How's My Network - Incentives and Impediments of Home Network Measurements

by
Alan Ritacco

A Dissertation
Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Doctor of Philosophy
in

Computer Science

December 2019

APPROVED:

_____
Professor Craig Wills, Dissertation Advisor

_____
Professor Emmanuel Agu, Committee Member

_____
Professor Mark Claypool, Committee Member

_____
Professor Radim Bartos, External Committee Member

Abstract

Gathering meaningful information from Home Networking (HN) environments has presented researchers with measurement strategy challenges. A measurement platform is typically designed around the process of gathering data from a range of devices or usage statistics in a network that are specifically behind the HN firewall. HN studies require a fine balance between incentives and impediments to promote usage and minimize efforts for user participation with the focus on gathering robust datasets and results. In this dissertation we explore how to gather data from the HN Ecosystem (e.g. devices, apps, permissions, configurations) and feedback from HN users across a multitude of HN infrastructures, leveraging low impediment and low/high incentive methods to entice user participation. We look to understand the trade-offs of using a variety of approach types (e.g. Java Applet, Mobile app, survey) for data collections, user preferences, and how HN users react and make changes to the HN environment when presented with privacy/security concerns, norms of comparisons (e.g. comparisons to the local environment and to other HNs) and other HN results. We view that the HN Ecosystem is more than just "the network" as it also includes devices and apps within the HN.

We have broken this dissertation down into the following three pillars of work to understand incentives and impediments of user participation and data collections. These pillars include: 1) preliminary work, as part of the How's My Network (HMN) measurement platform, a deployed signed Java applet that provided a user-centered network measurement platform to minimize user impediments for data collection, 2) a HN user survey on preference, comfort, and usability of HNs to understand incentives, and 3) the creation and deployment of a multi-faceted How's My Network Mobile app tool to gather

and compare attributes and feedback with high incentives for user participation; as part of this flow we also include related approaches and background work.

The HMN Java applet work demonstrated the viability of using a Web browser to obtain network performance data from HNs via a user-centric network measurement platform that minimizes impediments for user participation. The HMN HN survey work found that users prefer to leverage a Mobile app for HN data collections, and can be incentivized to participate in a HN study by providing attributes and characteristics of the HN Ecosystem. The HMN Mobile app was found to provide high incentives, with minimal impediments, for participation with focus on user Privacy and Security concerns. The HMN Mobile app work found that 84% of users reported a change in perception of privacy and security, 32% of users uninstalled apps, and 24% revoked permissions in their HN. As a by-product of this work we found it was possible to gather sensitive information such as previously attached networks, installed apps and devices on the network. This information exposure to any installed app with minimal or no granted permissions is a potential privacy concern.

Acknowledgements

I express my sincere gratitude to my advisor Professor Craig Wills for his guidance, support, time, and patience along the way. I am thankful for my committee members (Professor Agu, Professor Claypool, and Professor Bartos) and other WPI CS faculty for taking the time to work with me. I would also like to thank those who have been mentors to me over the years providing inspiration, including Dr. Johnson, Mr Q, Dr. Mahadev, Bob Rogers, and many others.

Ultimately this work is dedicated to my Father Nicholas Ritacco, Mother Nancy Ritacco, my dearest Wife Wendy, daughter Erin, son Michael, and my Uncle(s) Arthur, Jack, and Bob. I am for ever grateful to my brothers, nephews, nieces and the rest of my family and friends for without their continued support and encouragement this would not have been possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The prevalence of HNs and Internet usage has become common place in recent years as the cost for HN infrastructure and Internet-based connectivity becomes affordable for house-hold users. The review of the infrastructure in HNs, and the data surrounding these HNs along with Internet usage avails an acme of data collection points for these open environments. Collecting data from HNs affords researchers the ability to peek inside HNs and gather information on an overall big picture of what is going on inside HN environments, from behind the firewall that typically blocks HNs from external access. On the other hand peering into HNs also allows a view of both operation, collaboration of feedback and preference, as well as norms of data (e.g. comparisons to the local environment and to other HNs). We define local and global norms as, on average, point-in-time attributes (e.g. number of devices) and a rated status of operations (e.g. Wifi health) both for users local HN and in comparison to remote (global) HN Ecosystems.

Collecting data from Home Networking (HN) environments for statistics such as upload/download throughput, domain name service (DNS) and round trip time, DNS health,

network activity system level information, devices, apps, privacy and security, and services executing, while minimizing user impediments is a difficult task to achieve. We present a study to allow for the retrieval and review of these data points from the user's perspective as well as the research community, while minimizing impediments (e.g. minimal efforts to operate) and maximizing incentives to participation (e.g. providing robust set of results). We have the ability to peer behind HN firewalls to gather data and feedback as well as present results, and comparisons of HN Ecosystems to HN users using several approach types. We define the HN Ecosystem to include all entities that reside within and potentially across the HN environment, namely: apps, devices, Wifi, Internet access, configurations, etc. We believe that the integration of results, feedback and data points are helpful to both researchers and HN users.

Collecting this research data from HN environments gives researchers a purview of HN usage, traits, trends, and overall configurations. We believe that providing these results to assist HN users with understanding their environment, health of operations (e.g. is my Wifi operating normally), and to potentially tweak commonly used hardware and software for best performance as an incentive, with minimal impediments to participate, is both desired by users and advantageous for researchers. With 4.3 billion Internet users (˜ 56.8% of world population) [184] HNs have a vast number of environments for researchers to collect data from. The gathering and understanding of HN environments should start with the initial understanding of what data is important and how to gather data from users, and more importantly how to incentivize users to participate.

The motivation for HN data collection is primarily about gathering these HN measurements for research, with a focus on incentives and impediments for users to participate in HN studies. Researchers have interest into what is going on behind the HN firewall and inside of the HN Ecosystem to map and understand the HN Ecosystem, including local and global changes. In addition to researchers mapping and understanding the HN Ecosys-

tem, users are able to leverage these results to understand and potentially optimize their HN environments, while researchers can determine behavioural changes as mentioned. We believe that leveraging incentives for user participation can been accomplished using several methods, including: results (e.g. operational health status), financial (paying users to participate), and social aspects (e.g. how does my network compare to others), all of which we explore in this dissertation. On the other hand, minimizing impediments can allow for an increase in participation by lowering the barrier of entry required to take part in a study, and can be achieved using methods which require minimal user efforts for installation and operation (e.g. requiring only novice level skills). In this dissertation we provide a study around these user incentives and impediments, tied together with background work, a HN survey, an initial study leveraging a Java App, and a Mobile App for data collections within a HN.

Gathering these applicable data points from HNs should offer users the advantage of both high value of incentives (e.g. richness of results) as well as minimal impediments (e.g. ease of use) to operate and participate in a study. It should also provide a rich collection of data to research of the HN Ecosystem along with user results and a diagnostics health approach (e.g. feedback on operation status and norms) of HN environments to participants. The creation of a low impediment application to collect applicable and coveted data from HNs and provide valuable results for users is our goal. We conjecture that tools allowing for a multifaceted approach (e.g. results, norms and feedback) along with enticing users with incentives for participation will amass a robust HN dataset, while also satisfying HN users in terms of results (incentives). This approach includes the leveraging of an integration approach of data collection methods (Java, Mobile app, survey) and incentives, collection of user feedback, along with the detection of changes within HNs in and across disparate HN Ecosystems, all while minimizing participation requirements.

1.1.1 What is a Home Network

As shown in Figure 1.1 we define a Home Network (HN) as a residential environment that consists of an entry point device(s) serving up Wifi, modem, router, and switch services in a typically NAT'd (Network Address Translation) domain. This residential environment is typically served up Internet service by an ISP (Internet Service Provider) via DSL, Cable, or Fiber; although 5G and other telephony connections may exists they are all backbone'd via one of these methods, typically Fiber, and our studies are interested exclusively on HNs. An important factor to a HN is the Wifi services provided within the environment, which are typically home-branded Wifi services versus commercial or business grade hardware. Although some homes may have multiple entries (e.g. Internet Services) and multiple Wifi hot-spots (e.g. Wifi extenders), they are designated as a residential Wifi provider versus that of a business as the typical HN is not supported (internally) by professionally staffed Information Technology teams, although the user's ISP may provide fee-for-service; devices and apps are included as part of this HN Ecosystem and includes the functionality and features of both. In this work we view that the HN Ecosystem is more than just "the network" as it also includes devices and apps within the HN.

1.2 Focus and Timing of Work

The following is the focus and timing of work done as part of this dissertation. We have broken this dissertation down into the following three pillars of work to understand incentives and impediments of user participation and data collections. The pillars of this dissertation include the following three focus areas: 1) a Java applet approach (Chapter 2) to HN data collections offering low impediments to users to participate in a HN study, 2) a HN survey (Chapter 4) of to understand incentives around what HN users have interests

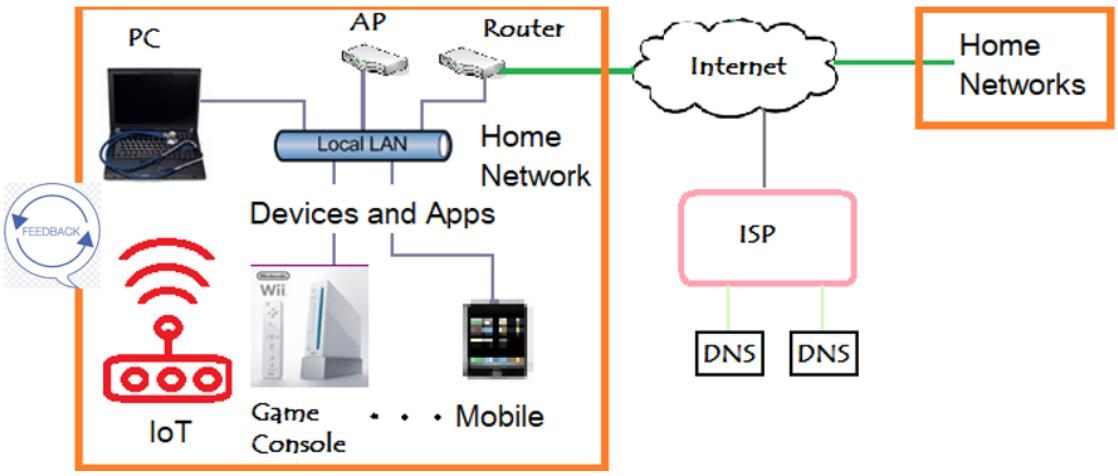

Figure 1.1: Example Home Network

in and around the HN Ecosystem, and 3) a Mobile app (Chapter 5) to collect a robust data set from the HN Ecosystem, while providing low impediments for participation and high incentives in terms of results, operational status, and comparisons to other HNs to users.

We started our discovery in 2009, and after a break re-engaged research and discovery in 2016. In this dissertation we have broken down our work into the following pillars or thrust chapters of focus. A short summary of these pillars include:

1. Low impediments - in 2009 a preliminary study of HNs was completed using a Java Applet, where we did an initial study to understand user incentives and impediments that are useful in gathering data for researchers.

2. Glean incentives - in 2018 we completed a HN survey to understand incentives around participation along with user experiences, comfort levels, and skills of managing HNs, as well as which approaches users prefer to gather data and review results from HN discovery. We also include Related approaches and background work completed in 2018, where we did an exploration into HNs and measurement points available to gather data. This background work includes approaches such as routers, apps, hardware, and Web/scripting tools. We look to understand data of

interest provided by each approach, and focused on creating a taxonomy of current HN functionality. ; and

3. High incentives - we conclude with an in-depth review, done in 2018-2019, of HN attributes, feedback, and configurations leveraging a Mobile app for discovery and results targeted to users. This work focus on a high incentive approach with user results driven around Privacy/Security, changes to the HN, and Norms across HNs.

In the next several sections we provide a deeper introductory review of these pillar areas (including background research), as well as our research questions and conclusions around HN collections, feedback, analysis of results, and implications we found during this research.

1.3 Research Questions

Previous Home Network (HN) studies have been targeted around gathering low level data and have provided minimal incentives with high impediments or barriers of entry to HN users to participate. These studies have tended to provide negligible information and feedback to and for HN studies, but tend to require expertise to operate. We believe that the landscape has changed and that users are now engaged and interested in an in-depth view of HN information, with minimal impediments and the preference of high incentives for participation. Information desired includes attributes such as devices, apps, health of operations (e.g. how is my HN operating), and global historical norms perspective for comparison; historical norms are data comparisons (over time) of the local HN as well as comparisons to other (remote) HNs (e.g. throughput, Wifi, etc). We have created and deployed a Java applet (via a Web browser) as a low impediment approach of collecting low level data in HNs, a survey to understand incentives for participation in HN studies, and a multifaceted high incentive Mobile app approach that combines HN device and app

discovery in order to discern the current HN Ecosystem. In addition, while devices and apps appear to be two distinctly different areas of functionality within HN, we consider them similar and part of the HN Ecosystem as they are additive entities and part of the HN. We believe that by providing HN users a low impediment approach with high incentives will allow us to collect a robust dataset from the HN Ecosystem and motivate users to participate. As part of this dissertation we pose the following research questions:

1. Can a Java Applet be an effective, low impediment, approach type for the collection of meaningful research data and provide meaningful results?

2. Are HN users interested in understanding privacy (e.g. data leaks) and security concerns (e.g. unauthorized devices, detection) in their HN Ecosystem?

3. Does a Mobile app provide the trade-offs users are looking for in data collections, results and range of data, and ease of operation?

4. Will users modify their set of apps, permissions, and devices when presented with potential privacy and security as well as health-rated results?

We test these research questions via our data collections, survey, and user feedback.

1.4 The Dissertation

The goal of this dissertation is to gather data for research and provide users results on operational health status of a HN Ecosystem, as well as understand how users react to the availability of these results, all using minimal impediment and incentive-based approaches for HN users to participate. In this vein, we created a Java applet, leveraged a HN survey, and a mobile app to gather data from a variety of HN areas including throughput, networking characteristics, health and historical norms, apps and permissions, while

collecting data from users via feedback of the effectiveness of the experience and results presented. The Java applet runs within a web browser and collected data periodical from HNs. We leveraged an HN survey methodology to glean incentives on participation of HN studies. Finally, we created a Mobile app that collects attributes of the HN and periodically requests feedback, via in app survey questions, and uses a star rating, graphing, and descriptive feedback system to understand effectiveness and overall HN user experience. We evaluate the success of this work by gathering data from the HN via the use of these pillar areas of research (Java, Mobile app, and a survey), running across a diverse range of HNs, to understand data collection, as well as the impact of results and feedback on user behavior, and understand incentives and impediments for user participation of HN studies. This work is also centered around monitoring users changes to the HN Ecosystem and health of operational status of the HN Ecosystem to understand changes, functionality, and comparisons of HNs.

1.5 Contributions

The main contribution of this dissertation is to understand incentives and impediments and approaches for HN studies. We also look to provide and improve the user HN experience, using both a Java and Mobile app-based approach, to Home Network measurement by providing high incentives and minimal impediments to data gathering within the HN Ecosystem. Our work makes the following contributions:

1. demonstrate that a Java applet and a Mobile app platforms are valuable approaches to collect data and features of the HN (e.g. devices, apps, privacy and security, as well as other [low level] information) and can display this HN Ecosystem information in an functional manner;

2. show that leveraging a HN survey we can understand skill level, preferences, and

preferred methods for access to HN data and results;

3. learn how users respond to HN Ecosystem information based on integrated feedback and actions taken (e.g. changes to apps and permissions) based off of HN results and user feedback;

4. measure HN wifi and apps operational Health status and functionality;

5. show that the collection of HN devices is valuable to users;

6. provide a method of research measurement that combines the collection and analysis of Mobile apps and devices as part of the HN Ecosystem; and

7. discover research implications that may impact HN users, including: access and enumeration of devices, apps and permissions, as well as exposing the risks of privacy and security of the HN Ecosystem.

1.6 Roadmap

In this section we review the three pillars previously reviewed, which include: 1) Java applet, 2) HN Survey, and 3) HMN Mobile App; we also includes related approaches and background in this section, which provides fodder and impetus across the entire dissertation.

1.6.1 A Java Approach to Home Network Measurement

Chapter 2 is work that was done, in 2009, to understand HN measurement, as part of the How's My network (HMN) project, leveraging a Java Applet for data collection. This chapter focuses in on an initial study completed around the development of a user-centered network measurement platform that limits impediments to participation using

using a signed Java applet for home network measurement, and provides (low) incentives for participation. We review the tool's capabilities and report the measurement methodology employed, as well as the results obtained from HN environments, and potential implications these results provided. Despite the sandbox-type restrictions in Java, the results include information about the configuration of the user's testing machine, wireless connectivity of the testing machine, available upload and download throughput, DNS performance and the number and type of devices on the user's network. The following are key takeaways from this work:

1. creation of a core measurement component of a future user-centric network measurement platform, as part of the How's My Network infrastructure, which offers incentives and minimizes impediments for user participation;

2. demonstrate the viability of using a Web browser for obtaining network performance information;

3. discovery of wired and wireless information in a HN;

4. measurement of DNS performance via a web browser; and

5. the ability to learn about networked devices on residential networks.

1.6.2 Peering into the Home Network

Chapter 3 peers into the HN environment reviewing devices, tools, and approaches used to collect data in HNs. This chapter is part of the research areas of this dissertation as it is the support of the related approaches and background work. This study was completed in 2018, after a hiatus, and resulted in complementary and extension of background work, completed by the Java applet, in the examination of HN measurements, privacy and security of apps, and devices in the HN Ecosystem, with incentives (and impediments)

as the cornerstone of the research. In this chapter we examine a variety of approaches that exist in the HN Ecosystem including collections from routers, apps, hardware, and Web/scripting tools. We are interested in understanding the complexity of these tools, and the required expertise to execute and configure as well as incentives and impediments to participation. We look to understand and compare measurement points (MPs), applications and expertise required across these four approaches, and examine data of interest provided by each approach. We show that focusing on a broad range of approaches, data of interest, and tools allowed us to create a new taxonomy of HN functionality.

In this work we review several approach types, and found that a hardware and router approach is costly and can be complex to setup and configure, and thus has a high impediment to entry, but provides access as the incentive. A web/scripting approach may not provide enough information the user is looking to understand, but has the lowest barrier to entry as it executes within a web browser. An apps approach may be free to users, and require minimal impediments to install, configure, and execute and thus has a lower barrier to entry for users. We also reviewed data of interest, from these approach types, in an an attempt to understand flexibility and incentives. We found that apps have the most flexibility in terms of data of interest, as well as historical norms (e.g. comparisons to others), and may provide the most optimal approach for users and potentially researchers as well. The following are key takeaways from this work:

1. a study that identifies software tools and their provided data targeted to HNs;

2. compare expertise required across four tool approaches: Routers, apps, Hardware, and Web/Scripting tools;

3. review the complexity of these tools, and the required expertise to execute and configure as well as impediments and incentives to participation;

4. create a taxonomy around data of interest provided by each approach; and

5. show that focusing on a broad range of approaches, data of interest, and tools allows us to create a new taxonomy of HN functionality.

1.6.3 Home Network Survey

Chapter 4 is part of the pillars of our work, and is a survey of HN user experiences, skill and comfort level, as well as preferences for data collections in the HN Ecosystem. This chapter focuses in on understanding incentives to attracting users to participate in HN studies as well as understanding HN user experiences, including their perceived skill and comfort levels with managing devices and services that reside within their HN. We look to understand the perceived value of information, from our HN survey, and focus in on a user-centric approach to understanding the user experience when managing and using a HN. This includes examining device types, Internet connectivity, management options, interest, and preferred method of management of user HNs.

This HN survey and study points toward HN interests in leveraging a Mobile app for review of HN data, across a wide variety of areas. These areas of interests expressed by users include privacy, security, norms of data, information discovery, as well as operational status information. We also found that HN users have an interest in data collection points that are directly tied to HNs, using targeted preference points for data collection and dissemination to HN users. We show the results from this survey and discuss how they fit into this dissertation. We found the following key takeaways from this work:

1. an understanding of perceived value of information and experiences of HN users when managing and using a HN;

2. examination of device types, Internet connectivity, management options, interest, and preferred method of management of user HNs;

3. examination of HN user skill level related to HN management and comfort;

4. show that HN users across all skills levels showed a high level of interest in understanding how to make changes to their devices, or HN to optimize their experience;
5. show that users are interested in an in-depth view of HN information;
6. show results and characteristics that incentives user participation; and
7. show that HN users have an interest in leveraging a Mobile app for data collections within a HN.

1.6.4 HMN Mobile App

Chapter 5 is a pillar of our research work and is a detailed review from the HMN Mobile app research, including: data collection, user feedback, and HN comparisons. This work focuses on (high) incentives and minimal impediments leveraging a Mobile app for data collections and data dissemination. In this chapter we provide details on the HMN Mobile app platform developed to measure information in the HN that includes: devices, apps privacy/security, networking, and comparison of norms, using a Mobile app approach to collect HN Ecosystem data as well as user feedback via a single app. We discuss the creation and deployment of this multifaceted Mobile app approach that combines HN device and app discovery, in order to discern the current HN Ecosystem as well as providing value, in terms of results, to help better understand the HN user experience. We provide a landscape view of the HN Ecosystem and details on the capabilities of the tool, measurement methodology, how it was deployed and tested, comparisons and norms of data, feedback results, impacts on user perceptions, and changes made to the HN ecosystem, as well as implications found from this work. We found that the following are key takeaways from this work:

1. demonstrate that a Mobile app is valuable platform to collect data from the HN Ecosystem (e.g. devices, apps, privacy and security, as well as other information)

and can display this HN Ecosystem information in an functional manner to HN users, with minimal impediments and a high level of incentives for participation;

2. show that users react to privacy and security concerns, based off of HN Ecosystem results, by making changes to their HN Ecosystem (e.g. apps, permissions, devices, as well as other information);

3. show that HN users are interested in understanding the operational Health status of their HN Ecosystem as well as comparisons to other HNs; and

4. show that a Mobile app can expose privacy and security issues of the HN Ecosystem (e.g. apps, permissions, connectivity, behaviors).

1.6.5 Conclusions and Future Work

Chapter 6.1 wraps up this dissertation with a summary, review of research questions and conclusions, and provides possible future work.

Chapter 2

A Java Approach to Home Network Measurement

As part of a project to develop a user-centered network measurement platform that limits impediments to participation, this work focuses on using the execution of a signed Java applet for home network measurement. We have developed a Java applet tool to understand the capabilities of such a tool for measuring characteristics of a user's network environment from the browser. This area focuses and reports on the capabilities of the tool, the measurement methodology employed, and initial results obtained for a set of residential users employing the tool. Despite the sandbox-type restrictions in Java, the results include information about the configuration of the user's testing machine, wireless connectivity of the testing machine, available upload and download throughput, DNS performance and the number and type of devices on the user's network. This work and study [139] was completed in 2009 and is a key resource to this dissertation as it ties together components related to incentives, impediments, toward a user centered focus and ultimately leveraging a Mobile app for data collection and dissemination of information to HN users.

2.1 Introduction

The work presented in this chapter is part of a larger project to develop a *user-centered* network measurement platform, called *How's My Network (HMN)*, which provides incentives via games and feedback on application performance, while limiting impediments so that the public perceives benefits in participation. Our initial work has focused on what performance measures can be obtained via a Web browser, which is a low-impediment platform for a wide variety of users. One of our projects examined network performance measures obtainable via JavaScript and Flash [78], while this work focuses on using the execution of a signed Java applet for home network (HN) measurement.

Traditionally, Internet measurement has been done from points in the network infrastructure or from well-connected research labs and universities. However, with the dramatic growth in Internet access from residences and out in public, often hidden behind Network Address Translation (NAT) boxes, the old measurement paradigm increasingly excludes the performance vantage points seen by the majority of Internet users. The size of this cadre of "invisible" Internet users is increasing as public wireless networking becomes more commonplace and home networking spreads further through the developing world.

The need for a new network measurement paradigm focusing on where users live and their specific interactions with the Internet has already been recognized. One outcome of the Community-Oriented Network Measurement Infrastructure (CONMI) Workshop Report was that "@home-style measurement" is needed to increase the number of Internet vantage points [87]. Desirable outcomes from an NSF Computing Infrastructure session on testing for the new Internet [143] include better representation of the user population, non-Linux performance tests and a "SETI@home" type mechanism for networking. Previous work [25] laments the widening gap between measurements for the visible and

largely invisible portions of the Internet community motivating the need for ”attractors” to provide incentives for user participation in measurement.

While existing network measurement platforms have several desirable features, they do not satisfy these needs. Platforms such as PlanetLab [18, 128, 156] and Archipelago [15, 71] provide flexibility for researchers in choosing metrics to collect, but their platform nodes are permanent, immobile and within a dedicated infrastructure. Alternative platforms such as NETI@home [149], DIMES [146] and DipZoom [176] allow measurements from any node in the Internet, but the scopes of their measurements are limited with currently little incentive for the general populace to participate. Finally, Gomez [57] and a variety of “speedtest” services [46, 154] include limited incentives for broader user participation, but are not designed to inform network research.

We have developed a HMN Java applet tool to better understand the capabilities of Java applets for measuring characteristics of a user’s network environment from within a Web browser. This work reports on the capabilities of the tool, the measurement methodology employed and initial results obtained for a set of residential users that used the tool. The results include information about the configuration of the user’s testing machine, wireless connectivity of the testing machine, available upload and download throughput, DNS performance and the number and type of devices on the user’s network. This work is important because it demonstrates that a wealth of information about the home network environment can be obtained via the ubiquitous Web browser and the work establishes a basis for understanding long-term trends in this domain.

This area makes a number of contributions for home network measurement:

1. demonstration of the viability of the Web browser for obtaining network performance information;

2. discovery of information about wired versus wireless connectivity of home ma-

chines;

3. the ability to learn about the number and types of networked devices on residential networks;

4. using JDIG, a Java-based DNS tool we developed, the ability to measure DNS performance obtained by users in a home network;

5. the ability to integrate upload and download throughput results, comparable to existing speedtest services, with other measures; and

6. a core measurement component of a future user-centric network measurement platform with incentives for user participation via feedback on applications and servers of interest to users.

2.2 Research Questions

The activities associated with the HMN platform can answer a number of research questions about home networks. The following list enumerates some of these questions that focused on in this work.

1. What is the nature of the home machine in which tests are run?

2. Do home machines use wireless connections and, if so, what can be learned about their wireless profiles?

3. What is the performance of networked applications running in a home environment?

4. What is the performance of DNS and what is its influence on application performance?

5. What is the nature of the home environment? What are the number and types of network devices in the home network?

2.3 Testing Framework

A testing framework tool was developed to allow for testing modules to be easily added. The specific testing modules used in this work are described in the following section. In addition, a custom client/server environment was designed to capture results from each test and store the results at the server for later analysis.

With the HMN-testing framework, our testing suite was provided to users via a self-signed Java applet. This approach allows a user's Web browser to execute our testing via the Java Runtime Environment (JRE). The signed applet provides range of access beyond the traditional Web browser sandbox level access and control, but is still constrained by the JRE. This approach is beneficial for users as it allows participation while not requiring the installation of any additional software on a user's machine.

The HMN testing suite is comprised of a simple graphical user interface with a single "Start" button along with a refresh time. Once the application is running it is set by default to repeat execution every five minutes. Re-execution of the tests has shown to be valuable for providing both longer-term information for the user and in for data in our repository. All data is stored based on the Internet Protocol (IP) address, although the use of cookies in the future can help correlate multiple tests and to anonymize results when made available to others. During the loading of the HMN applet a security certificate requests the user to accept the applications digital signature (the signed applet). Accepting the certificate allows the applet to perform operations outside of the sandbox. A user who does not accept the certificate will run the applet in the browser sandbox limiting its capabilities to learning some information about the testing machine as well as the upload

and download throughput to the origin server.

2.4 Methodology

The HMN applet used in this work consists of five modules that each obtain distinct types of information about a user's testing machine and networked environment. Each module is designed to obtain information about a research question posed in Section 2.2. These modules are run in phases as described in the following sections.

2.4.1 Test Configuration

Configuration information is first gathered about the machine performing the test. This information includes the type of browser and operating system, the internal and external IPs employed by the machine including whether the machine resides on a non-routable network (behind a NAT box) [136], and address of the primary DNS server. Local system property information such as Java version, paths, architecture type, CPU type and processor speed are also obtained.

2.4.2 Wireless Connectivity

Two types of information about a user's wireless connectivity are obtained. First, by querying the network configuration information the applet is able to determine whether the machine is networked via a wired or wireless connection. Second, when available, the applet can obtain the types of wireless network profiles employed by a user. Although not available for this set of tests, a future module is being developed to obtain the number and signal strength of wireless access points in range of a test machine.

2.4.3 Upload/Download Throughput

As a measure for comparison with other testing tools another model included a module measures upload and download throughput over TCP between the testing client and our origin server. This test provides a baseline for comparison with existing tools. Future modules will include similar tests to non-origin servers, which are allowed from a signed applet, where the chosen servers can be customized based on user preferences.

2.4.4 DNS Performance

DNS performance continues to be an important, if overlooked, aspect of service provided to home users. In order to test DNS performance we created a Java-based tool, called *JDIG*, with an interface similar to the public domain *dig* tool. Our tool can run as a standalone Java application, but for our work it is packaged as part of a module. JDIG performs a variety of DNS tests including the round-trip time (RTT) for obtaining a cached DNS entry. The tool also measures the average DNS RTT for a random set of *.edu* servers, the average DNS RTT for a set of popular servers [4] and the RTT to obtain a top-level domain (TLD) and generic TLD (gTLD) entry.

2.4.5 Local Network Environments

The final module in our set of tests determines information about the number and types of devices on the local network of the user's testing machine. This module is only invoked on networks with non-routable IP addresses that are behind a NAT box. The first step performed by this test is to determine the number of active devices on the network. It does so by issuing an ICMP request for the 255 IP addresses obtained by varying the low-order byte of the testing machine's IP address. A thread-based parallel can completes in 10-20 seconds. While not all active devices reply to the ICMP request an underlying

ARP request causes a reply for each valid IP address with the device's corresponding MAC address.

Once the scan of IP addresses is complete, the list and count of active devices is obtained by consulting the ARP table on the test machine. The type of each device is determined in two ways. A manufacturer of each device is obtained by matching the device MAC address with ranges assigned to manufacturers as done in *nmap* [115]. This approach works to determine special-purpose devices such as printers or game consoles. For general-purpose computers selected ports are scanned to fingerprint the type of operating system the machine is likely using.

2.5 Study

The Java testing applet resides on a quad-core server with 8GB of RAM running Linux located on WPI's campus network. The applet is downloaded via an Apache Web server. The server is also used for logging and throughput experiments. As experiments could be run at any time, a timestamp was created on the server for each applet result.

Once the applet was deployed, users on and off campus were invited to participate in testing. A total of 50 users (based on unique IP addresses) participated in the December 2008 timeframe. Using reverse DNS mappings each IP address was classified according to a commercial company, an educational institution or an ISP known to provide service to residences and public hot spots. Because our immediate focus is on residential and public users, tests from commercial and educational sites are not reported in this work. Thus, the results from 36 residential and public hot spot users are analyzed in this work. Based on the reverse DNS names all of these users are in the northeastern U.S. and can be classified into four ISPs, as shown in Table 2.1. Two of these ISPs are known to provide cable modem service, one provides DSL and one provides fiber optic service (FIOS).

Table 2.1: ISPs of Home User Tests Participating in Study

Provider	# Users	# Sessions	# Tests
Cable1	12	13	106
Cable2	12	25	109
DSL	6	9	23
FIOS	6	12	33
Total	36	59	271

The third column in Table 2.1 shows 59 unique sessions performed by our 36 users where additional sessions occur when the same user initiates the testing applet more than once. Finally, because the applet automatically re-executes its test after a five-minute sleep period, multiple tests are run within a session if the user allows the applet to remain active. Table 2.1 shows that a total of 271 tests were performed by our set of residential users.

All users accepted the digital certificate of the applet so in all cases it executed with signed applet privileges allowing for the full-range of data collection described in Section 2.4. A small number of data collection errors occurred because of insufficient local privileges even when using a signed Java Applet. These errors were specific to security privileges required by Windows-based operating systems, and occurred on non-residential networks so do not impact the results reported in this work. All phases typically take on the order of 40 seconds to execute within a user's browser.

2.6 Results

This section reports the results obtained by residential users in our study set for each of the five modules described in Section 2.4.

2.6.1 Testing Configuration

A summary of the testing configuration results for our 36 users are shown in Table 2.2. More than 94% of HMN residential users ran a Windows-based OS and of these users 45% were running Windows Vista. 61% of these users employed Internet Explorer as their browser while the remaining 39% of users employed Firefox. A small number of users with Linux-based systems ran our HMN tests and all of these users employed Firefox as their Web browser.

Table 2.2: Testing Configuration Highlights for Residential Users

94% of users run a Windows-based Operating System
45% of Windows-based OS are Windows Vista
61% of users run tests with Internet Explorer
39% of users run tests with Firefox
100% of users have a non-routable internal IP address
97% of users have a DHCP-assigned external IP address
47% of users have a non-routable primary DNS server

The internal (used by the testing machine) and external (used by the access point) IP addresses used by the residential testing platforms were examined. All home users running HMN had a non-routable internal IP address for their machine, meaning that the access point was using NAT. Based on examination of the reverse DNS name, almost all external IP addresses (97%) of were assigned by the ISP using DHCP with only 3% of HMN users with a static IP address

In looking at the DNS configuration, for 47% of HN users, the primary DNS server resided on a router/switch in the HN. Based on experience, these servers typically do not cache results, but simply "pass through" DNS requests to a caching DNS server managed by the ISP. All of the FIOS users employed such a DNS cache with mixed usage by users of the other ISPs.

2.6.2 Wireless Connectivity

Accessing the local network configuration information of the testing machine shows 38% of users ran from a wireless connected machine, while the remaining 62% used a wired PC.

Examining clients wireless caches shows that 56% of the testing machines have attached to at least one wireless network at some point in time. These wireless users have connected to five or more wireless networks at some time. These network types are in the range of: home, business, resort, and hot spot WiFi locations. Our set of users have attached to a total of 96 unique wireless networks. From the data, the most popular wireless networks are from Xerox, Cisco, and DLINK. Fifty-five unique types of hardware manufacturers used, with some overlap due to the convergence of MAC address space.

2.6.3 Upload/Download

As a measure of a user's connection performance upload and download throughput is determined to the server at WPI. Figure 2.1 shows a scatter plot of upload/download characteristics for each of the 36 HN users where each point is the average of all throughput tests for the given user.

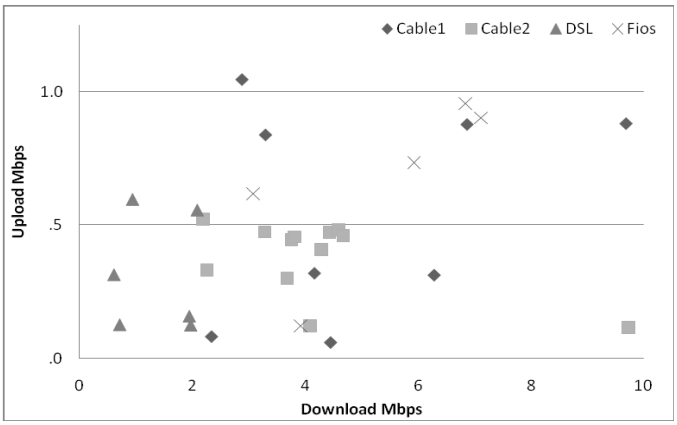

Figure 2.1: Scatter Plot of Average Upload and Download for Users by ISP

Each point in Figure 2.1 is characterized by the service provider from Table 2.1. Most of our HMN users fell in the 5Mb or less category for download and upload combined properties. The fastest upload/download throughputs were those using Cable1 and FIOS service providers. The distribution in Figure 2.1 shows how users with the same ISP have similar properties. All HN users have an asymmetric Internet connection where the download throughput is more than the upload throughput.

Figures 2.2 and 2.3 show a cumulative distribution function (CDF) for the download and upload throughputs of all 271 tests by our 36 users. The results in Figure 2.2 show variation amongst the download throughput, with DSL providing the lowest download throughput while FIOS and Cable2 providing higher download throughput and Cable1 providing the highest download throughput in our tests. On the other hand, the upload throughput values in Figure 2.3 show less distribution, with DSL and Cable2 users never receiving more than 0.5 Mbps in upload throughput. This clearly is a situation where users pay for download speeds and get nominal and/or obligatory upload speeds. ISPs can oversubscribe data lines by allowing lower bandwidth for upload than download.

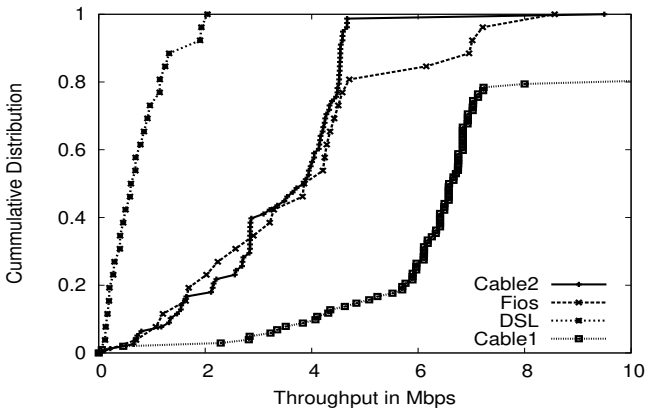

Figure 2.2: CDF of All Download Throughput Tests by ISP

Figure 2.4 shows the performance of popular speed testing services versus our HMN tests. The following speed testing services: `speedtest.com`, `speakeasy.com`, `DSLreports.com`, and `bandwidth.com`. Figure 2.4 shows the representative up-

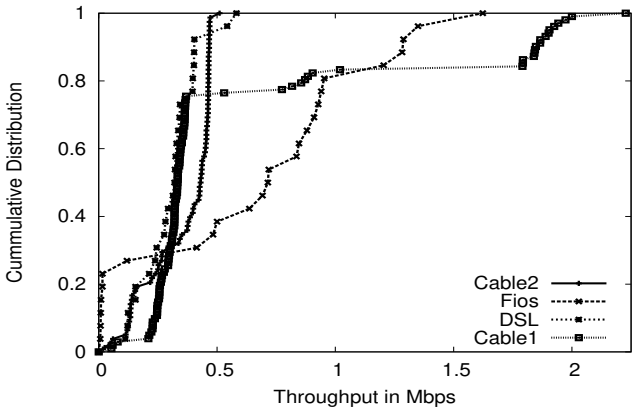

Figure 2.3: CDF of All Upload Throughput Tests by ISP

load and download throughput obtained for each service, each run from the same HN computer and cable provider. The HN system used has a known speed of 5Mb download and 512Kb upload. In each case, the nearest server was chosen for each of the speed testing services. The HMN results are similar to those of other speed testing services. While this is a sampling of data for one home network, similar results were found for other home network tests.

The amount of data each sent through for the download and upload measurements was examined using sniffer traces and other data analysis. HMN sent 1MB for download and 512KB for upload, `speedtest.net` sent 4.5MB for download, and 460KB for upload, `DSLreports.com` sent over 6MB of data for download, and over 800KB of data for download, `bandwidth.com` sent 4MB for download and 550 KB for upload.

2.6.4 DNS Performance

The availability of the JDIG tool within our test suite allows us to test DNS performance obtained by each of our users. Since our focus is on the DNS performance provided by the local DNS server of the ISP, the JDIG tool does not use OS resolver routines and therefore bypasses any OS-specific DNS caches, such as are present on Windows-based

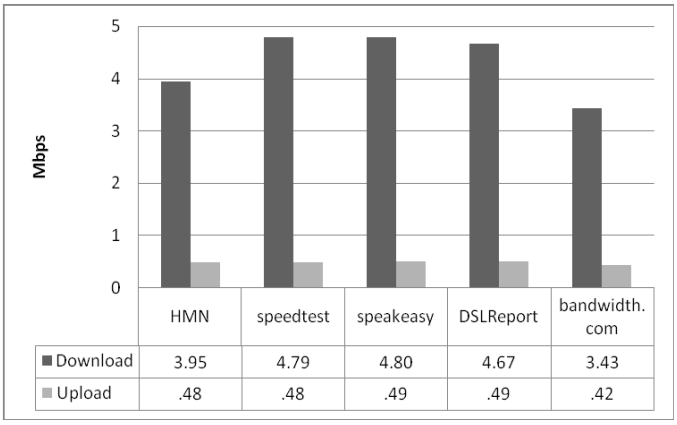

Figure 2.4: Throughput of HMN vs. Popular Speed Testing Services

Operating System machines. As described in Section 2.4 four types of DNS performance tests are conducted:

The first DNS test retrieved the A record for a server name then did a subsequent retrieval for the same name to measure the lookup time for a cached entry. Even for cases where the local network access point was configured as the primary DNS server, these requests are still passed through to a local DNS server of the ISP, which is caching the results of previous queries. Figure 2.5 shows the RTT results for cached queries of the 59 sessions in Table 2.1. The results in the figure show that the median lookup time for three of the ISPs is on the order of 20ms, although users of the Cable1 ISP typical provide the worst cached DNS performance. Worst case results are on the order of 150ms.

The DNS performance for a set of unlikely used servers is examined next. The servers are compiled from a list of 4000+ `.edu` sites. From this list, 25 random DNS requests are used as part of each user test. The average RTT is determined for the first test within each session with a CDF of these averaged results shown in Figure 2.6. As expected, these results show much higher RTTs than for the cached results of Figure 2.5 with median values between 100 and 150ms for the ISPs. 10-20% of the average values are over 200ms indicating much larger individual lookup times.

The DNS lookup times for 100 popular Web sites [4] was examined, with a CDF of

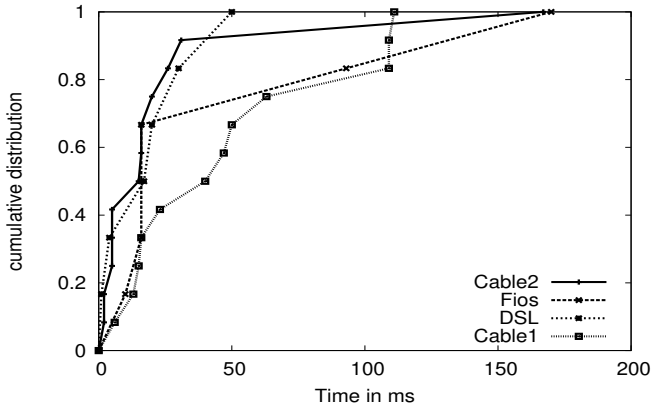

Figure 2.5: CDF of DNS Cached Entry RTT per ISP

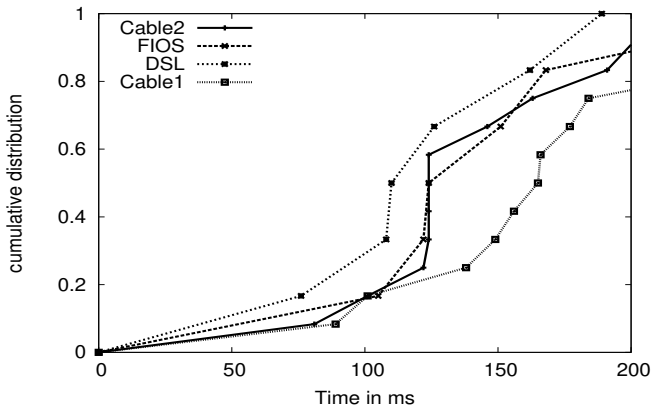

Figure 2.6: CDF of Average DNS RTT for 25 Random DNS Queries per ISP

the average for these results shown in Figure 2.7. Figure 2.7 indicates that many of these entries were already cached on the DNS servers as the median times are near the values for cached entries shown in Figure 2.5. Despite the relatively low lookup time in most cases, Figure 2.7 shows a small number of cases where the average across 100 servers is still large, again indicating some much larger individual lookup times. These results at the upper end require further study when all individual DNS results are recorded.

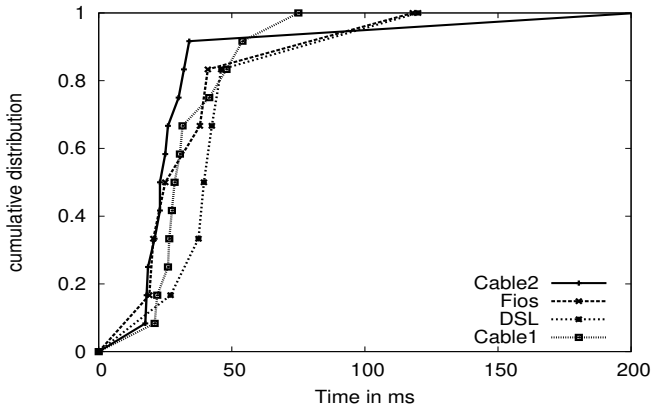

Figure 2.7: CDF of Average DNS RTT for Top 100 Queries per ISP

The final set of tests examined the performance of the local DNS servers to look up top-level, such as `.com`, and second-level, such as `wpi.edu`, domain names. These tests were conducted by generating invalid first- and second-level domain names that force a lookup to a root and a gTLD domain server. Figures 2.8 and 2.9 show CDF results for first- and second-level domain requests. These results show that the RTT from the client to the TLD and gTLD DNS servers is less than 100ms in most cases. While it is expected that TLD servers are not frequently queried, the gTLD servers must be queried for each new domain name that is encountered so performance is important.

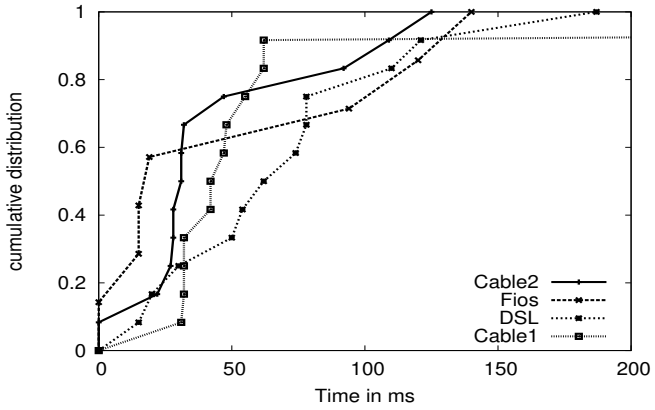

Figure 2.8: CDF of First-Level Domain RTT per ISP

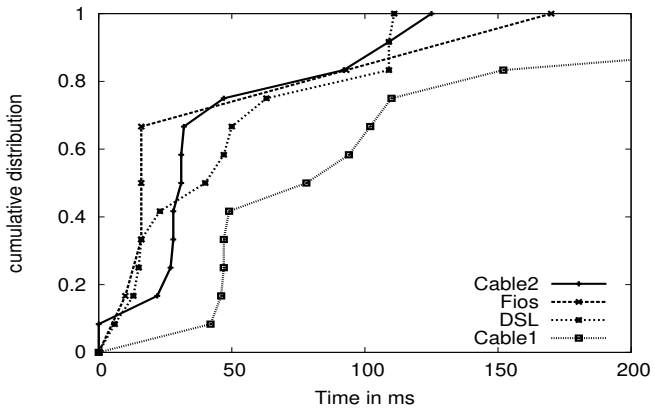

Figure 2.9: CDF of Second-Level Domain RTT per ISP

2.6.5 Local Network Environments

The last set of results use the methodology described in Section 2.4 to determine the number and type of devices on the 36 residential networks in our study. Home users have an average of three active devices. The typical scan returned the following devices: PC, router (where Cisco is the most popular brand), and a broadband modem (again where Cisco is the most popular brand.) Other HN devices detected ranged from gaming consoles (Nintendo Wii, PS3, etc.), video recording boxes (TiVo, Slingbox, etc), printers, hardware-based routers and switches (Cisco, 3Com), along with Windows and Linux-based PCs. Figure 2.10 shows the distribution of the most popular system types found (based on OS (Windows/Linux), and networking hardware).

