Peeking Behind the NAT: An Empirical Study of Home Networks

Sarthak Grover, Mi Seon Park, Srikanth Sundaresan,
Sam Burnett, Hyojoon Kim, Nick Feamster
School of Computer Science, Georgia Tech

ABSTRACT

We present the first empirical study of home network availability, infrastructure, and usage, using data collected from home networks around the world. In each home, we deploy a router with custom firmware to collect information about the availability of home broadband network connectivity, the home network infrastructure (including the wireless connectivity in each home network and the number of devices connected to the network), and how people in each home network use the network. Downtime is more frequent and longer in developing countries—sometimes due to the network, and in other cases because users simply turn off their home router. We also find that some portions of the wireless spectrum are extremely crowded, that diurnal patterns are more pronounced during the week, and that most traffic in home networks is exchanged over a few connections to a small number of domains. Our study is both a preliminary view into many home networks and an illustration of how measurements from a home router can yield significant information about home networks.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations—*Network monitoring*

General Terms

Measurement

Keywords

BISmark, Home Networks

1. INTRODUCTION

Home broadband Internet access is ubiquitous and rapidly evolving. There are now upwards of one billion broadband Internet users worldwide [1]; much of that usage is shifting away from conventional desktops and towards mobile devices, such as laptops, smart phones, and tablets [7]. Despite their pervasiveness, little is known about most home networks or how people use them.

IMC'13, October 23–25, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2504730.2504736 .

Indeed, to date, it has been surprisingly difficult to study home networks on a large scale, because network technologies like network address translators (NATs) present only an opaque view of the home network to the global Internet—specifically, without a monitoring device *inside* the home, traffic coming from any device in a home network appears to all be coming from a single device. This coarse granularity of visibility makes it impossible to observe the usage patterns of individual devices inside the home or observe characteristics about other parts of a home network (*e.g.*, the home wireless network). Observing home networks on a large scale requires developing—and deploying—an always-on monitoring device in the home that can capture information about individual devices, with the consent of the people that live in these homes.

To better understand home networks, we developed BISmark, a custom home router, and deployed it in more than 100 home networks in 21 countries around the world for more than one year. BISmark sits between the user's ISP access link and the rest of the home network and acts as a continuous monitoring device; as a result, it can observe all traffic entering and leaving the home network, and can attribute traffic flows to individual devices on the network. This device, if permitted, can also log other types of activity, such as the number of devices on the network at any given time and the amount of traffic that any particular device sends to a destination (*e.g.*, Google). It can also independently measure the performance of both the home wireless network and the access link. We study three aspects of home network usage:

1. *Availability.* How common are Internet connectivity outages in homes, and how reasonable is it to assume continuous connectivity?
2. *Infrastructure.* What networking technologies and devices do people use in home networks?
3. *Usage.* Is the capacity of a home network sufficient for the growing demands of users and applications? How does usage differ across individual devices in the home? What other patterns exist?

Although our study instruments a relatively small number of home networks, it nevertheless offers extensive visibility into aspects of homes that were previously opaque to researchers. Ultimately, regulators and Internet service providers may be able to use some of the techniques that we describe to perform similar studies on a larger scale. We believe that researchers, ISPs, policymakers, and users can use the home router as a measurement device to better understand various trends in availability, infrastructure, and usage. Data about availability can help regulators determine whether ISPs are delivering the service that they are promising to users. Data about infrastructure (*e.g.*, the amount of contention in the wireless spectrum) can help ISPs better understand and debug user perfor-

mance, and can provide evidence for regulators to release more spectrum when it becomes appropriate to do so. Information about usage can help ISPs better provision and plan, and it can help device designers better understand how (and when) people use various devices.

The *availability* of home broadband access networks lets developers of applications running over edge networks understand the connectivity conditions of a particular environment. Our study has yielded some surprising findings. For example, although in the Western world we often think of broadband connectivity as "always on", we found examples in China and India where users only power up their home network gateway when they intend to use it. We found that trends of availability held in general: only 10% of home networks in the developed world saw connectivity interruptions of more than ten minutes more frequently than once every 10 days, but about 50% of home networks in developing countries experienced such connectivity interruptions once every 3 days. Although it is difficult from our data set to ascertain the cause for this intermittent connectivity (*i.e.*, it could result from power outages, poor connectivity, or behavioral patterns), it is clear that it is not safe to assume that Internet connectivity is highly available in certain parts of the world.

The *infrastructure* of home broadband access networks lets us understand how users construct their home networks, and what types of devices they comprise. We explore various questions relating to infrastructure, such as the number of devices on each home network, and how the number of devices on these networks varies over time with usage. We find that in today's home network, there are more wireless devices than wired devices in general and this difference is greater in the developing countries. We also find households in more developed countries tend to have more devices compared to developing countries, and those devices are more likely to remain continuously connected to the home router. Connectivity shows a diurnal pattern across a week; the number of devices on the network on weekdays typically peaks during evening hours, but on weekends device usage is more consistent throughout the day. The median number of devices on a 2.4 GHz network is about five, whereas on the 5 GHz band, the median number of devices is two.

The *usage* of home broadband Internet access can shed light on the applications and devices that people tend to use on their home network, and the overall utilization of these networks. We analyze the periodicity of usage patterns in various home networks and observe the extent to which users saturate their access link and find that most home networks are lightly used, and do not saturate their downstream or upstream link most of the time. We analyze distributions of device usage and find that users normally have a subset of devices that they prefer to use for consuming most network data. We observe the diversity of domains visited and find that about 38% of the total volume of traffic is from a single most popular domain, among 200 whitelisted domains.

The rest of the paper is organized as follows. Section 2 overviews related work. Section 3 describes the process of collecting the various types of data that we use in our study. Section 4 presents results concerning the availability of broadband access in various home networks around the world, including the duration of outages. Section 5 describes the characteristics of the infrastructure in various home networks, including the number of devices that are in the network, whether the devices are connected via wired or wireless, and to what extent devices occupy different ranges of radio spectrum (*e.g.*, 2.4 GHz and 5 GHz). Section 6 profiles the usage patterns of users in different home wireless broadband networks and explores how usage patterns differ across devices, and how download speed affects usage patterns. Section 7 describes ongoing work and extensions to this study and discusses some broader implications of our results. Section 8 concludes.

