to-end QoS*, in *COIN-NGNCON 2006-The Joint International Conference on Optical Internet and Next Generation Network*, pp. 153–155, IEEE, 2006.

[229] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*. Academic press, 2010.

[230] H. Holma and A. Toskala, *WCDMA for UMTS: HSPA evolution and LTE*. John Wiley & sons, 2007.

[231] Ettus Research, "USRP B210." `http://www.ettus.com/all-products/UB210-KIT/`.

[232] MP Antenna, "SUPER-M ULTRA Mobile Antenna (25MHz–6GHz)."

[233] "Text2pcap - Generate a capture file from an ASCII hexdump of packets." `https://www.wireshark.org/docs/man-pages/text2pcap.html`, 2014.

[234] Wireshark, "Wireshark." `https://www.wireshark.org`.

[235] G. T. 23.401, *General packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access*, Sep, 2014.

[236] J. E. V. Bautista, S. Sawhney, M. Shukair, I. Singh, V. K. Govindaraju, and S. Sarkar, *Performance of CS fallback from LTE to UMTS*, *IEEE Communications Magazine* **51** (2013), no. 9 136–143.

[237] G. T. 23.272, *Circuit switched (CS) fallback in evolved packet system (EPS); stage 2 (release 10)*, Mar, 2012.

[238] SanDiego.org, "St. Patricks Day Parade." `http://www.stpatsparade.org/parade-festival-schedule.html`, 2019.

[239] ShamROCK San Diego, "ShamROCK Concert." `https://www.sandiegoshamrock.com/`, 2019.

[240] T. Group, "Thread specification." `https://www.threadgroup.org/`.

[241] J. Zheng and M. J. Lee, *Will IEEE 802.15.4 make ubiquitous networking a reality?: a discussion on a potential low power, low bit rate standard*, *IEEE Communications magazine* **42** (2004), no. 6.

[242] S. M. George, W. Zhou, H. Chenji, M. Won, Y. O. Lee, A. Pazarloglou, R. Stoleru, and P. Barooah, *DistressNet: a wireless ad hoc and sensor network architecture for situation management in disaster response*, *IEEE Communications Magazine* **48** (2010), no. 3.

[243] IEEE Standards Association and others, *IEEE standard for low-rate wireless personal area networks (WPANs)*, *IEEE Computers Society* (2015).

[244] Digi, "XBee3 Zigbee 3.0 datasheet." `https://www.digi.com/pdf/ds_xbee-3-zigbee-3.pdf`, 2019.

[245] M. Nekrasov, R. Allen, and E. Belding, "Aerial Measurements from Outdoor 2.4GHz 802.15.4 Network." `https://doi.org/10.25349/D9KS3W`.

[246] G. Judd, X. Wang, and P. Steenkiste, *Efficient channel-aware rate adaptation in dynamic environments*, in *Proceedings of the 6th international conference on Mobile systems, applications, and services*, ACM, 2008.

[247] Chang Hong Technology Co., "2.4G Dipole 2dBi Antenna." `https://www.sparkfun.com/datasheets/Wireless/Antenna/DA-24-04.pdf`, Feb, 2007.

[248] S. Basagni, F. Ceccarelli, C. Petrioli, N. Raman, and A. V. Sheshashayee, *Wake-up radio ranges: A performance study*, in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2019.

[249] C. Phillips, D. Sicker, and D. Grunwald, *A survey of wireless path loss prediction and coverage mapping methods*, *IEEE Communications Surveys & Tutorials* **15** (2013), no. 1.

[250] H. T. Friis, *A note on a simple transmission formula*, *Proceedings of the IRE* **34** (1946), no. 5.

[251] R. Gummadi, D. Wetherall, B. Greenstein, and S. Seshan, *Understanding and mitigating the impact of RF interference on 802.11 networks*, *ACM SIGCOMM Computer Communication Review* **37** (2007), no. 4.

[252] G. Gaertner and V. Cahill, *Understanding link quality in 802.11 mobile ad hoc networks*, *IEEE Internet Computing* **8** (2004), no. 1.

[253] G. Li, S. Zhang, W. Wei, and B. Yang, *An iterative multilateral localization algorithm based on time rounds for wireless sensor networks*, in *Networks Security, Wireless Communications and Trusted Computing, 2009. NSWCTC'09. International Conference on*, vol. 1, IEEE, 2009.