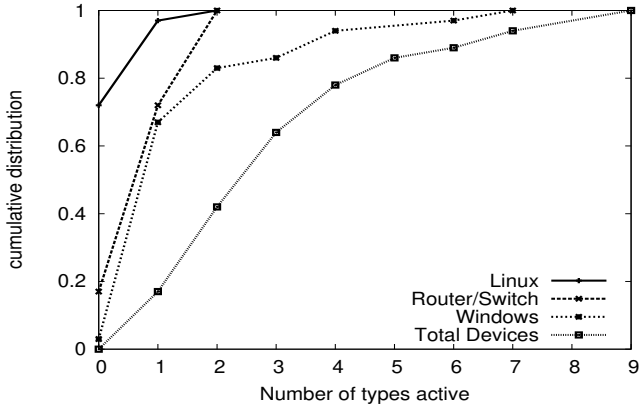

Figure 2.10: CDF of host types found during HMN scans

Table 2.3 shows the same data about devices as a percentage of the total number of devices and the total number of users. The results show that 52% of the devices are machines running a Windows OS with such a device present in 97% of networks for our users. Smaller percentages were found for specialized devices such as game consoles, digital-video recorders (i.e. Tivo) and printers, although these numbers are likely conservative as devices need to be active at the time of the tests in order to be detected.

Table 2.3: Devices of Home Users Participating in Study

Device Type	% Devices	% Users
Windows Machine	52	97
Network Device	33	83
Linux Machine	9	28
Game Console	2	6
Tivo	2	6
Printer	2	6

2.7 Summary

In this work, we have built and demonstrated the capabilities of using a signed Java applet for network measurement. This approach is particularly appealing for home network measurement as a signed Java applet can be run on a user's machine without permanent installation of any software, and thus minimizes impediments and leverages results as an incentives for participation. In combination with user-oriented feedback this approach can broaden the set of users employing such a tool and allow researchers to gain valuable insight into home network environments.

Our initial work has successfully deployed an applet and used it to gain information about the configuration of a user's testing machine, wireless connectivity, available upload and download throughput, DNS performance and the number and type of devices on the user's network. Results from an initial set of users both provide data about wireless connectivity of home network environments as well as the number and type of devices on these networks. The results also allow comparison of upload/download throughput and client DNS performance.

The takeaways from this work include the following:

1. creation of a core measurement component of a future user-centric network measurement platform, as part of the How's My Network infrastructure, which offers incentives and minimizes impediments for user participation;

2. demonstrate the viability of using a Web browser for obtaining network performance information;

3. discovery of wired and wireless information in a HN;

4. measurement of DNS performance via a web browser; and

5. the ability to learn about networked devices on residential networks.

Chapter 3

Peering into the Home Network

In this chapter we move into a review of entities that exist in Home Networks (HNs), by peering into the Home Network, and is key background work for the dissertation. This study was performed in 2018 and is a complimentary and additive to the Java approach research (Chapter 2), which leveraged low impediments and incentives toward participation. We look to understand approaches that exist, within the HN Ecosystem, along with their incentives and impediments. This work is a continuation of our HMN framework with a primary focus toward understanding incentives and impediments for HN users participation and how we can leverage different approach types along with their management, and configurations. In this, and the next several, chapter(s) we continue our research into incentive and impediments and describe how it is relevant to this research and overall dissertation.

3.1 Introduction

Previous studies have examined networking software approaches running on PCs and hardware, but there has not been a study that specifically identifies software tools and

their provided data targeted to Home Networking. We are interested in understanding the complexity of these tools, and the required expertise to execute and configure as well as impediments to participation. In this chapter, we look to understand and compare measurement points (MPs), applications and expertise required across four tool approaches: Routers, apps, Hardware, and Web/Scripting tools. We also examine the data of interest provided by each approach. We show that focusing on a broad range of approaches, data of interest, and tools allows us to create a new taxonomy of Home Networking functionality.

Over the past twenty years, the most widespread approaches to network discovery and research of Home Networks (HNs) have been to leverage physical hardware platforms to scan, and determine network flow using measurement points (MPs). MPs are the nodes, devices, or software where measurements are made [140]. Using protocols such as INM [23], Cisco Discover Protocol (CDP) [31], neighborhood awareness networking (NaN) [23] and others, have been leveraged to understand networking aspects of HNs.

In this chapter we review what information routers, software applications (apps), customized hardware, and Web/Scripting-based tools can determine in HNs. Understanding how we can leverage these approaches provides researchers, and more importantly HN users, results around local and global norms, as well as configuration of HNs and a new taxonomy system of information.

We refer to 'Tools' as software and applications running on a device. Tools can provide a plethora of localized information by peering behind the HN router using several approaches, in an attempt to determine and characterize configurations. These approaches range from modified and un-modified routers, apps and software tools, customized hardware, and Web tools; all which use active or passive techniques and applications while executing. These techniques include those that look to peer behind the HN router to understand layout, configuration, and historical norms of the HN.

As part of this study we look to understand how these techniques compare in terms of pros and cons of research, user incentives and impediments, and data of interest to HN users. This review focuses on the following approaches (exclusively): HN routers, Mobile and PC apps (running locally and in the HN environment that scan and analyze HN configurations), customized hardware residing in the network that is directly connected to the HN for analysis, and passive and active techniques used by Web and Scripting tools and which executes within the local HN environment. We have focused this review in these areas as they cover the range of hardware, customization, software, and active/passive techniques one can leverage for HN informational scanning, and also have had some review previously.

As a starting point, we examined HN usage and have found the following data points. A review from 2009 of HN usage, across the US, found that 63% of homes had broadband Internet connectivity, and 50% have a HN [49]. A PEW report from 2017 found that 73% of all homes now have broadband Internet connectivity, and most of these homes have a HN [129]. Projects such as How's My Network (HMN) [140] used a software approach to peer behind these HNs to understand devices, and characteristics of a HN using a low footprint and minimal impediment methodology to incentive end-users to participate. Other work in this area has focused on understanding and characterizing Internet access (bufferbloat) and device evaluation via a heavy-weight client [97][38], while other work focused on hardware approaches to understand HN configurations [90][22]. The work in these areas were primarily interested in local historical norms of information, and data gathering.

We examine data of interest that are part of the approaches we are interested in studying, along with the tools that classified within these approaches. The data of interest include: Throughput, Networking Characteristics, Health, and Historical Norms, as well as each underlying attribute. We have reviewed the data of interest that each of these

Approaches support, so that we can classify what each of tools cannot, could do, and do. We also look to understand the differences these approaches provide in terms of data collection and dissemination of information. We are focusing our review on these Approaches as they cover the range of hardware, customization, software (applications), and active/passive techniques leveraged in typical HN studies.

As part of this work we are also interested in health of devices (are they operating under normal parameters), applications, and protocols. This includes the health of DNS, security and privacy of devices and local configurations. The attributes of health include tools that examine configurations, normal operation, as well as the security and local device privacy. As a noted, apps running tools described in this review have implications in these areas for both the local and network users, including potentially users residing on the probed network; these concerns are typically in the form of security and privacy implications. As an example, a tool that determines network devices on a network has possible security and privacy implications as it can determine and potentially track local hardware (via MAC) and fingerprint services overtime. This data is readily available using the techniques we describe in this report. We have created a section around security, privacy and health to understand these impacts and the type of information gathered.

Our contribution to this area includes a taxonomy of tools that fall into the approaches studied (Routers, apps, Hardware, and Web/Scripting) and data of interest, difficulty level (incentive and impediments), and data availability (historical norms). The new taxonomies provide a classification of models around approaches and incentives, impediments, Source, Customization, and historical norms, along with data of interest; where norms has had little to no research up to this point. We conjecture that understanding local and global norms provided help users, and researchers, recognize the importance of distributed data sources and provide the incentive needed for participation in studies. In addition, we conjecture that understanding difficulty of applications usage and data

provided can be the impetus for users to participate in HN studies.

3.2 Background and Related Work

We provide background and a review of previous studies on the Approaches, Data of Interest, and networking as part of HNs. This includes approaches and data of interest that fall into commercial (pay for tools or advertised tools) and research tools, and we look to understand how they fit into the areas of security, privacy, health, and characteristics of devices in HNs.

3.2.1 Approaches

In this section, we provide background on approaches and data of interest we are looking to understand. This includes a review of studies on approaches (Routers, apps, customized hardware, and Web and Scripting tools), again we refer to the applications or the software running on these approaches as ”Tools”.

3.2.1.1 Routers

The following is a summary of information a typical HN maintains or includes (minimally) as part of its execution. A HN router has full access to the network and can gather information directly from Layer 0. A HN router can sniff traffic similar to apps running root privileges or a heavy weight application using Packet CAPturing (PCAP) libraries; note that PCAP requires root privileges to execute or a specific kernel and modules [144].

A router has full access to the network and devices that are plugged directly into the environment. Typically, a HN router saves the routing table that consists of: MAC address, the IP address that was assigned to your computer, and the lease time of your computer’s IP address; it also stores user-configurable items as well (port forwarding,

etc.) In addition, manufacturers are starting to create Mobile apps that control router access so that users do not need to login to the router via a web browser. These apps are still in the starting stage and provide local information with very little norm overview, and certainly do not provide a global purview of information. Routers need to have logging enabled to store even minimal information, and this setting can be disabled (in error) by users during setup. Deep packet inspection is not a feature routers typically support, out of the box, and need to be ”rooted” with tools such as WRT-DD [36] or similar software to allow these types of features.

As routers are responsible for making decisions about which of several paths network (or Internet) traffic will follow, there main purpose is that of flow-control. As an example, if more than one path is available to transmit data, the router is responsible for determining which path is the best path to route the information. The Function of a Router is to also act as protocol translators and bind dissimilar networks. Routers limit physical broadcast traffic as they operate at layer 3 of the OSI model [100]. Routers typically use either link state or hop count based routing protocols to determine the best path. The Role of a HN Router has not changed much over the past 20+ years and are still found to deal with layer three of the OSI model (network layer). This means the hardware device has full access to all devices flowing through its traffic control ports, but does little else for HN research, as found from [26][56].

A study on routers found that there is not much data stored on the router over time. However, different routers can and potentially does store different data. As an example, data consisting of the assigned IP address of the connected computer, the computer name (or nickname), the MAC address of the computer, and the total time that the devices have been connected to the router [141].

HN routers have similar properties where they serve up local network traffic, WiFi, and maintain lists of information about what is on the network. Hardware vendors differ

on what they provide for information, but typically contain the following services: DDNS (Dynamic Domain Name System) is a service that allows network clients to connect to the wireless router, even with a dynamic public IP address, through its registered domain name. As an example, a wireless router, embedded with the ASUS DDNS service and other DDNS services, also supports the mapping of hosts, ex: (showing client name, IP, MAC, and Interface type). The router provides network flow (netstat information, ping, traceroute, and name server lookups via command line tools). A router can and does provide information about the local HN, but does not (typically) have a purview into applications and types of tools running; these require a modified/rooted router boxes and expert knowledge of networking and IT infrastructure, as noted by [141].

Research using modified home router software and tools have been completed using the toolsets WRT-DD [36], and Tomato firmware and configurations (or similar). Research by [24] found that a user must have high level of domain knowledge to work in these challenging domains of rooted environments, and went on to claim that users with minimal experience should "stick with the stock router firmware." Other work in this area includes Bufferbloat analysis by [99], performance analysis of home routers [64], identifying lurkers in social networks [161], and throughput performance [85], where these service and activities have all been done at a local level. Other research using modified routers with WRT-DD (or similar tools [Tomato, etc.]) was research to help understand and control network flow, wireless access, and discovery [141][16][89][35].

Other studies of routers found that understanding the causal impact of the different performance metrics around network performance is the only quantitative way of making such trade-offs of providing valuable data [150]. This study showed that a range of routers provided the following information: system status (CPU, RAM, and logs), Wifi networks under its control (3G and 5G for example), and by using a "rooteed" router and DD-WRT firmware allowed researchers to control the devices similar to a Linux box (as it is

a Linux-based firmware), and collect, modify, and accept/reject streams via a very terse command line interface, and scripting tools.

An example screen-short of information from a commodity and stock router can be seen in Figure 3.1, and includes the following information: device type, client canonical name, local RFC1918 IP address, MAC address, TX/RX information, and amount of time on the network.

Figure 3.1: Example Screen Shot of a Router Configuration

3.2.1.2 Apps

What is an App? According to [133], the term Software Application first appeared in the early 1950s and a Software Application, or app, is a program or group of programs designed for end users [13] that executes on a device or piece of hardware.

3.2.1.2.1 Mobile App A mobile app, as described by [169] consists of a software program that is targeted for a specific hand-held or mobile device type, e.g. Android platform. Tools such as [27][86][98][116] provide security, password and threat prevention via a mobile app, and functionality for a specific niche or data of interest.

3.2.1.2.2 Java App A Java app or Java Applet is an application, which runs in the Java Virtual Machine, which interprets instructions and executes on the system (hardware) that the application resides on. A Java program executing in a native environment has full access to the system resources, depending upon user privileges of course. A Java Applet executes within a browser typically (or similar restrained environment) and depending upon the nature of the execution has limited access to the resources of the system and executes in a sandbox (similar to JavaScript code). In contrast, a signed Java Applet has a fuller set of access to the resource it is running on. Our previous study [140] on HNs dives into a methodology and framework using a signed Java Applet and had promising results.

3.2.1.2.3 Techniques Used by Apps The following are some of the popular techniques used in commercial and freely available tools. Commercial tools such as fing [54], are available via a mobile app and hardware to help understand device mapping and network layout via the hardware; an additional purchase of hardware and licenses are required for these features. Other commercial tools such as [114][179][181][63] provide users similar information as fing, where none of these tools provide historical global norms.

Custom and UNIX (and other OSs) networking tools such as netstat, NMAP, and similar have a high barrier to entry and include techniques such as [74], which requires dedicated hardware and in-depth system administration skills to operate. Approaches such as Kermit [30] and Microsoft's HomeOS project [14][41][40](established in 2010, now defunct) have attempted to use software models which requires a dedicated PC to run on. As an example, the HomeOS project was built with the premise that the Home needs an Operating System, this approach has been all but abandoned. Other examples of app approaches include Ph.D. thesis work from [188], which claims to have minimal

impediment to operate, but requires customized hardware running on a PC and extensive networking knowledge to operate.

There have been studies that have attempted to understand the layout (sketch) of the HNs, such as [131][61], as well as HCI (human computer interaction) studies around which techniques are useful to determine devices and resources in HNs, which have been unsuccessful as they have not provided a context of availability. An example of this is the Kermit [29] study, which attempted to map HNs broadband connectivity. A review by Grinter, et al. [62] attempted to use surveys to understand how HNs are setup and had similar results as [61], which found that participants needed technical knowledge to diagnose and deal with networked technologies, and that they turned to friends and family for help. This message has been re-iterated by the study from [130] where they found that the promise of future applications rests on the ability of house-holders to manage the home network, something that our collective research shows has not become easier since the first reports of connecting computers to the Internet. Furthermore, a previous study by [134] found that home networking is nontrivial for even the most qualified, and contend that these problems will not disappear over time as the networking industry matures, but rather are due to structural usability flaws inherent in the design of existing network infrastructure, devices, and protocols [147]. A study by Yiakoumis, et al. [186] looked at extending this type of work and came up with the concept of slicing the home network, in an attempt to understand the landscape of the HNs from an app and hardware perspective.

The dynamic adaptive streaming over HTTP (DASH) work and others in QOS (quality of service) and HN [96][159][158][174][187] have looked at performance of HN and content delivery, while [28] looked at benefits of Software Defined Networking (SDN) to see how they can help improve manageability. Studies have approached HN review to include load time of objects to determine under-performing content [55], while others have used DASH to understand video streaming in HNs [95].

Commercial tools (pay to use and requires a license), such as [145, 54, 112, 178, 182, 180, 105], are available that help understand the presence of devices on the local network, security, privacy, or use Wifi to determine least crowded channels or discovery using Bluetooth, and provide little historical data for local configuration and global norms. While these techniques provide specific data sets around BufferlBloat, DASH, throughput, or point and time information, none of the tools reviewed provide historical global norms of HNs nor configurations of these HNs.

There are commercial apps that are available for both Droid and iPhone that can scan Wifi networks (similar to HMN), and determine open ports, but this area not well researched nor does it have available data sets for researchers to review. These Mobile apps have limitations on accessing the TCP stack, and therefore cannot provide details that tools such as nmap (for the PC) can provide (active fingerprinting), noted by [94][125], and require elevated privileges to execute. The following are some apps reviewed and are either commercial or research and are the most applicable in terms of providing networking information:

- Netalyzr [97] provides information around Bufferbloat, which has been well studied, as well as general internet upload and download throughput.

- fing [54] is a mobile app which scans for local Wifi connected devices. An interesting feature is the ”Enable device recognition”, which requires remote best-match brand/model of device, assumedly via Mac address. This is an opt-in request that the user must click ”Enable” to allow access to the remote querying fingerpedia. This tool requires users to enter information about host, and can run a simple scan of open ports on the given host; it does not predict the type of host (ex: Linux RedHat, Microsoft Windows 10, etc.). Each scan is manual, and is required to be run explicitly by user requests [53]. There is an optional hardware device for lower

layer analysis scanning and reporting.

- Mobile NMAP [115] provides similar information to the PC counterpart, but it limited in terms of resulting scans (i.e. no OS fingerprinting, SYN scan, etc.). NMAP does not explicitly provide global norms, and requires a high level of knowledge and expertise to operate.

- Tools such as Wifi Analyzer, Wifi Master, and Wifi Connection provide Wifi channel information in an attempt to show the least crowded connection (channel) to the router [178][181][179]. These tools do not provide historical global norms, but may provide best practice information on management.

- Other tools such as Meshlium use Bluetooth and Wifi signals to identify devices in a given area; these are commercial products, which typically require a hefty upfront cost and monthly subscription, such as [105][117]. These tools are localized only and do not provide historical global norms.

- Rooted tools, such as from the review by [144], can provide a deeper view of things, but have a higher barrier to entry (requiring a fat client to be installed on a PC and customized hardware) and are not targeted the novice user. These closed source tools do not provide global norms.

- There are other commercial and open apps, which provide network scanning, Wifi, and upload and download information, but do not provide local and global norms nor provide the breadth of data across the range of spectrum users may be interested in.

3.2.1.2.4 Health and Apps

There are several apps which look to determine health (including security and privacy) by

examining the local host (PC or Mobile) run time nature, and configurations. Tools such as [5][20][153][34][163][122] look to understand mobile and PC security around virus protection, remote theft, safe browsing, SMS, encryption, proxy, and tracking. These tools use virus definitions, GPS location services, phishing definitions, encryption techniques (such as twofish, blowfish, and others) to encrypt applications and text messages, as well as triangulation (using Wifi and GPS) for health, security, and privacy.

3.2.1.3 Customized Hardware

Customized hardware consists of devices packaged (or not) with an app [117][54], network security devices, IoT devices such as sensors, automation devices, and network devices modified to allow for access control [36]. These devices use pass through features and remove the HN router from the network [117], or act as the primary Wifi connection for the network, thus routing all packets through the customized box. These devices serve as active and passive monitoring for security and remediation, device look-up and traffic control on the HN. As an example [117] algorithms to pre-execute control techniques to traffic flow, in conjunction with cloud-based inspection. These hardware devices are either highly customized routers/Wifi units [117], or expensive commercial products [43][105] targeted for specific tasks (e.g. security, sensors, discovery, etc.)

Most research-based hardware devices consist of modified routers, and software packages leveraging cloud-centric analysis. Commercial hardware approaches typically consist of devices packaged (or not) with an app [54][117], network security devices, IoT devices such as sensors, automation devices, and network devices modified to allow for access control [36]. These devices use pass through features and look to remove the HN router from the network and act as the endpoint security connection or as the primary Wifi connection for the network; thus routing all packets through the customized box. These devices serve as active and passive monitoring for security and remediation, device

look-up and traffic control on the HN. Each of the following hardware solutions require a monthly subscription to leverage security and privacy features, and thus have a barrier to entry.

- Bitfender [117] uses machine-learning algorithms to pre-execute control techniques to traffic flow, in conjunction with cloud-based inspection. The hardware is highly customized router/Wifi unit [84] and requires an additional hardware unit to replace HN units along with a monthly subscription [44][105] targeted for specific tasks (e.g. security, sensors, discovery, etc.). This device does not provide global norms.

- Cujo [45] is an inline device plumed into the Ethernet of the HN. The device monitors traffic for security threats, and looks to prevent sensitive data from leaving the HN. It is not clear if this devices require a switch that mirrors all ports to gather data, as it is plugged directly into the HN router. The device does not provide global norms.

- Dojo [42] is also plugged directly into the HN via a port on the router, but examines metadata versus full stream to determine actions. It is also not clear if this devices require a switch that mirrors all ports to gather data, as it is plugged directly into the HN router. The device does not provide global norms.

- Keezel [88] connects to the network purely via Wifi, acting as a hotspot and flow through using VPN-based technology. This device does not protect hardwired devices (unless direct mirroring is created on the router), and does not provide global norms.

- RaTTrap [135] is directly plugged into the HN modem, and the HN router is plugged into the device. This allows the device to examine all network flow inbound and outbound and it looks to block malware and other security threats.

- We also list Fing [54] device in this section as it is provided either as a software or hardware/software solution. As previously mentioned the FING tool does not provide distributed data from global norms, as it is a commercial closed app/hardware tool.

3.2.1.4 Web Apps and Scripting Tools

Tools running in a web browser and scripting tools, such as JavaScript, HTML, Python, and other similar languages are not required to be complied and are strictly speaking interpreted [183]. These tools, when executed from a Web Browse or similar environment, run in a sandbox and are allowed minimal access to system level resources.

Web or Script based approaches leverage either a browser or command line interpreted tools such as Perl, PHP, HTML5, and shells such as BASH. As an example, these tools look to query for: hardware devices, software running locally, OS and security settings, Active Directory (AD) configurations and settings, local web service settings, and local user and group information. To gather much of this information they must be installed via administration/root privileges locally or the collected information is minimal when running via a web browser. It should also be noted that unless these tools are run directly within the HN they would only provide scanning information from the edge of the network, as they cannot peer inside the HN and behind the router remotely.

- Open-AudIT [171] was released in 2002 and is targeted at system administrators who have deep knowledge of Linux or Windows systems to just install. The tool when run as root can collect network information such as hosts, MAC information, and when configured with NMAP a deeper dive of devices using NMAP fingerprinting. This tool does not provide global norms, and can be a challenge to configure and run.

- Spiceworks [155] is a Web-based port scanning tools, but can only determine edge information and cannot peer inside the HN.

- Pentest-tools provides TCP (and UDP) Port Scan with Nmap [162] via a web browser, and leverages the Nmap tool to collects data from their server running Nmap fat client. This tool does a minimal execution of Nmap or a bit more passive scanning. The tool leaks information around the location of the server running the scan by listing the time zone and time of the scan in GMT time. This tool does not provide global norms.

- Mxtoolbox's [109] provides similar functionality as Open-Audit and runs an open TCP connection via port request to the edge device on the HN. This tool does not provide global norms.

- How's My Network, predicating performance from within a Web browser sandbox [84] leveraged scripting tools run via a web browser for performance analysis in a HN.

3.2.2 Data of Interest

In this section we review data of interest and the attributes associated with each of these. The following are studies and other work that have done work similar to the data of interest we are most interested in. These tools determine the following high level of information as related to the approach and data of interest: Throughput, Networking Characteristics, Health, and Historical Norms. These data of interest individually provide a small amount of data, but tied together create a valuable picture of HNs and what is occurring in them.

3.2.2.1 Throughput

The area of throughput has had extensive studies in terms of: Upload, Download, Jitter, Network Flow, and Performance. Studies such as [140][84][74][183] looked to classify upload, download and Jitter by calculating changes in network traffic and differentials of time. While [97] looked at network flow and jitter in terms of bufferbloat and delta changes in traffic. Work done by [115][41][159] looked at local network performance of HNs to understand traffic flow and overall performance, other studies such as [150] looked to understand if performance matters in the HN. These and other studies provide background needed to tie together throughput of HNs to create a concrete picture of what is occurring in HNs.

3.2.2.2 Networking Characteristics

In the area of Networking Characteristics we are interested in the attributes such as: device discovery, network fingerprint, Wifi network discovery (online and previously attached to). These areas help bring together a picture of devices, and activity occurring in HNs. Studies such as [140][115][54][114] and others provide device discovery and network fingerprinting by using well known broadcast services, TCP evaluation, and port mapping to evaluate devices available on a network. Wifi tools such as [181][178][179] and others provide similar analysis for Wifi networks to understand a mapping of Wifi, radio and communication channels, as well as historical information of devices attached and previously attached to a Wifi network.

3.2.2.3 Health

Health of networks, applications, and devices includes the following attributes: DNS, apps, Security and Privacy (e.g. apps, location, monitoring). Studies such as [140][115][68][83][47] looked to understand 1st, 2nd, and 3rd tier DNS results, SOA requests, recursive requests,

as well as security and a variety of approaches to health and the local and remote DNS services. Research into apps (running on PCs or Mobile devices) have looked into options of security and privacy [45][42][88][135] on local devices. While research and tools such as [117][10][43] look at hardening security and privacy of the device and the network. While some of these tools require expert knowledge of networking and security, others look to harden and quarantine network flow, file access, network access, and app access.

3.2.2.4 Historical Norms

As previously discussed Historical Norms provide information on: Local Norms, Global Norms. While most apps and research provide local norms, there are only a few apps or research projects that provide global data norms for users (and researchers) to understand a big picture of what is occurring across disparate HNs. As an example, the work done by [140][84] provides results of throughput and networking characteristics of both the local and global norms for comparison.

3.3 Methodology

In this section, we provide the methodology used as part of this review. We start with the methodology that was used in this tech report, and then turn to why this research matters and look to understand optimal setup and how a user or researcher can mimic this in a HN.

We reviewed research papers, tech reports, and commercial products and then compared the methods used, and results provided by the given work. We have classified these, and looked for overlap and similarities, differences of results, as well as the methods used to collect and display the information. This research included an unbiased review of the Approaches, Data of Interest, Historical Norms, data collection techniques, and results.

From these results we turned our attention to what types of data results users and researchers are interested in and looked to combine these differing Approaches, Data of Interest and Historical Norms. The data of interest included data both currently collected and not currently collected by these works, along with how to display this data using different HCI approaches.

To understand the data of interest provided from these approaches we have download tools, reviewed papers, and dichotomized the results and methods of the app/tool. We looked at the results of these apps/tools and include a data of interest set discovery and attributes, including: throughput, devices, health, security and privacy, along with historical norms.

In addition, we look to understand the incentives, and impediments to each approaches and classifications around the areas. Incentives and impediments are focused on the user experience and more importantly the perceived value of the tool and barriers to entry respectively. An apps approach may be free to users, and require minimal impediment to install, configured, and execute and thus have a lower barrier to entry.

We have reviewed the following areas to understand why this research matters to the user and the research community. As previously mentioned, it is clear that the work done by commercial, discovery, and tech reports provide a clue to what users may be interested in. This includes the desirable areas of throughput, network characteristics, health, security and privacy, and most importantly how information collected compares at a local and global norms. An optimal setup for users to execute this work would be a collation of the data of interest into an app, along with information gathered from historical norms. With that said, these approaches can be completed using physical (inline) hardware or via an app.

As mentioned, the method used by this study included an in-depth review of research papers, commercial and openly available applications/tools, underlying protocols, imped-

iments and incentives, and a comparison across approaches and data of interest. We have examined a broad range of software and hardware as well as the Data of Interest that are part the Approaches that we are interested in. We first look to create a dictionary of terms and definitions for this study to clarify the method and have created the following definitions for this study, which include:

1. Approaches: Hardware and software that provide a plethora of localized information by peering behind the HN router using several approaches, in an attempt to determine and characterize configurations. We refer to the Software/applications running on these devices as 'Tools'.

    - Routers: Stock and Customized
    - Apps: Mobile, Java, executable(s)/binaries
    - Hardware: Customized hardware installed in the HN
    - Web/Scripting tools: Software running within a web browser or via a scripting run-time

2. Data of Interest: the gathering of desired data collection from the user perspective. We have included the attributes of each of these data of interest, which are the data points collected by the tools.

    - Throughput Attributes: Upload, Download, Jitter, Network Flow, and Performance
    - Networking Characteristics Attributes: Device discovery, Network fingerprint, Wifi Network discovery (online and previously attached to)
    - Health Attributes: DNS, apps, Security and Privacy (e.g. apps, location, monitoring(
    - Historical Norms Attributes: Local Norms, Global Norms

We used the following method to classify and understand the differences within and across each of these Approaches and Data of Interest. We compared each of the tools that execute across these Approaches and Data of Interest to understand the quality of information (high and low), incentives to execute (richness and quality of data), and impediments to entry (easiest to hardest). We next extended these comparisons, uniformly, across each of the approaches and data of interest (grouped by approach and the Tools), and used the following system to help understand the data of interest of each of the Approaches.

- Cannot be done: The data of interest does not have the access to this type of information.
- Could be done: The data of interest can do this.
- Done: The data of interest is supported by an application within this approach.

A comparison was also completed by reviewing the set of tools (arraigned by Approach type) and comparing by the following measurement of approaches, data of interest, and tools: incentives vs. impediments, sources and customization vs. data collection set, and historical local norms and historical global norms. We used the following method to understand these three measurement comparisons and differences, and used the data sets compiled as part of this study. We refer to this as user participation data points.

We compared incentives and impediments to each other and created a classification of paradigms to understand easiest and best vs. the hardest and least ranked tools. The following was used to understand Incentives, and Impediments of each Tool and thus Approach.

User participation: The following are the comparison areas of user participation data points: incentives and impediments, sources and customization vs. data collection set, and historical local norms and historical global norms. Incentives and Impediments: incentives and impediments are focused on the user experience and more importantly the

perceived value of the tool, and the barriers to entry:

1. Incentives:
    - None: No incentives are offered to participate in a study or are provided by the tool.
    - Low: Minimal amount incentives are offered by to participate in the study or information provide to operate the tool.
    - Medium: Incentives are offered that provide users a reason to want to participate or operate the tool.
    - High: There is a high amount of incentives offered to participate or the tool offers a wide range of information.
2. Impediments:
    - None: There are no impediments to operate or participate
    - Low: There are minimal impediments to entry
    - Medium: The impediment to entry is challenging and requires monetary or skill level to operate
    - High: Barrier to entry includes monetary or expertise to operate

We used the following for the comparisons of Sources and Customization vs Data Collections used the following classification.

1. Sources and Customization: is the platform open or restricted in terms of modifications and changes, including sources
    - Closed: No changes are allowed, and sources are not available.
    - Restricted: Minimal changes are available to the configuration or the sources

- Open: A wide variety of configuration options are available to modify, and sources are available.

2. Data Collection: the types of information provided by the given Approach/Data of Interest.

    - Restricted: A restricted view provides closed and minimal information
    - Flexible: A flexible view provides some modifications for request to wide range of information.
    - Open: An open view allows for low level modifications and access to configurations for customized set of information

Similarly, we reviewed and compared historical norms (local and global) in terms of data sharing and availability, and used the following classifications for this comparison. These are classified as Historical Norms: (either local or global)

1. Local Norms: The types of data the tools provided at the local network level, and if there is a long range or a point and time comparison of this information.

    - Closed: Information is not provided by this tool
    - Restricted: Data is gathered over a given set of time, and provided to users for review. A limited amount of Information is typically provided by the nature of the product, and is typically point in time.
    - Open: Data is available on a wide variety of areas, and is flexible for types of information provided.

2. Global Norms: What, if any, information is provided, by the tools, for comparisons to users running these Applications at the global level, across networks and users. Data that is used to compare to other environments, which is running on the same Approach/Data of Interest type.

- Closed: information is not provided by this tool.
- Restricted: Data may be gathered, but is not provided for review.
- Open: A wide range of information is available by this tool.

3.4 Approaches and Data of Interest

In this section, we organize approaches and data of interest by collections provided. We start by looking at how each of these approaches are tested and classified. We have created an approach taxonomy of routers, apps, hardware, and Web and Scripting, as shown in tables[3.1,3.2,3.3,3.4]. These tables provide information around: approach, source, incentive and impediments, and historical norms.

As part of the review of these areas we have also created four tables focused on approaches and data of interest, including what sources and customization, data collection, and historical norms they support. The tables provided show the measurement approaches that each of these areas cover, along with comparisons of like types. The objective of these tables are to help understand what types of information each approach provide (to users), along with incentives, impediments, as well as location and metrics. We use the classification shown earlier to describe these areas:

3.4.1 Routers

As previously discussed, routers can provide the richest set of information, but can require a high level of expertise to operate. Table 3.1 to show information around approaches, stock (out of the box) router and a customized rooted router, source, incentive, impediments, and historical norms.

1. Source: both methods are closed, with the exception that the modified router has updated firmware which exposes additional functionality (e.g. custom control points).

2. Location: both methods are targeted at home and commercial networks

3. Incentive: a stock router provides access as its major incentive, versus a modified router that provides users, and researchers, to customize their networking experience and data collection points.

4. Impediments: a stock router requires a medium level of expertise to install, and configure, as previously noted, and its major impediment is around cost of the unit. A modified router is also has the impediment of cost, but has the additional requirement of expertise to install, configure, and operate as it requires a high barrier to entry.

5. Historical Norms

    - Local Norms: both methods provide local information about the network to users.
    - Global Norms: neither method provides information from other users experiences or feedback.

Table 3.1: Classification of Router Approaches, Tools, and Data of Interest

Router Tool Approaches				
Tool Type	Source	Incentive	Impediment	Historical Norms
Stock	Restricted	Access	Medium purchase	Local Only
Rooted	Restricted Modifications to firmware	Access Custom	Equipment Expertise	Local Only

3.4.2 Apps

We have created a similar comparison for apps in Table 3.2 from the following:

1. Source: HMN and Nmap platforms provide are available as open source

2. Incentive: All of the platforms provide feedback as an incentive, with the exception of Fing which is driven as a pay for service and ad-driven tool.

3. Impediment: A review of impediments across the platforms shows the following:

    - HMN, and Fing are the easiest to install and require the least amount of impediment to entry
    - Nmap and Kermit require administrative access to run and a PC and customization to run, and thus have a higher barrier to entry

4. Historical Norms

    - Local Norms: all methods provide some information to the end user over a given set of time.
    - Global Norms: The HMN approaches is the only tool that provides both local and global norms for comparisons.

3.4.3 Customized Hardware

We next turn our focus to devices that are directly attached to the HN and are specifically targeted around discovery services. As discussed these devices use similar methodologies as routers, as they are directly connected into the network with layer 0 level access. The information gathered ranges from device types, machine names, internal throughput and throttling controls, WiFi troubleshooting, and network security. These devices collect local information, but do not share local or global norms, and are classified as heavy-weight. Tools such as fing require a hardware device to be purchase to extend the information available on the network. These tools can leverage both software interfaces

Table 3.2: Classification of Apps, Tools, and Data of Interest

Apps Measurement Approaches				
Tool Type	**Source**	**Incentive**	**Impediment**	**Historical Norms**
HMN Java	Open	Feedback	Medium Approve app	Local Global
HMN Mobile	Open	Feedback	Medium Install app	Local Global
nmap	Open	Feedback	High Expert	Local
Kermit	Restricted	Feedback	High Special HW Expert	Local
Fing	Restricted	Feedback via pay product	Medium Can require additional HW	Local
Netalyzr	Restricted	Feedback	Medium Install app	Local

and hardware as they directly plugged into the main network (similar to the router) to determine device characteristics using similar approaches to that of apps.

We have reviewed two approaches, as can be seen in Table 3.3, Fing Hardware, and HomeOS. Each of these approaches shown in Table 3.3 are classified as customized hardware, as they are specific to device scanning and HN tools. They are typically paired with software or run can be run directly via the hardware devices (Web Browser). Some key points of this include:

1. Source: Fing and the HomeOs approaches are both restricted and do not provide open sources

2. Incentive: Fing and HomeOS provide feedback as the major incentive, but both require custom hardware to run and a licenses is required from Fing to operate.

3. Impediment: Fing and HomeOS provide device information, but the HomeOS tool

looks to provide access and control of IoT-based devices. Both methods have a high barrier to entry, as they require customized hardware to execute and a license to operate.

4. Historical Norms

   - Fing and HomeOS both provide local norms over time.
   - Neither Fing nor HomeOS provide global norms of other (user) experiences.

Table 3.3: Classification of Customized Hardware Measurement Approaches

Customized Hardware Measurement Approaches				
Approach Type	**Source**	**Incentive**	**Impediment**	**Historical Norms**
Fing	Restricted	Feedback via pay product	Medium Require additional HW	Local
HomeOS	Restricted	Feedback	High Requires Custom HW	Local

3.4.4 Browser and Script-Based Tools

Browser and scripting tools run directly within a web browser, and do not require the user to download tools to execute. The Barrier to entry is low for the end user to execute, but the tools provides minimal information during executions, due to the sandbox that it executes within. Tools such as speedtest [183][121] and [84] run within a web browser and use point locations (throughout the country or localized) to understand throughput and jitter to know resources.

Table 3.4 is a taxonomy of these approaches, and includes a review of generic speedtest services, and HMN sandbox methodologies. HMN is an open source approach to testing

versus speedtest, and they are both targeted to HN and Commercial networks for testing. They both are free in nature and provide feedback and the major incentive, and require minimal impediment to execute via a Web browser. These approaches provide similar results, but HMN provides both local and global norms.

1. Source: Speedtest tools are closed source versus HMN, which is open source and can easily be modified.

2. Incentive: Both methods have the user incentive of feedback of information to execute, and are typically no cost to execute.

3. Impediment: Both of these methods have the most minimal of impediments to execute, but provide the least amount of information due to the nature of how and where they execute, e.g. via a web browser.

4. Historical Norms

    - LocalNorms: Speedtest services provide point and time executions vs. that of HMN which can provide a comparison based off of previous tests.
    - Global Norms: neither method provides global norms, but the HMN suite does provide the ability to understand longitudinal information from the data gathered.

Table 3.4: Classification of Web and Script Measurement Approaches

Web and Script Measurement Approaches				
Approach Type	**Source**	**Incentive**	**Impediment**	**Historical Norms**
Speedtest Services	Restricted	Feedback	Low Minimal info	Local
HMN Sandbox	Restricted	Feedback	Low Minimal info	Local

3.5 Data of Interest

The following is a review of each of the data of interest and tools reviewed in this study, along with the merits in their own area. These merits are classified at the Approach level, and include the following:

1. The router approach allows for the customization of information, but requires a high technical barrier to entry for customization. Data that is gathered is point-in-time and is typically cycled over X number of days, but is not made available in a longitudinal approach to users or researchers.

    - A stock router is a utility approach to computing, and networking as its main focus is access versus information.
    - A modified router can provide the deepest dive of information from method of approaches studied. This approach can provide information ranging from performance to information, but as previously discussed has a high barrier to entry.

2. The apps approach is the most flexible as it requires the least amount of efforts to entry for the user, and can be customized to allow for both practical user, research and technical information, without the use of hardware. With an app, users have the ability to understand data for global and local norm comparisons. This is important as an app can be customized and include a Human Readable Format, including:

    - Device listing
    - Throughput (Up/Down), jitter, etc.
    - Performance of the device
        - Legacy information about the devices attached previously (assumed it was scanned)

  - Device was present, or is now present, and is now gone

- What is shown and how it is shown, over time.
- Apps provide methods to push updates to devices (e.g. phones), with minimal impediment to end-users.
- Apps can flexibly be customized in terms of configuration and results to an end-users perspective.
- Apps can provide devices and configuration of networks, in a local and global approach.
- Apps provide and understanding of protocols and configurations of these domains.
- Apps also have information with activations and Activity of the HN and the device running scans.

3. Hardware measurement approaches can have similar success as a modified router method, as they have the ability to sit inline of the router, and can analyze data in a similar manner that a router. These devices serve a specific function, and have a license and cost as a barrier of entry. While the nature of these services are to minimize impediments, they have a high barrier to entry and are tech savvy approaches. Information gathered from these approaches are quite specific, and are targeted at a specific product approach. Data collection in these approaches are targeted, and provide feedback to the local vs global instance only.

4. Web and Script Measurement approaches have the least impediment for barrier to entry as they can run via a web browser, or similar. They provide minimal information in terms of data discovery, as compared to the other approaches, and only provide a local point-in-time data set.

We next turn our attention to the comparison of approaches and what types of results, device characteristics, platforms, components, and services they fall into. The classifications of user participation in terms of what can be discovered using a Router, commercial tools, and other hardware devices is shown in Tables 3.1,3.2,3.3, and 3.4. These tables look to understand how they fit into the data of interest. The tables provide information around the following classification areas that are the most important as shown from the work reviewed, and previous studies. We are looking to understand the following data points and sub-items, as they appear to be the most commonly studied across the set of tools and research papers reviewed.

1. Throughput
2. Network Characteristic
3. Health
4. Historical Norms

3.5.1 Throughput

We have created a classification of information that each approach type collects in terms of throughput. We have classified throughput to include the following characteristics of the areas we are reviewing (routers, apps, Customized Hardware, and Web and Scripting tools), and includes the following:

1. Internet and local network Upload and Download: Throughput of the Internet connection and internal throughput.
2. Jitter: Deviation from optimal performance of a given Internet connection or the fluctuation of latency over time, and includes ping spikes and lag.

3. Network Flow: Diagnostic of network layers, including TCP flow.

4. Performance of device (perf) is diagnostic information around the performance of devices attached to the network.

We can see from Figure 3.2, that a variety of information can be gathered using each of these approaches, and that the apps approach has at least one tool in the list that provides this data set.

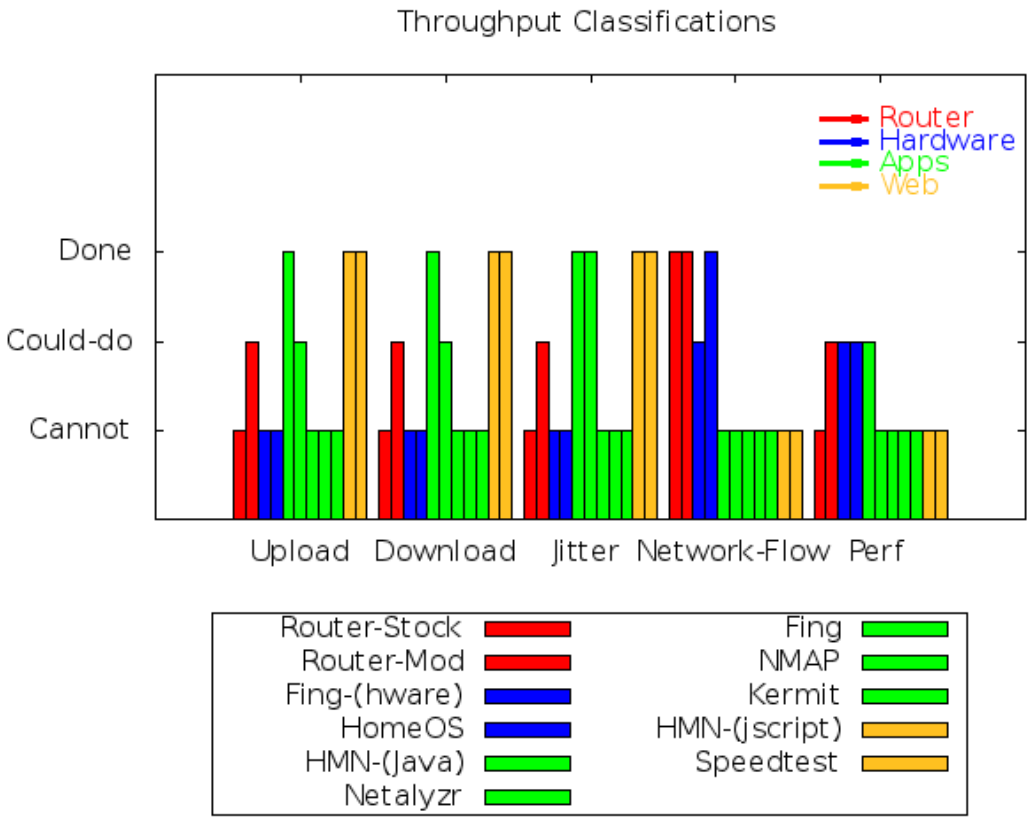

Figure 3.2: Throughput Classification

3.5.2 Network Characteristic

Tables 3.1,3.2,3.3, and 3.4 provide a review of network characteristics against the approaches we are reviewing. These include the following networking areas of review, and

what information can be collected.

1. Local device information: hardware, and software information for a given device. MAC address, name of host, networking information, and local software information (e.g. CPU, etc.).

2. Remote device information: fingerprint scan of network, including: TCP information, host and canonical names, and device types.

3. Application list: locally running software list

4. Network information: locally connected networks (e.g. Wifi, lans, etc.)

5. Wireless information: available and browse-able networks

In addition, Figure 3.3, provides a view of data points that are available using each approach. While a stock router may have access to most of this information, it does not collect or store these data points. Apps collect these data points across the set of characteristics reviewed. Hardware approaches are similar to the Router approaches, and typically do not collect all of these characteristics. Web and scripting approaches do not have access to gathering most of this information as they run within a sandbox.

3.5.3 Health

We move our focus toward health, and approaches used by the methods studied. Health includes local and remote networking, and Figure 3.4 shows health approaches by creating a classification of what can, and cannot be done in the following areas.