2. RELATED WORK

We briefly review related work on home broadband networks. We survey work that has performed "single shot" measurements of home broadband network performance and characteristics, Internet policy reports on performance in various regions, and qualitative studies that lend insight into how to better design home networks.

Measuring home network performance. There has been significant interest in measuring home and access networks, and many previous studies have tried to measure home networks in various ways. The studies use various methods for measuring home networks, from running measurements remotely from servers in the Internet to measure access link properties [18] to running tools on end-host devices [14, 16, 17, 25, 26, 28]. In contrast to these previous studies, we perform our measurements from gateway routers, not end-host devices, which enables *continuous* monitoring rather than one-shot measurements. Such continuous monitoring allows us to observe how usage patterns change over time, both on short and long timescales, as well as to report on other characteristics that require continuous monitoring, such as availability. Our work builds on our own previous work [32], which uses a deployment of custom home access points that conduct a unique set of performance measurements. In contrast to our previous work, this study broadly characterizes home network usage rather than focusing on access link performance.

National and regional Internet policy and measurement reports. Recently, there has been interest from national agencies to measure home networks for Internet policy and regulation purposes. The United States Federal Communications Commission, United Kingdom's Ofcom and the European Commission have all conducted large scale studies of access networks in conjunction with SamKnows [4–6]. Benkler *et al.* have a report on broadband transitions and policies around the world [10]. To date, these studies from regulatory commissions have focused exclusively on performance of the access network, rather than on properties of the home network itself or usage of the home network. The primary objective of such initiatives is to gain an understanding of access networks to enact better policies. They do not perform passive monitoring of the home network.

Qualitative design studies of home networks. There are previous studies on understanding home networks and attempts to design better systems [11, 15, 20, 21, 23, 27]. These previous studies are qualitative: they rely exclusively on human subject interviews and analysis on human interactions to identify problem areas and to suggest better designs. We offer a *quantitative* complement to these studies. We analyze passive measurement data that is automatically collected and reported, which are in turn used to derive meaningful observations about home networks that may not be obvious or even revealed through studies with human subjects. The automated nature of our data collection and monitoring allows us to observe longitudinal behavior and usage patterns, and it may in some cases result in more accurate data about human activity and network usage, since many of the questions that previous studies have asked in interviews could be more completely and accurately addressed by measuring the network traffic itself.

Previous studies have also built tools that improve interactivity and aid troubleshooting in a home network environment, for example measuring and displaying bandwidth usage and throughput in a home network [13, 14]. Our work analyzes data from *existing* networks, rather than trying to deploy new tools and observe how

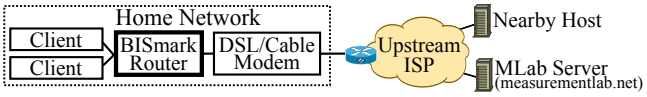


Figure 1: *The BISmark home router sits directly behind the modem in the home network. It collects both active and passive measurements.*

usage changes as a result of those tools. We analyze a wider variety of network features, including wired vs. wireless usage, number of active devices, diurnal patterns, and availability.

Behavioral studies of Internet usage in developing countries. Chen *et al.* studied the effect of sporadic and slow connectivity on user behavior and found a better Web interaction model for such environments [12]. Wyche *et al.* performed a qualitative study of how Kenyan Internet users adapt their usage behavior where Internet connectivity is a scarce resource in terms of availability, cost, and quality [33]. Smyth *et al.* performed a qualitative study on sharing and consuming entertainment media on mobile phones in urban India [31]. The data that we gathered in developing countries could help corroborate some of these studies.

3. DATA COLLECTION

Home routers can observe many aspects of home networks because typically all other devices in the home communicate both to each other and to the Internet via the router. Over the past three years, we have deployed routers in 126 homes across 19 countries. Each router measures the quality of the upstream Internet connection and collects limited information about device usage on the home network. This section introduces the router platform, the data we collect from the routers, and that data's implications for our study.

3.1 Collection Infrastructure

BISmark comprises gateways in the home, a centralized management and data collection server, and several measurement servers. We have instrumented the gateway with custom firmware that performs both passive and active measurements. Where appropriate, the firmware anonymizes certain aspects of the data before sending them back to the central repository for further analysis. Figure 1 shows a typical deployment in the home network, and how BISmark performs its measurements.

Firmware. BISmark is a custom home router firmware based on OpenWrt for Netgear WNDR3800 and WNDR3700v2 routers [2, 3]. Routers have a 450 MHz MIPs processor, 16 MB of flash storage, 64 MB of RAM, an Atheros wireless chipset, one 802.11gn radio, and one 802.11an radio. BISmark typically replaces a household's wireless access point and connects directly to the cable or DSL modem that provides Internet access to that household. Because the router sits on the path between the user's home network and the rest of the Internet, our software is uniquely positioned to capture information about both the characteristics of network connectivity and of home network usage (*e.g.*, usage patterns, applications). We expected routers to remain powered on almost all the time, since they provide the household's Internet connectivity; however, later in this paper we show that this assumption does not hold in several countries and regions.