[254] H. C. So and L. Lin, *Linear least squares approach for accurate received signal strength based source localization*, *IEEE Transactions on Signal Processing* **59** (2011), no. 8.

[255] P. Tarrio, A. M. Bernardos, J. A. Besada, and J. R. Casar, *A new positioning technique for RSS-based localization based on a weighted least squares estimator*, in *Wireless Communication Systems. 2008. ISWCS'08. IEEE International Symposium on*, IEEE, 2008.

[256] Z. G. Jun, L. Xin, X. Z. Long, and L. H. Chao, *Weighted least square localization algorithm based on RSSI values*, in *Instrumentation and Measurement, Computer, Communication and Control (IMCCC), 2015 Fifth International Conference on*, IEEE, 2015.

[257] Y.-T. Chan, H. Y. C. Hang, and P.-c. Ching, *Exact and approximate maximum likelihood localization algorithms*, *IEEE Transactions on Vehicular Technology* **55** (2006), no. 1.

[258] DJI, "Matrice 100." `https://www.dji.com/matrice100/info`.

[259] SciPy, "SciPy: Least-Squares." `https://docs.scipy.org/doc/scipy/reference/generated/scipy.linalg.lstsq.html`.

[260] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, *Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization*, *ACM Transactions on Mathematical Software (TOMS)* **23** (1997), no. 4.

[261] C. Goerzen, Z. Kong, and B. Mettler, *A survey of motion planning algorithms from the perspective of autonomous UAV guidance*, *Journal of Intelligent and Robotic Systems* **57** (2010).

[262] "Calculating the height of a tall building where only the number of stories is known." `http://www.ctbuh.org/HighRiseInfo/TallestDatabase/Criteria/HeightCalculator/tabid/1007/language/en-GB/Default.aspx`, 2018.

[263] Facebook, "Two billion people coming together on Facebook." https://newsroom.fb.com/news/2017/06/two-billion-people-coming-together-on-facebook/, June, 2017.

[264] Internet Live Stats, "Twitter usage statistics." http://www.internetlivestats.com/twitter-statistics/.

[265] Internet Society, "Global internet report 2019: Consolidation in the internet economy." https://future.internetsociety.org/2019/wp-content/uploads/sites/2/2019/04/InternetSociety-GlobalInternetReport-ConsolidationintheInternetEconomy.pdf, 2019.

[266] United Nations, "Universal declaration of human rights." http://www.un.org/en/universal-declaration-human-rights/, December, 1948.

[267] United Nations, "Resolution 32/13: The promotion, protection and enjoyment of human rights on the internet." https://http://undocs.org/A/HRC/RES/32/13, July, 2016.

[268] A. Peterson, *Turkey strengthens Twitter ban, institutes IP level block*, *The Washington Post* (Mar, 2014).

[269] *New internet shutdown in Turkey's Southeast: 8% of country now offline amidst Diyarbakir unrest*, *Turkey Blocks* (Oct, 2016).

[270] A. Dainotti, C. Squarcella, E. Aben, K. C. Claffy, M. Chiesa, M. Russo, and A. Pescapé, *Analysis of country-wide internet outages caused by censorship*, in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, 2011.

[271] Y. Breindl and J. Wright, *Internet filtering in liberal democracies*, in *Proceedings of the 2nd USENIX Workshop on Free and Open Communications on the Internet*, (Bellevue, WA), USENIX, Aug, 2012.

[272] D. Goldman, *Donald Trump wants to 'close up' the internet*, *CNN* (Dec, 2015).

[273] C. Riley, *Theresa May: Internet must be regulated to prevent terrorism*, *CNN* (Jun, 2017).

[274] P. Norris, *It's not just Trump. authoritarian populism is rising across the West. here's why.*, *The Washington Post* (Mar, 2016).

[275] H. R. Clinton, *Remarks on internet freedom*, *The Newseum* **21** (2010).

[276] B. Fung, *The FCC's net neutrality rules are officially repealed today. here's what that really means.*, *Washington Post* (June, 2018).

[277] *Freedom in the world 2017*, *Freedom House* (2017).

[278] G. Kasparov and T. Halvorssen, *Why the rise of authoritarianism is a global catastrophe*, *The Washington Post* (Feb, 2017).

[279] R. Williams, *The rise of authoritarianism*, *Psychology Today* (Mar, 2016).

[280] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, *et. al.*, *The science of fake news*, *Science* **359** (2018).