1. DNS Health: this includes a health of the DNS infrastructure in terms of networking, and reporting.

2. Legacy Information: what devices previously attached to it.

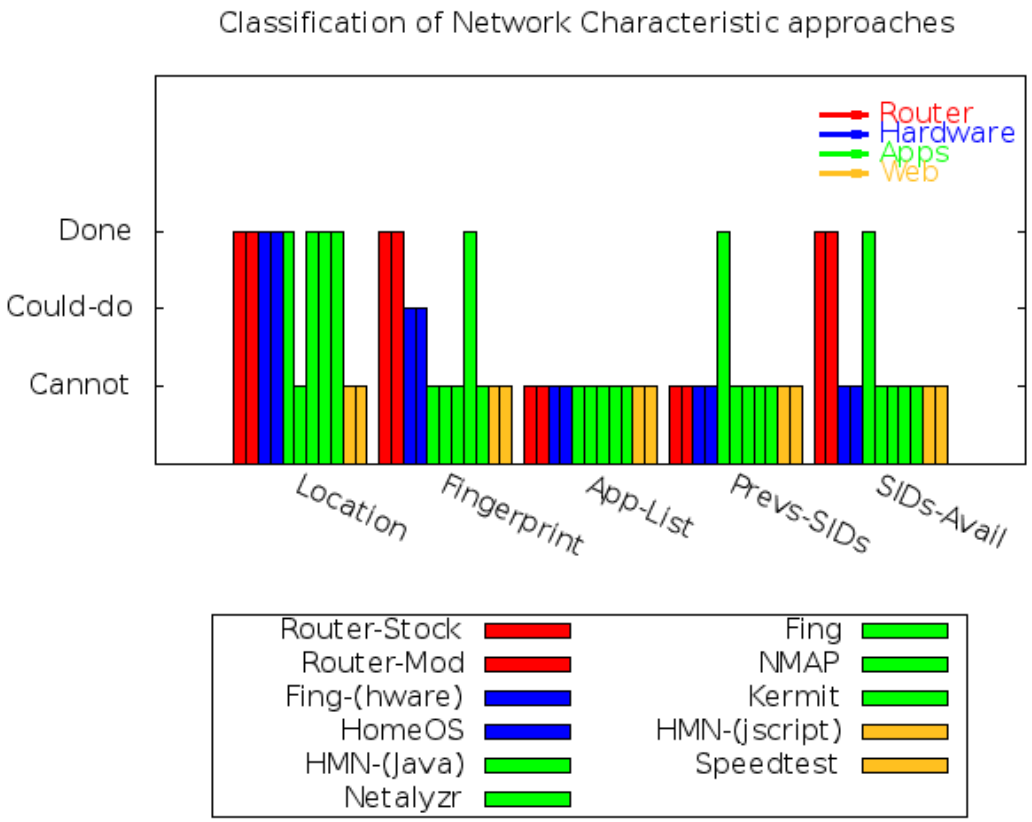

Figure 3.3: Network Characteristics Classification

3. Security review of device

4. Security review of other devices (via local connection)

5. Recommendation system for apps

6. Health check of device

7. Network app profiling

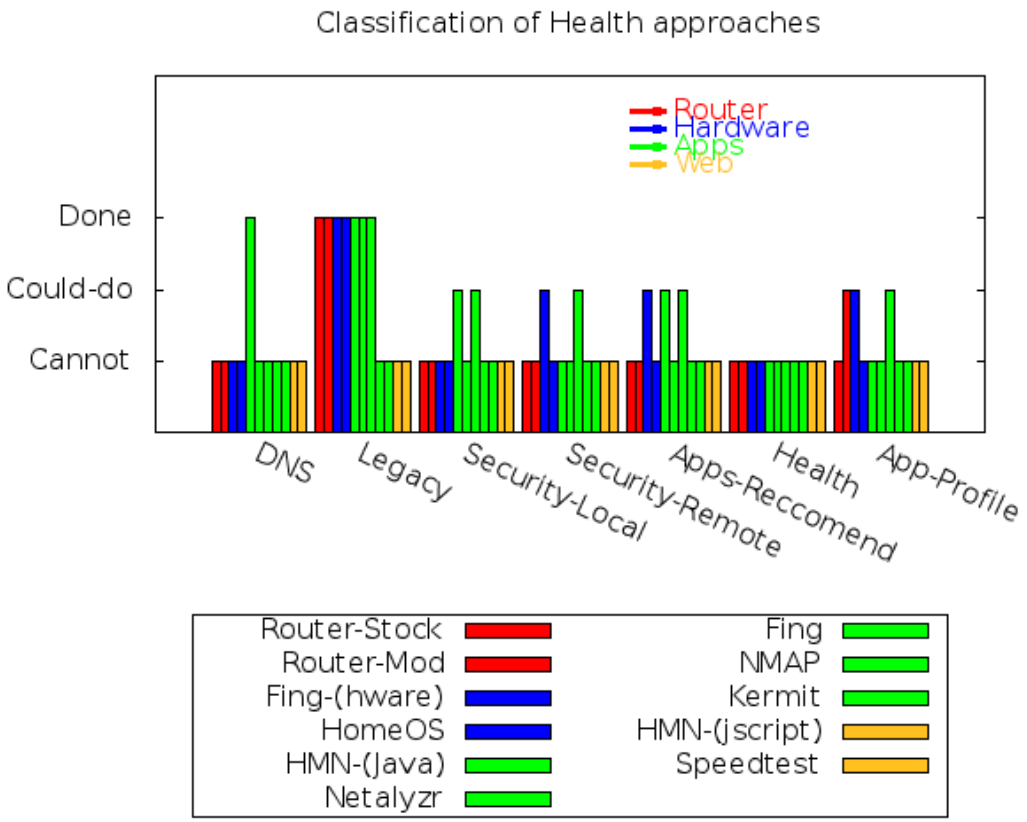

Figure 3.4: Health Classification

3.5.4 Historical Norms

The classification of historical norms in data includes gathering point-in-time data, long-term availability of local data, global review of comparison data between local and

other users experiences, legacy information using a longitudinal approach, and if the data is shareable to a wider community for research. We have created a classification of Norms in Figure 3.5, to help understand where there is overlap of methods in norms, and is a classification of what can, and cannot be done in the following items:

1. Local Norms: Local information over time.

2. Global Norms: Comparison of local and global scans.

3. Legacy Information: What devices previously attached to it?

4. Sharing of research data

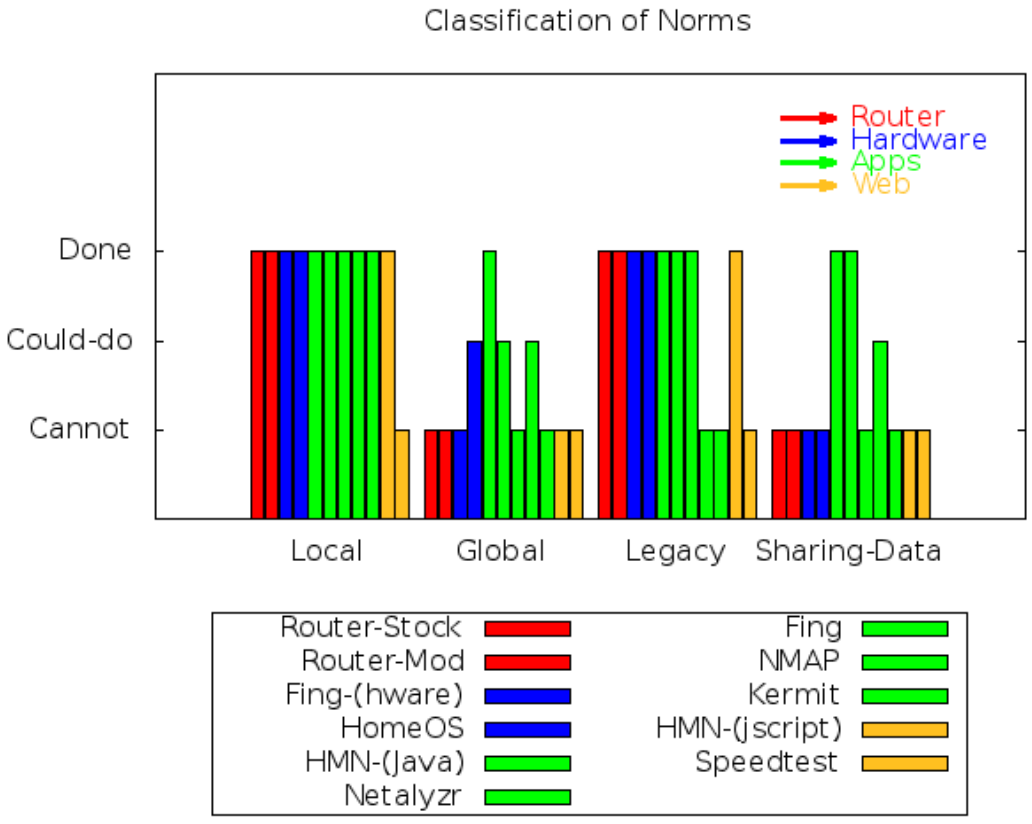

Figure 3.5: Historical Norms Classification

3.6 Comparison of Approaches

We have created plots from the previous classifications to understand user participation associated with: impediment vs. incentive, sources vs. metrics, and local vs global data availability. These graphs provide a view into the measurement approaches and where there is similarity, and differences. The objective of these graphs are to help understand what types of information each approach provides (to users), along with incentives, impediments, as well as location and metrics. As an example, we look to understand the differences between what a Stock Router, modified Router, hardware approaches, app approaches, and Web approaches provides in terms of information to the user. These graphs have data points that range between 1-4 (on both axis, starting at 1), which provide either easiest to hardest or low to high data information for the types plotted.

3.6.1 Incentives and Impediments

We have created an incentive vs. impediments Figure 3.6 to compare data of interest and tools, organized by approaches. Tools that reside in the upper left hand quadrant provide the most amount of data with the least amount of impediment. We can see that the cluster of tools reside in quadrants of the plot; As an example Web and Scripting have the least impediment, but provide the least amount of incentive (low).

3.6.2 Sources vs. Metrics

We have created a Sources vs Metrics scatter plot Figure 3.7, which compares sources vs data collections (or metrics) for each of the tools, organized by Approach. Figure 3.7 shows the flexibility of modification of sources, and configuration vs. the amount of quality data collected. The richest data and customization tools reside in the upper right hand quadrant. Tools such Nmap, and a modified-router provide the richest data sets,

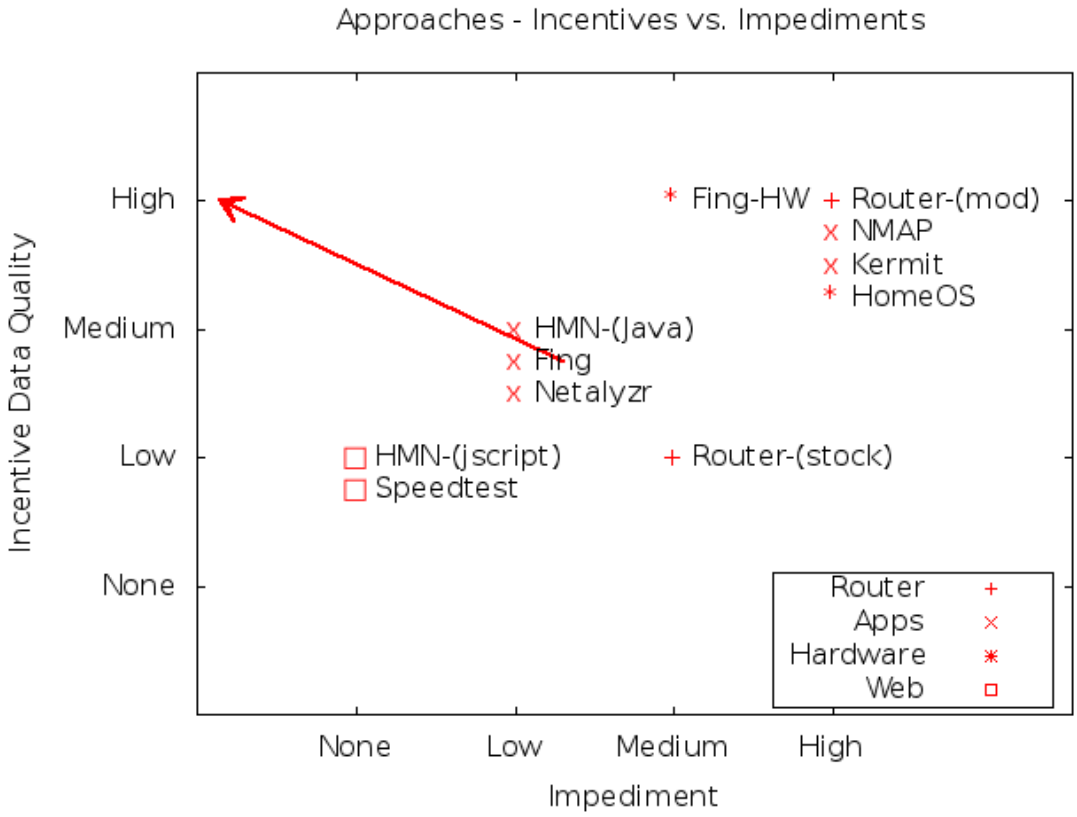

Figure 3.6: Incentive vs. impediments

along with the most amount of customization

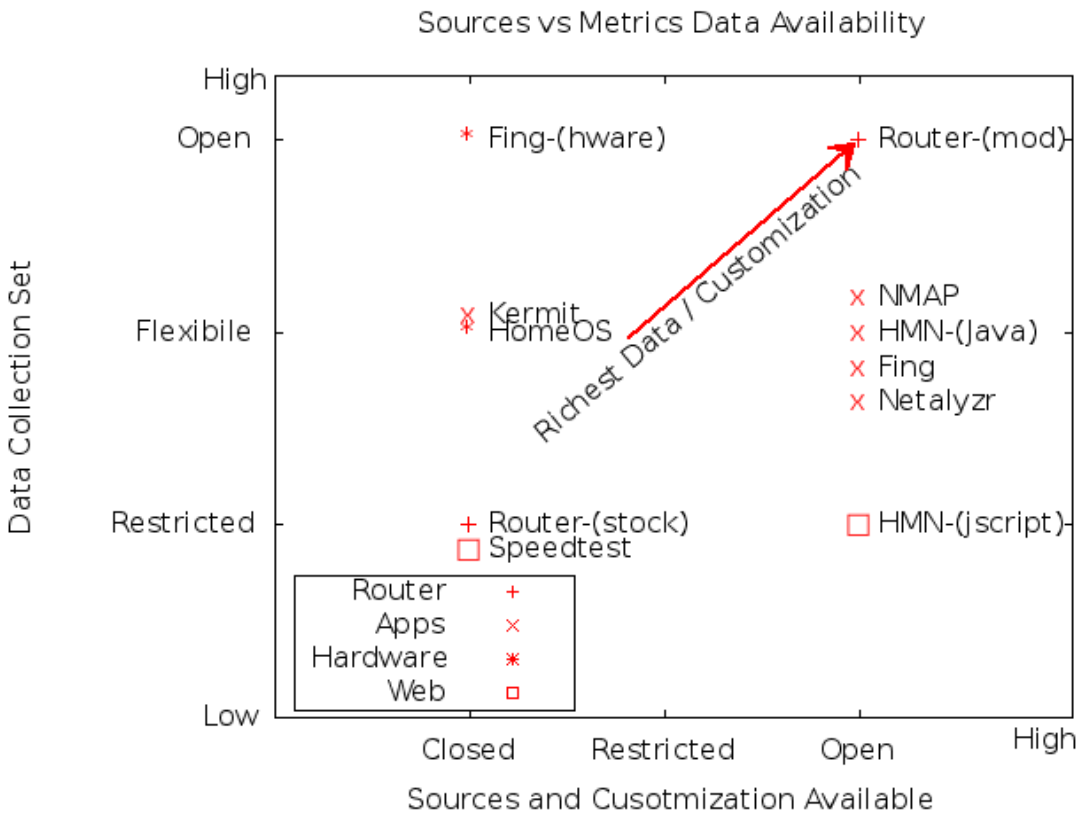

Figure 3.7: Sources vs. Metrics

3.6.3 Local vs. Global Norms

We have created a Historical Norms, local vs global, plot of tools organized by Approaches. Figure 3.8 shows tools classified by information availability, and historical norms of data sharing The richest data and customization tools reside in the upper right hand quadrant. HMN-Java provides the richest norms as it provides both local and global norms to users and researchers, while there are no other tools that provide Global historical norms.

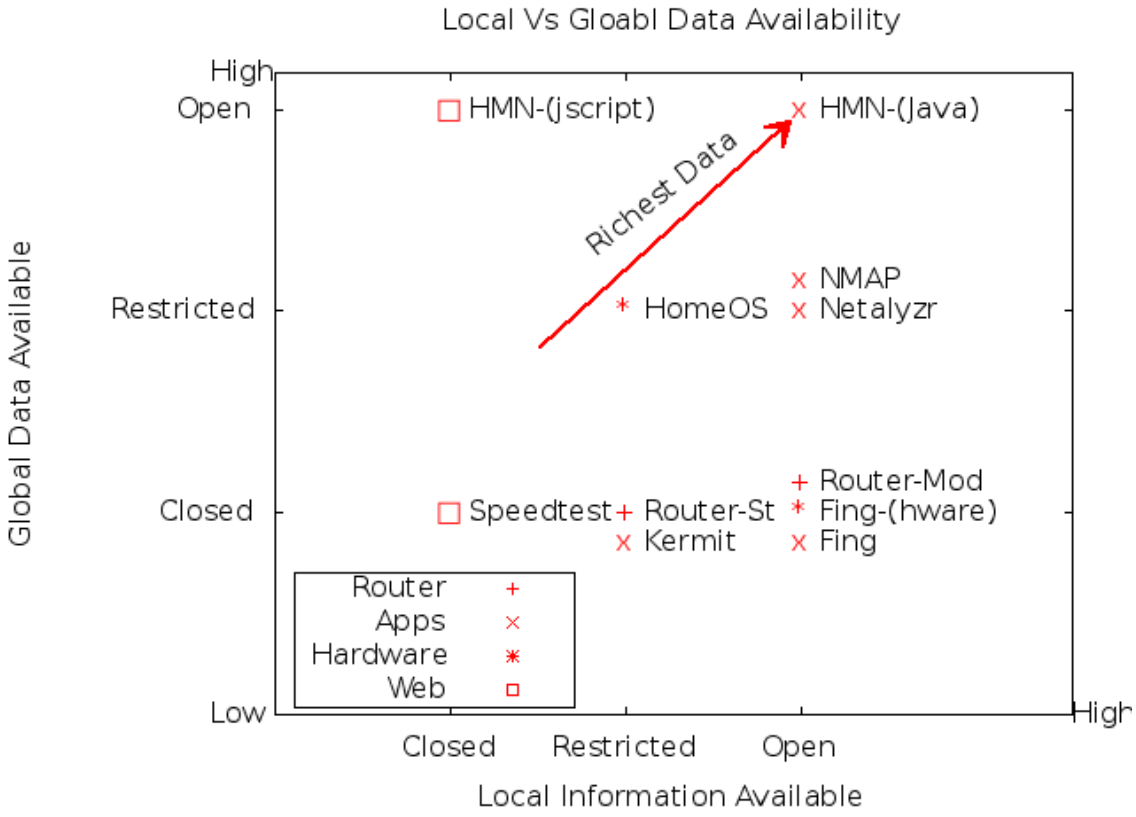

Figure 3.8: Historical Norms (Local Vs. Global)

3.7 Summary

We turn our attention to the tradeoff of these Approaches and data of interest, and provide a summary. We have reviewed four separate approaches to several data of interest and tools as part of HN access points: routers, apps, hardware, and Web and scripting tools. While each of these collection approaches have merit, there are pros and cons of each of these approaches, which we detail.

The router and hardware approaches provide the highest level of information (as they are plugged directly into the network), and allow for the deepest dive into the network layer. These approaches have a single focus and do not provide users and researchers with a ubiquitous approach to network measurement. It also has the highest barrier to entry, and they do not scale to provide this local and global norm methodology. Hardware tools such as Fing (and other commercial and research tools) can provide single point of information, but do not collect the breadth of data provide global norms.

The Web and scripting collection types have minimal impediment to run as they execute either in a web browser or via a scripting environment that resides on the local system. The down side to this approach is that these collections can only provide a small percentage of results as compared to a router, hardware or apps approaches, due to the sandbox they typically execute within.

Finally, the review of apps approaches found that there is a minimal impediment across the wide range of tools, as it can run on ubiquitously available mobile platform versus that of a fat client PC platform or a hard to configure hardware specific device. In addition, an app running in this space provides the most flexibility in terms of incentive of information vs. the impediment of execution, and can also provide similar information of the hardware counterparts in most cases.

The hardware approach is costly and can be complex to setup and configure, and thus

has a high impediment to entry, but provides access as the incentive. An apps research approach may be free to users, and require minimal impediment to install, configured, and execute and thus have a lower barrier to entry. A web/scripting approach may not provide enough information the user is looking to understand, but has the lowest barrier to entry as it executes within a web browser. To understand the data of interest provided from these approaches we have download tools, reviewed papers, and dichotomized the results and methods of the product/tool. These results found that apps have the most flexibility in terms of data of interest, as well as historical norms, and may provide the most optimal Approach, and thus tools, for users and researchers.

In this chapter we examined a series of approach types along with the incentives and impediments for users to leverage these tools, and laid the ground work for continued work as part of this dissertation. This preliminary work is important as provides details around related work and background of approach types, data of interest, and incentives and impediments of these categories reviewed.

The takeaways from this work include the following:

1. A study that identifies software tools and their provided data targeted to HNs;
2. compare expertise required across four tool approaches: Routers, apps, Hardware, and Web/Scripting tools;
3. review the complexity of these tools, and the required expertise to execute and configure as well as impediments and incentives to participation;
4. create a taxonomy around data of interest provided by each approach; and
5. show that focusing on a broad range of approaches, data of interest, and tools allows us to create a new taxonomy of HN functionality.

Chapter 4

Home Network Survey

In previous work we examined HNs via a web browser, and reviewed incentives/impediments for users to leverage approaches studied. In this chapter we look to understand specific traits of HNs via a survey. This work is a pillar or thrust of the dissertation as it is a dive into user HN data and results. We have created the How's My Network survey, completed in 2018, as the next phase of work with a focus on understanding user incentives to incentivize HN research. In this chapter we look to understand HN users traits, including: skills, comfort, and preferences when managing and understanding the layout of their and other HN users environments (e.g. local and global Norms), as well as a mapping of devices, apps, and general interests of HNs. This work also examines specific properties (e.g. devices, apps, privacy/security) and which approach types (e.g. Mobile app) users prefer to leverage to gather data and ultimately view results of HNs in an effort to understand incentives for participation in HN studies.

4.1 Introduction

Previous studies looking to understand user experiences around Home Network (HN) management, comfort levels and required skills have had minimal coverage up to this point. We are interested in understanding HN user experiences, including their perceived skill and comfort levels with managing devices and services that reside within their HN, which we believe will inventivize user participation in HN studies. In this chapter we look to understand the perceived value of information from a survey we completed around the project How's My Network (HMN), as focused on a user-centric approach to understanding the user experience when managing and using a HN. We also examine device types, Internet Connectivity, management options, interest, and preferred method of management of user HNs. We believe that the landscape has changed and that users are now interested in an in-depth view of HN information. An in-depth review of these areas has not been completed before, and the information from this review and survey, we believe points toward characteristics that users versus researchers are keen on today. In this chapter we show the results from this survey and discuss how they fit into potential future work.

Since the advent of the Internet Home Networks (HNs) have leveraged service provides for Internet network connectivity to their homes. HN users have options to connect to the Internet via a variety of ISPs types including A/DSL, Cable, FIOS/Fiber, and dedicated telephony services. With these offerings, and changes in technology, HNs access to the Internet is no longer a luxury, but rather an expectation. Understanding that the landscape has changed it is now important to understand what is happening inside of HNs.

A study done by the United States Census [165] (in 2013) on Internet usage found that 75% of census respondents had an Internet connection, compared to 78% of households

having a fixed broadband connection as of 2018 [19], 43% leveraging cable as their connection the Internet. An interesting aspect the researchers found is that 5% used a satellite connection as their primary connection. We found that 99.9% of respondents to the HMN survey reported owning a computer versus 84% from the 2013 Census report.

There has been a shift to faster connectivity, along with a device explosion within the HN, which is clear from a report done by [66] and w3. This study showed that the typical Home Network (HN) includes devices, such as hardware that allow for browsing, and connectivity to commodity Internet traffic. These devices include routers, modems, PCs, tablets and smart-phones, and specialized hardware devices. This work also delved into the complexity of these devices and how they impact HN users in terms of types, and hardware, and created a new taxonomy to classify HNs. This study also looked to understand interests, and comfort levels of HN users as related to configuration, setup, and experiences with HNs, devices, and Applications. In the remainder of this chapter we will continue to refer to 'comfort' as a user's ability to manage, configure, setup, operate devices and customized hardware, software/tools, and home networking hardware; where these areas also include Security and Privacy.

In this chapter, we show work we have done to designed a survey to take a focused examination of understanding HN user preferences, comfort levels and areas of interests in and around HNs. In addition we seek to understand user interests in wireless networks, mobile and networking devices, security, and privacy. This study and survey provides several contributions in understanding the user experience, comfort levels in HN users to ultimately glean incentives for participation in a HN study.

In particular, we provide:

1. an examination of user comfort levels, interest, and actions regarding; routers, mobile devices and Web-related tools;

2. an understanding of what type of information is important to HN users to operate their environment;

3. how security, privacy, configuration, and general operation health, are of interest to users; and

4. insight on how HN users would like to see data collected within their HN.

4.2 Related Work

In this section we provide background and related work in this area. A starting point is looking at work completed in commercial, research, and patents with an area of HN focus.

We start with work completed work done in the project How's My Network, where the focus was using a Java via a browser to understand HNs [140]; this research focused on measurements of HNs targeted at researchers. Work on Home Networks (HNs) was completed with a focus on a review of device types, and classification of methodologies 3. This work created a taxonomy of approaches used in HNs, along with specific areas referred to as data of interest, and tools (apps) used in HNs. Other studies focused on general information about HN Internet connections and demographics [165].

Studies such as [60] looked to understand digital competencies of university students, which would apply to how they leverage and manage HNs. This study found that educational readiness around digital competency is a key indicator to leveraging online tools. In addition a study by [142] looked at a variety of methodologies and research studies to determine skill levels in and around social and cultural boundaires. A study by [157] looked to understand Wifi connectivity in HNs using a hardware apparatus in the environment, for a period of time, with a focus on researchers versus HN users.

Bajpai, et. al. did a study [17] on legacy platforms which are all targeted toward throughput and provide researchers data versus HN users. A study focusing in on Wifi AP connection [127] looked to understand why APs are slow to connect, and provided some feedback to researchers as how to potentially improve a slow connect scenario. A study by [175] looked at user comfort level in and around specific background processing of apps on mobile devices, and found that designers should be improving the mobile-privacy that systems allow.

A review of similar work done in and around surveys was also completed. These include papers by [80], [81], and [17], which used a service model and installed a fat-client onto a local system (PC in this case) to determine preferences via pop-ups and other techniques (in-app) via survey questions. The method used in this body of work required a 'administrator' privileges and active techniques for monitoring and gathering information. Another study looked to understand Wifi congestion with neighbours, fitting nicely into HNs, and used heavily modified APs installed into test environments to measure overlap [126]. They found that their "metric" accuracy was confirmed when measuring congestion, and that it fits into an approach of remote management of Wifi connectivity.

A review of devices in HNs was also completed, an their impact in networking in-general. Cisco has predicted that by 2020 there will be 50 Billion IoT devices connected to the Internet, with a larger percentage of these devices residing in HNs [52]. Work done by [124] conceded that HN management is becoming increasingly challenging, and complex and argue for a vendor neutral API across all management tools, for configuration of Wifi networks. Various research and patents have been created around creating new low power and better efficiency of Wifi, connectivity, and IOT including [51], [91], [93], [92], and several others. While these approaches (and patents) target new technologies for commercial or research space, they provide little feedback for HN users.

Other work looked to provide approaches to models of security within the HN for IoT

devices, using in-hub hardware approaches (physical hardware or modified routers) [148], [3]. Privacy work in and around HN Internet traffic shaping, by [148], looked to leverage IoT and other devices, along with 802.3.ad (tunnelling, and VPNs) for privacy.

In this chapter we are taking into account approaches used for privacy, security, and health as well as user comfort, interest, and skill level when managing their entire HN, and believe that users are interested in easy methods of data gathering around their areas of interests. This review is HN user-centric and focuses on user interests to understand incentives and impediments toward participation in HN research. We present our findings from our current HN research survey, in the next several sections.

4.3 Research Questions

In this section we layout the research questions we are looking to answer as part of this study. A focus of this work, and as part of the HMN study, include the following questions as part of this study:

1. What is the landscape of how users view the set up current HNs?
2. What is HN users skill level of setting up HNs, devices, tools (apps, etc.) and configurations of HNs?
3. What interests do users have in the areas of information, security, privacy, setup, and maintenance of HNs?
4. What types of services are users interested in monitoring as part of HNs?
5. What is the preferred method to collect data about a HN?

4.4 Methodology

In this section we provide the methodology used as part of this review. We start with the methodology that was used and review the HN survey we completed, and then turn to why this research matters and look to understand HN comfort levels of managing their HN, devices, services, as well as user interests in and around HNs. In this survey we exclusively requested feedback from respondents in and around their permanent HN environment. Survey responses were collected from September to November 2018.

The survey includes 12 questions, as recommended by the APA and PEW best practices [129], to minimize survey fatigue and keep users engaged without overwhelming with too many questions. The survey is requesting information from an HN users exclusively around their experiences of their home network (HN) and their comfort level managing and using their HN. At the completion of the survey respondents are provided a summary of results. In addition, respondents that provide an email address will be sent a copy of this report.

The background for several of the survey questions include [33], [127], [126], and work shown in Chapter 3. The following are some of the definitions from the this work, which were used as a starting point to create this survey. These include approaches: hardware and software that provide a plethora of localized information by peering behind the HN router using several approaches, in an attempt to determine and characterize configurations.

We refer to the Software/applications running on these devices as 'Tools'. Approaches are classified to include: Routers, apps, Hardware, Web/Scripting tools. Each of these areas may have a subsection, which includes customized groupings. Data of Interest: the gathering of desired data collection from the user perspective. We have included the attributes of each of these data of interest, which are the data points collected by the tools.

The areas classified include:Throughput (upload/download, jitter, network flow, and performance), Networking Characteristics (discovery, Wi-Fi, and fingerprints), Health (including security and privacy), and Historical Norms (Local and Global norms).

4.4.1 Skill Level

The following are the questions and a review of how we are using the information from the survey in this chapter. We start with classification of respondents. Respondents are classified by taking answers from the categories of classification and split into skill levels, which we have modified from the NIH [33] competency scale:

- Novice: An individual that has limited experience and rely upon help. They have some common knowledge or an understanding of basic techniques and concepts.
- Intermediate: An individual that is able to successfully complete tasks with minimal help, but may need assistance from time-to-time.
- Advanced: An individual who can perform actions associated with a given skill without assistance. This group is recognized within their group as a person to ask when difficult questions arise regarding this skill.
- Expert: An individual who can provide guidance, troubleshoot and answer questions. They are known as an expert in this area, and provide guidance, troubleshoot and answer questions related to this area of expertise and the field where the skill is used.

4.4.2 Assessment

Respondents have been split by their skill level, which they have self selected. The areas we are slicing these respondents into are ability and comfort, Ability (including

management) and comfort (including preference), with the following categories: Self-assessment of skill level, Home router knowledge, Technology comfort, which includes mobile, PC, and hardware, Wifi literacy, and app literacy.
Home networks are broken down from the data collected, and the respondents information classifies their areas of ability and comfort level with HNs. The respondent is supplying information about their own home network which we previously classified. The data collected is used to classify HNs in specific areas as well, including: devices and connectivity.

4.5 Survey Questions

In this section we provide an overview of how we have used or plan to use these results as part of this work. We walk through each question, list possible answers after each question, and provide a short justification for each question.

Introduction to Survey

We start the survey with the following heading and basic information, and then move to each question.

How's My Network (HMN) Home Networking (HN) survey. This survey is student thesis work, under the department of Computer Science at Worcester Polytechnic Institute (WPI), on the project 'How's My Network'.

We are conducting a survey to learn about Home Networks, and help users fully access its functionality. As part of this survey we are exclusively looking for feedback on your Home Network environment, including: Wifi (wireless network), mobile and networking devices, security, and privacy. Your personal information and survey results will not be shared. This information will only be used in aggregate to better understand the current state of Home Networks.

The following is a review, Figure 4.5, of the current survey header description. The following URL is a link to the survey; `https://wpi.qualtrics.com/jfe/form/`

SV_3VqEIeisY3RUYYd

How's My Network (HMN) Home Networking (HN) survey.

This survey is student thesis work, under the department of Computer Science at Worcester Polytechnic Institute (WPI), on the project 'How's My Network'. We are conducting a survey to learn about Home Networks, and help users fully access its functionality.

As part of this survey we are exclusively looking for feedback on your Home Network environment, including: Wifi (wireless network), mobile and networking devices, security, and privacy. Your personal information and survey results will not be shared. This information will only be used in aggregate to better understand the current state of Home Networks.

Q1) To which gender do you identify?
Respondents have the option of selecting one of the following:

- ☐ Female
- ☐ Male
- ☐ Prefer Not to Say
- ☐ Other (open feedback)

We start by asking respondents to identify their gender. While we did not use this question to bucketize responses into classified genders, this information was used to help understand those responding to the survey. We have used the gender to help classify and slice data by selected categories.

Q2) How would you rate your Home Network skill level, this includes managing your Router/Mobile/PC devices and Apps/Software?

Respondents have the option of selecting one of the following:

- ☐ Novice (limited experience and rely upon help)
- ☐ Intermediate (you are able to successfully complete tasks with minimal help)
- ☐ Advanced (you can preform actions associated with this skill without assistance)
- ☐ Expert (you can provide guidance, troubleshoot, and answer questions).

Home Network skill level question is looking to have the user self-classify their abilities of managing their HN, devices, tools, and comfort level. The classification types have been aligned with the NHI's Competencies Proficiency Scale [33], NIH Competencies. This information is used as part of this study to classify users from what they have defined or selected along with how they answer questions related to their environment, comfort level, and abilities in this survey.

This information is used to classify users from what they have self selected, in terms of skill, along with how they answer questions related to their comfort level, and abilities in this survey. We use this selection to compare level of devices, and management across a variety of areas.

Q3) What type of home Internet connection do you currently have?
Respondents have the option of selecting one of the following:

- ☐ DSL/ADSL
- ☐ Cable
- ☐ FIOS/Fiber
- ☐ Dial up Via a Modem
- ☐ I do not Know, or Other; option allows open feedback from the respondent.

Internet classification type is asking basic information around their home network. We start with the basics of the Internet technology type to get an understanding of the landscape of today's Home Network Internet types being used. This question allows for

an open answer or to mark one oval. This information is being used to identify users by provider, and update of previous work done from [140] on HN Internet providers. This information is used to understand if HN users can identify their Internet provider.

Q4) Wifi and Mobile devices on my Home Network include:
Respondents have the option of selecting one or more of the following devices:

- ☐ I do not have any Mobile or Wifi Devices
- ☐ Android Phone
- ☐ Android Tablet
- ☐ Iphone
- ☐ Ipad
- ☐ Windows Laptop/Desktop
- ☐ Mac Laptop/Desktop
- ☐ Reading Devices (e.g. Kindle)
- ☐ Health and Wellness Devices (e.g. Fitbit)
- ☐ Game Console (E.g. PS, Xbox, Nintendo)
- ☐ Wifi Range Extender (e.g. Netgear)
- ☐ TV and Sound System (e.g. LG TV, Sonos Sound System)
- ☐ Streaming Devices (e.g. Roku)
- ☐ Smart Speakers Assistive Wifi Devices (e.g. Amazon Alexa, Google Home)
- ☐ Find Your Device (e.g. Tile)
- ☐ Thermostat (e.g. Nest)
- ☐ Home Security and Video Cameras (e.g. Ring, Nest, Lorex)
- ☐ Smart Lock (e.g. August Smart Lock)
- ☐ Smart Sprinkler or Home Control devices (e.g. Rachio)
- ☐ Other; option allows open feedback from the respondent.

In this question we are looking for standard devices in the HN, (e.g. Wifi, etc.), as well as peering into IoT-based devices. The list of devices from this list is similar to those done by the review of the most popular device list of 2018 [132]. These include Wifi, mobile, and other similar (Iot) devices the user has on their HN. Results from this question can be seen in Table 4.6. The question allows multiple selections and an open entry from the user.

This information is being used to classify devices a user has on their HN, and compare

against the mobile app collection data set. This is of interest as it points toward hardware commonality and global network and device norms. We have used this question to classify HNs using the device and connectivity classification we have created as part of this study. As an example in this question we can see that the most popular devices across all networks are Windows PCs (81%) and mobile devices (>65%), which more than 50% of all respondents reported having in their HN.

Q5) Which of the following best describes your abilities to install and configure a home router?
Respondents have the option of selecting one of the following abilities:

- ☐ I am not comfortable with setting up my Home Router / Home Network
- ☐ I am not comfortable with setting up my Home Router / Home Network
- ☐ I have minimal experience, and rely on friends and my Internet provider to help manage my Home Router / Home Network
- ☐ I have set up my Home Router / Home Network, but still need help from time to time
- ☐ I am very comfortable with setting up a Home Router / Home Network
- ☐ Other; allows open feedback from the respondent.

In this question we are looking to understand the abilities of the Home Network user when working with their HN router. This question is congruent with a series of future questions around HN experiences, and lines up with work done by [33], and the research shown in Chapter 3. This question allows the user to classify how they feel about their experiences with their HN. These classification of answers are similar to what Google Measurement labs [110] uses when looking for feedback. This question allows for an open answer or to mark one oval.

We are classifying the users home router knowledge, which is part of the abilities (including management) and comfort (including preference) classification we have created as part of this study. This question is used in conjunction with other questions to help classify areas related to skill level, and other functions in the HN. These answers are also quantifiable into novice, intermediate, advanced, and expert levels from the results, as

they lineup with the scale we have defined. We use respondents results to compare skill level versus management abilities in the 4.7 section.

Q6) Please rank the following in terms of your comfort level with using each type of device or application (lowest comfort [1] to highest comfort [4])
Respondents have the option of selecting one of the following per each of the comfort areas related to device/application type. Please note that this question allows one answer per row.

- ☐ I have no idea how to use
- ☐ I have used before, but typically need help
- ☐ I have downloaded software/updates and or run Apps (applications) on it
- ☐ I am very comfortable using.

The following Device types were listed (one per line) to select versus comfort level, and included:

- ☐ Mobile Device (such as a smartphone or tablet)
- ☐ PC (Mac or Windows) applications (word, etc.)
- ☐ Web Browser
- ☐ Router
- ☐ Purchased home networking devices that monitor and manage security and privacy (e.g. Bitfender, Luma, Dojo, F-securre, Fing, etc.)
- ☐ Customized hardware (specialized router, Linux Machine, etc.).

In this question we are interested in comfort level using devices or applications, and approached this area with work done by [21] and the terminology created around approaches shown in Chapter 3 in an effort to understand comfort level of each area related to: Mobile devices, PCs, Web browsers, Router, and customized / purchased hardware. Included as part of this question is purchased home networking device, and example hardware that fit, where they are types that monitor and manage security and privacy (e.g. Bitfender, Luma, Dojo, F-secure, Fing).

This information helps with understanding which approach type is desired for ease of management. We have used the responses to this question to classify the users technology comfort level of HN activities, which includes abilities (management) and comfort (including preference) to point to concrete areas of how users manage their HNs.

In addition, we used this question to classify the users technology comfort level of HN activities, which includes abilities (including management) and comfort (including preference) to help point to a concrete areas of how users manage their HNs.

Q7) Please select all actions you have made to your Wifi/Router or devices in the past)
Respondents have the option of selecting one or more of the following per each of the comfort areas related to device/application type

- ☐ Added a device to my Wifi
- ☐ Logged into my Router
- ☐ Updated my Router firmware
- ☐ Setup a new Router out of the box, including configuring Wifi and passwords
- ☐ Run an online speed test of my Internet connection (e.g. Google, Ookla)
- ☐ Logged into my router to review what devices have connected to my home network in the past
- ☐ Run a basic network scan to determine what devices are attached to my home network (e.g. NMAP, JNetMap, Network Scanner)
- ☐ Run advanced network diagnostics gathering: fingerprinting, topology, and Wifi layout. Using tools which typically require administrator or super-user privileges to execute (e.g. NMAP, Wifi Visual Analyzer, Internet Mapper Tool, Advanced IP Scanner).

This question is diving into the approach types, and data of interest as defined in Chapter 3. This question continues on the same line as the previous question (comfort level), and looks to understand Wifi/Router or devices the users have managed within their HN. This question includes generic areas executing a task or managing an application within the approach types and data of interest areas. The scope of questions directly correlate to interest of study and approaches done to this point in research and commercial applications. This question allows for an open answer or to mark a series of ovals. This information is used to classify the respondents Wifi/router literacy into the areas of abilities (including management) and comfort (including preference), that we have created as part of this study.

Q8) Please select all actions you have made to your Mobile or PC device(s)
Respondents have the option of selecting one or more of the following per each of the comfort areas related to device/application type

- ☐ Installed a new App (application) on my mobile device (e.g. smartphone or tablet)
- ☐ Used a recommendation App to install Apps on my devices
- ☐ Run an App to review what applications are running on my device (Mobile or PC)
- ☐ Run an App to review what networks my device (PC/Mobile) has connected to in the past
- ☐ Run an App to review the security or privacy of my Mobile/PC device (e.g. Verizon Security, Norton)
- ☐ Run an app to review what devices have connected to my home network in the past
- ☐ Run an App (Mobile or PC) to review the security of my home network (e.g. Home Network Security, Sophos, Check Point ZoneAlarm)
- ☐ Run an App to review the health of my home network and internet connection including tools that examine configurations, normal operation, security and local device privacy (e.g. Bullguard, Sophos, Cryptguard, Textsecure, Orbit)
- ☐ other ; option allowed for user open ended input.

This question is looking to understand the Actions HN users have performed on Mobile/PCs devices. This question also allows for an open answer or to mark a series of ovals. The information from this question is used to classify the respondents app literacy level, by the areas of abilities (including management) and comfort (including preference), that we have created as part of this study. This question has several check boxes that the user can select around apps, and we following a similar rating as other questions we calculate the respondents app Literacy. These areas were reviewed as part of previous research done, and include either generic questions on executing a task or a managing an application within an Approach type and data of interest.

The scope of this question directly correlates to interest of study and approaches done to this point in research and commercial applications. Work done by [80], [81], and as shown in Chapter 3, and was background for this questions. We have included examples to this survey to include software or tool types, e.g.: Google Speed test, NMAP software, etc.

Q9) Please select your interests around managing, or understanding in more detail your Home Network
Respondents have the option of selecting one of the following interest types; the selection allows for multiple options, these include:

- ☐ When changes happen to my Mobile (or similar) devices
- ☐ How to setup my home network correctly, Speed to the Internet
- ☐ Why my Wifi is slow, and how to fix it (connection to devices on your home network, e.g. printer, tv)
- ☐ How my network compares to others in terms of setup devices, and a Mobile Applications
- ☐ Is a networked device secure (e.g. is anyone trying to break into the device?)
- ☐ What devices have connected to my Home Network (my devices, and friends devices)
- ☐ Home network health (e.g. is my firmware up to date, are there open devices on my network)
- ☐ Device health (Is my device running at peak performance in terms of memory, apps, storage)
- ☐ How private are my network devices (e.g. are they leaking data that can be seen by the outside world?)
- ☐ Detecting unauthorized devices attaching to my network, and automatically disable them
- ☐ How to use my home network to control multiple devices from one location (e.g. Alexa Dot, tvs, audio, video, heat, kitchen, bathroom)
- ☐ How to control how much of my internet bandwidth can be used by each device
- ☐ other; allows for an open answer from the respondent.

This question is looking to understand HN users interests and management preferences. Management is centered around HN user interest in terms of (immediate and delayed) feedback of their HN or to manage their network. This question also allows for an open answer or to mark a series of check boxes. This question continues with the classification of the respondent's interest are in terms of gathering information and around their HN, and which approach type lines up with their interests. The results show that Mobile phone and PC results are desired for data collection at this point.

The data of interest, presented in Chapter 3, as part of this question includes security and privacy, throughput, as well as management of functionality. The question includes examples in each of these areas, where it makes sense, to show what type of tool would

potentially be used, or what it would provide. This question continues with the classification of the respondent's interest are in terms of self management or details in and around their HN. We use this information as fodder for the Mobile HN app, and which features respondents are most interested in seeing; although, some features may not be possible to implement. In addition, a review of what can and cannot be done from this question can be seen in the summary section of what can/cannot be done 4.9.

We use this question to classify users interests, along with each of the sub questions in this question are a result of the research from Chapter 3, [80], [81] and [132] and looks to understand how HN users have interest in retrieving and reviewing HN information. We are looking to understand where users have interest not only in areas of concern, but also how to gather and display this information. One approach is to leverage an app to provide this information versus using a Web browser. As an example, an app targeted at an Android specific Mobile HN app could gather and provide information to HN users, and fit into the mold of this question.

Q10) Which of the following would you prefer to use to gain access to information about your Home Network?
Respondents have the option of selecting one or more of the following per each the areas around preference of gaining access to HN information:

- ☐ My Mobile Phone/Tablet
- ☐ My PC/Mac
- ☐ A Web Browser
- ☐ My home network router
- ☐ A modified Router that allows for advanced diagnostics, and functionality
- ☐ A purchased unit of hardware specifically designed for Home Network data gathering (e.g. Cujo, Keezel, RaTTrap)
- ☐ I am not interested in gathering any information,
- ☐ Other; allows for an open ended answer.

This question continues with the classification of the respondent's interest are in terms of gathering information and around their HN, and which approach type from the Tech

Report lines up with their selection. We are looking for information to understand top approaches to gather information and how to most effectively continue our study in the HN.

Q11) What other information are you interested in understanding, in and around your Home Network? Let us know your thoughts.
Respondents have the option to enter an open ended thought in the provided box.

☐ Other ; Open ended feedback

In this question we are looking for additional information from respondents on things not covered and provide open ended feedback. This is direct user feedback of information not asked or touched on as part of this survey. We are looking to understand what users are interested, whether it Could or not. From all of the research done in the Tech Report we are interested to see which areas have not had research or commercial experiences at this point. This question allows for an open answer over several paragraphs, if needed. This information is used to understand the classifications of the user and the HN. We are looking for this information to point toward functionality or prompts to be added to the Mobile app we plan to deploy as part of our work.

This information is used to understand the classifications of the user and the HN. We are looking for this information to point toward functionality or prompts to be added to future research and how we can deploy these into our HMN study.

Q12) Email Address
Respondents are asked to enter an email address in the open ended text field provided.

☐ Email Address ; Open text for email entry

We use respondents email address to update them on study results and future work, via web links and downloadable tools.

4.6 Results from Survey

In this section we delve into the survey, including introductions and results. We have created a Home Network survey as part of the work from Chapter 3, which was targeted at Home Network (HN) users.

We used the following methodology to distribute and conduct the survey. How's My Network Home Network (HMN) Survey was released in September of 2018, and completed in Nov of 2018. The survey was sent out using email and social media links. The emails and social media included the following networking channels: an initial email to social and other connections asking them to forward to their social and other networks, survey Reddit posts, surveytandem utility (a share-able survey tool), and other social networks, the Worcester Polytechnic Institute (WPI) faculty and staff, and finally to a WPI Graduate and Undergraduate student email list. Of the 550 respondents we believe that roughly 12% where from the initial connections sent out, an additional 48% came in from the initial group forwarding to their social communities along with posts from Reddit and surveytandem, with a final 40% of responses coming in from the WPI community. We leverage these groupings (WPI and Social subgroups) to understand potential homogeneity as related to preference and skill level.

Figure 4.1 shows a timeline over the 60 days of the survey (29 date data points), with lower values removed for logical plotting purposes. A small percentage<2%) of respondents did not fill out each question, and sub-questions, of survey. Those respondents who did not fill out a specific question where removed from that question, and not included in the results.

As a by-product of the survey collection tool we were able to determine devices and browsers used by respondents to take the survey. The variety of devices include the following: 48% used a Windows system, 18% Iphone, 15% Mac, 14% Android, and 5%

Linux or other OS (including Chrome OS). 60% of all respondents used the Chrome browser, 23% used Safari (60% of all Mac/Iphone users), 11% used Firefox, and 7% used Microsoft IE/Edge. A data point of interest is the types of Operating Systems (OS) used by Mobile users, specifically Android users. The majority of users taking the survey (45%) were running version 8.0 (Oreo), which is currently one version behind the latest version of Android PIE (9.0). 19% running PIE (9.0), and 17% running Nougat (7.0-7.1), two versions behind. Almost 20% of respondents are running older versions of Android, Marshmallow (6.0x) or Lollipop (5.0x), which are deprecated or no longer supported.

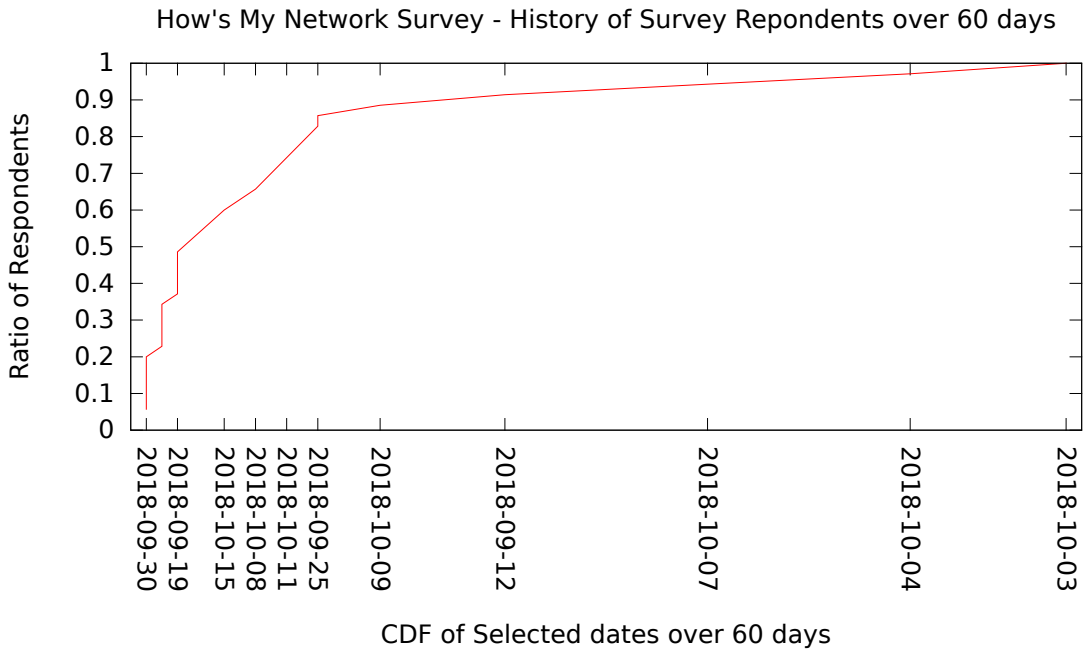

Figure 4.1: 60 Days of Survey Responses

The following shows results for each of the question.