Recruiting and deployment. Our deployment of routers across home networks has been organic: We have recruited most of our users by word-of-mouth, or through targeted advertisements for

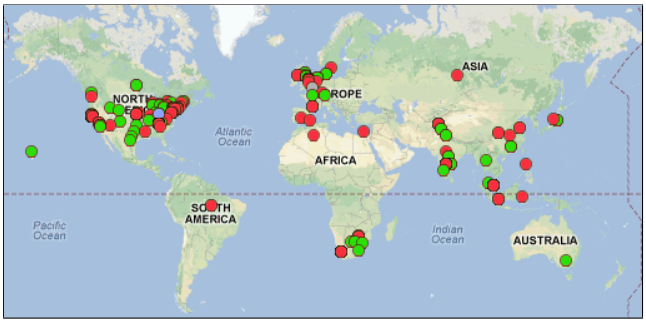


Figure 2: *The BISmark deployment as of May 2013. Each dot indicates a router. The green dots indicate routers that are currently reporting (156). Because we only use data from routers that consistently report data throughout the period of our study, we use data from 126 routers in 19 countries. The red dots include the full set of routers that have ever contributed data (295).*

Developed	Routers	Developing	Routers
Canada	2	India	12
Germany	2	Pakistan	5
France	1	Malaysia	1
United Kingdom	12	South Africa	10
Ireland	2	Mexico	2
Italy	1	China	2
Japan	2	Brazil	2
Netherlands	3	Indonesia	1
Singapore	2	Thailand	1
United States	63		
Total Routers	90	Total Routers	36

Table 1: *Classification of countries based on GDP per capita.*

specific experiments and projects that we have run as part of our research. For example, the router firmware performs continuous measurements of the performance of the home access link, which has garnered the attention of various policy and regulatory agencies. We have also performed smaller recruitment efforts in various areas for a usage cap management tool that we built on top of the firmware [24]. Depending on the experiments that different users have consented to (or not), we are able to collect different types of information. Most users have remained actively engaged in our experiments by virtue of the fact that they receive a free router as a result of their participation.

We classify the countries where we have deployed routers into two groups based on the GDP (Gross Domestic Product) per capita ranking in year 2011 [9]. We call countries for whom the per capita GDP falls within the top 50 *developed*; otherwise, we call them *developing*. Table 1 summarizes this grouping.

3.2 Data

We now summarize the data we collected from the BISmark deployment, then describe each data set in more detail. We will also highlight some factors that limit the conclusions we can (or cannot) draw from our data. Where possible, we have released the data collected from this study; the Capacity data (described below) is publicly available and is also continuously updated as the routers collect new measurements.[1] We have released all measurements that do not have personally identifying information (PII) (*i.e.*, everything except the Traffic data set).[2]

[1] `http://uploads.projectbismark.net`
[2] `http://data.gtnoise.net/bismark/imc2013/nat`

Dataset	Routers	Countries	Dates
Active measurements			
Heartbeats	126	19	October 1, 2012–April 15, 2013
Capacity	126	19	April 1–April 15, 2013
Passive measurements			
Uptime	113	19	March 6–April 15, 2013
Devices	113	19	March 6–April 15, 2013
WiFi	93	15	Nov. 1–Nov. 15, 2012
Traffic	25	1	April 1–April 15, 2013

Table 2: *Summary of data collected for this study. With the exception of the Traffic data set, which is subject to privacy restrictions, we will publicly release all data used in this study.*

3.2.1 Summary

Table 2 summarizes the data we collected from our deployment. BISmark routers perform both active and passive measurements. Many routers only collect performance measurements about the router's upstream Internet connection and basic diagnostic information about the number of connected devices; these measurements record no personally identifying information (PII) and, as a result, do not require written consent. In twenty-five homes where we have explicit consent, we collect additional information about the activity of users and devices on the home network; for those households we clearly explained the risks of PII exposure and obtained written consent.[3] Engineering and consent constraints dictated that we collect the data sets over different time periods. We now explain each data set in more detail.

3.2.2 Measurements

Heartbeats. Every router sends a "heartbeat" packet to the central BISmark server approximately once a minute. We use this data to measure router uptime. A heartbeat packet indicates that the router is on and online, but a lost packet could mean a router is powered off, offline, or has lossy connectivity to the server. These heartbeats can be lost, and the router makes no attempt to retransmit them. They are nonetheless frequent enough to provide reasonable confidence about the uptime of each BISmark router, which we analyze in Section 4. We consider heartbeats from 126 routers that were on for at least 25 days between October 2012 and April 2013.

Uptime. Starting in March of 2013, each router sends its uptime every twelve hours. This data distinguishes, at coarse granularity, between routers that are offline and those that are powered off.

Capacity. Every twelve hours, each router measures the capacity of its access link using ShaperProbe [30]. In Section 6.2, we jointly analyze the Capacity and Traffic data to determine the extent to which users fully utilize the capacity that their Internet service provider offers.

Devices. Every hour, most routers count the number of devices connected to their wired Ethernet ports and the number of associated clients on each wireless frequency. This data gives a broad view of users' device usage patterns, while still providing information at a coarse enough granularity to preserve user privacy. Section 5 uses this data to paint a broad picture of device usage throughout the world.

WiFi. Some routers collect data about the number of other access points (APs) in the vicinity. Each router only scans for other visible access points in the wireless channel that it is configured for; by default, the 2.4 GHz radio is configured for channel 11, and the 5 GHz radio is configured for channel 36. Each router attempts to scan for clients and access points every 10 minutes; unfortunately, the scanning process can sometimes cause wireless clients to disassociate from the router, so we reduce the scanning frequency if the router has associated clients.

Traffic. Because of potential privacy risks inherent in passively collected data, only 53 households consented to contribute detailed information about their devices and Internet activity. Of those, we consider data from 25 households that were active from April 1 to April 15 and exchanged at least 100 MB of traffic. To encourage users to consent to this data collection, we gave them access to a Web interface that allowed them to observe and manage their usage over time and across devices; this feature turns out to be quite useful for users who have Internet service plans with low data caps. We collect four types of information:

1. *Packet statistics.* We collect the size and timestamp of every packet relayed to and from the Internet. Although this data seems fairly innocuous, traffic analysis could in fact reveal very detailed information about user activity.
2. *Flow statistics.* We collect obfuscated IP addresses and MAC addresses, and application ports for a sample of Internet-bound network flows. At surface level, this data only reveals the kinds of applications participants use on their Internet-enabled devices (*e.g.*, HTTP, SMTP), but inference attacks similar to those on packet-level data are possible.
3. *DNS responses.* We collect a sample of A and CNAME records and obfuscate domain names unless they appear on a user-customizable whitelist of popular domain names; by default, the whitelist is the 200 most popular domains in the United States according to Alexa [8].
4. *MAC addresses.* We collect the MAC addresses of devices connected to the router and anonymize the lower half of each address, which allows us to identify manufacturers without identifying specific devices.