[281] H. Allcott and M. Gentzkow, *Social media and fake news in the 2016 election*, *Journal of economic perspectives* **31** (2017), no. 2.

[282] S. Timur and T. Arango, *Turkey seizes newspaper, Zaman, as press crackdown continues*, *The New York Times* (Mar, 2016).

[283] A. Taylor, *This single tweet got a Turkish journalist detained*, *The Washington Post* (December, 2014).

[284] *List of banned words on its websites and comments*, *Shuum.mn* (Mar, 2013).

[285] FIDH, *Turkey: Provisional release of human rights lawyer Mr. Levent Piskin*, *FIDH* (Nov, 2016).

[286] *Facebook, Twitter, YouTube and WhatsApp shutdown in Turkey*, *Turkey Blocks* (Nov, 2016).

[287] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau, *Tweeting from left to right: Is online political communication more than an echo chamber?*, *Psychological science* **26** (2015), no. 10.

[288] M. Lowen, *Is Gollum good or evil? Jail term in Turkey hinges on answer*, *BBC* (Apr, 2015).

[289] *Whoever criticizes Erdogan finds themselves in court; Here are the court cases!*, *LGBTI News Turkey* (May, 2016).

[290] K. Rogers, *The problem with insulting Turkey's President Erdogan*, *New York Times* (Dec, 2016).

[291] *Tor blocked in Turkey as government cracks down on VPN use*, *Turkey Blocks* (Dec, 2016).

[292] J. Zittrain and B. Edelman, *Internet filtering in China*, *IEEE Internet Computing* **7** (Mar, 2003).

[293] *Mongolia: Freedom of the press 2016*, *Freedom House* (2016).

[294] A. Dabrowski, N. Pianta, T. Klepp, M. Mulazzani, and E. Weippl, *IMSI-catch Me if You Can: IMSI-catcher-catchers*, in *Proceedings of the 30th Annual Computer Security Applications Conference*, (New York, NY, USA), ACM, 2014.

[295] M. Pizzi, *Isolated in camp, Syrians desperate to get online*, *Al Jazeera America* (July, 2015).

[296] M. Richtel, *Egypt cuts off most internet and cellphone service*, .

[297] M. Chulov, *Syria shuts off internet access across the country*, *The Guardian* (Nov, 2012).

[298] J. Cowie, A. Popescu, and T. Underwood, *Impact of hurricane Katrina on internet infrastructure*, *Report, Renesys* (2005).

[299] "Free basics by Facebook." `https://info.internet.org/en/story/free-basics-from-internet-org/`.

[300] *Wina justifies beating of Komboni radio owner*, *Tumfweko* (Oct, 2016).

[301] L. Parks and R. Mukherjee, *From platform jumping to self-censorship: internet freedom, social media, and circumvention practices in Zambia*, *Communication and Critical/Cultural Studies* (2017).

[302] P. K. Smith, J. Mahdavi, M. Carvalho, S. Fisher, S. Russell, and N. Tippett, *Cyberbullying: Its nature and impact in secondary school pupils*, *Journal of child psychology and psychiatry* **49** (2008), no. 4.

[303] S. Hinduja, "Doxing and cyberbullying." `http://cyberbullying.org/doxing-and-cyberbullying`, September, 2015.

[304] M. R. Hebl, J. B. Foster, L. M. Mannix, and J. F. Dovidio, *Formal and interpersonal discrimination: A field study of bias toward homosexual applicants*, *Personality and social psychology bulletin* **28** (2002), no. 6 815–825.

[305] J. Greenberg, *For Netflix, discontent over blocked VPNs is boiling*, *WIRED* (Mar, 2016).

[306] Tor, "Tor browser." `https://www.torproject.org/projects/torbrowser.html.en`.

[307] S. J. Murdoch and G. Danezis, *Low-cost traffic analysis of tor*, in *Security and Privacy, 2005 IEEE Symposium on*, pp. 183–195, IEEE, 2005.

[308] "What names are allowed on Facebook?." `https://www.facebook.com/help/112146705538576`.

[309] E. Galperin, "Changes to Facebook's "real names" policy still don't fix the problem." https://www.eff.org/deeplinks/2015/12/changes-facebooks-real-names-policy-still-dont-fix-problem, Dec, 2015.

[310] Facebook, "Government requests report." https://govtrequests.facebook.com/.

[311] M. Akyol, *Another Turkish witch hunt begins*, *US News* (Dec, 2014).