Q1) To which gender do you identify?

Respondents (%)	Gender
240 (44%)	Female
296 (53%)	Male
14 (3%)	Other

We can see from these results that there were more Male survey respondents versus that of Female or Other classifications from the selection.

Q2) How would you rate your Home Network skill level, this includes managing your Router/Mobile/PC devices and Apps/Software?

Respondents (%)	Skill Level
90 (16%)	Novice
230 (43%)	Intermediate
143 (26%)	Advanced
81 (15%)	Expert

The largest cross section of users were classified as 'Intermediate', with the lowest classification area of 'Expert'. We saw 50% of all female respondents classified as Intermediate versus 35% of males. In contrast 25% of Males where classified as Expert versus 1% of females.

Q3) What type of home Internet connection do you currently have?

Respondents (%)	Internet Media
68 (12%)	ADSL/DSL
273 (50%)	Cable
102 (19%)	FIOS/Fiber
8 (2%)	Dial Up via Modem
81 (15%)	I do not Know
8 2%	Other (fill in)

We saw that 85% of respondents reported that they had a home Internet connection, with an additional 15% reporting that they were unclear of connection type; although 5% percent of these reported they do have Internet connection but could not identify it. We estimate that >90% of all respondents have a home Internet connection. This number is an increase from a 2013 US Census (American Community Survey) on Internet and other areas of research, where 73% of households reported am Internet connection[165] and type.

Q4) Wifi and Mobile devices on my Home Network include: (All Device Types)

Respondents (%)	Device Types
3 (1%)	I do not have any Mobile or Wifi devices
292 (53%)	Android-Phone
123 (22%)	Android-Tablet
377 (67%)	Iphone
262 (48%)	Ipad-similar
446 (81%)	Windows Laptop/Desktop
230 (42%)	Mac Laptop/Desktop
193 (35%)	Reading Devices (e.g. Kindle)
127 (23%)	Health and Wellness Devices (e.g. Fitbit)
300 (54%)	Game Console (e.g. PS, Xbox, Nintendo)
110 (20%)	Wifi Range Extender (e.g. Netgear)
296 (54%)	TV & Sound Systems (e.g. LG TV, Sonos Sound system)
278 (51%)	Streaming Devices (e.g. Roku, Chrome)
28 (5%)	Digital Photo Frames
147 (27%)	Smart Speakers Assistive Wifi Devices (e.g. Amazon Alexa, Google Home)
31 (5%)	Find your device (e.g. Tile)
52 (9%)	Thermostat (e.g. Nest)
84 (15%)	Home Security & Video Cameras (e.g. Ring, Nest, Lorex)
9 (2%)	Smart Lock (e.g. August Smart Lock)
5 (1%)	Smart Sprinkler or home control devices (e.g. Rachio)
2 (1%)	Other Open Ended

Number of Homes (%)	Device Types
11 (2%)	1
34 (6%)	2
48 (9%)	3
72 (13%)	4
67 (12%)	5
75 (14%)	6
64 (12%)	7
65 (12%)	8
49 (9%)	9
23 (4%)	10
22 (4%)	11
1 (<1%)	12
8 (1%)	13
4 (1%)	14
2 (<1%)	15
1 (<1%)	16

This summary includes device types within the HN, where the respondents selected they have one or more of that type. We can see that Less than 1% of all respondents reported not having any devices on their network, and 95% of all respondents reported having between 2-11 devices in their HN. An additional 20% of respondents reported having more than eight devices. Overall respondents reported a total of 3395 devices, where the highest devices (on average) per household was six or 14% of households reported.

Q5) Which of the following best describes your abilities to install and configure a home router?

Respondents (%)	Ability to Install Router
74 (14%)	I am not comfortable with setting up my Home Router / Home Network
126 (23%)	I have minimal experience, and rely on friends and my Internet provider to help manage my Home Router / Home Network
156 (29%)	I have set up my Home Router / Home Network, but still need help from time to time
188 (34%)	I am very comfortable with setting up a Home Router / Home Network
3 (<1%)	Other Open ended

Novice users make up the largest percentage of respondents who are Not Comfortable with installing and configuring their HN router versus that of expert users who are very comfortable with these tasks.

Q6) Please rank the following in terms of your comfort level with using each type of device or application (lowest comfort [1] to highest comfort [4])

Device	I have No Idea how to use	I have Used Before but typically need help	I have downloaded software/updates and or run Apps (applications) on it	I am very comfortable using
Mobile Device (such as a smartphone or tablet)	1%	2%	9%	88%
PC (Mac or Windows) applications (word, etc.)	1%	4%	9%	86%
Web Browser	1%	2%	8%	89%
Router	9%	29%	22%	40%
Purchased home networking devices that monitor and manage security and privacy (e.g. Bitfender, Luma, Dojo, F-securre, Fing, etc.)	58%	12%	11%	19%
Customized hardware (specialized router, Linux Machine, etc.)	51%	17%	8%	24%

This table provides important feedback into users ability and comfort when working with their devices, or approach types, in the HN network. A high level of confidence working with a Mobile device, PC or Web browser was found with 88%, 86%, and 89% respectively. A sharp drop off in comfort was found when working with a HN Router, purchased and customized hardware. The skill areas of expert and advanced had the most comfort for these two approach types, these include: Router 63% and 95%, Purchased Hardware at 21% and 49%, and Customized hardware at 28% and 66% respectively.

Q7) Please select all actions you have made to your Wifi/Router or devices in the past)

Respondents (%)	Actions Wifi / Router
489 (89%)	Added a device to my Wifi
370 (67%)	Logged into my Router
239 (43%)	Updated my Router firmware
307 (56%)	Setup a new Router out of the box, including configuring Wifi and passwords
367 (67%)	Run an online speed test of my Internet connection (e.g. Google, Ookla)
240 (44%)	Logged into my router to review what devices have connected to my home network in the past
171 (31%)	Run a basic network scan to determine what devices are attached to my home network (e.g. NMAP, JNetMap, Network Scanner)
92 (16%)	Run advanced network diagnostics gathering: fingerprinting, topology, and Wifi layout. Using tools which typically require administrator or super-user privileges to execute (e.g. NMAP, Wifi Visual Analyzer, Internet Mapper Tool, Advanced IP Scanner)

While most users felt comfortable with adding a device to their HN Router, few ran a basic scan or network diagnostics. We found that 27% of novice HN users, 59% Intermediate, 81% of Advanced, and 91% of Expert logged into their HN router and also updated their firmware. We saw that 1% of Novice, 14% Intermediate, 42% Advanced, and 72% Expert HN users had a comfort with running a network scan, and reviewing their HN router for devices. While only 2% of Intermediate, 17% of Advanced, and 44% of Expert users where comfortable with all skills listed.

Q8) Please select all actions you have made to your Mobile or PC device(s)

Respondents (%)	Actions App / PC
526 (96%)	Installed a new App (application) on my mobile device (e.g. smartphone or tablet)
220 (40%)	Used a recommendation App to install Apps on my devices
287 (52%)	Run an App to review what applications are running on my device (Mobile or PC)
199 (36%)	Run an App to review what networks my device (PC/Mobile) has connected to in the past
250 (45%)	Run an App to review the security or privacy of my Mobile/PC device (e.g. Verizon Security, Norton)
138 (25%)	Run an app to review what devices have connected to my home network in the past
107 (19%)	Run an App (Mobile or PC) to review the security of my home network (e.g. Home Network Security, Sophos, Check Point ZoneAlarm
97 (18%)	Run an App to review the health of my home network and internet connection including tools that examine configurations, normal operation, security and local device privacy (e.g. Bullguard, Sophos, Cryptguard, Textsecure, Orbit)
12 (<2%)	Other (Open ended)

A majority of respondents reported installing apps on their PC/Mobile devices, while only a small percentage reported running an app to review Health (normal operations) of their HN in the past. Novice, Intermediate, Advanced, and Expert respondents reported 9%, 17%, 27% and 32% of interest respectively regarding review of application, network, and security/privacy of their Mobile device.

Q9) Please select your interests around managing, or understanding in more detail your Home Network

Respondents (%)	Interests Around Managing or Understanding HNs
271 (49%)	When changes happen to my Mobile (or similar) devices
290 (53%)	How to setup my home network correctly
338 (61%)	Speed to the Internet
380 (69%)	Why my Wifi is slow, and how to fix it (connection to devices on your home network, (e.g. printer, tv)
170 (31%)	How my network compares to others in terms of setup, devices, and a Mobile Applications
330 (60%)	Is a networked device secure (e.g. is anyone trying to break into the device?)
230 (42%)	What devices have connected to my Home Network (my devices, and friends devices)
279 (51%)	Home network health (e.g. is my firmware up to date, are there open devices on my network)
296 (54%)	Device health (Is my device running at peak performance in terms of memory, apps, storage)
325 (59%)	How private are my network devices (e.g. are they leaking data that can be seen by the outside world?)
322 (59%)	Detecting unauthorized devices attaching to my network, and automatically disable them
163 (30%)	How to use my home network to control multiple devices from one location (e.g. Alexa Dot, tvs, audio, video, heat, kitchen, bathroom)
160 (29%)	How to control how much of my internet bandwidth can be used by each device
8 (1%)	Other (Open ended)

A majority of respondents expressed an interest in Mobile changes when they happen, how to setup their HN, Throughput, how to report and fix Wifi issues, security/privacy and health of their devices and network. We also saw more than half of all respondents, across all skill levels, had an interest in the following areas specifically:

- ☐ Mobile device Changes
- ☐ HN Setup
- ☐ Speed of their Wifi and Internet connections
- ☐ Information about their Wifi Setup

☐ Devices within their HN, and when they change
☐ How to Control devices and Bandwidth

Q10) Which of the following would you prefer to use to gain access to information about your Home Network?

Respondents (%)	Preferred Access to HN Information
293 (53%)	My Mobile Phone/Tablet
364 (66%)	My PC/Mac
204 (37%)	A Web Browser
84 (15%)	My home network router
60 (11%)	A modified Router that allows for advanced diagnostics, and functionality
20 (4%)	A purchased unit of hardware specifically designed for Home Network data gathering (e.g. Cujo, Keezel, RaTTrap)
30 (5%)	I am not interested in gathering any information
5 (1%)	Other

Respondents across the board felt most comfortable with requesting to gather information from a Mobile or PC device.

Q11) What other information are you interested in understanding, in and around your Home Network? Let us know your thoughts.

Respondents (%)	Open ended / Other Interests
117 (21%)	Respondents provided feedback of some sort, ranging from requesting more information related to network security, to wireless setup.

We saw a strong set of feedback With more than 20% of users providing feedback. The feedback ranged from generic information to outlier requests in managing and servicing HNs. This information lined up with areas of interest using a Mobile or PC device to understand more information about HNs, along with areas of interests around throughput, optimization, security/privacy and health, comparison to other HNs, summary of information via an easily discernible app.

Q12) Email Address

Respondents (%)	Email Address
278 (51%)	Respondents provided an email address to contact them.

Devices and Browsers used to take Survey Tables 4.1 and 4.2 are the device types, Operating System (OS), and browser type used by respondents to take the survey. Table 4.3 and Figure 4.2 represent percentage and device count respectively of devices compared to Skill type, grouped from devices ranging between 1-3, 4-7, 7-9, and 10-16.

Table 4.1: Device Types Used to Take Survey (%)

Operating System	Percentage
Windows	48%
Iphone	18%
Mac	15%
Android	14%
Linux/other	5%

Table 4.2: Browser Type Used to Take Survey (%)

Browser	Percentage
Chrome	60%
Safari	23%
Firefox	11%
IE/Edge	7%

Table 4.3: Device Count Range Percentage vs. Skill Level

Skill	1..3	4..6	7..9	10..16
Novice	34%	43%	20%	2%
Intermediate	16%	45%	33%	6%
Advanced	11%	36%	39%	13%
Expert	12%	23%	35%	30%

4.7 Comparisons

In this section we provide results from the survey as related to comparisons of Skill level versus each of the questions. Results from this area are compiled by classifying

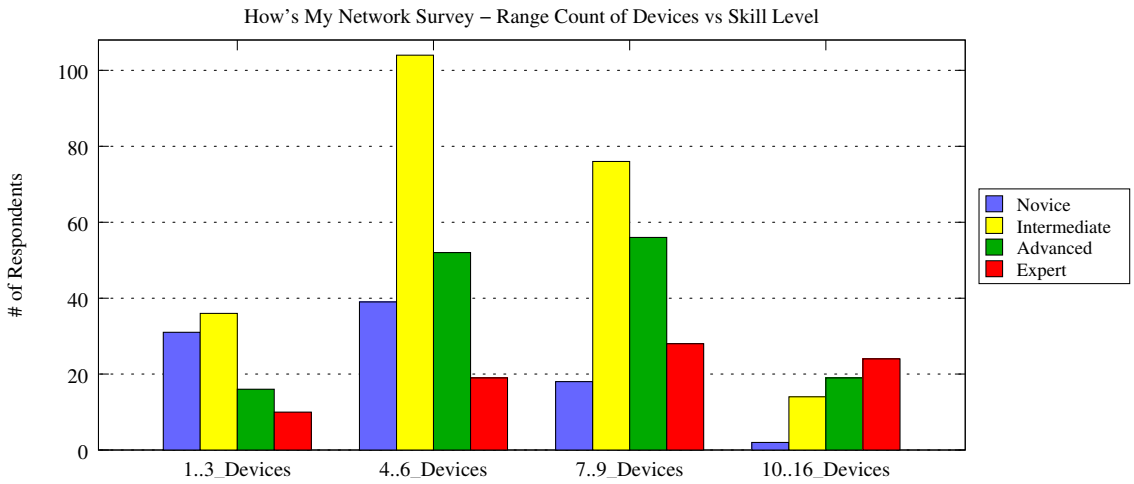

Figure 4.2: Device Count Range by Skill Type

the skill levels: novice, intermediate, advanced, expert. We calculate these comparisons juxtaposed by varying areas of these questions. We provide a short summary at the end of each question, where applicable, as a short review.

4.7.1 Gender vs. Skill

A comparison of Gender vs. Skill shown in Fig 4.3. In this area of summary Women respondents categorized themselves at a higher level in both novice and intermediate, and lower in advanced and expert skill level as compared to their male counterpart respondents. Table 4.4 shows the distribution as related to Female versus Male respondents across all skill levels.

Table 4.4: Gender Vs Skill Level percentages

Gender	Novice	Intermediate	Advanced	Expert
Female	23%	51%	22%	4%
Male	10%	35%	30%	25%

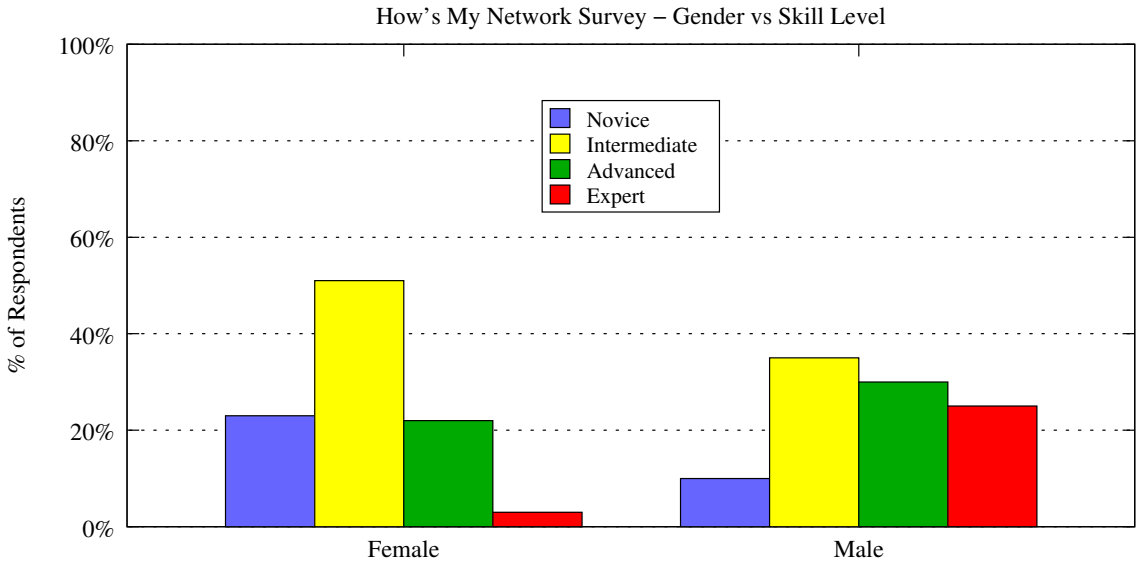

Figure 4.3: Gender vs Skill

4.7.2 Internet vs. Skill

A comparison of Internet type vs Skill is shown in Fig 4.4. 40% of novice users reported they did not know their Internet type vs that of 5% Advanced and Expert users. In all skills levels Cable was the dominate Internet connectivity media type.

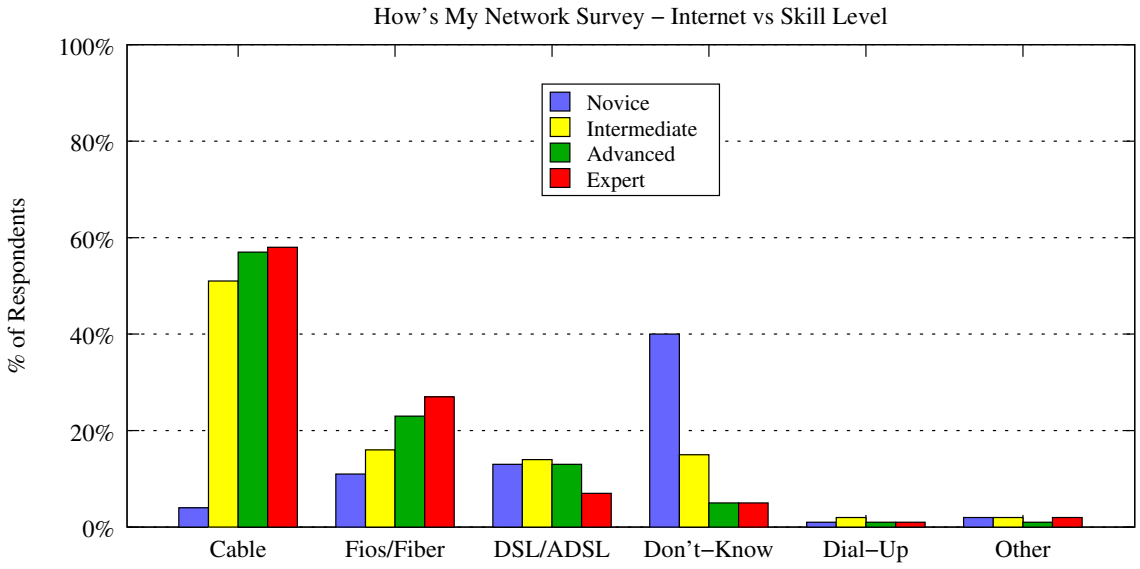

Figure 4.4: Internet type vs Skill

Table 4.5 provides some additional views into the differences between skill types and management; we have removed comments from this and rounded the percentages to the nearest digit. This figure follows the flow of the percentages shown from this question in terms of skill level across groupings. Table 4.5 also provides some additional views into the differences between skill types and management; we have removed comments from this and rounded the percentages to the nearest digit. This figure follows the flow of the percentages shown from this question in terms of skill level across groupings.

Table 4.5: Comfort level versus Skill managing HN Router

	Novice	Intermediate	Advanced	Expert
Not Comfortable	48%	12%	3%	0%
Rely on Friends	42%	31%	8%	4%
Need Help time-to-time	10%	45%	27%	2%
Very Comfortable	0%	11%	61%	93%

4.7.3 Devices vs. Skill

We next compare results between devices in the HN and Skill level, which can be seen in Fig 4.5. Figure 4.5 shows little change between Advanced and Export users in terms of device types in a network, save security and smart devices. Overall their is little difference between devices and skill level. In this figure Reading/Kindle device is one that is used for Ebooks or similar in users HNs. We removed outliers from this figure as the areas were too low in terms of values received.

4.7.4 Abilities vs. Skills

A review of Abilities to install/configure a Home Network or Home Router vs Skill level can be seen in Fig 4.6. We can see that Novice users had the least comfort in terms of installing/configuring a HN router, and was almost 4x less comfortable than their Intermediate counterparts. Intermediate users where 5x more likely to rely on friends

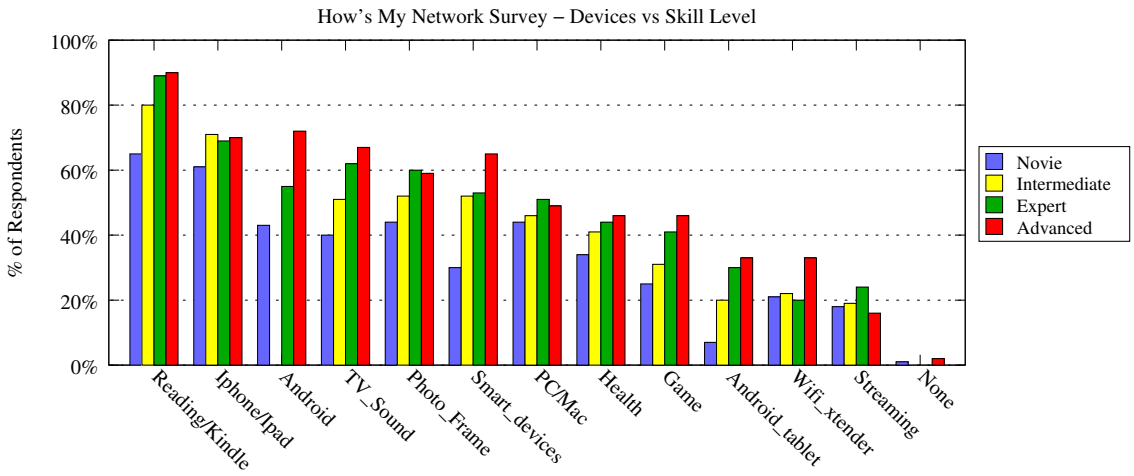

Figure 4.5: Devices in HN vs Skill

vs. their Advanced cohorts, and almost 2x more likely to need help from time to time. Advanced respondents are also ˜2x as their Expert counterparts to rely on friends, and ˜11x to need help from time to time. Expert users are 9x and 30x that of intermediate and advanced users in terms of being very comfortable managing their router.

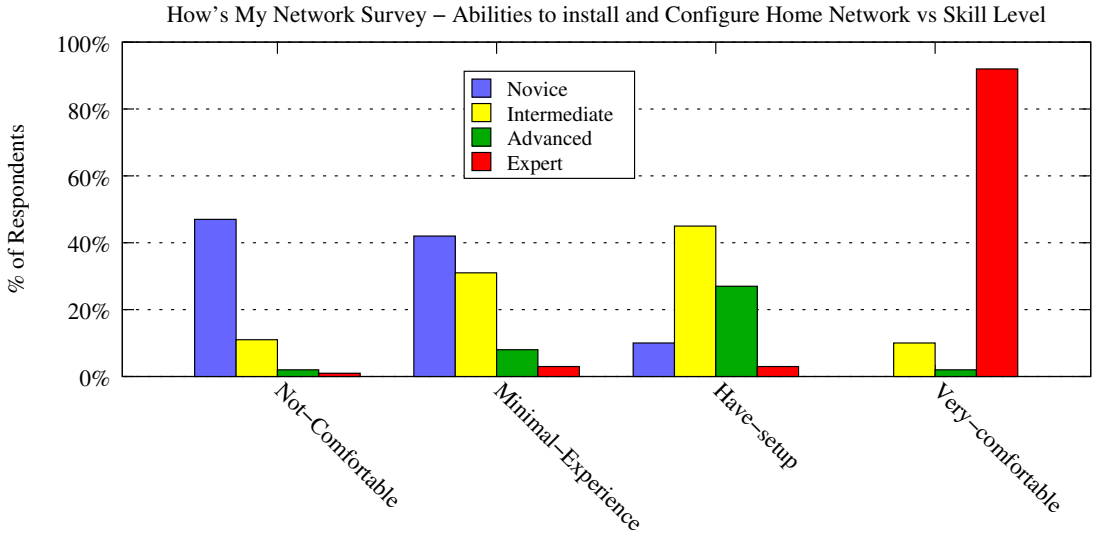

Figure 4.6: Ability to Install and Configure Router

4.7.5 Comfort Level vs. Skill level

In this area we are looking at comfort level with managing several types of devices or software/tools vs that of skill level. The areas we reviewed include: Mobile devices, PCs (Mac or Windows), Web Browser comfort, Router comfort, purchased hardware with specific tasks (e.g. Nest, etc.), and customized hardware/software solutions (e.g. Linux). We have created a graph for each of these comparisons against the skill level counterparts. These plots include the following: Mobile Fig 4.7, PC (Mac/Windows) Fig 4.8, Web Browser Fig 4.9, Router Fig 4.10, Purchased Hardware Fig 4.11, Customized Hardware Fig 4.12.

Starting with the Mobile vs. Skill level we can see from Fig 4.7 that most users are fairly comfortable using their Mobile device. Only a small percentage of Novice respondents reported they no idea how to use a mobile device. Similarly we can see that PC users, Fig 4.8, have a high level of comfort using their Windows/Mac system. ˜35% of Novice users reported that they need help in some capacity when using a PC vs that of only ˜10% of Intermediate cohorts.

A review of Web Browser comfort vs Skill level, Fig 4.9, shows that there is a high level of confidence overall when using. ˜25% of Novice users reported that they need help to use their web browser in some capacity vs that of only ˜12% of Intermediate, and ˜2% of Advanced users. An interesting fact in this case is that 100% of Expert users reported being very comfortable with using a Web Browser.

We next move to the review of Router Comfort vs. Skill level, Fig 4.10, and find that almost 7x more Novice users had no idea to use their router vs. that of Intermediate group, while there was there was only a ˜30% difference for this group in terms of needed help from friends, and an ˜10x difference in being comfortable with their router. Intermediate respondents reported at a ˜3x less being comfortable versus their Advanced counterpart, and ˜5x less than the Expert grouping. The grouping reported a 40% increase over In-

termediate in being very comfortable using their router. We can see a sharp increase of comfort as the skill level increases across each of these levels for the Router category.

A review of purchase hardware (e.g. Nest, etc.) vs Skill level, Fig 4.11, shows a geometric drop in need for help across all areas for the skill levels. There is a 5x increase between comfort level of Intermediate and Expert, vs that of 2.5x between Intermediate and Advance. A 2x+ increase is seen between Advanced and Expert respondents being very comfortable with these devices.

As a final comparison in this subgroup we reviewed Customized hardware/software (e.g. Linux, etc.), Fig 4.12, and Skill level. We can see similarities in terms of Novice, Intermediate, and Advanced users needed help using from friends. With Intermediate users ˜3x more likely needing help than their Advanced counterparts, and ˜10x more likely then Expert users. Expert respondents are 2x more confident vs. the Advanced group when using these tools/hardware.

In summary, the overwhelming majority of comfort, for each of the device types listed, grew at at a minimum of a linear rate (in some cases exponential) in terms of user classification and comfort; save the exception are of Web and Intermediate users.

4.7.6 Actions Wifi/Router vs. Skill level

We next move to a review of actions respondents have successfully accomplished using their Wifi/Router vs Skill level, Fig4.13, was completed. We break this area apart by Skill level and compare each of the sub categories. Each skill level reported adding a device to their network via their Router at similar levels. Novice users were 50% less likely to have logged into their router vs. Intermediate, and 2+x less likely vs advanced and Expert. $<$5% of Novice respondents reported to updating firmware, which was 5x less likely than intermediate, 12x less than Advances, and 18x vs that of Expert users. Intermediate users were 2x and 3x less likely to have run a device scan or logged into their router vs that of

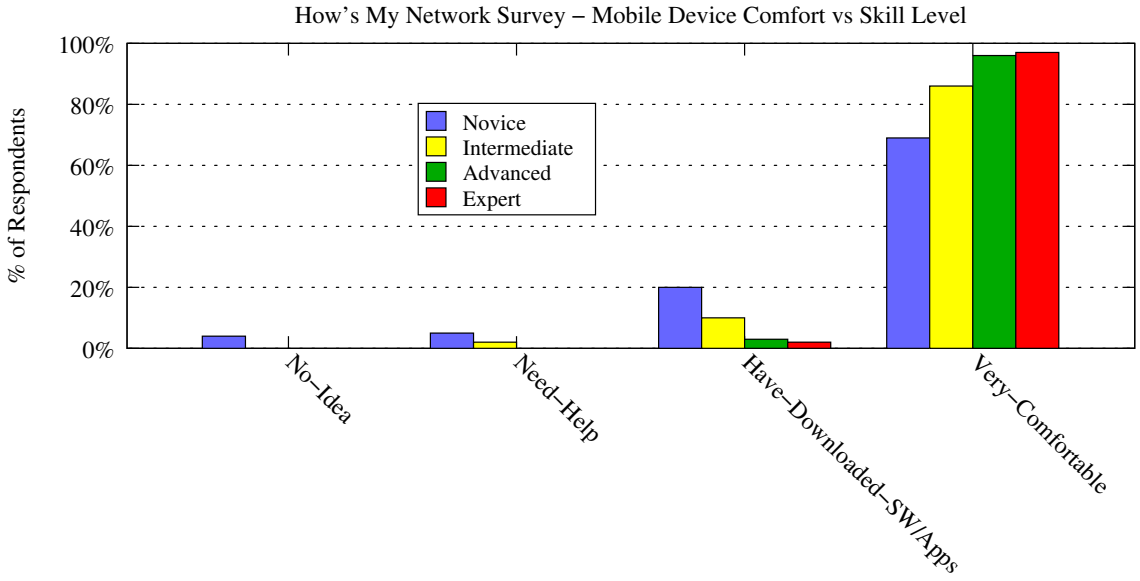

Figure 4.7: Comfort Mobile vs Skill

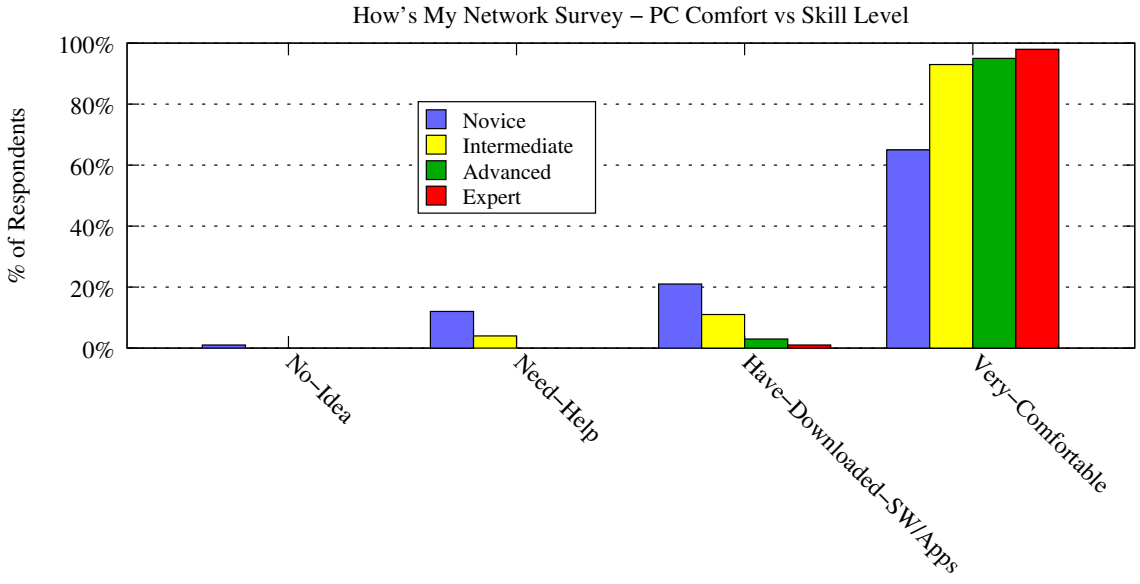

Figure 4.8: Comfort PC vs Skill

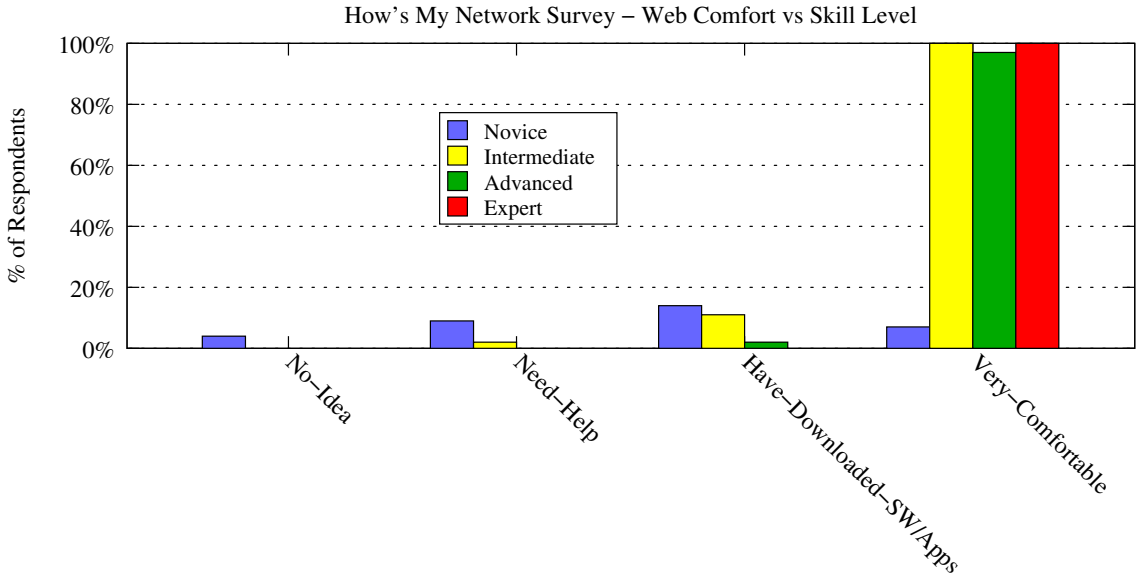

Figure 4.9: Comfort Web Browser vs Skill

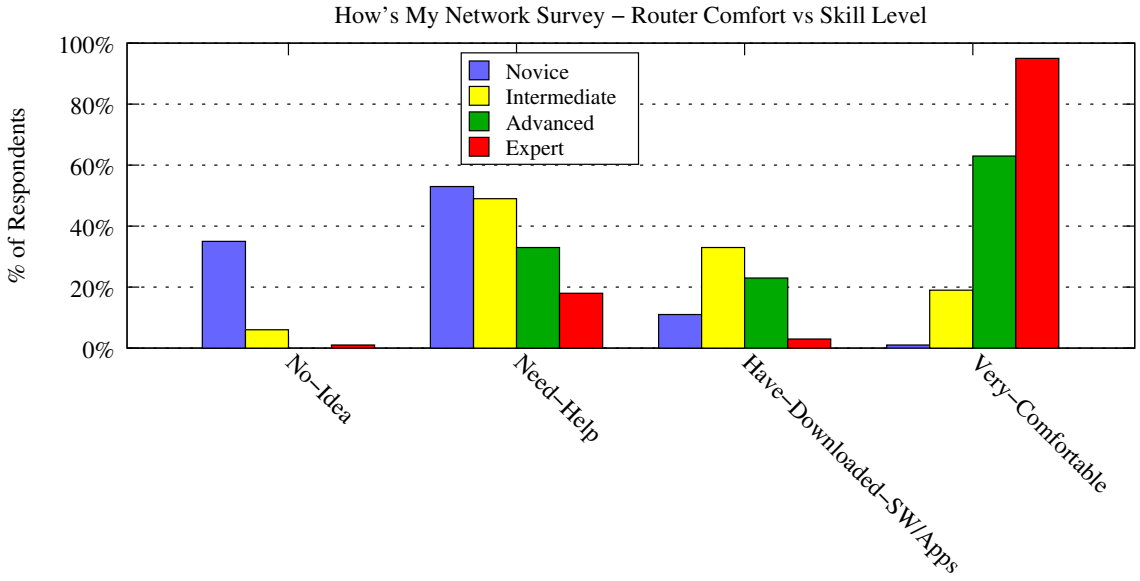

Figure 4.10: Comfort Router vs Skill

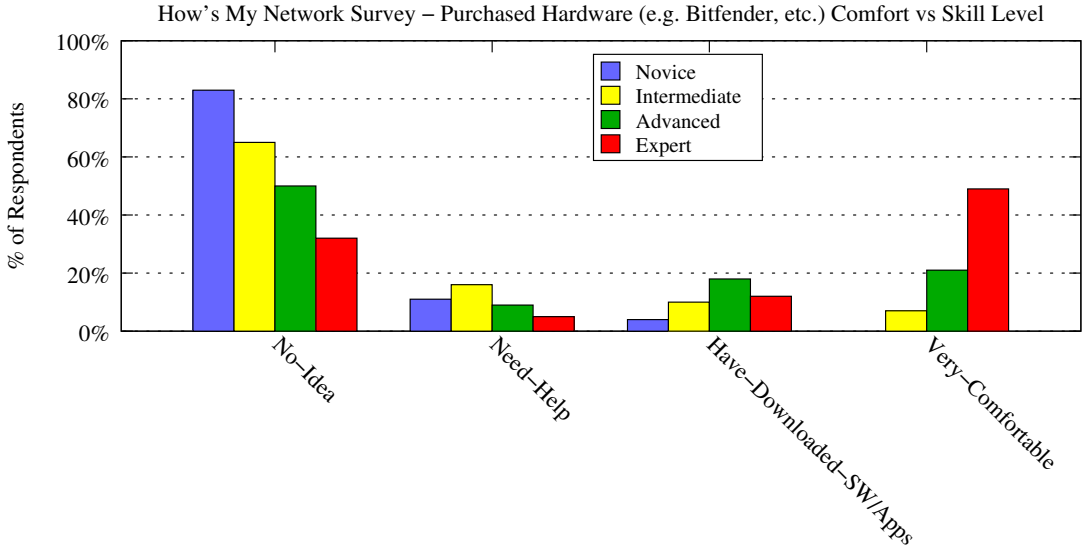

Figure 4.11: Comfort Purchased HW vs Skill

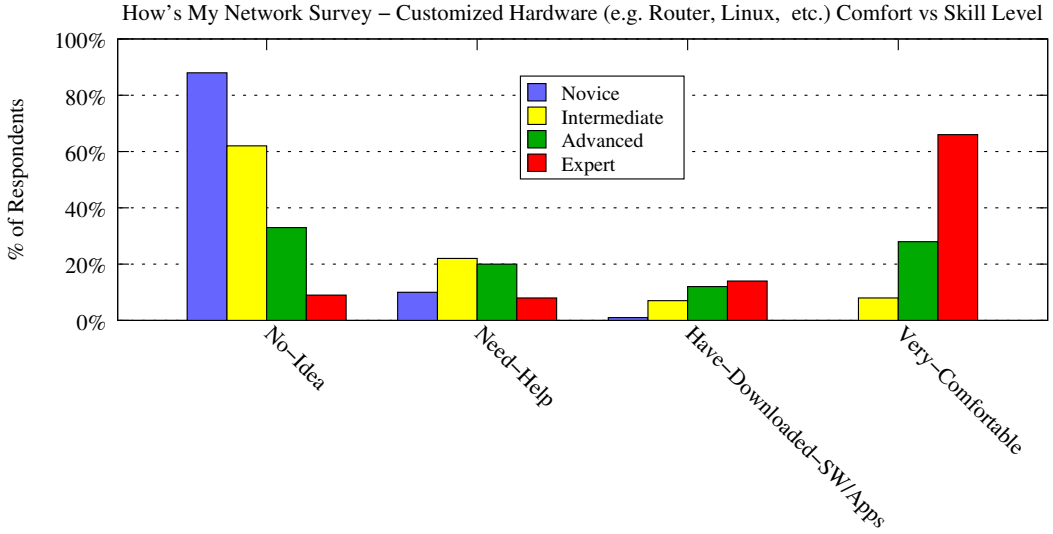

Figure 4.12: Comfort Customized HW vs Skill

their Advanced and Expert cohorts respectively. More than 50% of all Expert respondents ran simple or advanced network or device diagnostics in past. Intermediate respondents were 3x less likely to have run an advanced network diagnostic, and 50% to run a basic network scan versus the Advanced grouping.

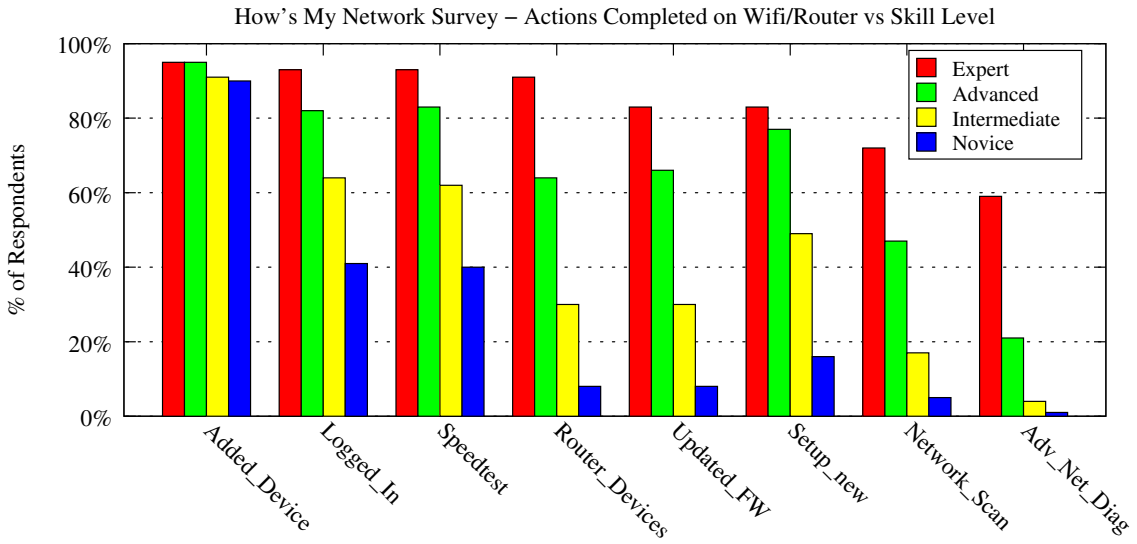

Figure 4.13: Actions Wifi vs Skill

In summary, users across all areas, of ascending difficulty levels, increased their interactions on their Wifi/Routers according to their skill level. This showed that expert users would have experience with running advanced skills versus that of novice users who did minimal actions across these device types, and rely heavily on assistance.

4.7.7 Actions Mobile/PC vs. Skill Level

Similar to the Wifi vs. Skill we compared respondents Actions completed on a PC/-Mobile devices vs. Skill level, Fig 4.14. An interesting item to note is that users across all skill levels downloaded, installed, and used a tool or app in each of the categories listed. The most popular areas across all skill levels were recommendation tools, tools that reviewed apps on the device, device Security/privacy, and network device discovery. ˜50%

of all respondents found interest in tools that reviewed apps running on their device/system, and overall 20% of all respondents had interest in reviewing device health. ˜50% of Intermediate and Advanced respondents had run a tool to determine which SSIDs their device had attached to in the past, while ˜20% and ˜40% respectively had used a tool to determine network security.

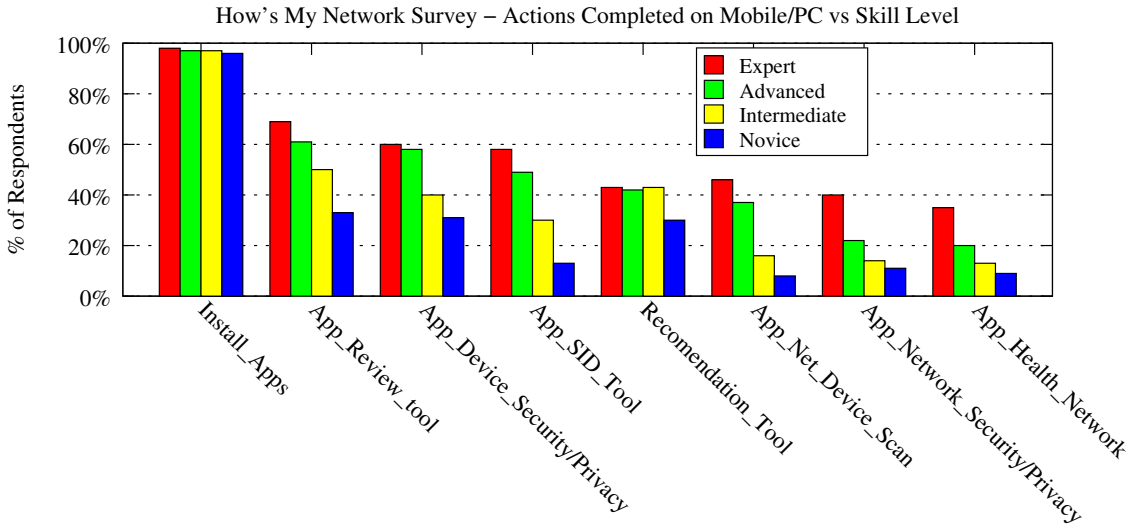

Figure 4.14: Actions Mobile/PC vs Skill

4.7.8 Interests Managing vs. Skill Level

We next move to examining respondents Interest in Managing or understanding in more details their HN vs Skill level, Fig 4.15. The following are some interesting data points from this review. ˜50% of all respondents, across all skill levels had interest in the following areas: Network device Scanning (e.g. when a device was added to their network), how to setup a HN, throughput of their local and Internet connection, why Wifi is slow (this had the highest overall rating across all skill levels), device privacy, and detection of unauthorized devices. ˜40+% of Intermediate and Advanced respondents had an interest in how their HN compares to other.

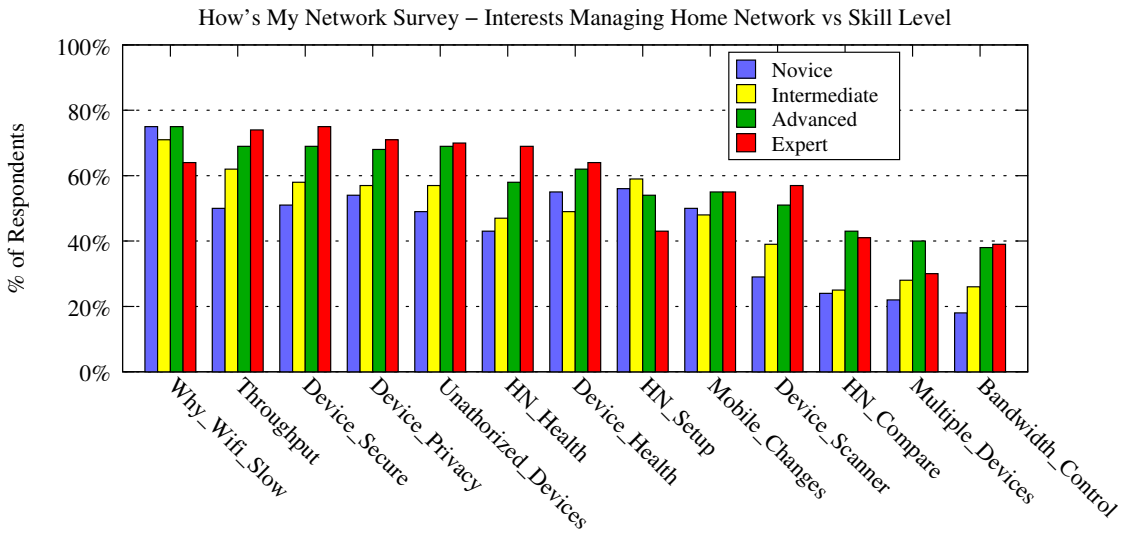

Figure 4.15: Interests Managing HN Vs Skill

4.7.9 Preferred Device to Review HN Info vs. Skill Level

In the final comparison we reviewed respondents preferred method to gain access to information about HN (e.g. using Mobile, etc.) vs Skill, Fig 4.16 shows all respondents preference. Across most categories, and groupings, of skill level respondents reported preference using Mobile or a PC to view HN information. Web was the third (3rd) preferred method to view HN information across all respondents and skill levels, and was ˜2x less of a preference to Mobile for Novice, Intermediate, and Advance users. Respondents also showed an interest in understanding comparisons across networks and mobile devices, with 43% of Intermediate 40% of Advanced HN users.