This data provides rich insights into user behavior inside the home. We quantify device ownership in Section 5.4 and characterize service usage in Section 6.

3.3 Limitations

This section discusses various limitations of our data sets.

Data collection may be interrupted. Various outages and failures—both of the routers themselves and of the collection infrastructure—introduced interruptions in our collection. Although we acknowledge that being able to collect all of our data sets over the same overlapping time period would be ideal, we needed to choose between using completely overlapping data sets from smaller time intervals or using the largest time intervals possible for each data set. As such, we used overlapping time periods where possible, and in other cases we used the largest possible time interval.

It is difficult to infer causes of downtime from Heartbeats. Although we are interested in measuring availability with the Heartbeats data, this data set really only indicates when the BISmark router is both powered up and online. Using this data to infer access link outages assumes that the BISmark router is always on, but as we will see in subsequent sections, some instances of "outages" are in fact users who power down the router for significant portions of the day. Additionally, these heartbeats are sent from

[3] Our university's Institutional Review Board (IRB) certified all aspects of our data collection and experiments.

the BISmark routers to a server at Georgia Tech only, so a loss of heartbeats might simply result from problems along the network path between the BISmark router and Georgia Tech. We assume that most persistent losses of heartbeats are due to either a failure at the access network itself, or due to the router being powered down; we are unable to distinguish these two cases. We started collecting uptime data at 12-hour intervals, but this coarse granularity means we can only confirm network outages in cases where routers are continuously powered on.

Privacy and ethics limit scalability. Although we would like to observe all passive traffic from the entire deployment, we are constrained by privacy and ethics concerns, which compel us to restrict the size of the Traffic data set and anonymize the data in ways that make certain types of analysis more difficult. For this study, we were only able to collect passive traffic traces from 25 homes in the United States; collecting these traces requires gaining institutional review board (IRB) approval, recruiting users who were willing to let us collect such traces, and obtaining consent from these users. Even for consenting users, we anonymize a significant portion of the passive traffic traces: We anonymize traffic to any domain name that is not in the Alexa top 200 or otherwise explicitly whitelisted by the user, and hash the bottom 24 bits of all MAC addresses. This anonymization prevents us from studying certain types of questions, such as the characteristics of the "tail" of user traffic volumes.

Households are not a representative population sample. Because we wished to minimize both capital risk and technical support investment, our deployment is biased toward close friends, family, colleagues, and technically-inclined volunteers in the United States and abroad. This may limit the generality of our results, particularly those drawn from only a few routers. Nonetheless, we believe our results shed light on important phenomena in home networks that could be verified in later, more targeted deployments; indeed, our ongoing research and recruitment efforts follow from data gathered on a smaller BISmark deployment.

BISmark enables its 5 GHz radio by default. Various countries have restrictions on the use of wireless spectrum; the use of the 5 GHz spectrum on wireless routers is only permitted in certain countries. For the most part, however, we do not disable the 5 GHz radio before shipping a BISmark router, and any user who flashes their own router with the firmware that we have published on the Web may have the 5 Ghz spectrum enabled by default. Therefore, although it might be interesting to explore the use of 5 GHz spectrum in countries where it is not legal, the way that we have deployed our routers necessarily biases our data set and does not allow us to answer this question accurately.

4. AVAILABILITY

We first study the extent to which Internet connectivity was available across the home broadband networks in our deployment, and whether any remarkable availability or usage patterns emerged. We observe the frequency and duration of downtime across the deployment and the extent to which these characteristics differ between home networks—we explore differences between developing and developed countries (and specifically, how downtime characteristics vary according to a country's GDP), as well as usage patterns that are unique to individual homes.

We measure the availability of Internet connectivity by recording periodic heartbeat messages which every router sends to our central server approximately once per minute. The arrival of a heartbeat signifies that the router is up and connected to the Internet. We define *downtime* as any gap in the heartbeat logs that lasts longer than ten minutes. The absence of a heartbeat could signify that the router is offline, has been shut down or rebooted, or that the heartbeat was lost. For reboots the router usually comes back up within a few minutes. In some cases, we can positively verify downtimes caused by powered off routers using the Uptime data set. All results in this section are based on the Heartbeats and Uptime data sets, as described in Section 3. Table 3 summarizes a few of the more interesting results.

Developed countries experience far less frequent extended downtime than developing countries: The median time between downtime in developed countries is more than a month while in developing countries it is less than a day.	§4.1, Fig. 3
The two countries in our deployment with the most frequent downtime are those with the lowest per-capita GDP (India and Pakistan).	§4.1, Fig. 5
In some cases in developing countries, broadband connectivity is not "always on" because users only turn on their routers to use the Internet.	§4.2, Fig. 6b

Table 3: *Highlights of Section 4 results.*

Figure 3: *Average number of downtimes per day that last at least ten minutes. Developed countries experience far fewer downtimes per day than developing countries.*

4.1 How reliable is home broadband access?

We study the "uptime" of home broadband Internet access across home networks in our deployment, in terms of the frequency and duration of downtime. We also explore the extent to which these characteristics vary by country.

Frequency of downtime. Figure 3 shows a distribution of the downtime over six months, for both developed and developing countries; we show a distribution of the average number of downtimes per day for each network, where downtime is defined as an interruption in connectivity of ten minutes or longer. As expected, the home networks in developing countries sustain substantially more downtime compared to those in developed countries. The median duration between downtimes for networks in developed countries is more than a month, whereas for developing countries, the median duration between downtimes is less than a single day.