[312] A. Hess, *On Twitter, a battle among political bots*, *The New York Times* (Dec, 2016).

[313] C. Miller, *Bots will set the political agenda in 2017*, *Wired* (Jan, 2017).

[314] S. Walker, *Salutin' Putin: inside a Russian troll house*, *The Guardian* (Apr, 2015).

[315] V. Shevchenko, *Ukrainians petition Facebook against 'Russian trolls'*, *BBC* (May, 2015).

[316] Electronic Frontier Foundation, "Privacy Badger." https://www.eff.org/privacybadger.

[317] Confide, "Confide." https://getconfide.com/.

[318] Guardian Project, "Orbot: Tor for Android." https://guardianproject.info/apps/orbot/.

[319] Electronic Frontier Foundation, "HTTPS everywhere." https://www.eff.org/https-everywhere.

[320] Ethnologue, "Summary by language size." https://www.ethnologue.com/statistics/size.

[321] D. Casale and D. Posel, *English language proficiency and earnings in a developing country: The case of South Africa*, *The Journal of Socio-Economics* **40** (2011), no. 4.

[322] International Telecommunication Union, "ICT Facts and Figures 2016." https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2016.pdf, May, 2016.

[323] Zambia Information and Communications Technology Authority, "ICT survey report - households and individuals." https://www.zicta.zm/Views/Publications/2015ICTSURVEYREPORT.pdf, 2015.

[324] V. Pejovic, D. L. Johnson, M. Zheleva, E. Belding, L. Parks, and G. van Stam, *The bandwidth divide: Obstacles to efficient broadband adoption in rural Sub-Saharan Africa*, *International Journal of Communication* **6** (2012).

[325] J. Zittrain and B. Edelman, *Internet filtering in China*, *IEEE Internet Computing* **7** (Mar, 2003).

[326] G. King, J. Pan, and M. E. Roberts, *How censorship in China allows government criticism but silences collective expression*, *American Political Science Review* **107** (2013), no. 2.

[327] J. R. Crandall, D. Zinn, M. Byrd, E. T. Barr, and R. East, *Conceptdoppler: a weather tracker for internet censorship.*, in *ACM Conference on Computer and Communications Security*, 2007.

[328] P. Winter and S. Lindskog, *How the great firewall of China is blocking Tor*, in *Proceedings of the 2nd USENIX Workshop on Free and Open Communications on the Internet*, USENIX, Aug, 2012.

[329] Y. Kou, B. Semaan, and B. Nardi, *A confucian look at internet censorship in china*, in *Human-Computer Interaction - INTERACT 2017*, (Cham), pp. 377–398, Springer International Publishing, 2017.

[330] S. Aryan, H. Aryan, and J. A. Halderman, *Internet censorship in Iran: A first look*, in *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet*, Aug, 2013.

[331] A. Chaabane, T. Chen, M. Cunche, E. De Cristofaro, A. Friedman, and M. A. Kaafar, *Censorship in the wild: Analyzing internet filtering in Syria*, in *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, 2014.

[332] Z. Nabi, *The anatomy of web censorship in Pakistan*, in *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet*, Aug, 2013.

[333] Freedom House, *Freedom in the world 2016 table of country scores*, .

[334] Internet Monitor, "Zambia." `http://thenetmonitor.org/countries/zmb/access`.

[335] High Commissioner of the Republic of Zambia , "Demography." `http://www.zambiapretoria.net/demography/`.

[336] OECD Better Life Index, "Turkey." `http://www.oecdbetterlifeindex.org/countries/turkey/`.

[337] B. Chilkhaasuren and B. Baasankhuu, "Population and economic activities of Ulaanbaatar." `https://www.ubstat.mn/upload/reports/ub_khotiin_khun_am_ediin_zasag_angli_ulaanbaatar_2012-08.pdf`, 2012.

[338] A. Strauss and J. Corbin, *Grounded theory methodology*, *Handbook of qualitative research* **17** (1994).

[339] SecurePost, "Securepost - safe, secure, social media." https://securepost.co.

[340] International Telecommunication Union, "2017 estimates for key ICT indicators." http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2017/ITU_Key_2005-2017_ICT_data.xls, 2017.

[341] International Telecommunication Union, "ICT Facts and Figures 2017." https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf, May, 2017.

[342] Alexa, "Top 500 global sites." https://www.alexa.com/topsites.

[343] Statista, *Global social media ranking 2017*, *Statista* (October, 2017).