To understand preferred device and skill levels we have reviewed two population groupings (WPI and Social) from the 550 respondents for comparisons of homogeneity versus the heterogeneous grouping. Figures 4.17 and 4.18 show a breakdown of skill levels across these two timeline populations (WPI and Social) groupings. The WPI and Social groupings are similar to what is shown in the entire pool of respondents, Fig 4.16. There are some minor differences between the WPI and Social groupings, these incldue:

social mobile preferences being slightly smaller than the WPI grouping. Overall we see minor differences between the heterogeneous set and the homogeneous WPI and Social groupings.

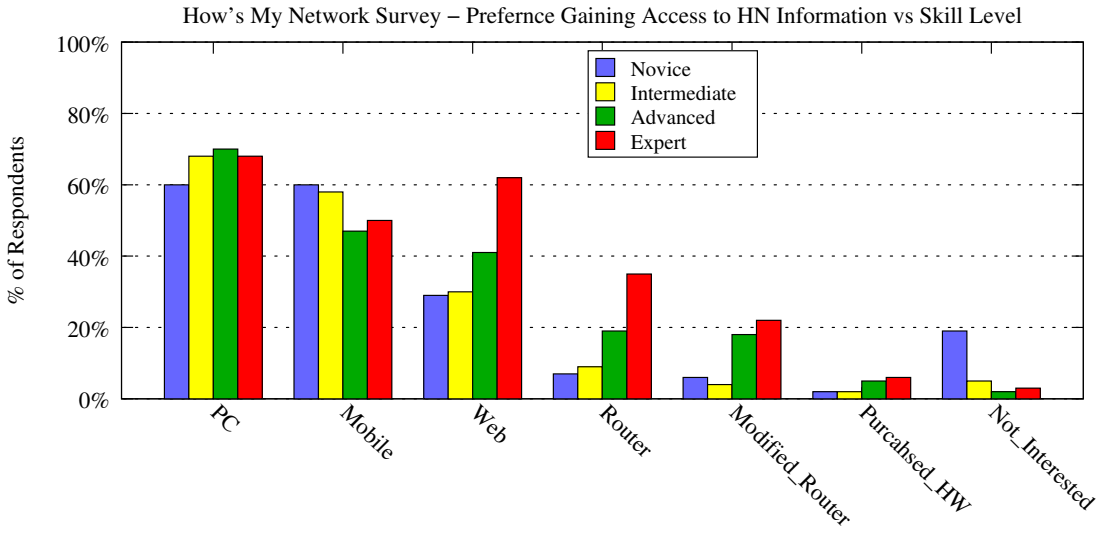

Figure 4.16: All Respondents - Preferred Method Gain Access vs Skill

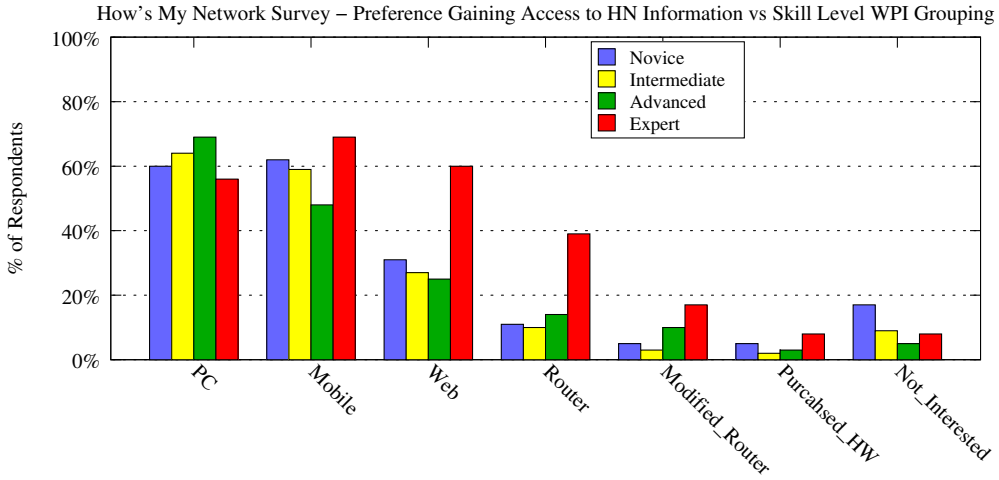

Figure 4.17: WPI Grouping Preferred Method Gain Access vs Skill

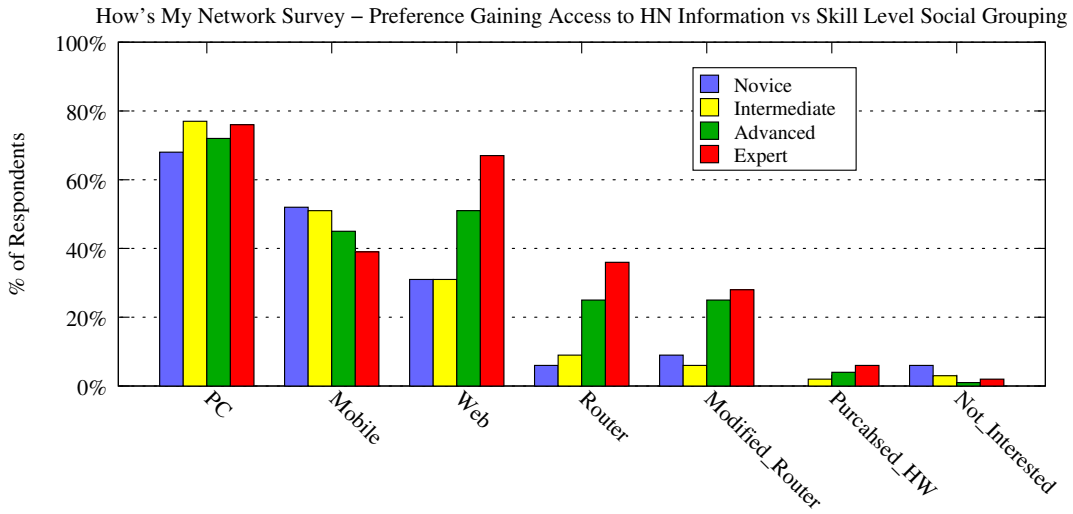

Figure 4.18: Social Grouping Preferred Method Gain Access vs Skill

4.8 Discussion

In this section we provide a discussion of the study, and changes that have occurred from previous studies that have done work in and around the HN areas.

Past work in this space has focused on high impediment approaches that may require users to have high skills to operate, and thus may excludes average HN users. As an example, previous work done by [80] [157] [50], and others, used restricted methodologies using tools that ranged from fat client software (manually downloaded and configured executable files), and customized hardware. Results from this HN survey generally support that users would prefer to gather data in an open and generic manner using ubiquitously available approaches with minimal impediments and a high level of incentives.

As part of this work, and the takeaways from this survey include the following:

1. HN users have an interest in understanding more information about HNs and prefer to leverage a tool running on a Mobile app.

2. Expert users are the only set of users who do not need help and felt completely comfortable with using and managing their HN, devices, or advanced tools.

3. HN users have an interest in understanding when changes occur in their HNs, such as devices, mobile security and privacy changes, and Wifi health.

4. HN users are more interested in gathering information using their Mobile device versus using a Web browser.

Next, we show the changes that are reflected from the review of this survey and how it relates to changes to the current HN landscape. We look at the results from the last two questions of the survey, which asked about user interest in managing or understanding HN and their preferences to gaining access to HN information. Comparing these questions with skill level, and examining the landscape as a whole via the survey we can see that users have a keen interest in a variety of information in and around their HN, including: network discovery, throughput (e.g. why is my wifi slow, etc), and other areas of HNs. The interests in this area is a key indicator that HN users are indeed looking for information to help optimize their environments. Users have shown an overwhelming interest in the gathering of data from their HN. In addition, respondents pointed toward using a mobile device as a data collection point when collecting data. This indicates the desire for a Mobile app in this space as it was highly desired by respondents.

The following are takeaways from this work based off of the comparisons and summary, where we find the following:

⋆ Until users reach an expert level they still feel they need help with basic functions, such as setting up a home router.

⋆ Novice users need the most support managing and configuring their HNs. This includes basic setup of routers, devices, apps/software, Web, customized hardware, and also includes areas such as security, privacy, and health of their devices and HNs.

- ⋆ Intermediate users fall into a similar category as their Novice counterparts in terms of support required across these range of services.
- ⋆ Advanced users claimed they had experience with specific areas, similar to their expert counterparts, but where 3x less likely to have these skills as compared to the expert groups (in some cases).
- ⋆ Expert users are willing to support other HN users
- ⋆ Expert users are the most likely (˜3x more than their closest counterpart) to use customized tools, such as a modified router, Linux, or have leveraged tools to scan their network.
- ⋆ A majority of users across all skills levels are interested in understanding HN specifics such as the following approaches: Wifi Speed, Internet throughput, device security/privacy, device detection, device and HN Health, and Mobile changes (e.g. apps, permissions and similar). This also includes changes, and norms around these changes and a comparison of their HN to others.
- ⋆ All users across all skills levels showed a high level of interest in understanding how to make changes to their devices, or HN to optimize their experience.
- ⋆ A majority of users across all skill areas preferred using a Mobile device to determine characteristics within their HN.

In addition to the takeaway points shown, a deeper dive into the desired preferences, along with approaches, as classified in Chapter 3, include the following desired areas:

- Wifi Speed: Identify health of connectivity, and attributes related to operation.
- Internet Throughput: upload and download, including comparisons to others.

- Device Security/Privacy: When changes happen to devices, and where to review these changes.
- Device detection: When new devices appear on the network.
- Device health: A fingerprint of how the device is operating, including a comparison to others.
- and HN Health: A fingerprint of how the HN is operating including throughput, devices, Wifi, etc. along with comparison to others.
- Mobile changes: A review and notification when apps or permissions change on the device.
- HN Norms: A review of how a HN and Mobile device compares at the local and global levels.

4.9 Summary

In this chapter we have provided details of the HN survey and study we completed, where we have shown HN user interests in wide variety of areas. These areas of interest point toward the creation and extension of HN studies that offer high incentives and minimal impediments for operation, and general points toward leveraging a Mobile app for this research. This work should provide options for security, privacy, information discovery, norms of changes, as well as health of the environment being reviewed. We have found that users have an interest in understanding the layout of their HN, Norms, privacy and security, and the leveraging of a Mobile app for data collection.

Takeaways from this work include the following:

1. an understanding of perceived value of information and experiences of HN users when managing and using a HN;

2. examination of device types, Internet connectivity, management options, interest, and preferred method of management of user HNs;

3. examination of HN user skill level related to HN management and comfort;

4. show that HN users across all skills levels showed a high level of interest in understanding how to make changes to their devices, or HN to optimize their experience;

5. show that users are interested in an in-depth view of HN information;

6. show results and characteristics that incentives user participation; and

7. show that HN users have an interest in leveraging a Mobile app for data collections within a HN.

Chapter 5

HMN Mobile App

In this chapter we provide details on work completed in 2018-2019 on the How's My Network project where we leveraged a Mobile app for research. This work is the final pillar or thrust of our dissertation. This work is complementary to the work we completed leveraging a web browser and Java, Chapter 2, for HN data collections, background work including attributes of HNs in Chapter 3, and a HN user survey to determine user incentives and impediments for participation, Chapter 4. The HN survey and background work pointed toward leveraging a Mobile app as it was general preferred by respondents for collection in the HN Ecosystem as it has the ability to provide high incentives for participation. We extend into the ubiquitous space of Mobile apps to delve into HNs to understand the HN Ecosystem and provide HN user-based results, as well as gather user feedback. We have developed a platform to measure information from the Home Network (HN) that includes: devices, apps privacy/security, networking, comparison of norms, and an approach to collect user feedback all via a Mobile app. We have created and deployed this multifaceted Mobile app approach, which combines HN device and app discovery, in order to discern the current HN Ecosystem residing behind the firewall, as well as providing value, in terms of results, to help better understand the HN user experience. While

devices and apps appear to be two distinctly different areas of functionality within HN, we consider them similar and part of the HN Ecosystem as they are additive entities and part of the HN. We can extract HN Ecosystem information by examining the running conditions of the environment that includes a review of the local network and the operational status of apps and local device configurations. This work looks to create a high incentive approach for data collection and results, using a low impediment approach (Mobile app) for participation.

This work provides details on the capabilities of the app, measurement methodology, how it was deployed and tested, comparisons and norms of data, feedback results, as well as impacts on user perceptions and changes made to the HN ecosystem due to this study.

5.1 Introduction

Over the past 10+ years Home Networks (HNs) have increased not only in prevalence, but also in terms of devices, apps, networking capacity, and overall users. The Internet has seen an major increase of world wide users, almost 2-fold in the past ten years, where we have seen (as of March 2019) ˜4.3B users (world wide) or roughly 56% of the world's population leveraging the Internet [185]. The explosion of the Internet for commerce, entertainment, and social connections has presented users with new devices and residing within their HNs. These devices include user applications (apps), and management tools that provide access and services to these new devices available. These apps (and devices) present security and privacy concerns as bad actors work to infiltrate and gain access to data. Techniques such as stalkerware [73, 2] allow developers to gain access to a user's device, location data and other sensitive information, under the guise of a non-intrusive tracking tool/app. As an example, an exploit discovered by Kapersky labs 'Triada' [166] found that 3rd party Android phones could exploit Google's administration permissions,

via a rooted Trojan app, to gain privacy/security access to the entire system, including all apps installed via Triada.

In addition, it has been estimated that $2/3$ of phone apps share data with third-parties[168], $3/4$ of apps [1] downloaded last year have vulnerabilities that could let hackers steal passwords and other sensitive data, and $2/3$ of antivirus [167] apps do not work properly. These Rogue 3rd-party apps [164] have collected passwords for 100k+ individuals and include network names, SSIDs, locations [6], plain-text passwords, and more [12] [151], and [11]. These new apps and services present HN users a litany of security, privacy, and overall general dilemma around gathering and providing information in a discernible way. We also reviewed research done around browser cache and privacy and security. Oren [123] looked at implications of privacy of exposing data via a Web browser, while Yaoqi [79] looked at exposure of Web cache for Geo-inference attacks, others looked at security and privacy implications of browser cache as related to uniqueness of Web browsing history patterns [120].

5.1.1 Previous Work

Past HN studies have focused on high impediment approaches with minimal information to HN users. Work done by [97] [80], [81], and [50] gathered results within a confined HN environment, and require a level of experience to operate, but provided minimal information for users to understand the overall and general HN operations of their devices, apps, and network environment. Work done by [101] reviewed specifics of app privacy and proposed a privacy model, but focuses on the LBE Privacy Guard tool-set, an app that requires a rooted Android phone, to control permissions. Other similar work done by [53], and Microsoft Corp [82] used surveys and or interviews to review app permissions on Android devices and proposed new run-time extensions for app privacy and security. Extending on this research others have looked to create new models and extend

security [65, 152]. We have seen other studies such as commercial work done by [54] include Wifi and network awareness, but provide minimal information to HN users around the overall nature of a HNs (e.g. apps, comparisons between other HNs, ratings of HN, etc). Other commercial work include vendor specific hardware and apps to identify devices on a HN (e.g. leveraging HN routers and Wifi hardware) for this assessment. Vendors and ISP (Internet Service Providers) such as [32] have started to provide generalized layouts to users, but are focused solely on device availability on the user network.

There have been attempts in this scope, in the PC and Mobile space, but they are also confined to a small area of focus and do not provide the results across the litany of areas HN users are accustom. Commercial and research tools such as [20][153][34][163], and [122] are targeted around virus protection, Web browsing, proxy services, app privilege elevation, or tracking of devices outside of the HN, but do not include a holistic view of HNs. Other studies have looked at security as related devices, Network, and apps [118] but are slow (or never complete executing) and provide minimal information in terms of results to users. Other studies have focused on IoT devices (in and out of the HN), but are challenging to manage for users as they either require administrative (root) permissions and a savvy user to install, configure and operate [172], [77][173], and have a high impediment to participants. Studies such as [118] have taken a more pragmatic approach to leveraging mobile devices for discovery, but provide a narrow review process and minimal generalized health or comparisons to other HNs. The study done by [97] extended their work and created a mobile app [113] (now defunct) that provided a review of Internet bloat and some other minimal aspects of HNs, but was targeted toward researchers versus the HN user. These and other studies have focused on producing raw or complicated results that are either too difficult to run or are challenging to understand in terms results, and have minimal or low incentives for participation.

5.1.2 How's My Network

This chapter is a continuation of the How's My Network (HMN) project and specifically on HN research, where we are looking to gather and provide information and a track for HN users to gain information to better understand their environments. We started work on the HMN and project and focused in on a user-centered measurement platform [140], where we looked to maximize incentives and minimize impediments to users by providing reasonably valuable information. This work is the next phase of research in and around HNs and leverages a study we completed on network types existing in HNs [137], as well as a survey [138] of HN user preferences of devices and access. We present a study and research in this chapter as part of the larger HMN project focused on high incentives and minimal impediments for participation. This work looks to examine a high level view of HNs, including HN user results, allowing for minimal impediment to the user in terms of form, function and usability. We also look to gather feedback and understand changes HN users will make due to leveraging a tool specifically targeted at security/privacy, and comparisons of HNs.

In addition to these areas of review it is important to note the distinction between Mobile app discovery (e.g. configurations, permissions, and setup) and basic HN analysis, as it points toward usage, security and privacy, and overall operational status of the local device. HN properties include network, devices, and apps residing within it and apps provide a rich amount of details about usage and connectivity HNs. Examining the apps running on a local device allows for a deeper understanding of category of apps, permission, and more importantly the privacy and security of the apps installed (including changes in usage). Mobile devices are tightly coupled to a HN and provides a link between access of resources and connectivity to the HN and Internet services. We believe that these devices and apps are simply an extension of the HN and are associated and can be homogeneously grouped. We think that HN users install these entities (devices and

apps) into their HN without having an understanding of what they are getting besides the advertised features (e.g. privacy and security issues). IoT and other IFTTT (if this then that) based protocols are now part of the HN, and are also congruent and integrated as part of the device layout and configuration of the HN. As part of this work We treat these devices and apps as part of the HN ecosystem and look to understand the entire HN which includes the physical network, devices, configuration, and apps residing within it. We believe we are the first to make the distinction that the HN Ecosystem is a combination of both devices and apps within the HN Ecosystem, and is more than just the network. We believe these to be the case and they they represent the modern HN Ecosystem, and we look to continue our work to examine both areas (Mobile devices and apps, and HNs) to provide a clearer picture of access, control, configuration, and help understand the functional standing of the entire HN environment.

5.1.3 Research Contributions

It is clear that there is a need for a new approach that focuses on gathering data and results for both users and researchers, as well as provides a holistic review of operational Health status and diagnostics of the entire HN. In our study we identify user ratings to help quantify and qualify those areas of HN are of interest to review [138]. This study focuses in on discovery and how user feedback is important in HN research across a generic set of HN Users. We believe that data collected, via a Mobile app, can provide the best trade-off between ease of use (minimal impediments) and the amount of data that can be collected and results (incentives for participation). This includes being able to collect data about all apps installed on a Mobile device, perform network analysis, and also gather valuable feedback from users around a wide array of areas of HNs. In this vein we have continued our work on HMN and have created a Mobile app that provides HN users a wider holistic view of important aspects of their HN, namely: Wifi health and rating, devices attached,

apps privacy and security with an overall rating, Internet throughput, and a how their HN compares to others (norms). We are working with a Mobile app approach to allow for the best trade-off of convenience and capability, while being able to manage the inclusion of apps, devices, IoT, and other functionality of privacy and security concerns of the HN ecosystem.

As part of this work we provide the following research contributions:

1. demonstrate that a Mobile app platform is valuable to collect data and features of HNS (e.g. devices, privacy and security, as well as other information) and can display this HN information in an functional manner;

2. learning how users respond to information based on integrated feedback and actions taken

3. the ability to measure HN wifi and apps operational Health status and functionality;

4. the ability to determine user actions (e.g. changes to apps installed) based off of HN results and user feedback;

5. provide a method of research measurement that combines collection and analysis of Mobile apps and devices, 3rd party entities that users are choosing to install in their HNs, into an associated and homogeneous grouping.

The rest of this chapter is organized as follows: Section 5.2 provides details of the app description; Section 5.3 outlines research questions of interest for this work; Section 5.3 provides research questions; Section 5.4 describes features; Section 5.5 provides details on data collection points; Section 5.6 is a background on how information is presented; Section 5.7 provides details on the study; Section 5.8 presents the results; Section 5.9 provides results from feedback; Section 5.10 is a discussion of research questions; Section 5.11 is a review of the areas we were surprised to find as part of the discovery process,

and Section 5.12 concludes with summary of this work. Appendix 6.2.2 provides additional information for review to support this work.

5.2 App Description

Continuing work in the HMN platform we have created a Mobile app for the Android environment, How's My Network Mobile app, which is available for most Android devices. We selected a Mobile app for this work as a majority of HN users, from the HMN Survey [138], generally reported a preference to use a Mobile app to gather data from within their HN. A Mobile device allows for access to both areas of interest within the HN ecosystem, namely apps and networking characteristics. We selected the Android Mobile environment as it comprises ˜75% of the marketplace worldwide [8], and is favorable to HN users in terms of usability and connectivity to the Google store for downloading, and operations.

An approach to gathering HN data was completed by the creation of the How's My Network Mobile app, that allows for flexible factors of data collection, while allowing for an agnostic approach across a set of Android devices. The approaches used are discussed in the following sections, along with the data sources. In addition, to capture these data we have created a highly flexible, customized, secure approach to data management and distribution using client/server methods, that is available for analysis and client communications. An apps-based approach lines up with a strategy of both local device configurations (e.g. security and privacy) as well as allowing for a robust method to access networking information and data.

The HMN Mobile app was designed to have a similar feel of other well known app categories, such as used by Entertainment (e.g. Social Media), Health and Fitness (e.g. Fitbit), and Games (e.g. Words With Friends). We also used best practices for Design as

shown by [111] [106] in terms of layout, color schema, and format. These best practices as shown by [107] also provide clean methods for the displaying of captured data in an easily discernible way. The HMN Mobile app runs as a 3rd party app, as defined by Google App/Play store [164], and requires minimal permissions to execute; it does not require administrative privileges to operate. An elevated Google app permissions for granular network access is requested (and required) to gain access to the locally connected SSID (network name) the HN user is running within; this elevated permission is only required in versions of Android $>$7.0, and thus is requested as needed during run-time; if this request is denied the HMN Mobile app operates normally without retrieving this information.

The HMN Mobile app provides the HN user a simple three tab layout, with a single start/refresh button and swipe down options on each tab. The user can start a new HN scan via the start/refresh button, and the HMN Mobile app was initially designed to attempt a new scan every 15 minutes or sooner, see Appendix A.1 for additional details on how to run a scan. Each new HN scan (user initiated or automatic) is designated with a unique identifier along with a unique device code. All data is encrypted and transferred to a secure server for future analysis. Users have the option of ”opting-out” of data logging or can delete all local data stored, via the settings features within the app; users can continue to use the app after opting out without issue. In addition the user can define which areas of the HN scan should execute or not via preference settings.

In the remainder of this section we review important aspects of a HN, including: defining and describing what a Home Network is, provide details on how we collected user feedback, describe health ratings of apps and networks, describe the classifications created for the gathering and Norms of data, and provide a review on integration of data collection feedback used in this work.

5.2.1 What is a Home Network

We define a Home Network (HN) as a residential environment that consists of an entry point device(s) serving up Wifi, modem, router, and switch in a typically NAT'd (Network Address Translation) domain. This residential environment is typically served up Internet service by an ISP (Internet Service Provider) via DSL, Cable, or Fiber; although 5G and other telephony connections may exists they are all backbone'd via one of these methods, typically Fiber. An important factor to a HN is the Wifi services provided within the environment, which are typically home branded Wifi services versus commercial or business grade hardware. Although some homes may have multiple entries (e.g. Internet Services) and multiple Wifi hot-spots (e.g. Wifi extenders), they are designated as a residential Wifi provider versus that of a business as the typical HN is not supported (internally) by professionally staffed Information Technology teams, although their ISP may provide fee-for-service. Devices and apps are included as part of what a HN consists of, and includes the functionality and features of both.

It should be noted that the How's My Network Mobile app can run against any network environment the Mobile device is connected, this includes hot-spots, business, and EDU-based networks. The app uses a setup of affinity to a local HN Wifi as accepted by the user, via pop-up questions, and internal settings; if a new network scan is attempted, on the non default Wifi, a pop-up will ask for permissions to change the affinity to this new HN, only then will a new scan begin. We look to include future work around the scanning and analysis of non HNs so we can compare differences in capacity, devices, and environments, but focus in on HNs for this study.

5.2.2 Feedback

We have created a series of 16 questions (see questions in Section A.2) as part of this research project, and the user is prompted with one at the conclusion of each HN scan (via a pop-up) and daily, if they have not run the HMN Mobile app. The first time the HN user runs the app they are asked the following question ”How has your perception of security/privacy in your Home Network changed within the past week?” This question is re-asked ever four days to provide a base line and on going metrics and feedback comparisons. We selected an approach type of immediate feedback via the HMN Mobile app versus that of an exclusively delayed survey approach, which we believe would be problematic as the information would not have been shown to the user for review. The point-in-time feedback method request allows for immediate and topical feedback from the user.

5.2.3 Health

We have created a Health rating scale for Wifi, and apps, to classify operational status as part of the HMN Mobile app. This includes a composite of each of these areas to create an overall rating provide via the HMN Mobile app. Others have created similar methods for health and ratings, as an example Amazon uses a health rating system Amazon EC2 to determine operational status checks for optimal performance and operational status of their Elastic Compute Cloud (EC2) console. Dell, and others, uses a Health Rating in software to determine battery Dell Health Rating health using techniques around optimal charge, and digital analysis and reporting of a red to green gauge for user review. We have extended the methodologies from [48] and [37] in this work where we use a star system along with a speed gauge to express health ratings. We define our scale to cover both the Health of apps and Networks, as shown in the subsequent subsections.

5.2.3.1 Apps

Ratings are taken from the risk level of of each permission (e.g. Security and Privacy of permissions) across all apps installed locally, and as an aggregate of all remote apps risk level for a remote comparisons.

5.2.3.2 Networks

Ratings include attributes such as link and signal speeds for the start rating and Internet throughput speeds for the graphing of data. A collection of local data is used for the local ratings, and an aggregate of remote data is leveraged for the remote comparisons.

5.2.3.3 Digital Fingerprint

Extending on the health and operational status of devices, apps and networks we have captured changes to these entities as the digital fingerprint for each HN. A digital fingerprint change includes the adding, removing, or modifying a device (e.g. configuration), app (e.g. new install) or an permission to an app (e.g. remove location permission).

5.2.4 Norms of Data

As part of the HMN Mobile app we have created a classification of Norm types of data (local and global) that includes gathering point-in-time data, long-term availability of local data, global review of comparison data between local and other users experiences, legacy information using a longitudinal approach, and in if the data is shareable to a wider community for users or researchers. We define Local norms to be that of what is happening on the same network the app is running (e.g devices, Wifi speed, etc.). Using a similar approach to Local Norms we define Global norms to be what is happening on similar or disjoint HNs across the entire study (e.g. average of Wifi speeds, average

number of devices, etc.)

5.2.5 Integration of Data Collection Feedback

We have used an integration of approach areas for this work that includes the following: Data Collection (e.g. devices in the HN), Results (e.g. display Wifi Speed via speedometer), and point-in-time User Feedback (e.g. questions related to each of the areas and results). We found a value in this integration of approaches as users are able to review results and respond immediately with feedback.

5.3 Research Questions

In continuing work on the How's My Network project, we look to expand upon on our previous study and add value for HN users. The activities associated with this new HMN paradigm can help answer a number of research questions about HNs; these are directly correlated to HN users. The types of questions around HNs help fill-out the HMN work, and include the following research questions. The following list enumerates some of these questions that focus on this work.

1. What HN ecosystem characteristics can be discovered using a Mobile app?
2. What HN information and features are HN users interested in?
3. How do users evaluate the Mobile app approach for collecting and sharing information about Home Networks?
4. Will users change their security and privacy perceptions of their HN by using a HN tool?

5. Which ”local HN” to ”remote HN” comparisons do users find the most and least useful?

6. What impact does a Mobile app have on a user? Will a user modify their HN Ecosystem based off of the HMN Mobile app?

In the next several sections we provide details on following: HN Info/Features (Section 5.4), Mobile app development (Section 5.5), how information is presented (Section 5.6), data collected from app use (Section 5.8), and feedback results (Section 5.9). In addition, we have created the following accompanying parallel structure flow across each of these sections (note there may be additional subsections depending upon discussions required):

1. Wifi and HN Throughput speeds

2. HN Device listing

3. App Security / Privacy

4. HN Wifi Health

5. Apps Health

6. Local Norms

7. Global Norms

8. Research Data and Other Characteristics Associated with HNs

5.4 HN Info/Features

In this section we provide a set of questions we have created to categorize our research across this work. In addition to our set of research questions from Section 5.3, we are

also interested in understanding what HN information and features users desire, along with using best practices for presentation. We look to answer questions posed as part of this work across the following areas and start with background for each of these areas of study.

5.4.1 Wifi and HN Throughput speeds

An overview and listing of Wifi specific and internal HN connectivity. This is a collection of points associated with link speeds, and link levels of the Wifi network. This also includes a view of Internet throughput; including download and upload speeds. These throughput speeds are associated with links inside and outside of the HN.

5.4.2 HN Device listing

Data collected from the local network on active, and non-active, devices along with a descriptive view of these devices (e.g. hostname or predicted name), and an option for providing a nickname for each device listed. In addition, when new devices are found (previously not seen) we provide a pop-up notification to the user that a new device has been detected.

5.4.3 App Security / Privacy

We look to understand app Security and Privacy, and provide an app listing, along with a review of moderate and dangerous permissions. We also include an overall rating for each app, in terms or privacy and security, as well as a review of each of the permissions requested or required by the app. In addition, we look to compare these required permissions per app versus permissions granted by the user, via the google framework. This work also includes reporting changes to users, via pop-up notifications, when permissions

and apps (e.g. a new or deleted app installed).

5.4.4 HN Wifi Health

HN Wifi Health, as described in Section 5.2, includes data collected in and around the HN Wifi via the HMN Mobile app (e.g. signal strength, and speed). These data is aggregated and uses the corresponding health rating system described to create a ranking.

5.4.5 Apps Health

An apps health rating is included as part of an overall view of apps operational status. A review of each permission is performed across all apps at a granular level to calculate a security/privacy rating per app, and an aggregate across all apps installed. This includes the type of permissions allowed/approached as well as overall view of all apps installed.

5.4.6 Local Norms

In Section 5.2 we described Norms of data, here we extend it to include Local Norms. Local norms is data that is specific to the environment where the data is located. As part of this work we have collected data that is specific to configurations of devices, apps, and the local HN (e.g. Wifi and Internet speeds, etc.) and have rated and classified these as Local Norms.

5.4.7 Global Norms

Section 5.2 provides details of Norms of data. We define Global norms as those data points and items that are similar across this study, and are aggregated from local Norms. We have created a review of all HN data and created Global Norms review of information,

in an attempt to understand and compare HNs. This includes throughput, devices, apps, permissions, privacy / security, and Health of the HNs.

5.4.8 Research Data and Other Characteristics Associated with HNs

In addition to the areas, covered previously, we have also collected a series of data sets and information including: DNS throughput and health, device information, networking information (ISP, connection type, SSIDs connected, etc.), locale and other comparison information. We do not provide these collected data sets to users directly as we feel that they are to granular of focus and out of range for most HN users in terms of skill level. We opted to create overall Health ratings with these and data sets, where applicable, to inform a wider base of HN user skill levels.

5.4.9 Data Collection Using a Mobile App

As part of the How's My Network framework we have looked at a variety of methods of collecting data from within a HN (e.g. a router). In this work we leverage a Mobile app approach, with normal user privileges (i.e. non-administration / non-root), for HN collections. Comparing that to other collection methods including a Web Browser, Routers, PC, or customized hardware that typically will require elevated privileges for collection, as shown in Chapter 3. We have created a comparison of approaches to understand data we can and cannot collect across these approach type (Router, Apps, Web, and Custom/HW), reviewed in Chapter 3. Table 5.1 describes data that we cannot collect (data of interest) across each of these approaches.

Table 5.1 includes what we can and cannot be collect using a Mobile app, and we can see that an apps approach allows for access to a large amount of data, but still would require admin access or integration of approaches (e.g. a remote agent) for collection of

data on the network (e.g. remote configurations); an agent installed on the remote systems would allow for access and synchronization between approaches offered. Similarly a (stock) router has access to the entire flow of traffic, but would require access changes (or custom configurations) along with an integration of approach types. Web allows for access to some information, and would require care to allow for access across systems, and even locally in some cases. Finally, Custom/HW may allow for direct access to the local network, but would still require integration of approach types for collections.

In building the HMN Mobile app we have used a non-elevated permission framework to minimize impediments to the largest set of users to allow for ease of install and the running of the app. While using this framework has also allowed for a robust amount of data collection other more advanced methods (leveraging administration/root privileges) allow for an in-depth review into system and networking features. These features include low level network data collections (e.g. traffic sniffing, shaping, etc.), system access (e.g. root level Linux access, file access, app access), and other methods that require customized hardware and specialized tools to manage and operate (e.g. customized router managing flow and access). While these advanced areas exists for capturing of data, there is a trade-off of between access, usability, and results which are paramount to ubiquity of use for HN users.

In an environment where Wifi is not available (e.g. mobile only) the HMN Mobile app will not execute a scan as it requires both Wifi and must be connected to the preferred HN. Potential future work would include allowing a scan in a non Wifi environment for throughput testing, and app discovery on the local device.

In subsequent sections we will describe the testing framework and leverage the list of questions in this section to examine and provide details of the HMN Mobile app research completed.

Table 5.1: Data we can and Cannot Collect - Data of Interest and Approach type Comparisons

Data of Interest	Router	Apps	Web	Custom/HW
Throughput	Requires modified Router	Yes	Yes	Yes
Network flow	Requires modified Router	admin access or integration with HW approach	Requires Approach integration	Admin access + custom Router
Network Control	Requires modified Router	admin access or integration with HW approach	Requires Approach integration	Admin access + custom Router
Wifi	Yes	Yes	Requires Approach integration	Yes
Devices on Network	Yes	Yes	Yes	Yes
Apps & Permissions on device	Requires Approach integration	Yes	Requires Approach integration	Requires Approach integration
Previous Wifi Attached	Requires Approach integration	Yes	Requires Approach integration	Requires Approach integration
Local Configurations	Requires Approach integration	Yes	Yes, minimal	Yes
Remote Configurations	Requires Approach integration	Requires integration of approaches	Requires Approach integration	Requires Approach integration
Historical Norms	Requires Approach integration	Yes	Requires Approach integration	Requires Approach integration
Health	Requires Approach integration	Yes	Requires Approach integration	Requires Approach integration

5.5 Mobile App Development

In this section we provide background on how data was collected using the HMN Mobile app. All data collected using the HMN Mobile app and is transferred securely to a server, located on the WPI network. The Server houses scripts for analysis, and runs two daemon processes for the collection and distribution of data. The two daemon processes run on a single CPU (quad core), with 8GB RAM, and operates as the 2nd and 3rd tier components for data collection for the How's My Network project. The daemons collect and aggregate data, perform throughput test, as well as all file operations.

As mentioned a series of scripts manage and coalesce the data for summary. This includes calculating of all device, throughput, networking, and app averages and summaries. The client server connection between the app and the WPI server is via custom written encrypted service using AES and RSA Public keys exchange. All communication to and from the client (HMN Mobile app) is completed using either the daemon's generated public key, or an on the fly created RSA public/private key pairing from the client; thus offering a high level of encryption when communicating to and or from the secured server via the app. We created privacy, and policy pages as part of this study (as required by Google) along with a detailed web site with How-To information (located at `http://hmn.cs.wpi.edu/` - How's My Network Web Page) on the WPI network.

We also programmaticly dropped to shell on the HN device to take advantage of the underlying OS (a modified version of the Linux OS), and leveraged system versus programmatic methods to gain access to information (e.g. networking, app information, device information) and speed of results. We would like to note that information leveraged using this shell method versus that of a secured a explicit framework may expose or leak privacy and security results about the user, device, or the HN Ecosystem. While most of this shell-based information is read only there may be additional privacy and se-

curity concerns allowing access to sensitive pieces of information without a framework for access to this data.

During a HN scan (the areas covered in the remainder of this section) processing is executed in either parallel or serial depending upon data collection type. All logging is completed asynchronously to minimize user impact, and are offload during non-doze times of the mobile device; log details include whether the scan was user initiated or an automatic scan, or other type of logged input. Additional details of how to run a scan via the app are covered in the Appendix A.1.

The HMN app went through the following process for validation of attributes in the HN Ecosystem. A local HN environment test-bed was setup that had specific attributes on/available during the "scans"; these included hardware (physical and virtual), apps, configurations, and settings. We validated that the app gathered these attributes or devices manually and via inspection tools (e.g. NMAP) as well as router cache. We also validated new devices/removed devices and other attributes (apps, etc.) as part of testing and validation process. A similar process was followed for app validation/testing and verification of permissions, and configuration discovery (e.g. installed apps). We validated the apps installed on the device by reviewing permissions on the local device and cross referenced on the Google Apps store. Configuration verification and validation was completed via a review of the devices and confirmation of known data parameters in the test bed of hardware and apps (e.g. ADB).

We next follow the parallel flow structure of the features of the HMN Mobile as shown from Section 5.3, and provide details over each of the sub-sections.

5.5.1 Wifi and HN Throughput Speeds

We collect Wifi data via the Mobile device using the HMN Mobile app, which includes signal and link speeds. Signal levels and link speeds are probed periodically within the

mobile app ($\approx$ 10mins), calculated, and securely shipped to the server for processing and review. These levels are then calculated using the following Wifi base values of dBm (decibel-milliwatts) levels (-70dBm - -50dBm), and Mbps (mega-bits-per-second) of 802.11 (b..ac or 11-866.5mbps) respectively. The calculations include:

$$AvgWifiLinkSpeed = \frac{\sum_{i=0}^{n} LinkSpeed_n}{n}$$

and

$$AvgsignalLevel = \frac{\sum_{i=0}^{n} LinkLevel_n}{n}$$

where n is the number of probes.

Internet upload and download throughput is measured via connectivity to the secure server at WPI, and the running of a series of tests to determine overall throughput in both directions. The server is mutlithreaded and was developed to handled multiple connections per second. The HMN Mobile app initiates the speedtest, which includes a payload of data sent to and from the mobile app and server. Measurements are completed during and after the tests and are saved to the log server for future analysis. TCP packets are sent to the server, and mobile device, in an attempt to track real world throughput versus that of bandwidth. The server reads 1500B packets (1MB), and calculates time of delivery, using the following functional method, where n is time to complete in ms, and frame sizes are 1500bytes.

$$Throughput(x) = \frac{\sum_{i=1}^{x} Frame_i * time}{n}$$

5.5.2 HN Device Listing

Network and device listings are determined by using several fast paced methods of discovery, where we examine network state on the local subnet, and review hosts alive or query-able. A complete network scan can be done in under three seconds, in most cases.

Device information includes: device name and host type (e.g. IoT, Printer, Router, etc.), along with a visual image of the device type associated (e.g. phone, PC, Router, etc.). The process includes a deterministic approach of host to name and device type mappings (e.g. Gaming Consoles, Streaming devices, Photo Frames, etc.). Device name and host types are determined by a series of networking probes, along with information available from DNS and MAC address information on the network. We then match these data sets against a heavily modified IEEE Organizationally Unique Identifier (OUI) database [102, 72], which includes manufacturer and typical host type. A complete list of categories and representative devices used in this study can be seen in Table 5.2.

We next move our attention toward the determination of devices status on the network, and during scans. We have created the following four categories to determine device detection on a HN: Always Present (e.g. found every-time a scan was run), Newly Visible (e.g. new device was found), No Longer Visible (e.g. removed from the network), and Transient Visibility (e.g. found, but not consistently). We use the following methodology for the determination of status across the four categories. Based upon testing and experimentation we use four consecutive scans, at the start and end, as our matching patterns, across these four types. The following are the types used and additional description: Always present - if a device is found during each network scan (e.g. Router). Newly Visible - a device is detected after not being present for the first four, or more scans, and is then always found. No Longer Visible - a device is determined to be no longer visible when it is no longer found present at least four or more times at the tail of the scans. Transient Visibility - a device which is found most of the time, but may not be found consistently (e.g. gaps between scans) or in a series of scans.

Table 5.2: Categories of Devices and Representative Types

Category	**Representative Devices**	**Additional Description**
Phones	Android, iPhone, Blackberry	
Tablets	Android and iPad	
Router/Wifi	Netgear, Belkin, including extenders, etc.	
PCs	Windows, Mac, Linux	
Game Console	XBOX, PS3, Nintendo	
Smart Home	Alexa, Dot, iRobot, IPTV, Video on Demand, etc.	A device that is connected to the Internet, and can remotely manage and monitor appliances, as well as provide virtual assistance
Streaming Devices	Roku, Chromecast, etc.	Devices to stream live and asynchronously broadcasts and media
IoT	Lights, Garage Door, Fitbit, Espresso Machine, Comfort Bed, etc.	Internet of Things (using specific IEEE 802.15.x protocols) for remote management of entities within a home
Smart TV	Vizio, Sony, Sound Bar, etc.	TV with streaming services and expanded audio
Printers	Epson, HP, etc.	
Video Security Cameras	Lorex, IP Cameras, etc.	
Other	Peloton, VOIP, Disk Storage Devices, etc.	Workout gear, Voice over IP, etc.
Unknown	Devices that could not be determined	

5.5.3 App Security / Privacy

A review of privacy and security is performed, in parallel, across all 3rd party apps installed on the device (not created or directly approved by Google). Information collected includes the full name of the app, designated name (google app name), and permissions granted by the user for the set of apps. A classification of app types and levels (ranging from normal, moderate, and dangerous) was compiled across known requested and granted permission types (e.g. Internet, Read my data, Write contacts, Read card data, and many others). Table 5.3 shows the top 10 (out of roughly 50) most popular app categories installed, as defined by Google Play Store. Other app categories includes areas such as: Business, Social, Maps, and many others.

Table 5.3: Top 10 Popular Google App Categories

App Category	**% Downloads HMN Testers**	**Examples of Popular Apps**
Entertainment	27	Netflix, Facebook, Twitch
Tools	11	Find My Device, Speedtest
Description NA	10	Adblock, Calculator
Productivity	6	Outlook, Dropbox
Finance	6	Turbo Tax, Credit Karma
Lifestyle	5	Pinterest, Calendar
Shopping	5	Amazon Shopping, Ebay
Health and Fitness	5	Fitbit, Runkeeper
Communication	4	Facebook Messenger, WhatsApp
Games and Puzzles	3	Candy Crush, Scrabble

We created a category of permissions and libraries to manage these transactions. These definitions were created in conjunction with Google's Android security model of familial permission types [9], described by [170], and shown in Table 5.4. The categories of permissions include dangerous areas such as location and camera, as well as normal permission access for networking requests (e.g. Wifi), and other general usage polices

related to Android devices. We have assigned levels of access according to the permission type and access request type for each permission level [9]. Permission categories for these areas include the following descriptions of: Normal, Moderate, and Dangerous permission requests [170]. Normal permissions are those that do not pose a risk to the user's privacy or the device's operation, according to Google. The device grants these permissions automatically to the requesting app at install and run-time. Normal permission types include: connecting to the internet, Networking, some Bluetooth operations, Wifi (and NFC information), alarms (setting) and wallpapers, and audio access (settings on a device).

Dangerous permissions are those that are deemed by Google [59] to potentially affect the user's privacy/security or the device's health and operation, and are defined by Google as "Dangerous permissions cover areas where Apps want data or resources that involve the user's private information, or could potentially affect the user's stored data or the operation of other apps. For example, the ability to read the user's contacts is a dangerous permission. If any app declares that it needs a dangerous permission, the user has to explicitly grant the permission to the app. Until the user approves the permission, your app cannot provide functionality that depends on that permission" [7]. There are 13 permissions (with photos and storage being in the same pairing as it is storage access) assigned to the dangerous grouping. At app install permissions are available to the user to peruse and understand which areas/permissions are being requested. Apps requesting access to dangerous categories of permission must explicitly receive approval from the user to be granted these permissions at run-time, and are dependent upon the SDK and OS version on the phone (e.g. version >7 require a pop-up request for some dangerous permissions). These categories include access to: camera, contacts, location, microphone, sensors, SMS, and storage. We created a new category 'moderate' for those permissions typically marked by Google as dangerous, but are just system resources to write out to

disk. An app may have many permission requests for access to a variety of device features, and some require user acceptance to provide such access (depending upon OS version). As an example access to location may require user approval (depending upon Android release version), via a manual pop-up window, which is added to the properties manifest after being approved.

The How's My Network Mobile app requests the following permissions as part of its manifest, as shown in Table 5.4, and includes Internet (including Wifi and network state), boot completion (used to start the app on reboot), vibrate (apply a notification), read phone state (used to gather information about the phone status), read and write external storage (used to read preferences and database information), location (only used to retrieve SSID on android versions $>$8.0), wake lock (other category and used to process data in the background), and foreground services access (used to run processes similar to background work). These permissions are requested at run-time and fall into the following categories: Wifi and network access (including Internet), remedial phone access, phone state, storage access, and location. Users can reject these permission requests via the google apps console (via their mobile device) at any time.

As a note, location awareness can be problematic for users as repositories exist consisting of "pre-calculated" AP (access point) locations, including latitudes and longitudes. An app can pinpoint a user's location ($<$100 Meters accuracy) simply by leveraging nearby AP information and without the need of GPS location access. Apps have the ability to request location or Wifi information with minimal security requirements or approval from users, depending upon SDK version targeted and release of the the OS version. This is a known security problem in the Android framework SDK where an app can target version 5.1 or lower SDK (which the HMN Mobile app does) and is allowed extended access as described by the google's bug tracking facility [58].

Table 5.4: Category of Google Permissions

Android Defined Permission Levels				
Category	**Description**	**Normal**	**Moderate**	**Dangerous**
Contacts	Find contacts on device, Read and Write Accounts			✓
Camera	Take pictures, Record video			✓
SMS / MMS	send, edit, receive messages			✓
Storage	Read and modify USB and other storage			✓
Location	Get approximate, precise, and access extra location information related to the GPS			✓
Photos	Read and modify Media/Files on USB and device storage, including mounting/unmounting and formatting storage.			✓
Microphone	Use microphone to record			✓
Phone	Make Calls, Write/Read Call Log, Route Calls, Modify Phone without intervention			✓
Wi-Fi / Network	Connect, Read, and Change Wifi / network	✓		
Device ID Identity	Read and Identity device information Find, read, remove, and modify			✓
Calendar	Read, Write, add/modify events, Send email w/o knowledge			✓
In-app purchases	Make purchases inside app.		✓	
Device history	Access and Retrieve sensitive log data. Read bookmarks and history	✓	✓	✓
Bluetooth	Broadcast / read data			✓
Wearable Sensors/Activity Data	Read/Write (e.g. heart rate monitors)			✓
Other	Gmail Access, Download Files, Receive Data, Network Access, Read Battery Stats, phone control (e.g. vibrate), and several others	✓		✓

An overall app rating is calculated using the following function,

$$f(x) = \frac{\sum_{i=0}^{n} rating_i}{n}$$

where n is the total number of permissions, and i is the sequence rating for the given. All of the HMN apps are re-scaled between $1..5$ to match the five star Google star rating system [104], via the following refactor-function, which we leverage across several areas of this work:

$$f(refactor) = \frac{Value - Min}{Max - Min} * (Max_n - Min_n) + Min_n$$

.