Duration of downtime. Figure 4 shows a cumulative distribution function of the downtime durations for both developing and developed countries. The plot shows that the median downtime dura-

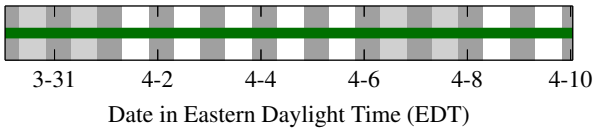
Figure 4: *Downtime duration for developing and developed countries. The median downtime duration is similar, but as we see in Figure 3, developing countries see downtime much more frequently.*

Figure 5: *The median number of downtimes across homes in each country during October 2012 – April 2013 vs. the GDP of the country where that home network was located. The marker size is proportional to the median downtime duration in that country. The vertical line separates developing and developed countries. We show only countries with at least three routers deployed.*

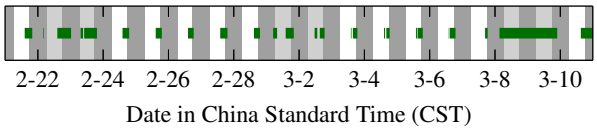
tion is approximately 30 minutes, and that downtime in developing countries tend to last longer. Downtime sometimes lasts several days.

Downtime characteristics vs. per-capita GDP. Figure 5 shows a scatter plot of the median number of downtimes experienced by home networks in each country vs. the *per-capita GDP* for that country in U.S. dollar equivalent (*i.e.*, the purchasing power parity) [22] for countries with at least three deployed routers. The vertical line divides nine developing from eleven developed countries. Although there is no discernible difference between any of the developing and developed countries in terms of the median number of downtimes, the two countries with the lowest per-capita GDP in our deployment—India and Pakistan—experienced significantly more downtime during October 1, 2012 – April 15, 2013. A home network in Pakistan experienced nearly two downtimes lasting at least ten minutes every day.[4]

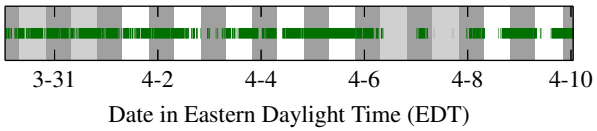
[4]These characteristics are are based on data from about 12 routers in India and about five routers in Pakistan for October 1, 2012 – April 15, 2013. In general, some country data from this scatter plot may be inconclusive, due to the small number of countries from which we collect data, but the trend appears to hold for at least the poorest countries.



(a) *This household never intentionally turns off its router, which is typical of routers in developed countries.*



(b) *This household often turns off its router when not using it. The router is available briefly in evenings and during weekends.*



(c) *This household's access link experienced sporadic ISP outages for several days in April 2013. Although the router was continuously powered on, we label this as downtime.*

Figure 6: *Examples of different modes of router availability. The thick horizontal green lines indicate intervals when the router is* available*. For reference, dark shaded regions indicate nighttime hours and lighter shaded regions indicate weekend daylight hours.*

4.2 Case Study: Router as Home Appliance

In addition to general trends, we also discovered some interesting cases of availability and usage patterns. We observed that in the United States, most users leave their routers powered on all of the time. The median US user has his router on 98.25% of time in the measured time period. Router availability in these homes resembles Figure 6a. In contrast, *home broadband is not "always on" in the developing world.* As a comparison, median routers in India and South Africa stay on only for 76.01% and 85.57% of the time, respectively. We observed cases where users powered their router on for specific time periods when they were using the Internet, much as someone would use any other appliance. We envision multiple possible reasons for this disparity. One reason could be *behavioral patterns*: in some cases, we can observe specific patterns where users power down their router when they are not actively using it (because of data usage caps imposed by certain ISPs). For example, Figure 6b shows one Chinese household that consistently keeps its router off except during the early evening. During the weekend, the router is on for longer periods, presumably with increased activity. Another reason could be *poor connectivity*, such as high loss or network outages caused by congestion, overload, or even poor infrastructure or equipment, as is possibly the case for the home network in Figure 6c.

5. INFRASTRUCTURE

This section examines the *infrastructure* that people in home networks use to access the Internet. We look at the typical composition of devices in home networks, be they wired or wireless, and oper-

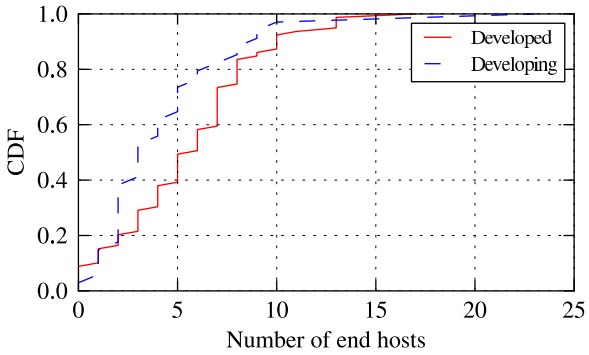


Figure 7: *Number of devices in each home network. More than half of the homes have at least five devices. Each home network has seven devices on average.*

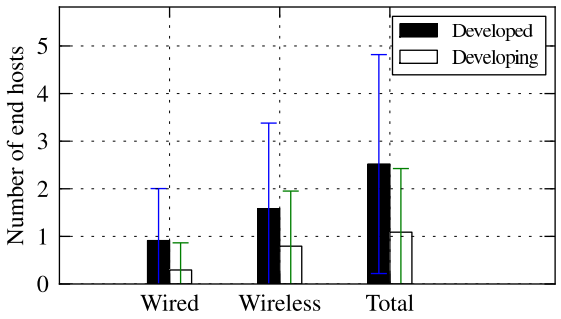


Figure 8: *Average number of devices connected to the access point at any time with error bars showing the standard deviation. We see that there are more wireless devices in both regions. Developed has more devices overall, and an even greater number of wired devices.*

ate either in the 2.4 GHz or the 5 GHz wireless spectrum. Table 4 summarizes the main results from this section.