[344] StatCounter, "Mobile operating system market share Zambia." http://gs.statcounter.com/os-market-share/mobile/zambia/#monthly-201302-201709-bar, Sep, 2017.

[345] StatCounter, "Mobile operating system market share Turkey." http://gs.statcounter.com/os-market-share/mobile/turkey, Sep, 2017.

[346] Statista, "Mobile OS market share 2017." https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/, May, 2017.

[347] A. Bhattacharya, *Android (goog) just hit a record 88% market share of all smartphones*, *Quartz* (Nov, 2016).

[348] Freedom House, *Turkey - country report - freedom of the press - 2014*, .

[349] C. Letsch, *Turkish composer and pianist convicted of blasphemy on Twitter*, *The Guardian* (Apr, 2013).

[350] I. Akwei, "Zambian opposition leader arrested over 'libelous' Facebook post." http://www.africanews.com/2017/04/14/zambian-opposition-leader-arrested-for-libelous-facebook-post//, Apr, 2017.

[351] *Zambia : Police arrest engineering student for 'insulting' President Lungu on Facebook*, *Lusaka Times* (Jul, 2017).

[352] Freedom House, *Freedom of the net - 2016*, .

[353] J. Sturcke, *Libel laws explained*, *The Guardian* (Aug, 2006).

[354] R. J. Krotoszynski Jr, *Defamation in the Digital Age: Some Comparative Law Observations on the Difficulty of Reconciling Free Speech and Reputation in the Emerging Global Village*, *Washington and Lee Law Review* **62** (2005), no. 1.

[355] US Department of State, "Mongolia country reports on human rights practices." http://www.state.gov/j/drl/rls/hrrpt/humanrightsreport/, 2016.

[356] L. Gardner, "Mongolia's media laws threaten press freedom." http://mediashift.org/2014/04/mongolias-media-laws-threaten-press-freedom/, Apr, 2014.

[357] Freedom House, *Mongolia - country report - freedom of the press*, .

[358] Twitter, "Information requests." https://transparency.twitter.com/en/information-requests.html.

[359] S. Hinduja, *Doxing and cyberbullying*, *Cyberbullying Research Center* (September, 2015).

[360] E. G. Ellis, *Doxing is a perilous form of justice—even when it's outing Nazis*, *WIRED* (Aug, 2017).

[361] IFLA, "How to spot fake news." https://www.ifla.org/publications/node/11174.

[362] E. Kiely and L. Robertson, *How to spot fake news*, *Fact Check* (Nov, 2016).

[363] Z. Yang, C. Wilson, X. Wang, T. Gao, B. Y. Zhao, and Y. Dai, *Uncovering social network sybils in the wild*, *ACM Transactions on Knowledge Discovery from Data (TKDD)* **8** (2014), no. 1.

[364] J.-P. Verkamp and M. Gupta, *Five incidents, one theme: Twitter spam as a weapon to drown voices of protest*, in *Proceedings of the 3rd USENIX Workshop on Free and Open Communications on the Internet*, Aug, 2013.

[365] H. Allcott and M. Gentzkow, *Social Media and Fake News in the 2016 Election*, *Journal of Economic Perspectives* **31** (2017), no. 2.

[366] S. C. for Freedom, "Demirtaş's lawyer accused of joining HDP's WhatsApp group ." https://stockholmcf.org/demirtass-lawyer-accused-of-joining-hdps-whatsapp-group/, Apr, 2017.

[367] W. Hatti, "Attorney Levent Pişkin is being charged for meeting his client Selahattin Demirtaş." https://washingtonhatti.com/2017/04/11/attorney-levent-piskin-is-being-charged-for-meeting-his-client-selahattin-demirtas/, Apr, 2017.

[368] Twitter, "Twitter Help Center: Country withheld content ." https://support.twitter.com/articles/20169222.

[369] J. Conditt, *Turkey shuts off internet service in 11 Kurdish cities*, .

[370] L. Parks, H. Goodwin, and L. Han, *"I Have the Government in My Pocket": Social Media Users in Turkey, Transmit-Trap Dynamics, and Struggles Over Internet Freedom*, 08, 2017.

[371] Google, "Android developers dashboard." https://developer.android.com/about/dashboards/index.html.

[372] R. Rivest, A. Shamir, and Y. Tauman, *How to leak a secret*, *Advances in Cryptology—ASIACRYPT 2001* (2001).