5.5.4 HN Wifi Health

We aggregate and average Wifi data collection levels, and then re-factor them to be between a rating of one and five, using the previously discussed refactor-function. The data collected is used to calculate an overall Wifi Health, with a rating created between one and five is used.

5.5.5 Apps Health

Apps Health is an aggregate and average apps data collections levels, and then re-factor them to be between a rating of one and five, using the refactor-function. The data collected was also used to calculate apps Health, here we refactor the values (using the refactor function previously shown).

5.5.6 Local Norms

Local Norms of data is collected by leveraging and averaging the data sets shown in this section. These data collection points and stored remotely for processing on the secure server, and the aggregated data is securely shipped to the app when requested (e.g. re-loading). This includes Internet throughput, device Count, apps Health, and Wifi Health. The remote data throughput is an overall average for both upload and download speeds respectively. Device count is an aggregate of all devices averaged over each execution of the app. Apps and Wifi health are calculated using the average app ratings, signal level and link speeds (calculated within the HMN Mobile app).

5.5.7 Global Norms

Global Norms include data across all participants of the HMN Mobile app. These data collection points are stored remotely for processing on the secure server, and the aggregated data is securely shipped to the app when requested (e.g. re-loading). This also includes the average of all HN collection points: Internet throughput, device count, apps Health, and Wifi Health. The remote data throughput is an overall average for both upload and download speeds respectively. Device count is an aggregate of all devices across all HNs and averaged (hourly) via scripts on the server. Apps and Wifi health are calculated using the average app ratings, signal level and link speeds (calculated within the HMN Mobile app).

5.5.8 Research Data and Other Characteristics Associated with HNs

As part of the current state of the app default data is collected, and stored, around configurations. Data is read from the local device and network, and we include a unique identifier for each HN scan, and includes: device information (type, name, SDK, OS Ver-

sion and Code Rev, Security Patch version, and Manufacturer type), and Wifi information (current Wifi, networks attached to, strength, and connectivity type). In addition to these areas networking information is gathered (IP, DNS, netmask, MAC, etc.), as well as DNS information using our custom created DNS tool (jDIG for Android); we preform DNS testing (e.g. google.com) for both performance and overall health of the network as well as external IP verification. DNS performance is calculated using RTT and average connectivity to first level servers, primary DNS, and external DNS servers. As part of the overall execution of the app a background process periodically logs HN Wifi speed along with an app and a HN health rating. These results are collected and calculated (including RTT/throughput) using highly customized databases, and scripting tools in the Mobile app and offloaded to the server.

5.6 How Information is Presented

In this section we provide background on how the information collected is displayed in the HMN Mobile app. The HMN Mobile app is split into three distinct elements, via well known UI/UX-based tabular (Tab) layouts, Figure 5.1. These Tabs allow the user to navigate between device listing, apps Security/Privacy and Summary comparisons. Each HN provide unique information about their networking, device, and privacy and security. These tabs provide information from the posed questions, from section A.2.

The HMN app provides the user with a conveniently located Wifi health speedometer that includes a continuous scan of the local networks Wifi speed, and throughput, and can be seen in Figure 5.1. The app also provides a 'pink question mark', Figure 5.2, that allows users to provide immediate feedback from a series of questions available 5.2; note that questions will still pop-up via notifications if the user does not click the question mark, once daily. In addition the app provides "popup" notifications to the HN user when

a new device is found, or an app is added, modified, or deleted.

Figure 5.1: Default/Network view of app

Figure 5.2: Add a Comment

5.6.1 Wifi and HN Throughput Speeds

We display results of the Wifi throughput speed indicator via the customized progress view speedometer, as shown in Figure 5.1, leveraging the Chart library [108]. The color coding for the Wifi throughput health includes Red (bad connection), Yellow (good connection), and Green (excellent connection). These and other color coding(s) were selected as they follow generic international stop, slow, and go colors, and fall into research done by [160] for color coding palette.

The local and remote HN throughput (Internet speeds) is displayed in a graph with colored and labeled bars along with a numerical representation of the average throughput

achieved, for both upload and download speeds. These data is collected and stored remotely for processing on the secure server, and the aggregated data is securely shipped to the app when requested (e.g. re-loading). The remote data throughput data is an overall average for both upload and download speeds. The calculations for the rating system uses the re-scaled metrics described previously in Section 5.5, and are an aggregate of Internet overall throughput.

5.6.2 HN Device Listing

A HN device listing is displayed in the 1st tab of the HMN app, and provides a listing of device information, sorted by IP (typically fourth octet) range, as shown in Figure 5.3. The HMN app uses a training method to notify when a new device is found. When a new device is located a notification will be displayed with the device information for review.

The HMN app locally stores attributes and information about devices including: hostname, IP, MAC, last time available, and the wifi network attached. The app allows users to view this information, and add a ”nickname” for a given host type; the nickname feature was added early on as part of user requests for management. Using best practices as shown by [107] the device is highlighted in the listing for ease of recognition. As part of the generic UI/UX design users can run a network list refresh scan by using the generically accepted swipe down feature within the tab’d window.

5.6.3 App Security / Privacy

Apps Security and Privacy is displayed in the second tab of the HMN app, as shown in Figure 5.4. The HMN app uses built in logic and training to understand if new apps are found or are changed (e.g. a security permission is modified). When a change is found a notification will show the changed entry for review.

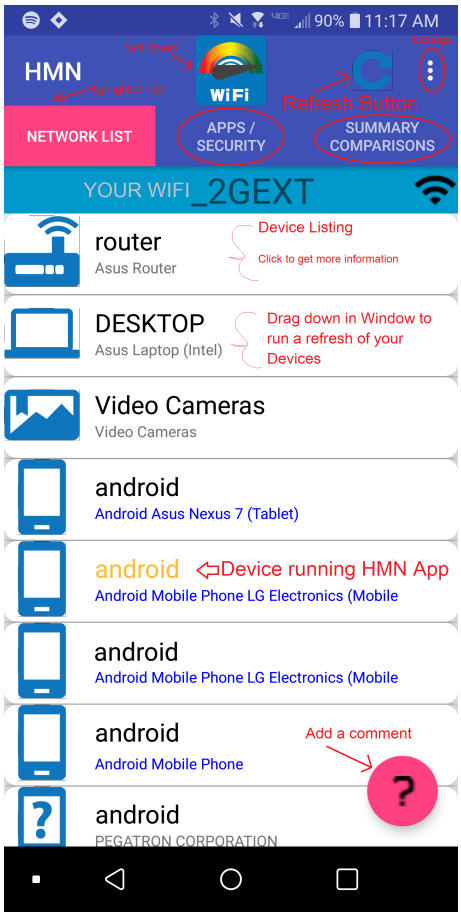

Figure 5.3: Devices on HN

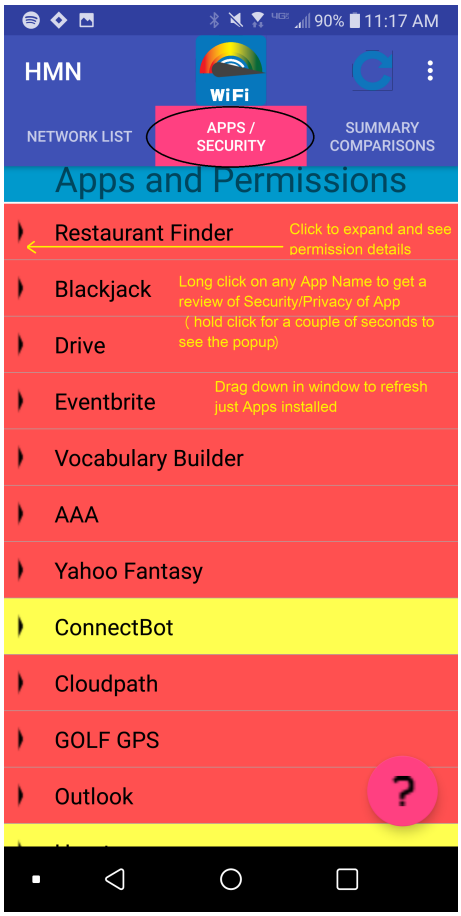

Figure 5.4: Apps

The in app color coding was completed using a compliment to the Google app rating system [59] Dangerous, and Normal. We assigned values to each of these levels: 1:dangerous, 3:normal and added a moderate:2 rating for those permissions that are simple system resources (e.g. writing to disk).

Users can select/click on package/app names to expand the listing of permissions for a given app, Figure 5.5. Users can long-click (1 second or longer) on the app name or click on a given permission to find out details of the requested permission, Figure 5.6; we also provide the level of permission and details on each severity. Users have the option of ”long-clicking” on the app name to load Google-based permissions via the Google apps permission tool. The user is notified when an app change is detected via pop-up notifications. These changes include new app installs, un-installs, or permission changes for existing apps. These notifications include a literal digital fingerprint update in the app, as well as stored in the log services. The user is shown which app or permissions were modified, deleted, or added via a ”+” (added) or a ”-” (deleted) sign in-front of the app name.

5.6.4 HN Wifi Health

HN Wifi Health is displayed in Tab-3 in the HMN app, and consists of the results as shown in Section 5.5. We display an overall Wifi Health rating using a star rating system; ranging between one and five stars, where one is poor and five is excellent. We refer to this rating as Wifi Health, and assign a color coding of green, yellow, and red accordingly to the rating shown described in Section 5.5 and shown in Tab-3, Figure 5.7.

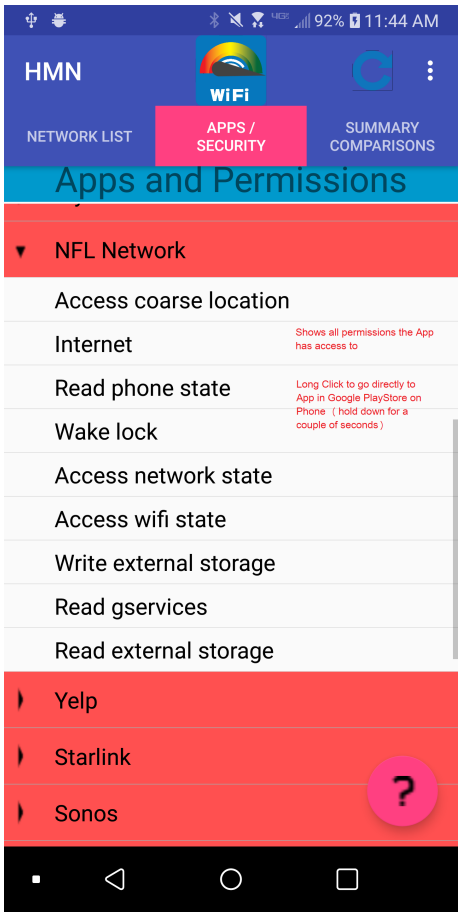

Figure 5.5: Apps

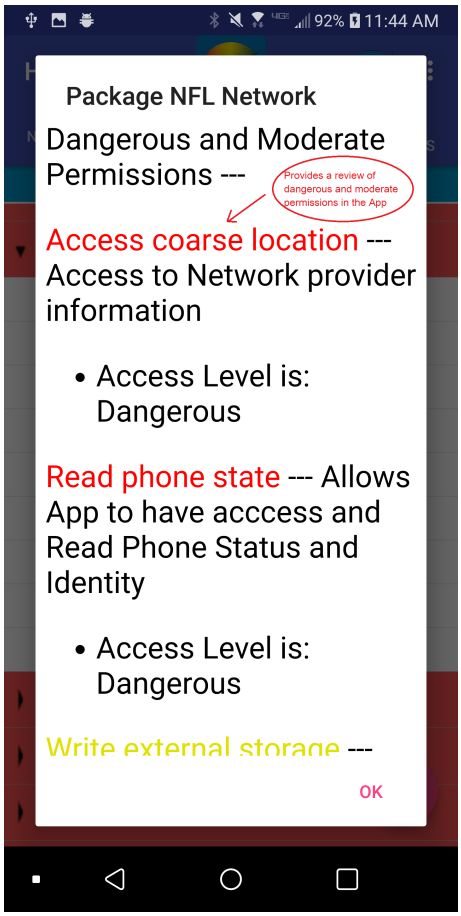

Figure 5.6: apps Expanded

5.6.5 Apps Health

An apps Health rating is displayed in Tab-3 in the HMN app, using the calculations from Section 5.5, where we display an overall apps health rating using a star rating system; ranging between one and five stars, where one is poor and five is excellent. We refer to this rating as apps health, and assign a color coding of green, yellow, and red accordingly to the rating shown described in Section 5.5 in Tab-2 (and Tab-3) as shown in Figures 5.4, 5.5, 5.6. These ratings are then used for an overall health rating of apps 5.7.

5.6.6 Local Norms

The calculations for the rating system uses the re-scaled metrics described previously, and are an aggregate of apps or Wifi overall operation. A yellow color star is added to the each of the ratings (Wifi and app health) according to the final average value calculated for the Wifi, and app health. A histogram is shown for the comparison between local Wifi throughput (download and upload), and device information is displayed numerically, Figure 5.7.

5.6.7 Global Norms

The calculations for the rating system uses the re-scaled metrics described previously, and are an aggregate of apps or Wifi overall operation, referred to as health. A yellow color star is added to the each of the ratings (Wifi and app health) according to the final average value calculated for the Wifi, and app health. A histogram is shown for the comparison between remote Wifi throughput (download and upload), and device information is displayed numerically.

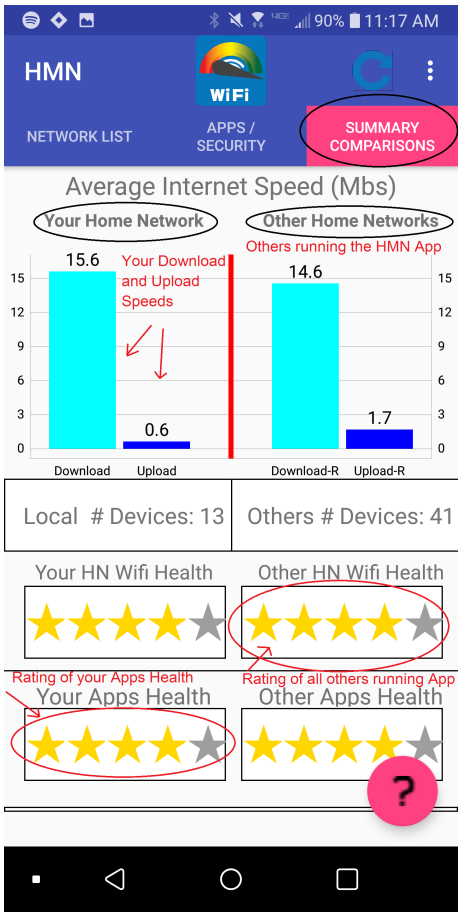

Figure 5.7: Tab 3 Norms

5.6.8 Research Data and Other Characteristics Associated with HNs

Research data and other characteristics within the HN, or the default configurations, shown in the top panel of the app (Fig 5.1). We have collected Wifi, Networking, DNS, and device versioning information (phone, OS, etc.) to help provide information to the HN user for both comparison and optimal setup. We display these results and summarize these data in the Section 5.8.

5.7 Study

The HMN Mobile app is available in the Google apps store How's My Network [70], and is linked via the How's My Network web page (How's My Network Web Site). The secure server is a single core, quad core system with 8GB of RAM, running Linux and located on WPI's secure campus network and data center. As previously mentioned the server maintains and collects all logs via secure communication to the mobile app, and process files on a periodic manner (hourly).

The HMN Mobile app was deployed over three testing phases, and included over 100 testers, where we leveraged usable data from 100 plus unique users across the entire study. Each of the testing phases helped with development and discovery of valuable data. Users installed the HMN Mobile app for 7, 21, or 30 days, and we found that all users uninstalled the app after testing. Table 5.5 is a summary of the phases along with time frame, participation, length and number of participants. Phase one ran from Jan-Feb and was the initial-testing phase. This phase included 10 testers (determined by IP and device type) and saw ˜75 comments/responses over the 30 days of testing. Phase one received beneficial feedback on operational changes, and helped narrow down minor bugs, and final user requests and feedback (e.g. addition of a nickname to the device list).

Phase two consisted of 44 participants (mostly from New England), over 21 days

(Feb-March) and included ˜1200 comments/responses; phase two included testers who previously participated in the How's My Network Home Survey [138], the WPI community, and social media release sites (e.g. Reddit). Phase two ran from February into March, 2019, and was a long term study where minimal changes were made to the app. This phase helped with data gathering over a larger group of users and differing HNs and ISPs.

Phase three consisted of an ˜14+ day study using Amazon's Mturk Mechanical Turk and ran during March 2019. We started with a test phase to generate interest, with 10 participants, which ran over several days. We then opened the testing up to a set of 40 participants, running over seven days. This entire phase had 50 unique users (mostly from North America), where we received ˜1000 usable comments. Users were paid to participate in the study (via Mturk) and required only to be a quality Mturk user (rated by Amazon's Mturk engine) and have ownership of an Android phone and an HN to participate, we did not apply or request any skill level for participation. The users participated in a pre and post survey and also provided feedback daily via the HMN Mobile app.

At the conclusion of phase three all participants were asked follow-up or final questions, which matched the initial questions posed; we have used this set of questions for comparisons across users. The third phase received valuable feedback from the users across the board and we saw ˜95% of users, whom initially joined the Mturk group, stay with the study for the entire duration. Participants answered at a minimum two questions daily, along with the running of the HMN Mobile app. As the Mturk engine is mostly used for online survey we used a programmatic approach to working with the Mturk community, where we created a tool-set for communication and distribution via the HMN Mobile app to the Mturk users, along with manual methods and scripts for connections to the results of the engine.

Table 5.5: Summary of HMN Mobile App (All Phases)

Phase #	Time Frame	Participation	length	# Participants
Phase 1	Jan-Feb 2019	Social Community	30 Days	10
Phase 2	Feb-March 2019	Social and WPI Communities	21 Days	44
Phase 3	March 2019	Mturk	7 Days	50

5.8 Data Collected from App Use

In this section we will review the results obtained by the HMN Mobile app study, across all phases (unless otherwise noted) as described in Section 5.5, and 5.7. We collected data from over 57K user and automatic refresh/runs across all participants (with an average of 25 manual refreshes per user), and amassed ˜600MBs of raw data. We will move across each of the areas of focus and provide the following review of each item: how data was collected and presented, information found via the HMN app, and HN user reactions to these results. While the primary goal of the study is to understand user interactions the by-product and artifacts of the HMN Mobile app included the following collection points and data-sets. Additional app collections details can be found in Appendix A.4.

5.8.1 Wifi and HN Throughput Speeds

These items have been coalesced into the Local and Global Norms sub-section.

5.8.2 HN Device Listing

5.8.2.1 Devices

A network list and view of what resides in the local HN uses the methods and features from this and Section 5.5, as part of discovery of the local network, and as described in Table 5.2. Of the 100+ HNs participating in the study we found that the average HN had

14 devices, with a variety of device properties / types. We found over 243 unique devices, across the 13 categories shown in Table 5.2, and 4087 total devices across all HNs; we paired down devices which were found from edu networks. The categories include Networking, Phones, PCs (laptop/desktop), Gaming Consoles, Tablets, TVs, Streaming devices (e.g. Roku), Printers, Smart Home Devices (e.g. Alexa), IoT (e.g. Nest), Video/Security Cameras (e.g. Lorex), Storage Units (inXtron hard drive unit.), Workout Equipment (e.g. Peloton), Active Scanning Network hardware, and cleaning devices (e.g iRobot). We have extended the classification of device listings done by [69, 67], and shown in Table 5.6 (sorted by percentage of HN), with Phones, Routers and PCs making up the largest percentage of all devices found. The other categories of devices include disk storage, workout gear, and other devices not identified (combined into a single category).

Table 5.6: Categories of Devices and Percentages

(sorted by % of HNs)

Device Category	**Percentage of HNs**	**Percentage All Devices**
Router/Wifi	100	12
Phones (iPhone/Android)	98 (34,66)	53
PCs (Windows/Mac/Linux)	87 (95,4,1)	11
Printers	71	4
Smart Home	50	1
Smart TV	36	2
Streaming	34	4
IoT	20	1
Video / Security Cameras	20	1
Game Console	18	2
Tablets	8	4
Other / Unknown	7	<1

Android devices make up more than 70% of the worldwide phone market and was the most popular mobile phone found (66%) versus the iPhone (34%), less than 1% was

made up of other phone types (e.g. BlackBerry). Android made up 90% of the tablets, with iPads (Apple) making up the remaining 10%. More than 1% of all devices found were IoT units, and included Light, plugs, and controlling units (e.g. Nest). 18% of all networks were found to have a gaming console (Nintendo, PS, Xbox), and 50% were found to have a smart device (e.g. Alexa, Dot, etc.) Almost 90% of all homes had an additional Wifi unit (extender) or second, with Netgear (40%) and Asus (20%) being the most popular of the 10 types of Networking gear found across all HNs. 71% of all HNs were found to have a network printer, and almost 34% had a streaming device (e.g. Roku). 36% of all homes have a smart TV, 20% have either security/network cameras, and almost 10% of homes had a disk backup system. We found six workout devices (e.g. Peloton), five Active Networking Scanning devices (hardware) and two cleaning devices (e.g. iRobot) across all HNs.

We also reviewed HN median income level and device types across device categories to understand differences of devices dependent upon income levels. A study done by HUD [39] found that lower income households have lower rates of Internet and devices when compared with other groups of income (e.g. higher). We have reviewed HN devices by zip code and their corresponding mean income levels. Data was compared against mean income levels as provided by the US IRS and the Population Studies Center at the University of Michigan [189] across US zip codes. We have compared these results across three areas of mean HNs: high (top 20% [72K and up]), middle (60% [45K-72K]), and low (bottom 20% [less than 45K]). Table 5.7 shows categories of devices versus these mean income levels. We see minor changes across most categories and levels of income.

5.8.2.2 Transience of Devices

We reviewed device state as they existed on the network, and classified them, as described in Section 5.5: always found, no longer visible, newly visible, or transient visibil-

Table 5.7: Comparison of High, Middle, and Low Mean Income and Device Types in HN (Mean Income levels by % of HNs)

Device Type	**High (top 20%) Mean Income**	**Middle (60%) Mean Income**	**Low (Bottom 20%) Mean Income**
Phones	40	43	31
Wifi/Router	9	10	13
PCs	19	15	18
Tablets	7	7	5
Streaming Devices	6	5	7
Printers	5	5	8
IOT	6	3	2
Other	3	4	3
Smart Home	0	0	1
Game Console	1	3	2
Smart TV	3	4	5
Video Security Cameras	1	3	0
Unknown	1	1	7

ity. Table 5.8 represents all phases (see Table 5.5), and is a review by classification/device Category across the four transient states. These include: category states along with four selected groupings of devices types (Networking, Phones, IoT, and Streaming) and their respective percentages each. From Table 5.8 we can see that most devices fall into a transient visibility category, with newly visible and no longer visible devices being the least found. We found that 16% of users had a no longer visible devices and 61% of all no longer visible devices were Phones, and may in-part be due to concerns around device security as noted.

We have also analyzed phase 2 and phase 3 groupings of transient of devices across our data collection phases to understand grouping of users by similar install times. We have reviewed phase 2 and phase 3 as shown in Tables 5.9 and Table 5.10 which represent states of devices across both phases. We can see that in phase 2 devices have a higher percentage of transient visibility versus that of phase 3 where we see devices having a higher percentage of being always found.

Table 5.8: Categories of Devices Per HN - All phases

Classification / Device Category	**Always Found %**	**No Longer Visible %**	**Newly Visible %**	**Transient Visibility %**
Phones	11	1	0	88
Tablets	40	2	0	58
Router/Wifi	35	1	1	64
PCs	23	1	1	75
Game Console	20	0	0	80
Smart Home	65	12	0	24
Streaming Devices	29	4	0	67
IOT	46	0	0	54
Smart TV	17	0	3	80
Printers	24	0	0	76
Video Security Cameras	31	0	0	69
Other	27	3	3	66
Unknown	50	0	25	25
All Devices	19	1	0	79

Table 5.9: Categories of Devices Per HN - Phase 2 - 21 Days

Classification / Device Category	**Always Found %**	**No Longer Visible %**	**Newly Visible %**	**Transient Visibility %**
Phones	3	3	1	93
Tablets	5	5	0	90
Router/Wifi	5	3	0	92
PCs	3	7	1	90
Game Console	8	0	0	92
Smart Home	0	0	0	100
Streaming Devices	6	2	0	92
IOT	20	0	0	80
Smart TV	11	0	5	84
Printers	8	0	0	92
Video Security Cameras	8	0	0	92
Other	31	5	0	64
Unknown	0	0	29	71
All Devices	5	3	1	91

Table 5.10: Categories of Devices Per HN - Phase 3 - Mturk - 7 days

Classification / Device Category	Always Found %	No Longer Visible %	Newly Visible %	Transient Visibility %
Phones	14	1	1	84
Tablets	46	2	0	52
Router/Wifi	41	0	1	58
PCs	23	0	1	76
Game Console	41	9	0	50
Smart Home	79	0	0	21
Streaming Devices	34	5	0	61
IOT	55	0	0	45
Smart TV	14	0	5	81
Printers	18	0	0	82
Video Security Cameras	31	0	0	69
Other	31	5	0	64
Unknown	0	0	0	100
All Devices	24	1	0	75

5.8.3 App Security / Privacy

5.8.3.1 Installed Apps

In this section we review what apps are installed, and describes details and categories associated. Examining apps, along with the privacy and security of the device running the HMN Mobile app has provided valuable information in terms of categories of usage, and overall levels of device/app health. Of the 3096 unique apps installed, the average user had 60 apps installed on their devices across all phases. We next turn our attention to the state of these app, and the permissions requested by each of them, and provide a timeline for each of the collection points.

Taking into account apps permissions we have broken down the discovery into the following categories, which Google and others have created, and categorizing of apps from work done by [119]. As a background on publishing an app on the Google App

store, the app publisher first selects the category that best fits the app. The following are the list of approved category areas (we have combined these as a generic grouping, e.g. games, where logical): Entertainment, Tools, Productivity, Finance, Lifestyle, Shopping, Health and Fitness, Communication, Game Puzzle, Travel and Local, Photography, Game Casual, Business, Education, Music and Audio, Sports, News and Magazines, Food and Drink, Social, Personalization, Game Simulation, Game Arcade, Video Players, Game Adventure, Game Card, Maps and Navigation, Game Role Playing, Medical, Books and Reference, Game Strategy, Game Board, Game Action, Game Trivia, Game Word, Game Casino, Casino, Weather, Game Sports, Game Educational, House and Home, Dating, Comics, Auto and Vehicles, Parenting, Game Racing, Libraries and Demo, Game Music, Events, and Beauty.

The top 10 apps installed across all users is shown in Table 5.11, and includes category, number of user installs, and the current Google Ranking associated with the app. These apps include categories in social, and production tools such as Facebook, Netflix, and Outlook Mail client; note we removed the How's My Network app from this list as all users had our app installed.

Table 5.11: Top 10 Popular Installed Apps

Apps	**Category**	**% Users**	**Google Ranking**
Google Play Games	Entertainment	53	16
Facebook Messenger	Communication	45	2
Netflix	Entertainment	36	6
Instagram	Social	32	4
Google Pay	Finance	25	NA
Twitter	News & Magazines	24	45
Snapchat	Social	24	10
Facebook	Social	24	1
Spotify	Music & Audio	23	11
Outlook	Productivity	23	NA

5.8.3.2 Detected Changes to Apps

As noted in Section 5.7 we examined the permission level of all 3rd-party apps, and created (and storing of) a digital fingerprint for ease of notification when changes are detected to apps in terms of (deletion, additions, or changes). As part of users leveraging the HMN Mobile app we found that an average of 6 (median of 1.5 or 3.5 for non-zero uninstalls) apps, per user, were uninstalled, and 32% of all users uninstalled one or more apps across all phases; 27% phase 2 and 36% phase 3. The top ten categories of baseline app category percentage versus those uninstalled during the testing are shown in Table 5.12. We can see that 23% of uninstalls were defined as Description-NA, vs 10% at the start, of which category the app fell into. Almost 27% of all un-installed apps fell into entertainment, which includes Social Media and similar areas. A few examples of uninstalled apps include: Yelp, Uber, Ted, Reuters, NFL Network, Fitbit, Forge of Empires, CNN, Battery Indicator Free, and Solitaire Story. Many of these uninstalled apps require dangerous or moderate level permissions as previously discussed.

Table 5.12: Top 10 Popular Installed Vs. Uninstalled App Categories

App Category	Baseline Percentage	Uninstalled Percentage
Entertainment	27	27
Description NA	10	23
Tools	11	6
News and Magazines	2	5
Game Board	1	5
Game Casual	2	3
Health and Fitness	5	3
Shopping	5	2
Game Role Playing	1	2
Game Puzzle	3	2

We have also reviewed phase 2 and phase three groupings for Installed Vs. Uninstalled App Categories across our data collection phases to understand grouping of categories by

similar install times. We have reviewed phase 2 and phase 3 as shown in Table 5.13 and Table 5.14, which app installed vs uninstalled across each phase respectively. The uninstall percentages vary by phases and across the categories, which may be indicative of types of apps installed and concerns around privacy and security of apps and categories installed.

Table 5.13: Top 10 - Phase 2 - Popular Installed Vs. Uninstalled App Categories

App Category	Baseline Percentage	Uninstalled Percentage
Tools	10	20
Entertainment	24	13
Health and Fitness	2	9
Description NA	12	7
Travel	4	7
Game Casual	1	6
Communication	1	4
Productivity	6	4
Lifestyle	4	4
Shopping	4	4

Table 5.14: Top 10 - Phase 3 (Mturk) - Popular Installed Vs. Uninstalled App Categories

App Category	Baseline Percentage	Uninstalled Percentage
Entertainment	20	19
Description NA	13	11
Tools	10	8
Productivity	6	5
Shopping	6	5
Lifestyle	6	5
Finance	6	5
Health and Fitness	6	5
Game Casual	3	5
Travel	3	2

5.8.3.3 Permissions and Changes

We next turn our attention to reviewing permission changes of apps. As previously discussed, apps have properties including permissions and users have the option of disabling permissions, such as location, microphone access, etc., on their mobile devices, via the device or app settings. We have captured these (app) permission changes while users were running the HMN Mobile app. These changes include those (users) disabling or revoking requested access to category areas such as location, microphone, and other (dangerous) permissions, as defined by permission categories shown in Table 5.4. We examined all permission changes across all category types shown, focusing on the dangerous category, and found that 24% of users revoked one or more permission, across all apps installed (per user); 27% phase 2 and 34% phase 3.

As part the apps and permissions collection points we have split this work into distinct groupings for convenience: apps install (initial install of apps and their requested permissions along with the total installs requesting permissions), and End Testing (end of users HMN Mobile app testing). The apps Install point is a grouping of all unique apps installed (multiple times), along with their requested permission. The requested permission consist of a union of all permissions, across all app installations. This step allowed us to generate a per app requested permission set that is similar to the initial install configuration on the Google Play Store. The End Testing is permissions granted to apps installed (per user) at the conclusion of user testing.

Table 5.15 provides a view of permission categories and percentages, including: a union of permissions/apps (as a baseline) of unique and total installs, granted permissions (at the end of testing) of total installs, and the percentage of change between the start and end of testing; in this grouping we have combined the Other category into one grouping of like permissions. We found changes in permissions from the initial install of apps (baseline) to the end of user testing in notable permissions areas such as: Location,

SMS/MMS, and Microphone categories. We can see that the ratio of app installs to End Testing shows the percentage of time that users grant permissions when requested, and find that categories that have a relatively low percentage indicate that users are less likely to grant these permissions.

Table 5.15: % of Permissions (All Phases) per Categories from Start to End for Unique and Total Installs

Category	**% Unique Apps Requesting Permission**	**% Total Installs requesting permission**	**% Total Installs user granting permissions**	**% of requested permissions (start to end)**
Camera	39	39	38	98
Storage	71	72	70	98
Photos	71	72	70	98
Wi-Fi/Network	88	89	86	98
Calendar	2	3	3	98
Contacts	42	46	45	97
Location	51	55	54	97
Microphone	24	28	27	97
Phone	42	45	44	97
Device ID Identity	40	44	43	97
In-app purchases	21	21	20	97
Device history	40	44	43	97
Bluetooth	21	33	32	97
Other	2	2	2	96
SMS/MMS	10	10	9	91
Wearable	1	1	1	91

We also reviewed phases 2 (21 days), Table 5.16, and phase 3 (Mturk), Table 5.17, permission categories and percent of change between the start and end of testings for these phases. Table 5.16 and Table 5.17 show the granting and requested permissions between categories, a subset of apps from all users. Similar to all users, we found changes in permissions from the initial install of apps (baseline) to the end of user testing in permissions areas sch as location, and microphone. We again find that categories that have

a relatively low percentages indicate that users are less likely to grant these permissions. We find that these results vary between the phase length in terms of both categories and granted permissions.

Table 5.16: Phase 2 - % of Permissions per Categories from Start to End for Unique and Total Installs

Category	**% Unique Apps Requesting Permission**	**% Total Installs requesting permission**	**% Total Installs user granting permissions**	**% of requested permissions (start to end)**
Contacts	46	45	45	99
Microphone	25	26	26	99
Bluetooth	32	32	32	99
Storage	72	72	70	98
Location	50	50	49	98
Photos	72	72	70	98
Phone	46	46	45	97
Wi-Fi	58	90	88	97
Device ID Identity	44	44	43	97
In-app purchases	28	21	20	97
Device history	44	44	43	97
Camera	40	36	35	96
Calendar	9	10	9	93
SMS/MMS	11	12	11	92
Wearable	9	9	9	92
Other	2	2	2	90

We examined a set of popular apps permission changes, at the end of user testing, where we reviewed the following popular apps: Facebook Messenger, Netflix, Spotify, Reddit, Uber, Yelp, Fitbit, Zillow, Flashlight, and Outlook. We found users revoking permissions for categories such as location, microphone, and other areas. As an example we can see that 2% of users revoked location for apps such as Uber, Yelp, and Outlook, and almost 5% of users revoked Microphone access for apps such as Netflix, Spotify, and Uber. These users revocations line up with privacy and security concerns users have ex-

Table 5.17: Phase 3 (Mturk) - % of Permissions per Categories from Start to End for Unique and Total Installs

Category	**% Unique Apps Requesting Permission**	**% Total Installs requesting permission**	**% Total Installs user granting permissions**	**% of requested permissions (start to end)**
Location	61	66	62	93
Microphone	28	33	31	93
Wi-Fi	71	73	68	93
Contacts	49	56	52	92
Camera	47	51	47	92
Storage	77	80	74	92
Photos	77	80	74	92
Bluetooth	33	38	35	92
Wearable	10	12	11	92
Phone	47	50	46	91
Device ID Identity	44	48	44	91
In-app purchases	25	23	21	91
Device history	44	48	44	91
SMS/MMS	9	9	8	90
Other	1	1	1	90
Calendar	9	10	8	87

pressed as part of this study. Additional permission details can be found in Appendix A.5.

We next review revocation of categories across all users. Table 5.18 is a listing of revoked permissions users applied during the running of the HMN Mobile app, across the categories shown in Table 5.4; this table is a listing of dangerous permissions, where we have included two additional non-dangerous permissions types (Wifi/Network, and Gmail), and we have excluded the non-dangerous device history category from this listing. We can see that 9% of users disabled phone access, 7% disabled Wifi access (non-dangerous), and 5% disabled Location access. In addition, 5% of packages had SMS/MMS permissions disabled, and 4% of all apps had location disabled.

We were surprised that permissions were revoked by users at this rate, across apps/users/permissions, and believe that this is indicative of privacy and security concerns. Across all of these areas a commonality of dangerous permissions such as location, Microphone, and others, as well as non-dangerous permissions such as Wifi (and others), which again we believe is in-part due to concerns of privacy and security. It is clear that users have made the steps needed to help with privacy and security, by disabling several categories of permissions (such as location, phone, and wifi), across numerous apps, that they feel should not have these permissions and more importantly should not have this access.

5.8.4 HN Wifi Health

A review of HN Wifi Health on average shows a result of a four star rating, with the overall ratings as shown in percentages in Table 5.19. We can see that 73% of all HNs had a rating of 3 or better, and only 27% had a rating below 3, with only 1% having a rating of 1. It should be noted that there were no reports of a five star rating overall.

Table 5.18: Dangerous Permissions revoked across all Apps and Users

Category	% dangerous perms deleted, Apps	% dangerous perms deleted, Users
Contacts	2	5
Camera	2	3
SMS/MMS	5	5
Storage	2	5
Location	4	5
Photos	2	5
Microphone	3	2
Phone	2	9
Wi-Fi/Network (Other)	2	7
Device ID Identity	3	1
Calendar	2	1
Bluetooth	2	2
Wearable	4	2
Other / Cellular	2	1
Other / Gmail	3	1

Table 5.19: Wifi Health Ratings

Percentage	**Rating**
37%	3
36%	4
26%	2
1%	1

5.8.5 Apps Health

A review of HN Wifi Health on average shows a result of a four star rating for 97% of HNs, with the overall ratings as shown in percentages in Table 5.20. We can see that only 1% of all users had a rating of 5, and 2% had a rating of 3.

Table 5.20: Apps Health Ratings

Percentage	Rating
97%	4
2%	3
1%	5

Figure 5.8 shows a CDF of those apps that all users had installed (baseline) and the dangerous permissions these apps requested versus those apps users uninstalled and their dangerous apps requested. While the CDF in Figure 5.8 follows a similar line pattern, we can see that 50% fell on 5 and 4 dangerous permission requests respectively. We also can see that the overlap for permissions cross boundaries at two dangerous permissions, and roughly stayed at one permission delta from this point on. The results shown are surprising as one would have assumed that these results would be reversed in terms of dangerous permission types deleted/revoked.

Table 5.22 shows a list of unique (distinct) apps installed versus total installed apps with dangerous permissions. This table shows that the more popular apps (installed more) are those that require more permissions.

We have created an Android permission category request, shown in Table 5.21, to highlight both the top requested app permission types (e.g. Internet) and the unique, as well as a breakdown of the top permissions requested (e.g. read, write, access, etc.) and installed access type requested (e.g. Dangerous, Normal, Moderate), based off of table 5.4 access levels. Table 5.21 shows Normal and Moderate permission, along with the 13 Dangerous permission levels (possibly requested) and defined by Google. We can

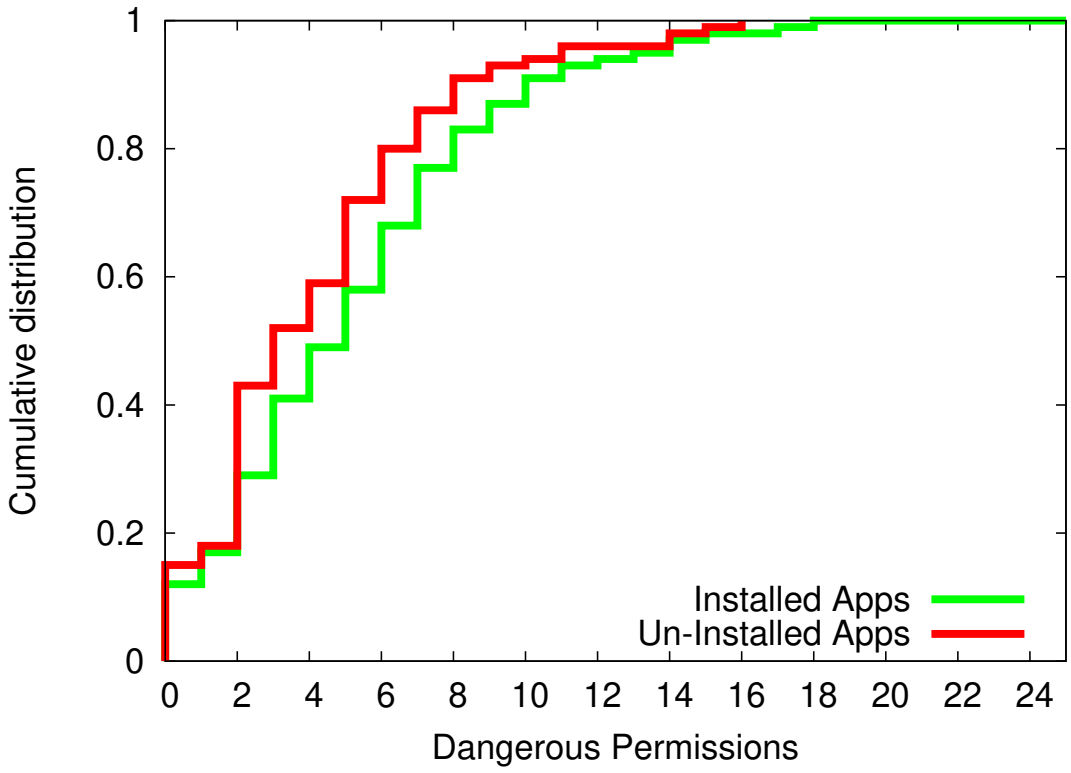

Figure 5.8: Installed vs Un-installed Apps Dangerous Permissions CDF

see that the location permission was requested by 32% of all apps installed across all users, and 34% of unique apps. Other notable dangerous permissions requested include: camera, record audio, and read SMS (text messages). Of the 3095 unique apps installed (total of 6092 apps across all users) the read SMS apps was requested by 150+ of these apps, and included: Amazon Shopping, Uber, Facebook, Amazon Alexa, and others. The location permission was requested by 1347 apps that ranged from games, social media, tools, and included apps such as Facebook, NPR News, Pokémon Duel, CA Lottery, Wendy's Mobile app, and many others.

We have created a table of the top apps installed and their respective levels of access requests (dangerous, moderate, normal) to provide background on number of permissions requested in each category. Table 5.23 shows popular apps such as Facebook, Spotify, Netflix, and others along with the respective number of permission category requests. As an example the Facebook app requested the following 10 dangerous permissions: Call Phone, Location (including granular access), Read and Write Contacts, Get Account Information, Read Phone Status, and Read and Write Calendar.

We focused in on the dangerous permission location and examined all packages installed, across every HN, to see which HNs have disabled this permission. We found that 5% of HN users, 4% of all apps, explicitly disabled the location permission for a set of apps. Only one app was found to be in the top 10 most downloaded list (Google Instant apps). Other apps in this list include: fitbit, Yelp, Flip, AAA, and others. As previously noted we are not surprised that users have made these changes to location, as it indicates that HN users are concerned with Privacy and Security of their HNs, and are not trusting permissions set by the app developers.

Table 5.21: Categories of App Permissions Granted

Permission Group	% Unique / installed	Breakdown % top perms	Level
Contacts	34 / 45	34 Access, 16 Read, 5 Write	Dangerous
Camera	33 / 40	33 Access	Dangerous
SMS / MMS	11 / 12	11 Receive/Read, 3 Send	Dangerous
Storage	73 / 75	73 Read, 67 Write	Dangerous
Location	34 / 32	34 General, 32 Fine, 33 Coarse	Dangerous
Photos	73 / 75	73 Read, 67 Write	Dangerous
Microphone	20 / 27	20 Record	Dangerous
Phone	42 / 45	9 Call, 3 Read, 2 Write, 42 State	Dangerous
Wi-Fi	51 / 60	51 Access	Normal
Device ID / Identity	63 / 71	63 Access (R)	Dangerous
Calendar	10 / 8	10 Read, 4 Write	Dangerous
In-app purchases	12 / 14	12 Access (R/W)	Moderate
Device / app history	63 / 72	63 Read	Dangerous
Bluetooth	25 / 31	25 Access	Dangerous
Wearable Sensors / Activity Data	9 / 8	9 Access, 4 Motion	Dangerous
Other	80 / 100	80 Network, 43 Vibrate	Normal

Table 5.22: Dangerous Permissions (distinct and installed)

Grouping	Qty
Dangerous permissions per distinct app	5
Dangerous permissions per installation	7

Table 5.23: Top App Permissions Installed and Categories

App Name	Dangerous	Moderate	Normal
Google Play Games	1	0	5
Facebook Messenger	14	2	23
Netflix	2	2	12
Instagram	8	2	15
Google Pay	4	2	13
Twitter	7	2	13
Snapchat	6	2	14
Facebook	10	2	25
Spotify	3	4	18

5.8.6 Local and Global Norms

We have combined the Local and Global Norms section into one subsection for this summary data, and have included the Wifi and HN throughput speeds as well as upload-/download throughput speeds in this section. The Summary of Local and Global Norms and comparisons collects Internet throughput statistic and provides local and remote listed devices found (on average), as well as provides a star rating of the HN's Wifi and apps heath. Starting with Internet connectivity throughput (download and upload), Figure 5.9 shows 9 of the 22 distinctively grouped providers in the US, which we have labeled as Cable1..9; additional throughput measurements via Fiber, DSL, and Cable can be viewed in A.3 Section. Figure 5.9 shows the cumulative distribution function plots (CDF) for US Cable providers with Cable8 having the overall best ratings for download capacity.

The spread across all of CDFs provide a clear picture of throughput for both download and upload. We can see that the top providers, shown in Section A.3, Fiber4, DSL9, and Cable 8 provide the best download throughput overall with Fiber 2 and DSL5 hav-

ing the worst performance. Cable 8, Fiber 4, and Cable 3 provided the best overall upload throughput, whereas cable 4, DSL 5, and DSL 7 provided the lowest overall upload throughput. We found that the average download Inet speed to be 11Mbps, and 3Mbps for upload, across all users. We can see from the charts that the lower end clearly pulled the throughput numbers down, and that not all providers guarantee and or offer similar throughput numbers (for either download of upload). See Appendix A.3 for additional Internet throughput details.

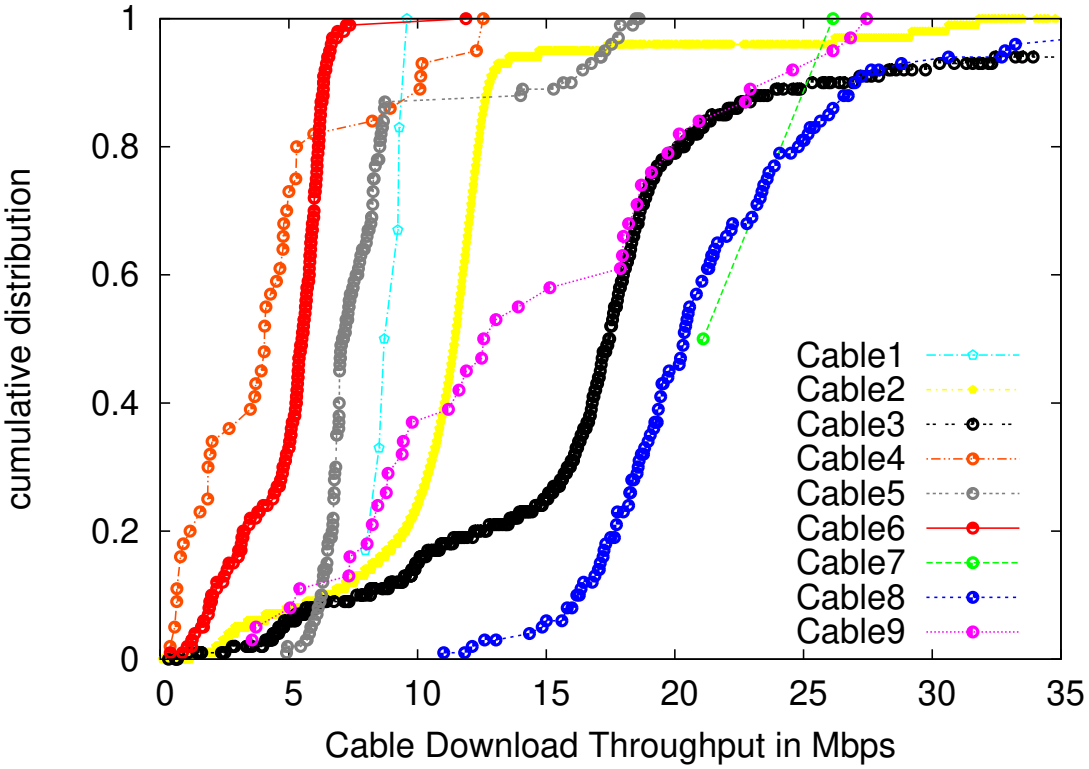

Figure 5.9: Cable Inet Providers Download Throughput CDF

5.8.7 Research Data and Other Characteristics Associated with HNs

We mapped IP locations, and ISPs and provide a map and table for each of the 130 usable (out of 148 total) participants; the remaining 18 users did not fully complete the

running of the HMN Mobile app. A list of Internet types (Cable, Fiber, DSL) that used can be seen in Fig 3, with additional details available in Section A.4. A map of states, and countries includes users from US states as shown in Fig 5.10, and from countries outside of the US Fig 5.11 (generated via the [103] free online tool). In terms of Wifi connectivity, and APs attached, HN user's connected to ˜ 24 Wifi networks (SSIDs) outside of their HNs, via their Mobile device. A by-product of these SSIDs being available is the potential for tracking of locations devices have been in the past, without location or other advanced permissions required.