5.1 How many devices?

More devices in developed countries. Figure 7 shows a CDF of the number of devices seen in each home; twenty percent of networks had at least two unique devices, and more than half of homes had at least five unique devices. Figure 8 details the number of devices for developed and developing countries. Developed countries have, on average, one more device connected to the access point at any given time than developing countries. All in all, households in countries with higher economic standards tend to have more devices in their network, and this number is more pronounced for wired devices. We assume this is because gaming consoles (*e.g.*, XBox, Playstation, Wii) or entertainment devices (*e.g.*, Apple TV, Google TV, Squeezebox) are more common in developed countries.

More "always-connected" devices in developed countries. Table 5 shows the number of households that have at least one wired or wireless device that never disconnects from the home gateway router for over five weeks and its percentage against the total number of households. The portion of households that have at least one "always-connected" device in developing countries is significantly lower than for developed countries. Media entertainment boxes are

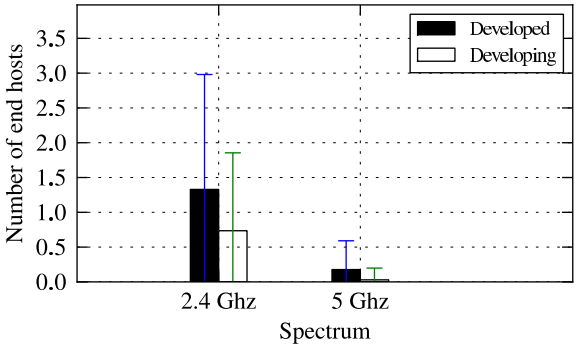


Figure 9: *Average number of wireless devices connected at any given time per spectrum with error bars showing the standard deviation. There are significantly more devices on 2.4 GHz than on 5 GHz.*

In developed countries, 43% of home networks have at least one always-on wired device; only 12% of home networks in developing countries have such as device.	§5.1, Tab. 5
The 2.4 GHz spectrum is significantly more crowded; the median number of devices seen on the 2.4 GHz spectrum is five, whereas on the 5 GHz band, the median number of devices is two.	§5.3, Fig. 10
The median number of access points seen from a home network in developed countries is about 20; in contrast, home networks in developing countries see a median of about two access points	§5.3, Fig. 11

Table 4: *Highlights of Section 5 results.*

an example of a wired device that never disconnects from the router, and a wireless VoIP phone is an example of an "always-on" wireless device. Some households may never turn off their desktop or laptop. Although we do not explore the reasons for such connectivity in depth, we assume that households in developing countries tend to power off devices when not in use, possibly to minimize electricity or data usage. Unexpected and frequent power outages in these countries also affect device connectivity.

5.2 Wired or wireless?

More wireless than wired. Figure 9 shows the average number of wired and wireless devices attached to (and associated with) the home router over two weeks, measured hourly, for developed and developing router groups. There are generally more wireless devices than wired devices. Although this observation could reflect limitations of the router, which has only four available wired ports, the average number of wired ports used is less than one in both groups; this shows that many households use wireless even though wired connections generally provide better throughput, latency, and stability. This result also confirms the trend of moving away from wired communication [7] and towards primarily using wireless devices access the Internet. For many households, wireless technology has developed to the point where it is good enough for day-to-day Internet usage, and wired communication has little more to offer.

On the other hand, our results also suggest that wired devices are still in fairly widespread use in some homes. Chetty *et al.*

Group	Total houses	Houses with always-connected wired device	Houses with always-connected wireless device
developed	79	34 (43%)	16 (20%)
developing	34	4 (12%)	4 (12%)

Table 5: *Number of households that has one or more wired or wireless device which never disconnects from the home gateway router for over five weeks.*

stated that home users still desire wired communication for reliability, speed, and security [15]. Due to the physical constraint of wired communication (*i.e.*, cabling) devices using wired ports are likely stationary ones such as desktop machines, network printers or media entertainment gadgets like Apple TV. Even in homes with wired devices, our analysis suggests that a typical home gateway router could likely suffice with only two ports. Surprisingly, only a few households use all four Ethernet ports (9% for both developed and developing countries).

5.3 How much is each spectrum used?

Spectrum contention is an important problem in home wireless networks. Many devices talking to many access points in the vicinity causes contention and interference problems, which in turn reduces the available bandwidth of the wireless channel. Our results confirm that the 2.4 GHz spectrum is quite crowded, especially in developed countries, which could create bottlenecks as access link throughputs continue to increase. The 5 GHz spectrum, on the other hand, is less crowded (at least for now).

Wireless spectrum usage. Figure 9 shows that there are more devices active on the 2.4 GHz spectrum than on the 5 GHz spectrum at any given time. This phenomenon may result from the fact that dual-band devices are more expensive, and single-band devices default to the more popular 2.4 GHz spectrum. Phones are equipped almost exclusively with only 2.4 GHz radios. Figure 10 shows a CDF of the total number of *unique* devices seen per household. We see that the median number of devices on the 2.4 GHz spectrum is five, while the median number of devices on the 5 GHz spectrum is two.

We see a similar trend in the number of other access points seen on both spectrums. The median number of other access points is only about one device in the 5 GHz spectrum, while it is higher for the 2.4 GHz spectrum, as expected. Figure 11 shows a CDF of the total number of unique access points seen per household. Scanning is done only in the channel the access point is configured in (channel 11 for 2.4 GHz by default, though the user can configure it), so this does not tell us all the access points available, but it does tell us how likely it is that interference occurs due to competing access points. We see that the 2.4 GHz spectrum is more densely occupied in developed countries, and interestingly, we see that there are two modes in both sets; either there are very few access points in that channel or there are a lot (more than ten in developed and more than three in developing countries).

5.4 Which device vendors are most common?

For the 25 homes in the United States in the Traffic data set, we observed the frequency of various types of devices connected to the home network. When collecting the Traffic data set, we obfuscate the bottom 24 bits of the MAC addresses of all devices seen by the gateway. The first 24 bits allow us to look up the manufacturer, and though this does not always tell us if the device is a phone, tablet, or laptop, it still provides us enough information to distinguish network gateways from smart devices and wireless cards.