[373] D. Chaum and E. Van Heyst, *Group signatures*, in *Advances in Cryptology—EUROCRYPT'91*, pp. 257–265, Springer, 1991.

[374] Zetetic, "SQL Cipher." https://www.zetetic.net/sqlcipher/sqlcipher-for-android/.

[375] Apple, "iOS Security." https://www.apple.com/business/docs/iOS_Security_Guide.pdf, Mar, 2017.

[376] Google, "Language and locale." https://developer.android.com/guide/topics/resources/multilingual-support.html.

[377] S. Parvini, *Northern California gets its wettest winter in nearly a century*, *LA Times* (Apr, 2017).

[378] J. Serna, *The Oroville Dam spillway was wrecked months ago. Here's where the repairs stand as rain season looms*, *LA Times* (Nov, 2017).

[379] Cal Fire, "Top 20 destructive California wildfires." https://www.fire.ca.gov/media/5511/top20_destruction.pdf, Aug, 2019.

[380] V. Martinez, *Here are the 5 largest California wildfires*, *LA Times* (Dec, 2017).

[381] Cal Fire, "Top 20 largest California wildfires." https://www.fire.ca.gov/media/5510/top20_acres.pdf, Aug, 2019.

[382] J. Gettleman, *More than 1,000 died in South Asia floods this summer*, *New York Times* (Aug, 2017).

[383] J. Amy and M. Sedensky, *Rescuers seek anyone — alive or dead — left behind in Harvey's floodwaters*, *Chicago Tribune* (Aug, 2017).

[384] J. Yeung, *Australia wildfires: Here's what you need to know about the deadly blazes*, *CNN* (Jan, 2020).

[385] R. Hughes, *Amazon fires: What's the latest in Brazil?*, *BBC* (Oct, 2019).

[386] S. Lovgren, *Brazil amazon fires threaten river habitat of thousands of fish species*, *National Geographic* (Sep, 2019).

[387] R. Westrate, *How bushfires will affect Australia's security*, *Lawfare* (Feb, 2020).

[388] M. Dilley, R. S. Chen, U. Deichmann, A. L. Lerner-Lam, M. Arnold, J. Agwe, P. Buys, O. Kjekstad, B. Lyon, and G. Yetman, *Natural Disaster Hotspots: A Global Risk Analysis*, *The World Bank* (2005).

[389] E. Toksabay and T. Gumrukcu, *Turkey moves to oversee all online content, raises concerns over censorship*, *Reuters* (Aug, 2019).

[390] J. Griffiths, *Myanmar shuts down internet in conflict areas as UN expert warns of potential abuses*, *CNN* (Jun, 2019).

[391] M. Reardon, *Women's march overwhelms mobile network in dc*, *CNET* (Jan, 2017).

[392] A. Perrin, "Digital gap between rural and nonrural america persists." `https://www.pewresearch.org/fact-tank/2019/05/31/digital-gap-between-rural-and-nonrural-america-persists/`, May, 2019.

[393] M. Anderson and A. Perrin, "17% of teens sometimes can't finish homework because of digital divide." `https://www.pewresearch.org/fact-tank/2018/10/26/nearly-one-in-five-teens-cant-always-finish-their-homework-because-of-the-digital-divide/`, Oct, 2018.

[394] S. Mihelj, A. Leguina, and J. Downey, *Culture is digital: Cultural participation, diversity and the digital divide*, *New Media & Society* **21** (2019), no. 7.

[395] E. V. Estacio, R. Whittle, and J. Protheroe, *The digital divide: Examining socio-demographic factors associated with health literacy, access and use of internet to seek health information*, *Journal of health psychology* **24** (2019), no. 12.

[396] S. T. Jin, H. Kong, R. Wu, and D. Z. Sui, *Ridesourcing, the sharing economy, and the future of cities*, *Cities* **76** (2018) 96–104.

[397] FCC, "Wireless resiliency cooperative framework." `https://www.fcc.gov/wireless-resiliency-cooperative-framework`, Dec, 2016.

[398] C. Mills, *How a Wireless Network Prepares for Hurricane Harvey*, *BGR* (August, 2017).

[399] K. Namuduri, Y. Wan, and M. Gomathisankaran, *Mobile ad hoc networks in the sky: State of the art, opportunities, and challenges*, in *Proceedings of the second ACM MobiHoc workshop on Airborne networks and communications*, pp. 25–28, ACM, 2013.