We received ˜28K log input points around Wifi speed and throughput, and saw an average Wifi speed of 22MBps on average or 14MBps median across all HNs. The average Received Signal Strength Indicator (RSSI) across all participants was ˜3.5 (scale 0-4) with a median of ˜4, which indicates that most HNs had a reasonably high Wifi speed and connectivity (e.g. Excellent >= -50 dBm). The link speeds indicate that most HN participants used a 802.11abgn wifi (84%), and 16% using 802.ac (866.5Mbps), overall the vast majority (˜63%) used an 802.11n speed wifi unit. Additional device information can be found in Section A.4.

Percentage	Inet Type
65%	Cable
23%	DSL
8%	Fiber
2%	T1
2%	Comp

Table 5.24: US ISP Connection Types

A comparison on number of devices found within each HN ranged from a high of $40+$ to a low of 5, with the average of 14 found to be in a HN. Each environment had

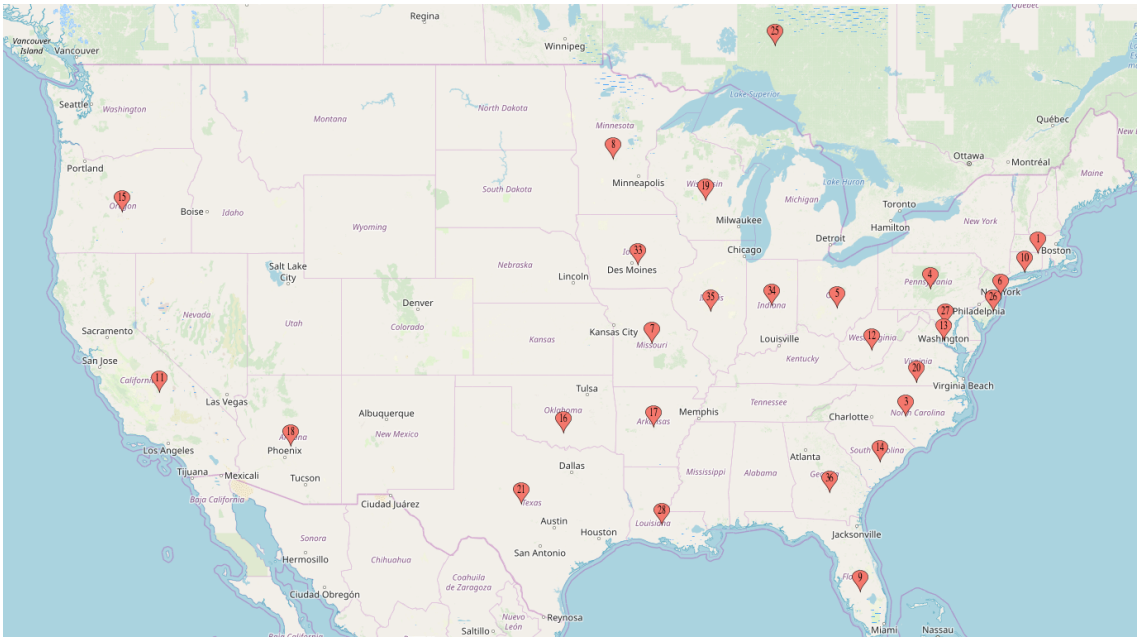

Figure 5.10: HMN Usage US

Figure 5.11: HMN Usage Outside of the US

at a minimum a Wifi/Router, a Mobile device, and a PC. The basic device comparison information was well received as shown in the research questions section 5.3.

Our final area of comparisons includes HN Wifi and apps Health. As previously reviewed the Health indicators provides a comparable rating (1:5) using yellow color coded stars, and is inline with survey, and app rating systems across the Internet. We start with the HN Wifi health, where we found that the average Wifi HN was rated as four stars. This rating indicates that there was strong RSS, and throughput measurements across the most HNs. Fig 5.12 shows a Cumulative Distribution Function (CDF) of the 102 collection Wifi health points (not all HNs reported Wifi health), where we see that most Wifi HN Health fall into the range of 2-4, and we did not see any 'perfect' 5-star ratings; note that a point in time rating may have been seen in the app, but over time this was not the case.

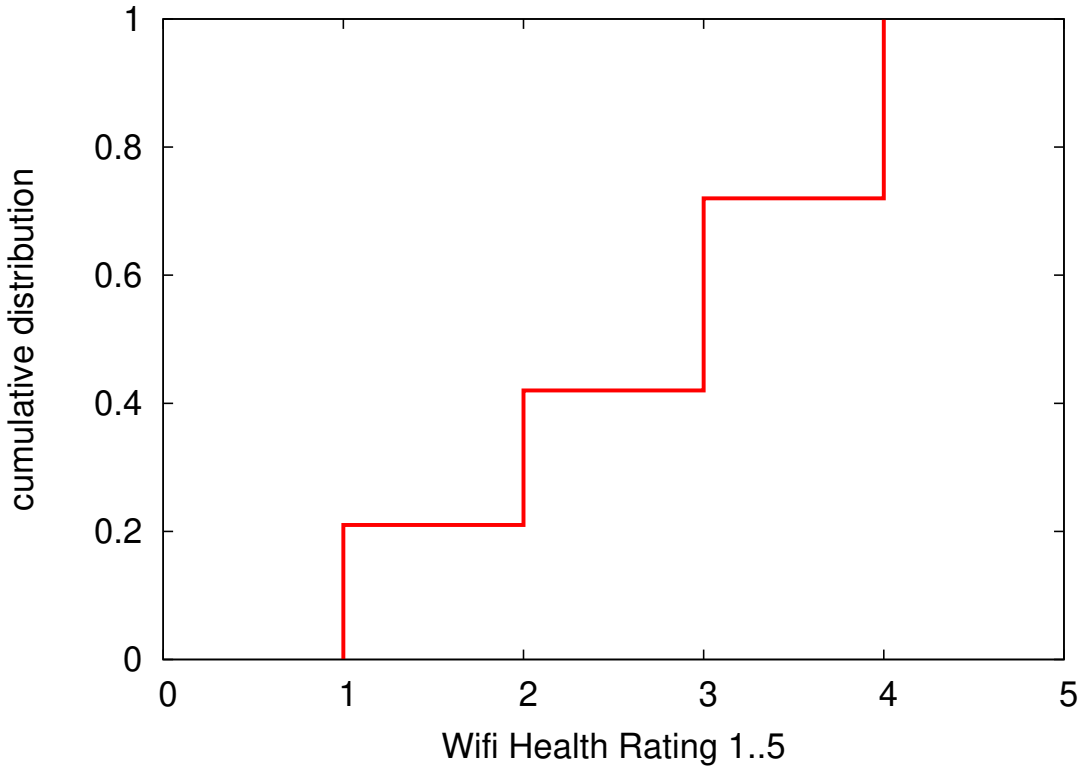

Figure 5.12: Distribution of Wifi Health Rating

We next move to the app Health rating, and find that the average rating is four stars. We can see that most app ratings fall into this range from Fig 5.13. We did not find any apps with a one star rating, and only one with a five star rating. These outliers indicate that a large percentage of apps require dangerous or moderate levels or permissions. We also that there there were minimal ratings of two and three, where a 4-star rating is roughy 75% of all app Health scores. While these are high numbers in terms of ratings, there are a swath of apps that require potentially unneeded and elevated privileges that users should review on a regular basis as they can cause security and privacy issues.

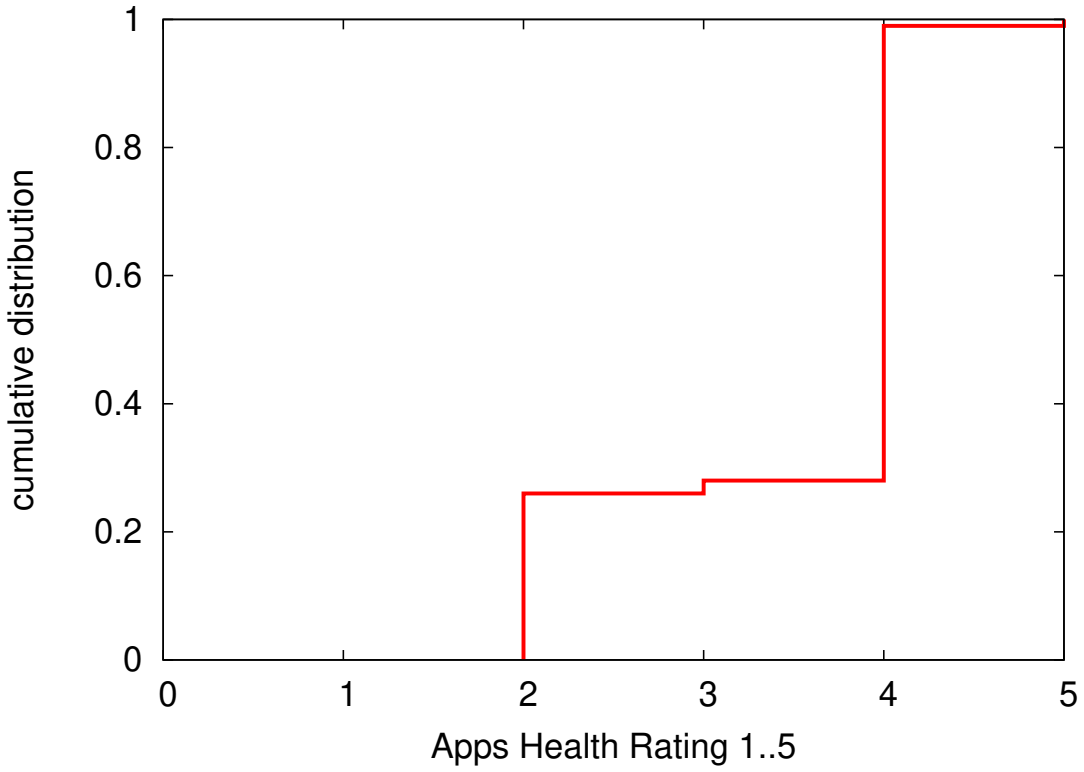

Figure 5.13: Distribution of Apps Health Rating

5.9 Feedback Results

In our final area of review we look to user comments and survey results for comparisons. In previous work [138] we leveraged a survey to HN users looking to understand experiences, devices, and comfort level in and around their HN. We have extended this survey work and include a series of questions built into the HMN app, and inline with Section 5.5. See Appendix A.2 for a complete list of questions used. We used key word and sentence matching techniques, as part of all comparisons and results, to gather the statistics shown in this section as well as scripts, Linux shell commands, and sheet-based analysis. This analysis consisted of breaking each of the areas down into a binary response, along with categorization ares (e.g. Help, Security/Privacy, Usage, and Comparison), where appropriate, for results shown in this section. We move through each subsection area designated in Section 5.4 and provide binary and categorical results, where appropriate, along with discussion for each question.

5.9.1 Wifi and HN Throughput Speeds

We found that 87% of users responded 'yes' that providing information on Wifi "speed" was helpful, with 13% expressing no interest. Users provided additional feedback, which included comments associate with not feeling as if they are receiving their fair share of Internet bandwidth from their ISP after running this test, and could leverage this to negotiate for a difference in price.

A review of network connectivity and throughput found that the common requests could be broken down into four areas, including: Help and Tips, Which devices are using the bandwidth, Ping Times, and Compare local and remote devices and networks. The following four notable comments were provided by testers *"It's useful to see how my connection is compared to others. Maybe when I next renew with Comcast I can use my*

slow connection to push for a discount", *"information on what could be done to speed things up if possible"*, *"tips on how you can make your speedometer faster or if there is weather or other things slowing it"*, and *"I like to monitor how the bandwidth in my network is used, and this tool gives a very easy way to check that. I would add alerts for when there is an abnormal use of the bandwidth, and suggestions on how to improve the performance of our network"*. Table 5.25 shows a breakdown of these four categories, with the top 75% of feedback looking for Comparison or Help.

Table 5.25: What information would be needed for Speed-tests and Wifi Speedometer?

Is Wifi and Internet Throughput Helpful?	**Percentage**
Compare (local and remote)	44%
Tips / Help	31%
Device Breakdown Which device is using throughput, where, and when	13%
Ping Times	11%

5.9.2 HN Device Listing

93% of users overwhelming found the device listing helpful, with only 7% expressing that this information was not helpful. Those who listed no as an answer provided feedback that they were looking for more information to understand the listings, and what it meant to the over HN. The following is a notable comment from the list of responses: *"The network scan tab gives me way more information than I ever knew about my network. I feel like this app is a good education tool to help you evaluate your network and activity"*.

Table 5.26 provides a variety of details from the feedback around metrics (comments) on how to manage privacy and security (and general help) for devices, IoT management, if (other) Wifi devices and Networks are interfering, and how HNs compare. A wide range of responses centered around an in-depth knowledge of HN devices, and privacy/security and general remediation. We fond that 71% of all responses were embodied in areas of:

In-depth details, and Privacy /Security and help. The follow are two specific comments in the ”in-depth” and ”Privacy/Security” areas: *”If possible I would like to see which devices are using the most data on the network and maybe get a warning if any of the devices have known vulnerabilities”.*, and *”I am interested in learning how the things that are not secure can harm my network and what to do to fix it, how dangerous are the things that aren't secure, I can see the devices that are linked to the network but not sure if they are secure. I did like it showed a new device was added to the network as my daughter got a new computer yesterday”.*. These, and other comments, pointed towards the categories shown in Table 5.26.

Table 5.26: What other information would you like to have included as part of a Device Network Scan?

Classification of Response	**Percentage**
An in-depth exploration into areas of (devices): Access, outages, vulnerabilities, attempts	43%
Privacy / Security and general help	28%
No Changes Needed	12%
How to Manage IoT devices	9%
When Other networks interfer or devices attach	5%
Alerts	3%
How my Network Compares	1%

5.9.3 App Security / Privacy

A review of results of the apps privacy and security features divulge that 83% of users felt that this was helpful in showing permission, and possible issues with apps (in one location). As an example we found the following specific comment as part of the responses: *”Yes it's very useful because I found apps that had permission to certain aspects I wasn't aware of. This allows me to better control my apps and what they can do”.* This comment was similar to the yes responses, with slight variations. The remaining users felt they either had a good handle on security/privacy or that it was too much information. Users

who responded no provided similar messages as the following comment *"not really, as I usually know this info already, I guess it is good for an at a glance look"*. We also found users who uninstalled apps responded with comments such as: *"I did realize that I should uninstall apps and devices I no longer use."*, and *"I only checked the apps list that are red and uninstalled the ones I do not need or use"*.

Users requested an in-depth exploration into issues, and possible remediation steps, across both yes and no responses, where we classified responses around four categories: Security/Privacy of apps, Breakdown of apps (timing, usage, taxonomy of permissions, limitations, etc.), Hints and Tips, and Alerts to changes (although this one already exists it was pointed out by 4% of users). The top 72% of responses fell into the Security/Privacy Risk, and an in-depth review, as shown in Table 5.27. Notable comments in this area include *"I would like to see security risks, attempts to access from outside the home, connectivity time, detailed health meanings"*, and *"it might be nice just to have it sectioned out by who has permissions for the camera and then list the apps who has permission to see my contacts and then list the apps. And then show which apps have actually used their permissions"*.

Table 5.27: What other information would you like to have included as part of apps Security and permissions Home Network Scan?

Classification of Response	**Percentage**
Security/Privacy Risk	36%
An in-depth exploration into areas of (apps, devices, HN): Access, outages, vulnerabilities, attempts	36%
How to help privacy and Security	24%
Alerts	4%

In addition to this we provide a deeper dive into details users had interest in as part of the privacy and security of their HN. Table 5.28 shows results from this question, and breaks down to the following four areas: Review of usage, Remediation, Alerts, and how things compare to others (devices, frequency of usage, connectivity over time, etc). The

top 75% of responses fell into an in-depth review and help/remediation of privacy and security. We found the following notable comments from this area of feedback, specifically (in-depth, and privacy/security remediation): *"I would like to see a summary of how the security my networks offers compares to those of other providers. I would also like to have a summary of unsuccessful attempts to breach my security barrier, to give me an idea of how many times people have tried to access my information. More over, I would like to see the details of what security measures are taken, how they work, how they're implemented and why its important"*, and *"I would love to see how to fix any flaws in my home network. is there something on another device that is causing issues? I would like to know details that I can help my network be strong and supportive"*.

Table 5.28: What type of details would you like to see as part of a security and privacy review of your Home Network

Classification of Response	Percentage
An in-depth exploration into areas of (apps, devices, HN): Access, outages, vulnerabilities, attempts	42%
How to help privacy and Security	33%
Alerts	23%
Compare	1%

5.9.4 HN Wifi Health

HN Wifi Health received feedback similar to other ares, where responses fell into help of information related to Wifi, and the rating system. The following comments provided narrative around Wifi health, and the norms that they provided: *"It has been very helpful, but I would like for some information on how to improve my wifi health"*, and *"yes because i didn't realise how many devices really used my wifi you tend to set them up and forget about them. i guess now i know why my power bill is always so high...lol"*. Categories of responses fell into feature requests, breakdown of results, comparison, and general help for HN Wifi health.

5.9.5 Apps Health

Apps Health fell into responses to the question ”What types of information would you like to see a Privacy / Security Mobile app provide?”. This question centered around the health and privacy and security integrating into an app that works on a Mobile device. Notable comments in these areas include: *”it would be interesting to see what exact permissions each app is using each time”*, and *”suggestions for other apps to use based on the apps that people with similar apps as mine us”*, and *”if apps are using my internet, or installing viruses”*. We found that users had interests ranging around requests for integration into anti virus definitions, and tools, as well as how to fix and share information to remedy issues.

We found that 62% of users reported that operational Health (Wifi and apps) helped them understand current status and was indeed helpful to them, and found comments such as: ”a very useful tool in determining the health of your devices and apps. You can see what could be potentially risky and it compiles all the information neatly” and others looking for details on how to improve their environment “I would like for some information on how to improve my Wifi health”.

5.9.6 Local Norms

A review of results around star rating found that 76% of users felt that they saw value, with 19% (from both yes and no users) requesting more details be provided on background of metrics. Notable (yes and no) comments included the following *”yes. I love star ratings. They’re universal and easy to use and understand. I wish more things would use star ratings, so simple and effective”*, and *”I’d love to have suggestions on how to improve the performance of my network, with maybe some projections of how the rating of the network would change if I tried any of the proposed solutions”*.

5.9.7 Global Norms

Results from this question included the following categories from overall responses: breakdown of which apps, devices, etc. are in need of attention, a comparison between historical data or other networks, Tips and Help to optimize, and how to secure my apps and HN, as shown in Table 5.29. Users provided feedback and comments such as *"I would like to see a little more detail about what it actually means and possibly what we could do to improve any area that is lower than it should be. also maybe how each score is based"*.

Table 5.29: What information would you like to see added to a Star Rating System for comparing your Home Network/apps/Devices?

Classification of Response	**Percentage**
Breakdown Which apps, devices are in need of attention	30%
Compare (local and remote)	28%
Tips / Help	20%
Security / Privacy Concerns	22%

We found that 80% of all users reported that the information provided in the Summary Comparison Tab to be helpful. Users stated that they used the information in this tab to help diagnose and to determine that their Internet speeds or Wifi is slow. We also found feedback and questions looking for in-depth exploration on why other HNs had better Internet throughput, better ratings, and overall better scores. Overall the recurring theme here is that users are looking for information and remediation techniques for devices, apps, and security and privacy.

Table 5.30 provides details on what HN users would like to see included in a comparison chart, and how they are compared. We can see that 31% of users expressed interest in finer details on security/privacy, and 23% are looking for help on improving their HN and fine-tuning when problems arise. 25% looked for information on comparisons to others

(including apps, and types), while 21% looked for information on local comparisons of information and ratings of throughput per device, and hot running apps and devices.

Table 5.30: What other information would you like to see provided when comparing your Home Network/apps/Devices versus Others?

Classification of Response	Percentage
Security/Privacy Risk	31%
Help and Tips on improving HN	25%
An in-depth exploration into Comparison to other HNs including (apps, devices, Health):	23%
Usage of devices, and apps Including (per device, point in time, hot devices, speed, etc.)	21%

5.9.8 Preferred Method to Gather HN Information

In our HMN survey we asked device types users would prefer to use when collecting HN data. We also asked this same question during the HMN Mobile app study. The results from this can be seen in Figure 5.14. We can see that 55% would prefer to use a mobile device, while 24% prefer a PC, 14% Router, and 8% a customized piece of hardware. These percentages are similar to the initial HMN survey results reported 4.6, with the PC numbers being slightly lower in the survey.

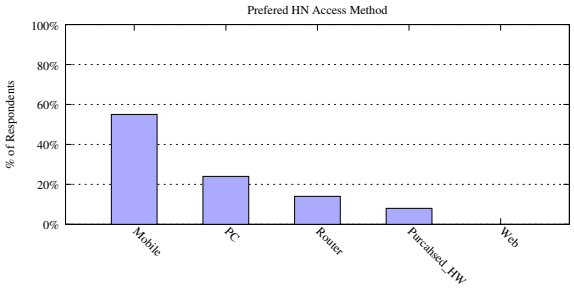

Figure 5.14: Device Preference gathering information within HNs

5.10 Discussion of Research Questions

In this section we follow-up on the research questions from Chapter 5.3, and how this study has impacted the areas of discovery, information, and preference as well as provide what information we found surprising as part of this work. We work through each research question, labeled as RQ1..5, and provide discussion around each of the areas of focus.

RQ1: What HN ecosystem characteristics can be discovered using a Mobile app?

We have shown that leveraging a Mobile app, with non-administrative privileges, for discovery of HNs is an effective method for the determination of sources (e.g. devices, apps, feedback, etc.) that also provides a easy method for users to operate with minimal impediments. Characteristics areas include: device discovery, app information, privacy and security, throughput speeds, Wifi speeds, user feedback, as well a method to calculate and display a Health Rating. We can also determine user patterns (as related to privacy and security) by reviewing installed apps, and un-installed apps over time to understand potential privacy and security concerns, and other types of information. Discovery of pinpoint information can be challenging without administrative privileges, but an app method clearly falls into the areas that are well fit for both convenience and discovery, while minimizing impediments and maximizing incentives for users to leverage.

We learned that a Mobile app can successfully gather a large array of HN information, which also may be of interest to both researchers and users, without the need for administration privileges. These areas include the following HN ecosystem details: devices and identification types, network speeds and throughput, app privacy and security details, local configurations (e.g. Wifi networks attached to), collecting user feedback, and leveraging data-sets for comparison to other users. However, using this approach has limitations for control of data outside of the mobile app domain, as administration privileges are needed for access to devices within the HN to control flow or enact changes on

the network. These include controlling bandwidth to specific devices, controlling what type of flow is allowed to devices (e.g. Netflix, Twitch, etc.), and other types of actions that require hardware interactions with administration privileges

A review of what cannot be successfully be completed or the data that cannot be gathered via the How's My Network app pointed toward limitations in two main areas. The HMN Mobile app has constraints in terms of data collection points, as well as functionality as compared to other methods. As an example a Mobile app typically does not have the ability to easily collect software, and permissions running on remote devices without direct access. Other areas that an app running with non-administrative privileges have include collecting network information at Layer-0 or Layer-1 (not possible without administrative privileges and a hi-jacked operating system); even with these privileges data collection would still be limited to the network segment (versus that of a router) and remote control would be minimized to that of the local device. In addition, control of other devices or remote tools would be a challenging proposition, even with the advent of IoT-based protocols.

Overall, we were surprised that we can pull as much data as we could (e.g. location, Wifi Connectivity, and apps information) without having to request a series of what would be assumed elevated privileges for this access. With this distinction, these potential security and privacy areas may be prone to hackers [177], privacy leakage, and leveraged for dubious reasons. Although there are restrictions to a variety of areas and permissions required, via the Google security model (e.g. granular network access and specific OS versions), it still may not be clear to HN users on exactly what these services are providing and or requiring. Access requested versus access allowed in some case do not overlap in terms what details are actually provided by the environment.

RQ2: What HN information and features are HN users interested in?

We have seen a recurring theme of interests across this, and previous work, including:

Security and Privacy information of apps and devices, help and tips on how to optimize devices and a HN, usage information including metrics and point and time data, and a comparison of other HNs and devices. These areas also included requests for in-depth exploration into "when changes occur" across the HN and or devices (e.g. apps, permissions, and new devices); these distinctions and changes can all be classified as privacy and security changes. The feedback, received from HN users, point toward an interest in the type of data provided by this app, as well as more details in these areas.

Feedback across the board pointed toward a variety of areas that have been grouped into a series of classifications, but all have a common theme of an in-depth exploration of information on how to remedy issues (e.g. ping speeds was mentioned by gamers). As an example users provided the following comments *"I still think having a feature allowing you to see how much wifi each device was using would be good"* and *"what I could do to get max speed. also it doesn't list what my possible speed for.my router is. also evaluate my usage over time"*. Feedback also pointed to areas where users had little or no interest in as part of this work, and included the following lower ranked / commented areas: data usage over time, which devices are using the most bandwidth, IoT usage, and Wifi comparison. As an example we received feedback and comments such as *"some notifications indicating anything extreme changes. A battery life diagnosis"*., and *"I would like to see usage by each item listed. that would be really helpful"*.

RQ3: How do users evaluate the Mobile app approach for collecting and sharing information about Home Networks?

In this question we provide several areas of information and analysis, including user input and data gathered. We found that 76% of respondents indicated that this app has helped with understanding privacy and security, with the remaining 24% (and including a smaller percentage of the yes responses) looking for more details, alerts, help, and mostly remediation of issues.

In our initial survey of the HN users we asked which device types users would prefer to use when collecting HN data; we asked this same question during the HMN Mobile app study where we found that 55% would prefer to use a mobile device, while 24% prefer a PC, 14% Router, and 7% a customized piece of hardware. We found these numbers to be similar to the initial survey we performed in our previous HMN Survey study 4.6, with the PC numbers being slightly lower in the survey.

Throughout this work we have seen a keen interest from HN users around the area of privacy and security. The general theme of results have shown that users found that the HMN Mobile app provided valuable information in terms of apps, and devices. The apps information was of interest to users as it yielded information about permissions as well as an overview of general rating of the app.

We were not overly surprised to see that users reported an interest in understanding comparisons between HNs, as we saw interest in cross functional areas as part of feedback. As an example we found comments and feedback, including: *"i would like to see what the average person has for devices and security. am i more or less secure than others"*?, and *"while there is basic information for my home network, others seem to have faster uploads, so I would be curious to see if they are using the same internet service as me"*. These and other responses show an interest in the norms of services provided in this study as well as a comparisons to how they stack up in terms of ratings.

RQ4: Will users change their security and privacy perceptions of their HN by using a HN tool?

This research question is interested in understanding perception changes of security and privacy. We have included the control question results and user feedback received from the HMN Mobile app in this section, as it closely matches other user feedback in terms of Security and Privacy of apps and thus operations of the HN ecosystem. We

asked the following question to all testers as the first question, and subsequently as the final question during testing.

”How has your perception of security/privacy in your Home Network changed within the past week?”

We used this question as the pre and post-amble question to this study, and Fig 5.15 shows a review of results, taken from initial responses. The overall responses from the initial responses show that most respondents felt that their network was secure, and that their perception had not changed in the past week or more. Of the approximately 150 responses received to this initial question we found that 17% felt that they saw a change (even if minimal), and 84% of the testers reported that they are more security conscious after using the HMN Mobile app, and that their perception of security had changed, as shown in Figure 5.15. We can see that there was an almost flip of security / privacy concerns by the users from the start to the end of the study.

Users supplied comments such as *”I am considering seeing if I can remove some of the devices on the network it looks like maybe that us holding me back”*, *”yes. I have shut down access to certain devices such as my iot security access and garage door Wi-Fi access”*, and *”yes I have shut off devices to save my electric bill and bought software security for my phone”*. These, and other comments, reinforce that users have modified their privacy and security nature due to using the HMN Mobile app.

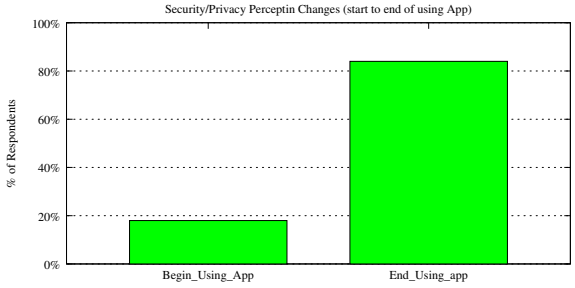

Figure 5.15: Security Preference Question at Start and End of Using App

We were surprised to see that HN user perception of privacy and security changed as much as it did between the testing time-frame. We had assumed there would be a change, but we saw an almost 60% swing in perception change related to Privacy and Security by users (starting at 17% having no change to almost 85% to perceiving a change) from beginning to the end of the study. This is of interest as we had expected a shift, but did not realize it would be this significant. Users provided feedback and comments such as *"I'm still concerned over Facebook's keeping plain text passwords. that freaks me out. makes me think of my overall security including my home network"*, and *"I've deleted a couple apps that were of concern. I think I'm going to update my Wi-Fi router after this so it's more secure"*. These and other comments pointed toward a change in perceived security and privacy behaviour related to HNs.

RQ5: Which "local HN" to "remote HN" comparisons do users find the most and least useful?

In this study we found that users expressed interest in understanding how their networks, apps, and privacy and security compared to others. This ranged from operational health ratings, to device information, as well as a request for an in-depth exploration of information not already provided in the HMN app. We found that 58% of users expressed an interest in understanding comparisons of HN norms. We found this across comments, including *"i would like to see what the average person has for devices and security. am i more or less secure than others"?*, *"Any detection of the network going down at any time"*, *"What accounts for a higher or lower star rating? Are there steps I can take to improve my rating"*?, *"[local and remote] security threats"*, *"why is my network working horribly or maybe even is it up to the standard it should be at"?*, *"which phones or devices are using the most data"*, and ratings around *"My performance compared to a 'standard/optimal' one would be great"*.

Overall the general consensus was that there is an interest in understand a local view of how their HN is operating versus a global view for comparison, which we refer to as Local and Global norms of data.

RQ6: What impact does a Mobile app have on a user? Will a user modify their HN Ecosystem or digital fingerprint based off of the HMN Mobile app?

We found that the HMN Mobile app did impact users perception of HNs and Mobile usage. We discovered that HN users modified their HN Ecosystem as related to apps, and devices on their network. Starting with feedback, we found that 46% of users felt that they have changed the way the use their device, apps, or HN because of using the HMN Mobile app. A common recurrence across both yes and no responses was that users were looking for more details on help for remediation. In addition, we can see that users uninstalled almost six apps, while using the HMN Mobile app. The top types of uninstalls fell into social media (Entertainment), or similar, and had dangerous permissions listed across all apps.

In addition, HN users provided feedback in and around health scores and star rating system provided by the HMN Mobile app. The following are a few specific comments around health scores: *"I think this could be a very useful tool in determining the health of your devices and apps. You can see what could be potentially risky and it compiles all the information neatly"*, *"Yes, because it provides useful data and statistics and also provides ratings to help you visual the data and understand where your devices and apps stand with overall health"*, *"it would be nice if there was a more detailed breakdown that offered tips to improve the ratings"*, *"the star method is helpful and I can't think of a better method to communicate that information"*, and *"Yes, it is helpful as I can view how each day progresses and see what I feel the results were"*. Overall the feedback pointed toward a resounding yes that the star rating and health system was helpful to understand

how a HN system is functional. We can see that the overall value was shown to be a majority of users agreeing with this system being useful. A small percentage of users looked understand the rating system (which was available in the help and on the How's My Network Web page) and or found no value in this method as shown.

While the feedback provided directly from users was not overwhelming 46% stating they changed the way they use their HN, the number of uninstalls found via the HMN Mobile app logs (Avg 6 per user, median of 1.5, or 3.5 for non-zero un-installs) during the study generally supports that users have changed their behavior due to reviewing their privacy and security of apps installed on their device. These results point towards a greater awareness of security and privacy concerns, which users pointed out throughout this work.

5.11 Research Implications

In this section we provide a review of the items that we were surprised to be able to gain access to, and discover, as well as possible implications of these surprises. We were pleased that we were able to gather, present and share this information with users and researchers, however we have concerns that anyone who creates an app has the ability to also gather this sensitive information so freely, and with minimal notification to the user. The following are implications found by this work:

1. SSID Access:

   As a by-product of this work, and during the discovery phases, we found that an app can gain access to previously attached SSID information, without any permissions required. This is surprising as this data may be able to identify previous locations and allow for tracking of users, all without requesting dangerous location permissions. This is problematic in terms of privacy and security for the user.

2. Devices in the HN:

In addition, and with no dangerous permissions requested (Wifi/Networking only required), we were able to enumerate devices residing on these HNs. We were also surprised that information about 3rd party apps and their permissions are so easily accessible and also do not require dangerous or restricted permissions to gain this access. it is clear that the privacy and security are the major implications for both of these areas.

3. Apps and Permissions:

   3rd party apps installed can gather sensitive information about apps, sets of permissions, devices, and potential (previous) locations. In addition, there is the concern around determination of suggestive habits and trend tracking of the user. These implications and concerns are similar to those of tracking users via web browsing history, configurations / settings, as well as other methods, and fall into exposing sensitive and potentially (leaking) dangerous information about users. These are all centered around user privacy and security concerns and the gathering and potentially exposing of sensitive data for tracking.

   In addition, users reported a 55% preference in leveraging a Mobile app for HN Ecosystem data collections and reporting. This lower majority percentage could be in part due to the implications found as part of this research, specifically as related to privacy and security concerns.

While these features of the Android framework allow developers to create a wide range of applications on an open platform, they are cause for alarm to users in terms of privacy and security. These privacy and security concerns may not just be relegated exclusively to the Android OS. As an example the iOS (iPhone) permissions framework may have similar and potentially dangerous implications around user privacy and security issues. Research done by Check Point [75] found that the IoS platform exposes Apple's

Contacts app to allow the device to exposing information and run malicious code. A review of IoS Permissions, similar to the Android platform, is available at [76].

While these manufactures' security models differ in terms of programmatic and top level usage, the framework hold similar in terms of granting and revoking of permissions, however the IoS framework does appear to have a stronger coupling to dangerous permissions and restriction of base functionality requirements.

5.12 Summary

We have created a minimal impediment and high incentive Mobile app, which requires minimal permissions to measure privacy, security, and general health of HNs. The results shown in this work generally supports that using a Mobile app is advantageous and preferred by users to gather data, and provides a method that is capable of providing information HN users are interested in understanding over other approach types. The data and responses collected from this work points to value in information that HN users yearn for in terms of security, privacy and general health of operations.

We have obtained a significant amount of data from the distribution of the HMN Mobile app and have determined a wide range of information about HNs, which have not been available or reviewed in the past; including app health and Network health. These areas, posed with device input and response data and feedback generally support and point to an app-based methodology providing valuable information for users, while still providing a robust dataset for researchers in terms of both research and methodologies. We have shown that the set of popular apps (installed the most) tend to require more dangerous permissions (e.g. storage, location, camera, contacts, etc.), and can be problematic in terms of privacy and security.

An app method works well for a large swath of usage (device types) in a HN as

well as collection of user feedback around privacy, security, and HN information. These include gathering of permission types, and devices available and access times, changes to apps, overview of app and permission levels to the user, and a health rating system from system operations. The Health (apps and Wifi) rating methodology we created was well received, as was the Throughput graphing information. However, there are limitations to this approach in terms of data collection, including: network sniffing and flow-control, precise device mapping, and other areas of network access and application support. In addition, dissemination of the data needs further study, as we did not dig into different styles of displaying data to the user and there are several methods that can be used (e.g. UI/UX approaches).

We believe that this work is instrumental in providing the research community a view of what a Mobile app is capable of gathering characteristics of in HNs, and more importantly that the data collection is indeed valuable information areas (e.g. privacy and security). The areas of Privacy and Security of interest includes changes to the set of apps (e.g. permissions), apps being uninstalled, as well as new devices being added or removed from a HN. We found that 32% of users un-installing apps, 24% de-authorizing permissions, and 16% having no longer visible devices in their HN, which generally supports that users are concerned with Privacy and Security in their HN.

The development of the local and global norms as well as the health rating system we have developed are flexible and can be applied to other areas of study. In this study we used these areas to point toward a methodology to compare HNs, devices, and tools, as well as privacy and security issues. We also believe that the calculation of the digital fingerprint (changes) in the Mobile app, for quickly detecting changes, helps point to user actions/re-actions (e.g. uninstalls/no longer visible, report finding new devices or apps, etc.), and demonstrates providing useful information on what types of methods users are interested in privacy and security, and the overall health of HNs.

As part of this overarching work, we have successfully evaluated and leveraged a Mobile app for the gathering of data, in and around HNs, as a low impediment and high incentive tool to provide and report results to HN users. We used an integrated approach to gathering HN data that also includes reporting point-in-time feedback directly within the app on a variety of HN focus areas (e.g. security and privacy concerns, Wifi and Internet speeds, device Information, Health operations of their HNs, etc.). We have analyzed user feedback and found that the compromise of approaches and leveraging a Mobile app works well for both data collection and dissemination of data. We have found that combining the areas of apps and devices into an association that is homogeneous has been successful in terms of both feedback and determination of features of a HN.

The following are takeaways from this work:

1. demonstrate that a Mobile app is valuable platform to collect data from the HN Ecosystem (e.g. devices, apps, privacy and security, as well as other information) and can display this HN Ecosystem information in an functional manner to HN users, with minimal impediments and a high level of incentives for participation;

2. show that users react to privacy and security concerns, based off of HN Ecosystem results, by making changes to their HN Ecosystem (e.g. apps, permissions, devices, as well as other information);

3. show that HN users are interested in understanding the operational Health status of their HN Ecosystem as well as comparisons to other HNs; and

4. show that a Mobile app can expose privacy and security issues of the HN Ecosystem (e.g. apps, permissions, connectivity, behaviors).

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this chapter we provide a summary of the work we have completed, research contributions, conclusions of the dissertation, and future work. We start first with a review of the dissertation work, and move to research questions posed, and provide details and conclusions around these questions in our hypothesis. We conclude and provide future work as part of this chapter.

This dissertation provides several contributions in and around the HN Ecosystem space, including device preferences, health of operations, and detection of changes HN users made based off of potential privacy/security concerns. We have successfully evaluated a Java applet and Mobile app for the gathering of data in the HN Ecosystem and created an integrated approach of Data Collections and User Feedback; reporting point-in-time results directly within an app (e.g. security and privacy concerns, Wifi and Internet speeds, device information, Health of operations, etc.). We found that the Java applet approach had minimal impediments to users for participation, but may not provide a robust data collection dataset and results to foster participation. We next leveraged a survey, to

glean incentives, and after analyzing user feedback we generally found that a Mobile app is a both a good trade-off and the best platform for the collection of attributes and reporting of information (high incentives) in the HN Ecosystem. We then built and distributed a Mobile app for distribution and data collections, and found that this approach was generally advantageous both for the collection of a rich dataset while minimizing impediments and providing high incentives for participation. We have found that combining the areas of devices and apps into an association that is homogeneous and part of the HN Ecosystem has been successful in terms of both feedback and determination of features of a HN. In general we have found that users have modified their HN Ecosystem when presented with results from their HN and norms of other HNs. The data collected supports that users are concerned with Privacy and Security in their HN, which was evident from the change in perception via feedback.

In addition, we have found research implications around privacy and security concerns of using installed app. These research implications have exposed that with minimal or no access required/approved by the governing framework apps have access to sensitive information such as: previously attached SSIDs, apps and permissions, and network scanning. These implications are a concern as an app can be used to determine factors of suggestive habits (e.g. proclivities) and trend tracking of the users. These implications and concerns center around the possibility of exposing sensitive and potentially dangerous information about users. These are all centered around user privacy and security concerns and the gathering and exposing of sensitive data (e.g. tracking, and habits).

The following is evidence that points toward conclusions from results of the research questions in this dissertation:

1. *Can a Java Applet be an effective, low impediment, approach for the collection of meaningful research data and results?*

   We found that this work suggest that a Java approach is an effective low impediment

approach and allows for a reasonable collection of data and results. While the collection and results are not overwhelming, in-comparison to other approach types, they did provide value and overall worthwhile methodology.

2. *Are HN users interested in understanding privacy (e.g. data leaks) and security concerns (e.g. unauthorized devices, detection) in their HN Ecosystem?*

   We found that the results of the HMN survey generally supported that users, across all skill levels, had interest in understanding HN specifics of the HN Ecosystem, including attributes such as Wifi Speed, Internet throughput, device security/privacy, device detection, device and HN Health, and Mobile changes (e.g. apps, permissions and similar). We also found that this generally held true from the data, results, and feedback from the HMN Mobile app, where 84% of users reported that they are more security conscious. This also held true for the detection of changes to the HN Ecosystem, Privacy and Security of Apps/permissions, Norms of results and comparisons locally and globally to other HNs.

3. *Does a Mobile app provide the trade-offs users are looking for in data collection, results and range of data, and ease of operation?*:

   Based on user our survey and feedback we found general support that the HMN Mobile app did impact users perception of HNs and provided the trade-offs users are looking for. We found that a majority of users prefer to use a mobile device to gather HN Ecosystem information and provided positive feedback on leveraging an app and the results, across the testing. The privacy and security implications found in this dissertation may contribute to the hesitance of selecting a Mobile app approach type for data collection and thus the reporting of a higher percentage for this approach type.

4. *Will users modify their set of apps, permissions, and devices when presented with*

potential privacy and security as well as health-rated results:

During the lifespan of the HMN Mobile app we found that 24% of all users revoked one or more permissions, and 32% uninstalled an app during testing. Based on HMN Mobile feedback 84% reported that they are more security conscious, and 46% reported a change in the way they use their devices, apps, or HN Ecosystem because of using the HMN Mobile app. These results generally support that detection of changes has helped point to user actions/re-actions (e.g. Uninstalls of devices/apps, finding new devices or apps, and permission changes, etc.)

In conclusion, we have continued our HMN measurement environment leveraging first a Java applet and ultimately a Mobile app (HMN Mobile App) to allow for an expand and robust test suite of high incentives and low impediments. The measurement platform has been able to elicit data from a HN ranging from basic information to network related classes. Using this approach we have offered a high level of incentives required by the end user and yet still offer the needed data gathering vital to measurement platforms. We have shown that an approach type of integrating data collection with user feedback is helpful to gathering HN Ecosystem attributes. In summary, the results of this study suggest that users have an interests in understanding their HN environment in depth, and that they will make changes based off of privacy and security concerns, health of operations provided, and results of devices, apps and permissions.

6.2 Future Work

In future work we look to add several features for discovery, reporting, and tips for end users around remediation and improvement of services, along with combining approaches in the HMN work. We have broken the future work into two areas: immediate (continuing with the current HMN app work) and bigger picture/larger scale efforts that would be new

research.

6.2.1 Immediate

We look to include the following immediate work leveraging the current HMN Mobile app framework. We look to add in features to the HMN Mobile app that allow for a wider view of Privacy and Security by tying together virus protection and vulnerability signatures with a simplified front-end for techniques of defense and blocking aimed at both devices and apps. We also look to add in additional results and information around real-time monitoring and functionality to allow for simple and complex rules to be processed across apps, networking, and devices. Additions of detailed information via the UI and an updated UX design to allow for more details and information into specific areas, including: health of apps, devices, and Network. These additions look to allow users to drill down into specific details, and more importantly provide beneficial methods for optimizing services. We also look toward adding an in-app options to allow for the scanning of the device and throughput when not connected to the preferred Wifi and HN.

To allow for a deeper study of this space we look to add the following methodology for data collections, this approach would cover devices and apps. The app would be deployed in two parts, with participants agreeing ahead of time to complete both. The app would gather identical types of data for the same amount of time, but it first would be blinded to the user so that any actions they take regarding install/uninstall of other apps would not be expected to be the result of data they have gleaned from our app. The app deployed would then be unblinded, with users being able to access and understand the privacy/security concerns of each other app they have installed. In this way actions dependent upon knowledge gleaned from our app could be parsed from random actions.

In addition, approaches that provide auto-repair and recommendation options for security, privacy, and health are proposed. Combining user expertise (or lack thereof) and

experiences into the feedback to allow for cross app help, and optimization of environments. While the feedback was useful we feel that in a larger scale this would get dropped and not be as pivotal in the bigger picture of the work. This would also allow for collection of how users are leveraging the app, and will help to infer if users indeed making changes, learning or other behavioural results from using the HMN app. This future work would help with understanding user data patterns as well as the provide the ability of the app to offer advice, to the overall user-base, in areas such as privacy/security, app usage, and devices in their HN and device.

6.2.2 Bigger Picture

In bigger picture of future work we look to focus in on several areas of research including: creating a framework or safeguards for permissions of apps, and research into combination of approaches to allow for a robust set of results for HN users, these include:

1. Create an operational framework for permissions of apps as a starting point for both Google and Apple apps. Review popular apps residing in the Play stores (Google and Apple) to determine permission sets. Leverage this framework and permission set for operational status or safeguards within each of these constructs to permissions needed to operate versus those permissions that are requested, comparing Play store methods and permissions. We believe that this will allow for a tighter privacy and security model and thus create greater protections for the HN Ecosystem.

2. Use a combination of approaches (e.g. Mobile and customized hardware) for research of real time protections and services, while enforcing the requirements of minimizing impediments and extra incentives of information and results to the user. This framework would require an IoT or inline device that is plug-and-play for the HN user to operate. Although this is a slight deviation from an exclusive Mobile

app combining approach may provide value, in terms of results, and comprise, in terms of preferences, for HN Ecosystem.

3. Understanding how other apps make use of techniques employed to gather sensitive or private information.

4. Adding in HN measurements that can leverage approaches with embedded points of discovery that are installed in backbone vendors (e.g. Google, Akamai, and others). Adding these partnering and allow for remote agents we look to gather both high and low level external testing of network characteristics.

Bibliography

[1] *3 out of 4 mobile apps downloaded by consumers last year have vulnerabilities that could let hackers steal your passwords and other sensitive data*. https://amp.businessinsider.com/iphone-and-android-apps-have-security-weaknesses-from-simple-oversight-2019-6. [Accessed June 2019].

[2] *A New Threat: Stalkerware*. http://thedataist.com/a-new-threat-stalkerware/. [Accessed June 2019].

[3] Abbas Acar et al. "Peek-a-Boo: I see your smart home activities, even encrypted!" In: *arXiv preprint arXiv:1808.02741* (2018).

[4] *Alexa Top 500 Sites*. http://www.alexa.com/site/ds/top_sites.

[5] Alibba. *Alibaba Mobile Security*. https://play.google.com/store/apps/details?id=com.eset.ems2.gp&hl=en. [] 2015.

[6] Hazim Almuhimedi et al. "Your location has been shared 5,398 times!: A field study on mobile app privacy nudging". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM. 2015, pp. 787–796.

[7] *Android Dangerous Permissions*. https://developer.android.com/guide/topics/permissions/overview. [Accessed Nov 2019].

[8] *Android Market Share - 2019*. http://gs.statcounter.com/os-market-share/mobile/worldwide. []

[9] *Android Permissions*. https://developer.android.com/guide/topics/permissions/overview.html#normal-dangerousl. [Accessed June 2019].