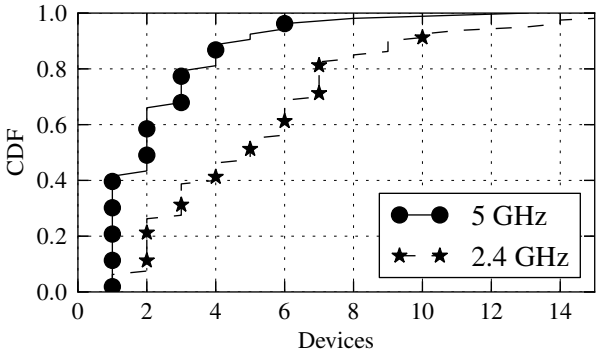


Figure 10: *Number of unique devices seen on the two wireless spectrums. The median is about five devices on 2.4 GHz and two devices on 5 GHz.*

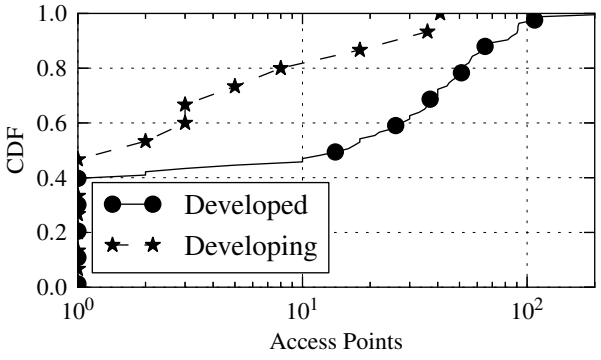
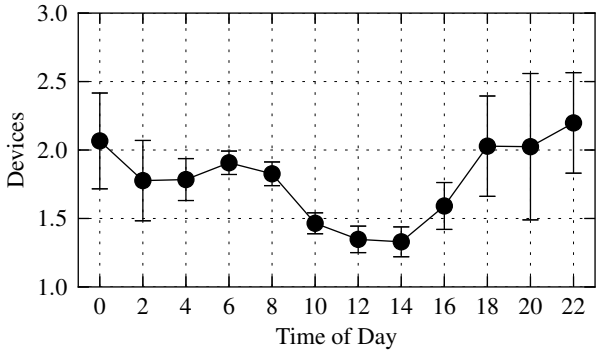


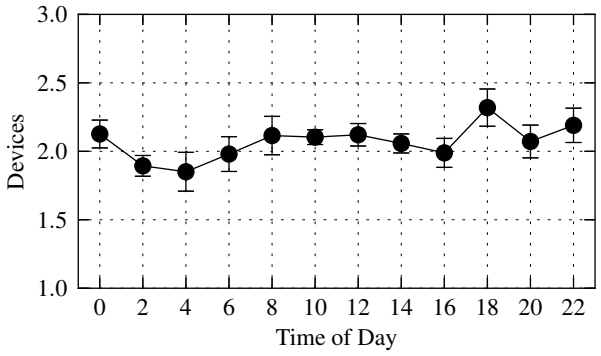
Figure 11: *Number of access points seen on the 2.4 GHz spectrum in developed and developing countries. There are many more access points visible in developed, where we see modal behavior with either very few or a lot of neighboring access points.*

Figure 12: *The number of devices seen in the Devices data set across all homes in the Traffic data set (25 homes in the United States). We only considered devices which transferred at least 100 KB.*

Based on passive monitoring of 25 homes, Figure 12 plots the device types seen in the Traffic data set. We have removed all references to Netgear originating from our BISmark routers. The most common device was manufactured by Apple, followed by Intel. Other Samsung and smart phones were also reasonably com-



(a) *Weekday usage is diurnal.*



(b) *Usage on weekends is more constant.*

Figure 13: *Diurnal effect on wireless device usage. There is a clear weekday diurnal effect on the number of devices online.*

mon.[5] We recently started gathering Traffic data in several developing countries, so we will soon be able to compare the distribution of device manufacturers in developed and developing countries.

6. USAGE CHARACTERISTICS

In this section, we identify notable characteristics of home network usage by analyzing data from the WiFi, Traffic, and Capacity data sets. Table 6 highlights our findings.

6.1 Which usage patterns are diurnal?

Figure 13 uses the WiFi data set to show the mean usage of wireless devices during each hour of the day, partitioned into weekday (Monday–Friday) and weekend. We observe a diurnal pattern in the number of unique devices at various times of day on weekdays; usage peaks during the evenings, and it is lowest during the afternoons, when users are usually at work. We see that the number of devices dips only slightly at night (compared to the dip during the day), which may result from cellular devices that remain on at night, as opposed to laptops that are more often switched off at night when users are asleep. Diurnal patterns are less pronounced on weekends.

[5]The *Printer* manufacturer was an Epson. *Hardware* includes Giga-Byte and Microchip. *VoIP* is a UniData device. *Internet TV* includes Roku, TiVo, and ASRock home theatres. *Wireless Card* includes AzureWave and GainSpan. *Gaming* includes Nintendo and Mitsumi (which manufactures controllers for Playstations, Xbox, and Wii). Microsoft (possibly Xbox) is shown separately. Gateway includes TP-Link, Realtek, Liteon, D-Link, Cisco-Linksys, Belkin, and Askey. *Smart Phones* includes HTC, LG, Motorola, Nokia, and a confirmed Samsung Galaxy S II (Murata Inc.). Other Samsung devices, including phones and tablets, are shown separately. *Original Device Manufacturers (ODMs)* include Compal, Hon Hai Precision, Quanta, Universal Global Systems, Winstron Infocomm. Misc. includes Polycom (a telecom product manufacturer), Prolifix (which makes Internet-enabled thermostats), and Pegatron (which manufactures a variety of products including notebooks, laptops to gateways).