[400] K. Namuduri and A. Soomro, *Reliability, throughput and latency analysis of an aerial network*, in *Ad Hoc Networks*, pp. 382–389. Springer International Publishing, 2017.

[401] S. Rohde and C. Wietfeld, *Interference aware positioning of aerial relays for cell overload and outage compensation*, in *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, IEEE, 2012.

[402] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, *Drone small cells in the clouds: Design, deployment and performance analysis*, in *Global Communications Conference (GLOBECOM), 2015 IEEE*, IEEE, 2015.

[403] K. Gomez, A. Hourani, L. Goratti, R. Riggio, S. Kandeepan, and I. Bucaille, *Capacity evaluation of aerial LTE base-stations for public safety communications*, in *Networks and Communications (EuCNC), 2015 European Conference on*, IEEE, 2015.

[404] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, *On the number and 3D placement of drone base stations in wireless cellular networks*, in *Vehicular Technology Conference (VTC-Fall)*, IEEE, 2016.

[405] A. Dhenke, M. Gowda, and R. Choudhury, *Extending Cell Tower Coverage through Drones*, in *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, HotMobile '17, 2017.

[406] J. Temperton, *Drones and TV white space: the future of flood defence*, *Wired* (Aug, 2015).

[407] K. Fall, *A delay-tolerant network architecture for challenged internets*, in *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '03, pp. 27–34, ACM, 2003.

[408] Qualcomm, "Everything you need to know about 5G." `https://www.qualcomm.com/invention/5g/what-is-5g`.

[409] Verizon, "What is 5G?." `https://www.verizon.com/about/our-company/5g/what-5g`.

[410] *When is 5G coming to the U.S.?*, *Lifewire* (October, 2019).

[411] FCC, "Broadband progress report." https://fcc.gov/reports-research/reports/broadband-progress-reports/2016-broadband-progress-report, January, 2016.

[412] N. Mendoza, *Many US rural areas still suffer slow internet speeds*, *TechRepublic* (September, 2019).

[413] FCC, "E-Rate: Universal Service Program for Schools and Libraries." https://www.fcc.gov/consumers/guides/universal-service-program-schools-and-libraries-e-rate.

[414] P. Flahive, *Inside The Movement to Improve Access to High-Speed Internet in Rural Areas*, *NPR* (September, 2019).

[415] USDA, "USDA to Provide $150 Million to Help Rural Communities Affected by Natural Disasters." https://www.usda.gov/media/press-releases/2019/09/10/usda-provide-150-million-help-rural-communities-affected-natural, September, 2019.

[416] Cisco, "5G RuralFirst." https://www.5gruralfirst.org/.

[417] B. Jansen, *NYC firefighters use drone to help battle blaze for first time*, *USA Today* (Mar, 2017).

[418] M. Kutner, *London firefighters used drone to battle grenfell tower blaze*, *Newsweek* (Jun, 2017).

[419] M. Moon, *American red cross is launching a drone disaster-relief program*, *Engadget* (Sep, 2017).

[420] M. Blood and E. Knickmeyer, *'Can't confirm. Dark': Officials couldn't see danger at dam*, *Associated Press* (Sep, 2017).

[421] J. Brodkin, *Tropical Storm Harvey takes out 911 centers, cell towers, and cable networks*, *Ars Technica* (Aug, 2017).

[422] M. Reardon, *Hurricane Harvey: How the wireless carriers fared*, *CNET* (Sep, 2017).

[423] A. Lertsinsrubtavee, L. Wang, A. Sathiaseelan, J. Crowcroft, N. Weshsuwannarugs, A. Tunpan, and K. Kanchanasut, *Understanding Internet usage and network locality in a rural community wireless mesh network*, in *Proceedings of the Asian Internet Engineering Conference*, pp. 17–24, ACM, 2015.

[424] L. Chiaraviglio, N. Blefari-Melazzi, W. Liu, J. Gutierrez, J. V. D. Beek, R. Birke, L. Chen, F. Idzikowski, D. Kilper, J. Monti, and J. Wu, *5G in rural and low-income areas: Are we ready?*, in *ITU Kaleidoscope: ICTs for a Sustainable World*, November, 2016.

[425] M. Hall, *When network connectivity can mean life or death*, *Network Computing* (May, 2019).

[426] J. N. Ralph Ellis, *Drone, social media make flood rescue happen*, *CNN* (January, 2018).

[427] M. Koren, *Using Twitter to save a newborn from a flood*, *Huffington Post* (August, 2017).