[10] Frank Andrus. "Beyond scan and block: an adaptive approach to network access control". In: *Network Security* 2011.11 (2011), pp. 5–9.

[11] *App Fraud*. https://www.buzzfeednews.com/amphtml/craigsilverman/how-a-massive-ad-fraud-scheme-exploited-android-phones-to. [Buzzfeed Article].

[12] *App Wifi Passwords*. https://gizmodo.com/if-you-used-this-app-the-password-of-the-private-wifi-1834217718/amp. [Gizmodo Article].

[13] *Application and Software*. https://www.techopedia.com/definition/4224/application-software. Accessed: 2018-03-17.

[14] Abdullahi Arabo, Ian Brown, and Fadi El-Moussa. "Privacy in the age of mobility and smart devices in smart homes". In: *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE. 2012, pp. 819–826.

[15] *Archipelago Measurement Infrastructure*. http://www.caida.org/projects/ark/. [Web Site].

[16] Andres Arcia-Moret et al. "Intelligent network discovery for next generation community wireless networks". In: *Wireless On-demand Network Systems and Services (WONS), 2016 12th Annual Conference on*. IEEE. 2016, pp. 1–7.

[17] Vaibhav Bajpai and Jrgen Schonwlder. "A survey on internet performance measurement platforms and related standardization efforts". In: *IEEE Communications Surveys & Tutorials* 17.3 (2015), pp. 1313–1341.

[18] Andy Bavier et al. "Operating System Support for Planetary-Scale Network Services". In: *USENIX Symposium on NSDI*. San Francisco, CA, USA, Mar. 2004.

[19] *BroadBand Usage 2018*. https://www.dailywireless.org/internet/usage-statistics/. [Accessed Nov 2019].

[20] BullGuard. *BullGuard Mobile Security*. https://play.google.com/store/apps/details?id=com.bullguard.mobile.mobilesecurity&hl=en. [] 2015.

[21] Michael Butkiewicz et al. "Klotski: Reprioritizing Web Content to Improve User Experience on Mobile Devices." In: *NSDI*. Vol. 1. 1. 2015, pp. 2–3.

[22] Kenneth L Calvert et al. "Instrumenting home networks". In: *ACM SIGCOMM Computer Communication Review* 41.1 (2011), pp. 84–89.

[23] Daniel Camps-Mur et al. "Enabling always on service discovery: Wifi neighbor awareness networking". In: *IEEE Wireless Communications* 22.2 (2015), pp. 118–125.

[24] Jacob Carlsson. *Comparison in functionality between a closed and two open source distributions in a router*. 2016.

[25] Martin Casado et al. "Opportunistic Measurement: Extracting Insight from Spurious Traffic". In: *Proceedings of the Fourth Workshop on Hot Topics in Networks*. College Park, MD USA, Nov. 2005.

[26] Ranveer Chandra, Christof Fetzer, and Karin Hogstedt. *Adaptive topology discovery in communication networks*. US Patent 7,366,113. Apr. 2008.

[27] *Cheetah Mobile Clean Master*. http://www.cmcm.com/en-us/clean-master/. Accessed: 2018-03-17.

[28] Marshini Chetty and Nick Feamster. “Refactoring network infrastructure to improve manageability: a case study of home networking”. In: *ACM SIGCOMM Computer Communication Review* 42.3 (2012), pp. 54–61.

[29] Marshini Chetty et al. “Why is my internet slow?: making network speeds visible”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2011, pp. 1889–1898.

[30] Luca Chittaro. “Visualizing information on mobile devices”. In: *Computer* 39.3 (2006), pp. 40–45.

[31] *Cisco Discovery Protocol*. https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/cdp/configuration/15-mt/cdp-15-mt-book/nm-cdp-discover.html. Accessed: 2018-03-17.

[32] *Comcast Devices On Network*. https://play.google.com/store/apps/details?id=com.comcast.cvs.android. Accessed: 2019-05-06.

[33] *Competencies Proficiency Scale*. https://hr.nih.gov/working-nih/competencies/competencies-proficiency-scale. Accessed: 2018-08-20.

[34] Crypt4All. *Sophos Mobile Security*. https://play.google.com/store/apps/details?id=com.codewell4.Crypt4AllLite&hl=en. [] 2015.

[35] Anita D’Amico et al. “Integrating physical and cyber security resources to detect wireless threats to critical infrastructure”. In: *Technologies for Homeland Security (HST), 2011 IEEE International Conference on*. IEEE. 2011, pp. 494–500.

[36] *DD-WRT*. https://www.dd-wrt.com/site/index. Accessed: 2018-01-05.

[37] *Dell Battery Ratings*. ttps://www.dell.com/support/article/us/en/04/sln143156/checking-battery-health-status-on-dell-laptops-and-notebooks?lang=en. [Dell Battery Ratings].

[38] Lucas DiCioccio, Renata Teixeira, and Catherine Rosenberg. “Measuring home networks with homenet profiler”. In: *International Conference on Passive and Active Network Measurement*. Springer. 2013, pp. 176–186.

[39] *Digital Inequality and Low-Income Households*. https://www.huduser.gov/portal/periodicals/em/fall16/highlight2.html. []

[40] Colin Dixon et al. “An operating system for the home”. In: *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*. USENIX Association. 2012, pp. 25–25.

[41] Colin Dixon et al. “The home needs an operating system (and an app store)”. In: *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*. ACM. 2010, p. 18.

[42] *Dojo hardware security and privacy device*. https://dojo.bullguard.com/. [] 2008.

[43] Jack Poller Doug Cahill. *An Adaptive and Layered Approach to Endpoint Security*. https://www.bitdefender.co.th/media/wysiwyg/gravityzone/elite-security/ESG-White-Paper-Bitdefender-Jun-2017.pdf. [DYN Networks]. 2017.

[44] Jack Poller Doug Cahill. *An Adaptive and Layered Approach to Endpoint Security, June 2017, ESG White Paper*. https://www.bitdefender.co.th/media/wysiwyg/gravityzone/elite-security/ESG-White-Paper-Bitdefender-Jun-2017.pdf. [] 2008.

[45] Jack Poller Doug Cahill. *Cujo hardware security monitoring device*. www.getcujo.com. [] 2008.

[46] *Broadband Reports.com Speed Test*. http://www.dslreports.com/stest.

[47] DYN. *Measuring DNS Performance with Open Recursive Name Servers*. https://dyn.com/blog/dns-performance/. [DYN Networks]. 2015.

[48] *EC2 Amazon Ratings*. https://aws.amazon.com/premiumsupport/faqs/. [EC2].

[49] W Keith Edwards et al. "Advancing the state of home networking". In: *Communications of the ACM* 54.6 (2011), pp. 62–71.

[50] Diana Joumblatt and Renata Teixeira. *End-Host Measurement Survey, Web based survey*. 2009.

[51] Grant M Erickson et al. *Efficient communication for devices of a home network*. US Patent 9,629,193. 2017.

[52] Dave Evans. *The Internet of Things. How the Next Evolution of the Internet Is Changing Everything*. https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf. US Patent 9,872,240. 2011.

[53] Adrienne Porter Felt et al. "Android permissions: User attention, comprehension, and behavior". In: *Proceedings of the eighth symposium on usable privacy and security*. ACM. 2012, p. 3.

[54] *Fing-Network Tools*. https://www.fing.io/. Accessed: 2018-03-17.

[55] Marcel Flores, Alexander Wenzel, and Aleksandar Kuzmanovic. "Oak: User-Targeted Web Performance". In: *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on*. IEEE. 2017, pp. 2654–2655.

[56] *Flying Squirrel is the Department of Defense standard wireless discovery and mapping application. It runs on Windows and Linux using commercial laptops, wireless cards, and GPS devices.* https://www.nrl.navy.mil/itd/chacs/5545/flying-squirrel/FAQ. Accessed: 2018-03-17.

[57] *Gomez Peer Community*. http://www.porivo.com/.

[58] *Google Android Version Issue Tracker*. https://issuetracker.google.com/issues/63125228. [Accessed June 2019].

[59] *Google Permissions*. https://developer.android.com/guide/topics/permissions/overview#normal-dangerous. []

[60] Olena Hrebeshkova Grebeshkov and Oleksandr Vostryakov. "Surveying digital competencies of university students and professors in Ukraine for fully online collaborative learning". In: (2017).

[61] Rebecca E Grinter et al. "The ins and outs of home networking: The case for useful and usable domestic networking". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 16.2 (2009), p. 8.

[62] Rebecca E Grinter et al. "The work to make a home network work". In: *ECSCW 2005*. Springer. 2005, pp. 469–488.

[63] Lucas Guardalben et al. "A cooperative hide and seek discovery over in network management". In: *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*. IEEE. 2010, pp. 217–224.

[64] Michael Hall and Raj Jain. "Performance analysis of openvpn on a consumer grade router". In: *cse. wustl. edu* (2008).

[65] Stephan Heuser et al. "{ASM}: A Programmable Interface for Extending Android Security". In: *23rd {USENIX} Security Symposium ({USENIX} Security 14)*. 2014, pp. 1005–1019.

[66] *Home Broadband Speeds*. https://broadbandnow.com/guides/how-much-internet-speed-do-i-need. Accessed: 2018-12-05.

[67] *Home Networking Devices Glossary - Most Common Devices in your Home Network Setup*. https://www.fing.com/news/what-devices-in-home-network-setup. [Accessed June 2019].

[68] Hadrien Hours et al. "A study of the impact of DNS resolvers on CDN performance using a causal approach". In: *Computer Networks* 109 (2016), pp. 200–210.

[69] *How-to Guide for Classifying Devices - ForeScout*. https://www.forescout.com/company/resources/classify-devices-how-to-guide-8-0/. [Accessed June 2019].

[70] *How's My Network Mobile App*. https://play.google.com/store/apps/details?id=wpi.howsmynetwork. [HMN Mobile App].

[71] Young Hyun. "The Archipelago Measurement Infrastructure". In: *Proceedings of the CAIDA-WIDE Workshop*. San Diego, CA, USA, Nov. 2006.

[72] *IEEE OUI Database*. http://standards-oui.ieee.org/oui.txt. [Accessed June 2019].

[73] *Inside the 'Stalkerware' Surveillance Market, Where Ordinary People Tap Each Other's Phones*. https://www.vice.com/en_us/article/53vm7n/inside-stalkerware-surveillance-market-flexispy-retina-x. [Accessed June 2019].

[74] *Internet Speed Meter Lite*. https://play.google.com/store/apps/details?id=com.internet.speed.meter.lite&hl=en. [] 2008.

[75] *IOS Check Point*. https://appleinsider.com/articles/19/08/10/apples-ios-contacts-app-claimed-to-be-vulnerable-to-sqlite-hack. [Accessed Nov 2019].

[76] *IOS Permissions*. https://stackoverflow.com/questions/29894749/complete-list-of-ios-app-permissions. [Accessed Nov 2019].

[77] *IoT Inspector: Studying Smart Home IoT Device Behavior*. https://freedom-to-tinker.com/2018/04/23/announcing-iot-inspector-a-tool-to-study-smart-home-iot-device-behavior/. [Blog Post].

[78] Artur Janc, Craig E. Wills, and Mark Claypool. *Network Performance Evaluation within the Web Browser Sandbox*. Submitted for publication. Feb. 2009.

[79] Yaoqi Jia et al. "I know where you've been: Geo-inference attacks via the browser cache". In: *IEEE Internet Computing* 19.1 (2014), pp. 44–53.

[80] Diana Joumblatt et al. "HostView: Annotating end-host performance measurements with user feedback". In: *ACM SIGMETRICS Performance Evaluation Review* 38.3 (2011), pp. 43–48.

[81] Diana Joumblatt et al. "Peeking without Spying: Collecting End-Host Measurements to Improve User Experience". In: ().

[82] Jaeyeon Jung, Seungyeop Han, and David Wetherall. "Short paper: enhancing mobile application permissions with runtime feedback and constraints". In: *Proceedings of the second ACM workshop on Security and privacy in smartphones and mobile devices*. ACM. 2012, pp. 45–50.

[83] Jaeyeon Jung et al. "DNS performance and the effectiveness of caching". In: *IEEE/ACM Transactions on networking* 10.5 (2002), pp. 589–603.

[84] Murad Kaplan et al. "How's My Network? Predicting performance from within a Web browser sandbox". In: *Local Computer Networks (LCN), 2012 IEEE 37th Conference on*. IEEE. 2012, pp. 521–528.

[85] Zolidah Kasiran and Juliza Mohamad. "Throughput performance analysis of the wormhole and sybil attack in AODV". In: *Digital Information and Communication Technology and it's Applications (DICTAP), 2014 Fourth International Conference on*. IEEE. 2014, pp. 81–84.

[86] *Kaspersky Lab Internet Security*. https://usa.kaspersky.com/internet-security. Accessed: 2018-03-17.

[87] kc claffy kc et al. “Community-oriented network measurement infrastructure (CONMI) report”. In: *SIGCOMM CCR* 36.2 (2006). ISSN: 0146-4833. DOI: http://doi.acm.org/10.1145/1129582.1129594.

[88] *Keezel hardware security and privacy device*. https://keezel.co/. [] 2008.

[89] Athar Ali Khan, Mubashir Husain Rehmani, and Yasir Saleem. “Neighbor discovery in traditional wireless networks and cognitive radio networks: Basics, taxonomy, challenges and future research directions”. In: *Journal of Network and Computer Applications* 52 (2015), pp. 173–190.

[90] Hyojoon Kim and Nick Feamster. “Improving network management with software defined networking”. In: *IEEE Communications Magazine* 51.2 (2013), pp. 114–119.

[91] Ryan Yong Kim and Venkata Subba Rao Pathuri. *Setup of multiple IOT devices*. US Patent 9,210,192. 2015.

[92] Ryan Yong Kim and EuChong Son. *Network device source entity triggered device configuration setup*. US Patent 9,872,240. 2018.

[93] Ryan Yong Kim and Ohad Zeira. *Setup of multiple iot network devices*. US Patent App. 15/363,051. 2017.

[94] *Kismet Wireless Scanner*. https://www.kismetwireless.net/android-pcap/. Accessed: 2018-03-17.

[95] Jan Willem Kleinrouweler. “Enhancing over-the-top video streaming quality with DASH assisting network elements”. In: *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM. 2017, pp. 113–116.

[96] Jan Willem Kleinrouweler, Britta Meixner, and Pablo Cesar. “Improving Video Quality in Crowded Networks Using a DANE”. In: *Proceedings of the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video*. ACM. 2017, pp. 73–78.

[97] Christian Kreibich et al. “Netalyzr: illuminating the edge network”. In: *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM. 2010, pp. 246–259.

[98] *LastPass*. https://lastpass.com/. Accessed: 2018-03-17.

[99] *LEEDE project, 2014*. https://www.bufferbloat.net/projects/cerowrt/wiki/. Accessed: 2018-03-17.

[100] Yadong Li et al. “Research based on osi model”. In: *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*. IEEE. 2011, pp. 554–557.

[101] Bin Liu, Jialiu Lin, and Norman Sadeh. “Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help?” In: *Proceedings of the 23rd international conference on World wide web*. ACM. 2014, pp. 201–212.

[102] *MAC Address Block Large (MA-L)*. https://standards.ieee.org/products-services/regauth/oui/index.htm. [Accessed June 2019].

[103] *Maps and Location, Free*. https://www.mapcustomizer.com/. Accessed: 2019-04-28.

[104] Stuart McIlroy, Nasir Ali, and Ahmed E Hassan. “Fresh apps: an empirical study of frequently-updated mobile apps in the Google play store”. In: *Empirical Software Engineering* 21.3 (2016), pp. 1346–1370.

[105] *Meshlium IoT Gateway Hardware*. http://www.libelium.com/products/meshlium/. Accessed: 2018-03-17.

[106] Miriah Meyer, Michael Sedlmair, and Tamara Munzner. “The four-level nested model revisited: blocks and guidelines”. In: *Proceedings of the 2012 BELIV Workshop: Beyond Time and Errors-Novel Evaluation Methods for Visualization*. ACM. 2012, p. 11.

[107] Miriah Meyer et al. “The nested blocks and guidelines model”. In: *Information Visualization* 14.3 (2015), pp. 234–249.

[108] *MPAndroidChart*. https://github.com/PhilJay/MPAndroidChart. [Github P.Jahoda Opensource tool].

[109] *Mtoolbox web port scanner*. mxtoolbox.com. [] 2008.

[110] Hendrik Muller and Aaron Sedley. “HaTS: large-scale in-product measurement of user attitudes & experiences with happiness tracking surveys”. In: *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*. ACM. 2014, pp. 308–315.

[111] Tamara Munzner. “A nested model for visualization design and validation”. In: *IEEE transactions on visualization and computer graphics* 15.6 (2009), pp. 921–928.

[112] *Net Mapper*. https://play.google.com/store/apps/details?id=com.wwnd.netmapper&h. Accessed: 2018-03-17.

[113] *Netalyzr App*. https://play.google.com/store/apps/details?id=edu.berkeley.icsi.netalyzr.android&hl=en_US. [Netalyzr App is now defunct, March 2019].

[114] *NetMapper App*. https://play.google.com/store/apps/details?id=com.wwnd.netmapper&h. Accessed: 2018-03-17.

[115] *Nmap - Free Security Scanner for Network Exploration & Security Audits*. http://nmap.org.

[116] *Norton Mobile Security 3.15*. http://norton.com. Accessed: 2018-03-17.

[117] Mihai Novitchi. *Anti-malware emulation systems and methods*. US Patent 8,407,797. Mar. 2013.

[118] *NYC Secure*. https://play.google.com/store/apps/details?id=gov.nyc.zips. Accessed: 2019-05-07.

[119] Babatunde Olabenjo. "Applying Naive Bayes Classification to Google Play Apps Categorization". In: *arXiv preprint arXiv:1608.08574* (2016).

[120] Lukasz Olejnik, Claude Castelluccia, and Artur Janc. "On the uniqueness of web browsing history patterns". In: *annals of telecommunications-annales des télécommunications* 69.1-2 (2014), pp. 63–74.

[121] Ookla. *OoklaSpeedtest*. https://play.google.com/store/apps/details?id=org.zwanoo.android.speedtest&hl=en. [] 2014.

[122] Orbot. *Orbot Proxy*. https://play.google.com/store/apps/details?id=org.torproject.android&hl=en. [] 2014.

[123] Yossef Oren et al. "The spy in the sandbox: Practical cache attacks in javascript and their implications". In: *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM. 2015, pp. 1406–1418.

[124] Ashish Patro, Srinivas Govindan, and Suman Banerjee. "Outsourcing home AP management to the cloud through an open API". In: *Traffic (in Mbps)* 2 (2013), p. 4.

[125] *PCAP for Android*. https://play.google.com/store/apps/details?id=jp.co.taosoftware.android.packetcapture. Accessed: 2018-03-17.

[126] Changhua Pei et al. "How Much Are Your Neighbors Interfering with Your WiFi Delay?" In: *Computer Communication and Networks (ICCCN), 2017 26th International Conference on*. IEEE. 2017, pp. 1–9.

[127] Changhua Pei et al. "Why it takes so long to connect to a WiFi access point". In: *INFOCOM 2017-IEEE Conference on Computer Communications, IEEE*. IEEE. 2017, pp. 1–9.

[128] Larry Peterson et al. "Experiences Building PlanetLab". In: *Proceedings of the 7th USENIX SOSDI*. Seattle, WA USA, Nov. 2006.

[129] *PEW Home Networking Review 2017*. http://www.pewinternet.org/fact-sheet/internet-broadband/. Accessed: 2018-01-05.

[130] Erika Shehan Poole et al. "Computer help at home: methods and motivations for informal technical support". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2009, pp. 739–748.

[131] Erika Shehan Poole et al. "More than meets the eye: transforming the user experience of home network management". In: *Proceedings of the 7th ACM conference on Designing interactive systems*. ACM. 2008, pp. 455–464.

[132] *Popular Devices 2018*. https://www.postscapes.com/internet-of-things-award/winners/. Accessed: 2018-11-04.

[133] R Kelly Rainer et al. *Introduction to information systems*. John Wiley & Sons, 2013.

[134] Dave Randall. "Living inside a smart home: A case study". In: *Inside the smart home*. Springer, 2003, pp. 227–246.

[135] *RaTTrap hardware security and privacy device*. www.myrattrap.com. [] 2008.

[136] Y. Rekhter et al. *Address Allocation for Private Internets*. RFC 1918. Feb. 1996.

[137] Alan Ritacco and Craig Wills. "Peering into the Home Network". In: (). URL: https://hmn.cs.wpi.edu/.

[138] Alan Ritacco and Craig Wills. "Survey of Home Networks". In: (). URL: https://hmn.cs.wpi.edu/.

[139] Alan Ritacco, Craig Wills, and Mark Claypool. "How's My Network? - A Java Approach to Home Network Measurement". In: *Proceedings of the 18th International Conference on Computer Communications and Networks (ICCCN)*. San-Francisco, CA, USA: IEEE, 2009. DOI: \small\verb$http://www.cs.wpi.edu/~claypool/papers/hmn-java/$.

[140] Alan Ritacco, Craig Wills, and Mark Claypool. "How's My Network? A Java Approach to Home Network Measurement". In: *Computer Communications and Networks, 2009. ICCCN 2009. Proceedings of 18th Internatonal Conference on*. IEEE. 2009, pp. 1–7.

[141] *Router Interrogation. Colby Lahaie, David Paradise. Research Project work done as part of the Senator Patrick Leahy Center for Digital Investigation, Champlain College, 2013*. https://www.champlain.edu/Documents/LCDI/archive/POST_ME_Router-Interrogation-ReportPDF.pdf&usg=AOvVaw2tHbiS9vxwV3saq2tvq7Qa. Accessed: 2018-03-17.

[142] Anique Scheerder, Alexander van Deursen, and Jan van Dijk. "Determinants of Internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide". In: *Telematics and Informatics* (2017).

[143] Henning Schulzrine. "The New Internet – Goals, Testing and Infrastructural Needs". In: *Breakout Session of the NSF Computing Research Infrastructure PI Meeting*. Boston, MA USA, June 2007.

[144] *Security and Apps on Android*. http://resources.infosecinstitute.com/security-hacking-apps-android/. Accessed: 2018-01-05.

[145] *Security and Apps on Android*. http://resources.infosecinstitute.com/security-hacking-apps-android/. Accessed: 2018-03-17.

[146] Yuval Shavitt and Eran Shir. *DIMES: Let the Internet Measure Itself*. http://www.arxiv.org/abs/cs/0506099v1. 2005.

[147] Erika Shehan and W Keith Edwards. “Home networking and HCI: What hath God wrought?” In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2007, pp. 547–556.

[148] Anna Kornfeld Simpson, Franziska Roesner, and Tadayoshi Kohno. “Securing vulnerable home IoT devices with an in-hub security manager”. In: *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*. IEEE. 2017, pp. 551–556.

[149] Charles Robert Simpson Jr., Dheeraj Reddy, and George F. Riley. “Empirical Models of TCP and UDP End-User Network Traffic From NETI@home Data Analysis”. In: *In PADS*. Singapore, May 2006, pp. 166–174.

[150] Ramesh K Sitaraman. “Network performance: Does it really matter to users and by how much?” In: *Communication Systems and Networks (COMSNETS), 2013 Fifth International Conference on*. IEEE. 2013, pp. 1–10.

[151] *Sketch Apps*. `https://lifehacker.com/delete-these-sketchy-android-apps-that-are-tracking-you-1834148357/amp`. [LifeHacker Article].

[152] Stephen Smalley and Robert Craig. “Security Enhanced (SE) Android: Bringing Flexible MAC to Android.” In: *Ndss*. Vol. 310. 2013, pp. 20–38.

[153] Sophos. *Sophos Mobile Security*. `https://play.google.com/store/apps/details?id=com.sophos.smsec&hl=en`. [] 2015.

[154] *Speedtest.net*. `http://www.speedtest.net/`.

[155] *Spiceworks Scanner*. `community.spiceworks.com/tools/port-scan?blog_2484`. [] 2008.

[156] Neil Spring et al. “Using PlanetLab for network research: myths, realities, and best practices”. In: *SIGOPS Oper. Syst. Rev.* 40.1 (2006), pp. 17–24. ISSN: 0163-5980. DOI: `http://doi.acm.org/10.1145/1113361.1113368`.

[157] Srikanth Sundaresan, Nick Feamster, and Renata Teixeira. “Measuring the performance of user traffic in home wireless networks”. In: *International Conference on Passive and Active Network Measurement*. Springer. 2015, pp. 305–317.

[158] Hengky Susanto. “Congestion control with QoS through network utility maximization”. PhD thesis. University of Massachusetts Lowell, 2014.

[159] Hengky Susanto, Byung Guk Kim, and Benyuan Liu. “User Tolerance and Self-Regulation in Congestion Control”. In: *arXiv preprint arXiv:1706.03632* (2017).

[160] Michael J Swain and Dana H Ballard. “Color indexing”. In: *International journal of computer vision* 7.1 (1991), pp. 11–32.

[161] Andrea Tagarelli and Roberto Interdonato. “Who’s out there?: identifying and ranking lurkers in social networks”. In: *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ACM. 2013, pp. 215–222.

[162] *TCP and UDP Nmap port scanning web tool*. pentest-tools.com. [] 2008.

[163] TextSecure. *TextSecure Private Messenger*. http://download.cnet.com/TextSecure-Private-Messenger/3000-2150_4-76145455.html. [] 2014.

[164] *The risks of third-party app stores*. https://us.norton.com/internetsecurity-mobile-the-risks-of-third-party-app-stores.html. [Accessed June 2019].

[165] Thom and Camille Ryan. American Community Survey Reports, ACS-28, U.S. Census Bureau, Washington, DC, 2014. *Computer and Internet Use in the United States: 2013*. https://www.census.gov/history/pdf/2013computeruse.pdf. Accessed: 2018-11-17. 2013.

[166] *Triada Exploit*. https://security.googleblog.com/2019/06/pha-family-highlights-triada.html. [PHA family highlights Article from Google].

[167] *Two thirds of Android antivirus apps don't work properly*. https://www.engadget.com/2019/03/17/most-android-antivirus-apps-dont-work/. [Accessed June 2019].

[168] *Two-Thirds Of Phone Apps Share Your Data With Third Parties*. https://patch.com/us/across-america/two-thirds-phone-apps-share-your-data-third-parties. [Accessed June 2019].

[169] *Understanding Mobile Apps*. https://www.consumer.ftc.gov/articles/0018-understanding-mobile-apps. Accessed: 2018-03-17.

[170] *Understanding Permissions in the Android World*. https://clevertap.com/blog/understanding-android-permissions/. [Accessed June 2019].

[171] Mark Unwin. *Open-AudIT*. http://www.opmantek.c. [] 2008.

[172] *User Perceptions of Smart Home Internet of Things (IoT) Privacy Behavior*. https://iot-inspector.princeton.edu/blog/post/getting-started. [Blog Post].

[173] *User Perceptions of Smart Home Internet of Things (IoT) Privacy Behavior*. https://freedom-to-tinker.com/2018/10/22/. [Blog Post].

[174] Bjorn J Villa. "Enhancing Quality Aspects of Adaptive Video Streamingin Home Networks". In: (2014).

[175] Daniel Votipka et al. "User comfort with Android background resource accesses in different contexts". In: *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*. 2018, pp. 235–250.

[176] Zhihua Wen, Sipat Triukose, and Michael Rabinovich. "Facilitating Focused Internet Measurements". In: *Proceedings of the ACM SIGMETRICS*. San Diego, CA, USA: ACM Press, June 2007.

[177] *White, Black, and Grey Hat Hackers*. https://us.norton.com/internetsecurity-emerging-threats-what-is-the-difference-between-black-white-and-grey-hat-hackers.html. [Accessed June 2019].

[178] *Wifi Analyzer*. https://play.google.com/store/apps/details?id=com.farproc.wifi.analyzer&hl=en. Accessed: 2018-03-17.

[179] *Wifi Analyzer App*. https://play.google.com/store/apps/details?id=com.farproc.wifi.analyzer&hl=en. Accessed: 2018-03-17.

[180] *WiFi Connection Manager*. https://play.google.com/store/apps/details?id=com.roamingsoft.manager&hl. Accessed: 2018-03-17.

[181] *WiFi Connection Manager*. https://play.google.com/store/apps/details?id=com.roamingsoft.manager&hl=en. Accessed: 2018-03-17.

[182] *WiFi-Master-Speed-Test-Booster*. https://play.google.com/store/apps/details?id=com.leo.wifi&hl=en. Accessed: 2018-03-17.

[183] *Windows Speedtest*. http://www.speedtest.net/apps/windows. Accessed: 2018-03-17.

[184] *World Internet Stats*. https://www.internetworldstats.com/stats.htm. [Web Page].

[185] *World Wide Internet Statistics*. https://www.internetworldstats.com/emarketing.htm. Accessed: 2019-05-05.

[186] Yiannis Yiakoumis et al. "Slicing home networks". In: *Proceedings of the 2nd ACM SIGCOMM workshop on Home networks*. ACM. 2011, pp. 1–6.

[187] Shuai Zhao et al. "Study of user QoE improvement for dynamic adaptive streaming over HTTP (MPEG-DASH)". In: *Computing, Networking and Communications (ICNC), 2017 International Conference on*. IEEE. 2017, pp. 566–570.

[188] Xuzi Zhou. *Understanding home networks with lightweight privacy-preserving passive measurement*. University of Kentucky, 2016.

[189] *Zip Code Characteristics: Population Studies Center UMICH*. https://www.psc.isr.umich.edu/dis/census/. []

Appendix

A.1 How to Run a Scan

The following is an excerpt from the Web site How's My Network Mobile App on how to run a new Home Network Scan, and is targeted primarily at end user support:

First set your Home Network Wifi by clicking the Blue Semi-arrow. If this is your first time running it will ask if this is your Home Network: Click Yes, and new Scan will begin to run. Across most user testing phases a new scan will run automatically every 15 minutes or less, or can be run manually via the refresh button.

If you attempt to run a future scan on a network that is not your Default Configured HN selected, the pop-up will ask if your HN has changed before you can run a scan Specifics about each tab operation include the following:

1. Network List Tab: Scan for devices that exist on your Home Network, items on this tab include:

    - Image of the Device type detected (e.g. iPhone)
    - Device Name and IP Address (a numerical number Assigned to each device connected to your Home Network, like a phone number). Your device will be color-coded Gold
    - Clicking on the name of the device will show information such as: Last time

device was found, MAC address, and Wifi connected to

- User request Add: You can enter in a Nickname for the Device in the pop-up. To revert back to the scanned name click ”Add a Nickname” with no entry to clear the nickname, and revert.

2. apps / Security Tab (privacy and security): Scan your device and create a fingerprint when things change, such as apps and permissions

   - A list of all user apps installed on your device
   - The apps are Color coded by security and privacy level:
     - Green - Safe
     - Yellow - Caution (some permissions used may have privacy/security concerns),
     - Red - Danger (app is using permissions that are directly related to security/privacy of your device [e.g. tracking, access to camera, etc.])
     - Gray - app has been deleted
   - Clicking on the arrow will open the list of permissions per app. Long-clicking on the app name will open the app settings on the Device
   - Clicking on one of the permissions will show details of permissions type, along with the access level (red color coded is danger)

3. Remote (Global) Norms Tab: (privacy and security) How your network compares to others in terms of Health, privacy, and devices

4. Wifi Health: A view of Wifi connection Strength connected to your Home Network. The health of your

5. Wifi is shown via a speedometer view color coded as:

- Green: Excellent
- Yellow: Moderate
- Red: Danger

Features of the app include:

1. Swiping down on the ”Network List” tab will re-scan your network
2. Swiping down on the ”apps and Permissions” tab will re-scan apps installed on your device
3. Swiping down on the ”Remote Norms” tab will re-load comparisons with your network to others using this app in their Home Network
4. Clicking the Refresh button in the top right corner will re-run a series of tests (as set by settings)
5. Clicking the preferences in the top right corner will open the configurations available, which include:
    - Settings - settings which are used when clicking the refresh button (scan network, scan
    - apps. throughput scan, disable all)
    - About - This window
    - Disclaimer - Shows the main disclaimer splash screen
    - Opt Out - Allows you to opt out of logging all data, but you still can use the app without logging

A.2 App Questions

The following are the list of daily or weekly questions users see as part of the HMN Mobile app. The first question shown (in bold) is posed at the start of the testing phases within the app, and typically at the end, where applicable. Responses and results from these questions are shown in section 5.9.

1. **"How has your perception of security/privacy in your Home Network changed within the past week?"**,
2. "Do you find the Internet speedtest and Wifi Speedometer helpful? If so, how is it helpful?",
3. "What information would you like to see added to Speed test results and a Wifi Speedometer?",
4. "Is the Star Rating System shown in the Summary/Comparison tab helpful in understanding your Devices and apps overall health?",
5. "What information would you like to see added to a Star Rating System for comparing your Home Network/Apps/devices?",
6. "What type of details would you like to see as part of a security and privacy review of your Home Network?",
7. "Have you changed the way you use apps or Devices, in your Home Network, because of using the HMN App?",
8. "Does this app provide enough detail to help you better understand privacy and security in apps and your Home network? If so, how has it changed your perception?",

9. ”Which of the following methods is preferred to gain access to information about your Home Network, apps, and Devices: Mobile, PC, Web, Router, or customized hardware?”,

10. ”Do you find the information provided in the Network Scan tab helpful in understanding your Home Network?”,

11. ”What other information would you like to have included as part of a Device Network Scan?”,

12. ”Do you find the information provided in the Apps/Security and permissions tab helpful? If so, how is it helpful?”,

13. ”What other information would you like to have included as part of apps Security and permissions Home Network Scan?”,

14. ”Do you find the Summary Comparison Tab information of your Home Network/Apps/Devices versus others helpful? If so, how is it helpful?”,

15. ”What other information would you like to see provided when comparing your Home Network/Apps/Devices versus Others?”,

16. ”What types of information would you like to see a Privacy / Security Mobile app provide?”

A.3 Internet Throughput Graphs

The following are Internet Throughput graphs for Download/Upload of DSL, Fiber, and Upload for Cable. We have provided the following Figures for these CDFs: 1,2,3,4, and 5; where we have labeled these as Cable1..9, DSL1..9, and Fiber1..4 respectively.

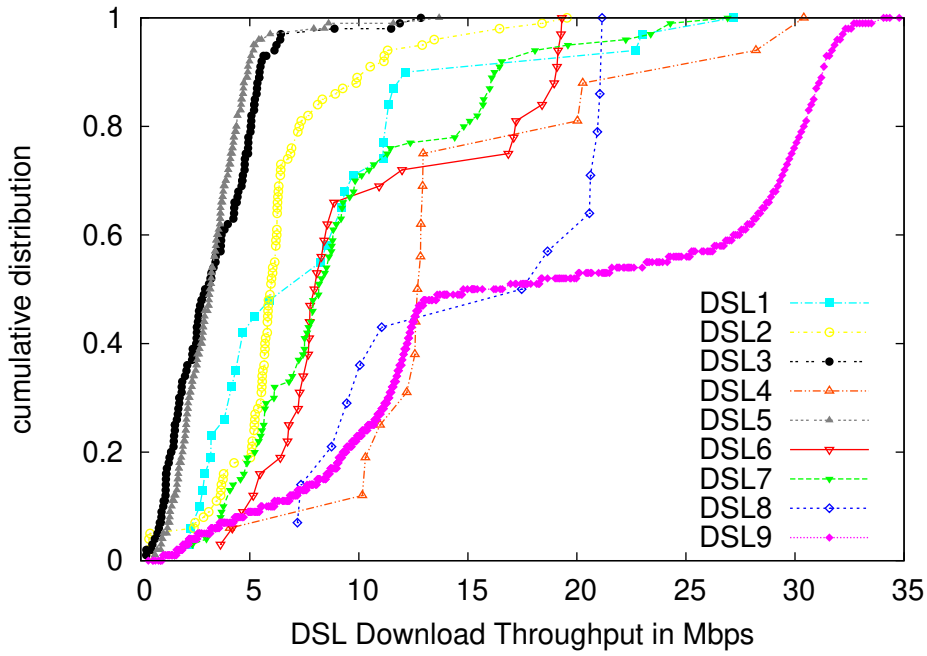

Figure 1: DSL Inet Providers Download Throughput CDF

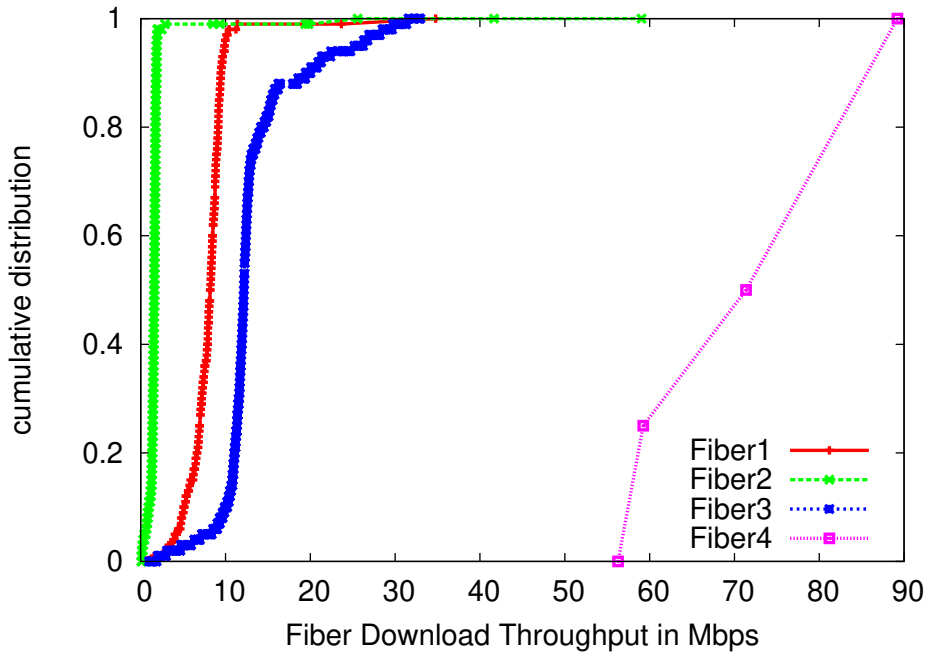

Figure 2: Fiber Inet Providers Download Throughput CDF

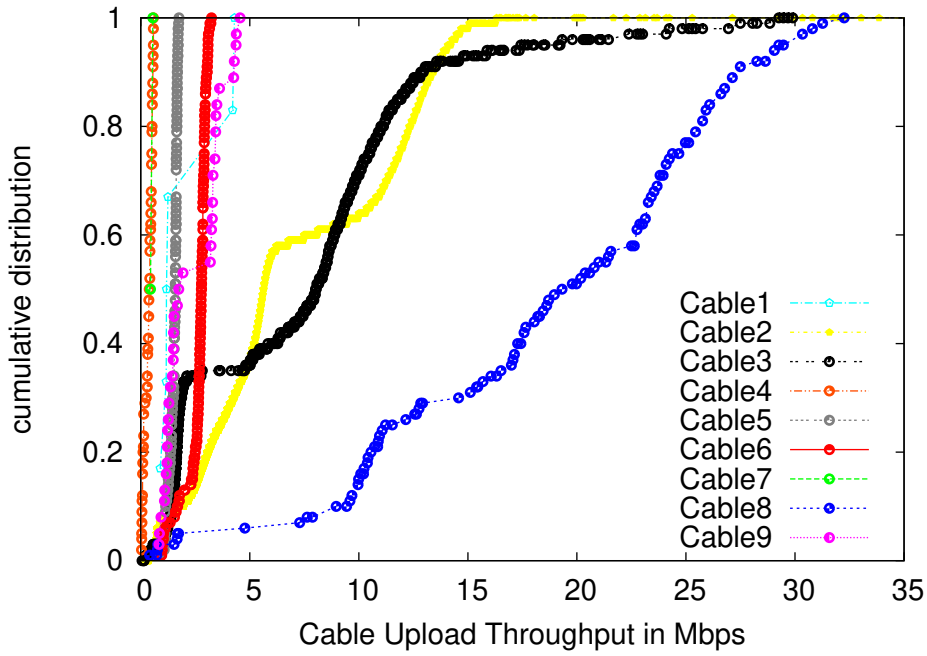

Figure 3: Cable Inet Providers Upload Throughput CDF

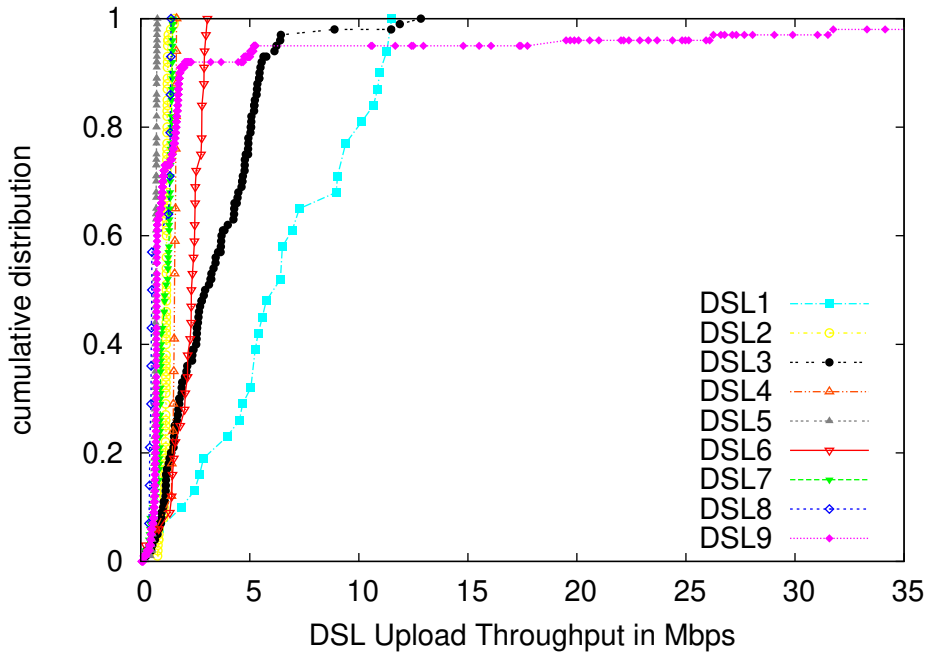

Figure 4: DSL Inet Providers Upload Throughput CDF

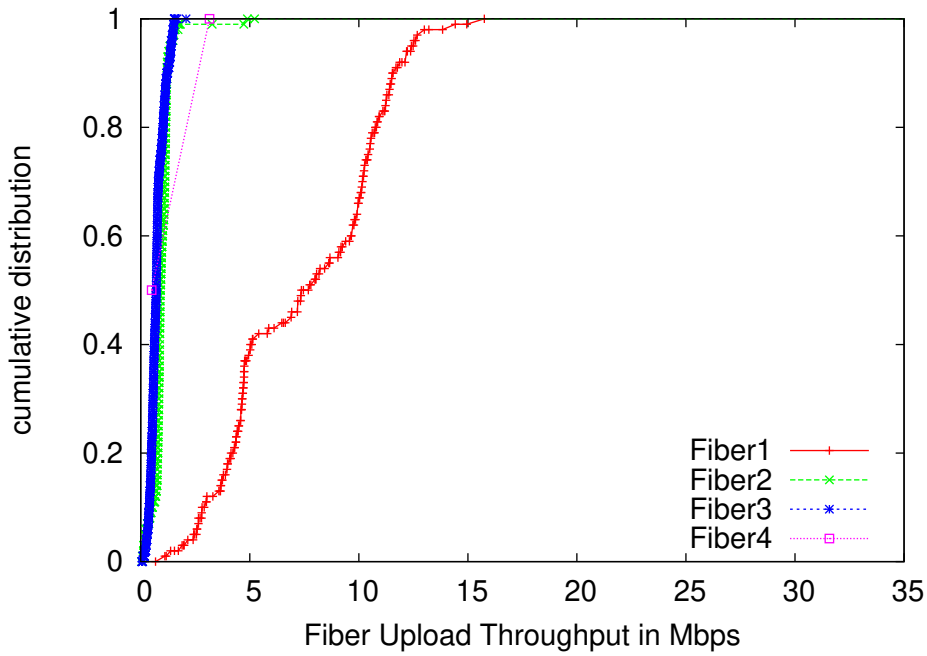

Figure 5: Fiber Inet Providers Upload Throughput CDF

Percentage	OS Type
62%	Windows PC
36%	Mac
2%	Other (e.g. Linux)

Table 1: PC and OS Types

A.4 Data Collection from App

Figure 1 shows the most popular PC types (combined PC/Laptops), along with their accompanying OS.

We examined each app and matched the location the app is categorized into, and found that the top 10 most popular categories fall into what is in Figure 2, across all participants. We were not able to discover the categories for roughly 9% of apps installed, as they were not available or listed.

Table 5.24 shows a listing of the top 12 (of 82) US ISPs. We have not listed location as these providers are nation wide in most cases.

Table 2: Top 10 Popular App Categories & Percentages

App Category	Percentage
ENTERTAINMENT	27%
TOOLS	11%
Description NA	10%
PRODUCTIVITY	6%
FINANCE	6%
LIFESTYLE	5%
SHOPPING	5%
HEALTH AND FITNESS	5%
COMMUNICATION	4%
GAMEs and PUZZLE	3%

Percentage	Service Providers	Inet Type
30%	Charter Communications Inc	Cable
23%	Comcast Cable Communications LLC	Cable
6%	Verizon Wireless	DSL
5%	Suddenlink Communications	Cable
4%	Verizon Communications Inc.	Fiber
4%	AT&T Corp.	DSL
2%	Worcester Polytechnic Institute	Comp/Fiber
2%	RCN	Cable
2%	Mediacom Communications Corp	DSL
2%	Cox Communications LLC	Cable
2%	CenturyLink Communications LLC	DSL
2%	Amazon.com Inc.	T1

Table 3: ISP Percentage of Users Using Provider(s) and Connection Types

A summary of devices used by participants to run the HMN Mobile app is shown in Table 5, and we can see that 45% of all participants were using Oreo (Version 8), and only 17% using the latest Android OS (9), which was released in Aug, 2017.

A review of DNS providers 4 shows that 53% HNs rely upon their ISP for DNS, with 44% using Google's DNS services (8.8.8.8). We created a rating of all DNS health across all ISPs and HNs and saw that the median lookup time (RTT) was 23ms, and the overall health was calculated to be 40ms, using an average metric across all DNS requests

Percentage	DNS Provider
44%	Google (8.8.8.8)
53%	Internal Router DNS
3%	External Direct

Table 4: DNS Providers

Percentage	OS Release	OS Version
45%	Oreo	8
19%	Nougat	7
19%	Marshmallow	6
17%	Pie	9

Table 5: Android OS Types

aggregated and taking the median score. in terms of subnet connectivity we see that 96% of all participants resided in a Class C subnet, with the remaining 4% used a class B, of the total 148 connections (a small subset of users did not run the service and only connected which resulted in minimal or no data); 98% use an RFC1918 non-routable address (e.g. 192.168.x, 10.x, etc.) for their internal network connected devices.

A.5 Dangerous Apps and Permissions

Table 6 is a listing of popular apps and % of permissions granted at the end of testing across all users. We can see that there is a 7.7% standard deviation of revocations across all of these apps/permissions.

Table 6: Popular 10 Apps % of Permissions granted, per category, end of testing

Category / App	FB-Messenger	Netflix	Spotify	Reddit	Uber	Yelp	Fitbit	Zillow	Flashlight	Outlook
Contacts	100	3	96		95	92	100	100		96
Camera	100		96		95	92	100	100	100	83
SMS MMS	100				95		100			
Storage	100	94	96	100	95	92	100	100		96
Location	100			100	95	92	100	100		96
Photos	100	94	96	100	95	92	100	100		96
Micro-phone	100	97	96		89	92				
Phone	100	97	9		95	100	100	100	100	
Wi-Fi	100	97	96	100	84		100	100	100	96
Device / Identity	100	97	9		95		100	100	100	
Calendar	100				16					96
In-app purchases	100	97	96	100						
Device history	100	97	9		95		100	100	100	
Bluetooth	100		96		89		100			
Wearable						100				
Other	100	100	100	30	100	100	100	100	100	96