(a) *Upstream traffic, and the measured upstream capacity.*

(b) *Downstream traffic, and the measured downstream capacity.*

Figure 14: *Diurnal pattern of link utilization for one home in our deployment. Capacity remains fairly constant, but utilization levels track daily cycles.*

Weekday traffic is much more diurnal than weekend traffic.	§6.1, Fig. 13
Some home networks consistently oversaturate their uplink; they are likely able to do this because of the "bufferbloat" phenomenon.	§6.2, Fig. 15
The single most usage-hungry device consumes about 65% of the total home network traffic, on average.	§6.3, Fig. 17
The most popular domain by volume in a home network is responsible, on average, for about 38% of the total wide-area traffic from that home network, but only 19% of the connections.	§6.4, Fig. 19

Table 6: *Highlights of Section 6 results.*

Figure 15: *Link utilization as a function of the measured throughput. Downlink saturation varies between 0 and 1. Uplink saturation is under 0.5 for most homes except for 3 cases. 2 homes over utilize their uplink*

Figure 14 shows an example of diurnal traffic patterns for upstream and downstream traffic from one home in the Traffic data set. Capacity measurements (shown as the dotted line at the top of each plot) remains fairly constant, while utilization follows a roughly diurnal trend. Many homes in Traffic exhibit patterns similar to those in Figure 14.

6.2 Do users saturate their access links?

Access link speeds continue to increase, but it is not known whether users actually take advantage of this additional capacity. We measure utilization by computing the maximum per-second throughput every minute for users in the Traffic data set. We then compare the utilization with the access link throughput as estimated by Capacity. We only consider instances when there is some device exchanging traffic with the Internet. Figure 15 plots 95th percentile link saturation against capacity estimates for uplink and downlink in bits per second on a log scale. In most cases there is plenty of spare capacity. At the 95th percentile, only two homes saturate the link and most homes use less than 50% of the available capacity. Thus, even when the link is used, utilization does not come close to capacity most of the time. If we include cases when the router is on and connected to the Internet but has no traffic then these numbers drop even further.

The circles on the scatterplot show the equivalent case for upstream traffic. We expect upstream usage to be less than downstream, because popular Internet services download more traffic than they upload; as expected, upstream utilization is less than downstream. Figure 15 shows that uplink utilization exceeds capacity for certain homes. Figure 16 shows a timeseries of capacity estimates and utilization for both of these homes. This user consistently saturates the uplink; when combined with "bufferbloat" problems problems endemic in home networking hardware, this behavior causes overestimation of the upstream throughput [19].

6.3 How much is each device used?

Figure 17 shows the fraction of traffic that individual devices contribute to overall home data consumption, ordered by the traffic consumption. Every household has at least three unique devices, but one dominant device typically consumes the bulk of the traffic (60% on average; the next most dominant device consumes about 20% of the traffic). Although we do not have ground truth about the types of devices in each home, it is obvious that even if people have multiple devices, they prefer to consume most data from a single device. Differentiating devices types (*e.g.*, laptops, tablets, phones, media players, etc.) from network traffic is ongoing research.

(a) *In this home, utilization always exceeds the measured upstream capacity. Upon further investigation, we discovered that this user continually uploads scientific data from his home network.*

(b) *Diurnal bursts in traffic can sometimes result in utilization exceeding the network's upstream capacity.*

Figure 16: *In some cases, uplink utilization exceeds estimate of capacity. Buffering in customer premises equipment ("bufferbloat") can result in situations where utilization exceeds the measured capacity. These two homes likely experience significant latency and performance problems due to constant uplink saturation.*

6.4 Which domains receive the most traffic?

We now examine the popular domains that users in home networks visit and how those sets of popular domains vary across different home networks. As described in Section 3, we measure statistics for network traffic to domains that match a whitelist of domain names based on the 200 most popular domain names from Alexa, plus any domains that users add to this list using a Web interface built into our router firmware [8]. The router anonymizes DNS lookups to all other domains.

Which domains are consistently popular? We first explore the domains that are the five and ten most popular domains in a significant number of home networks. Figure 18 shows this distribution; the dark bar indicates the number of times a particular domain appeared in the top five domains in a home network, and the lighter bar shows the number of home networks for which a domain appeared in the top ten. The most consistently popular domains on this list are as expected: Google, YouTube, Facebook, Amazon, Apple, and Twitter. Unsurprisingly, the tail is also quite

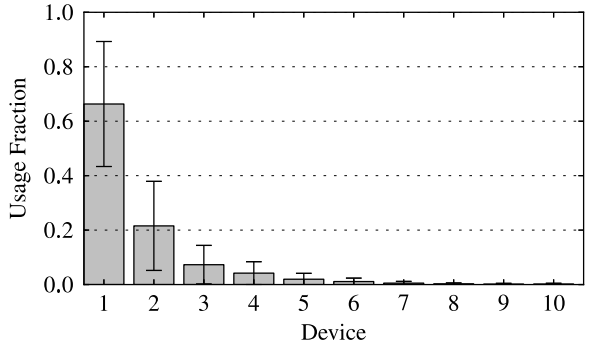


Figure 17: *Breakdown of data usage by device. We see that even in homes with multiple devices, there is usually a dominant device responsible for most of the traffic.*

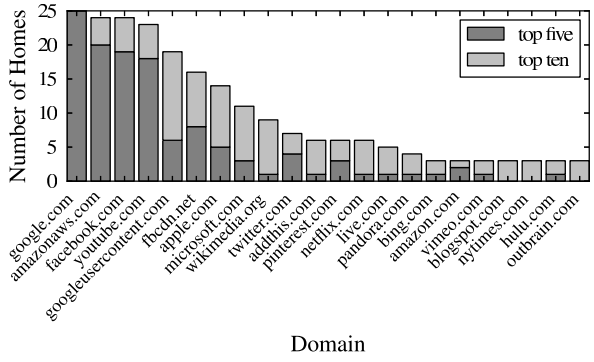
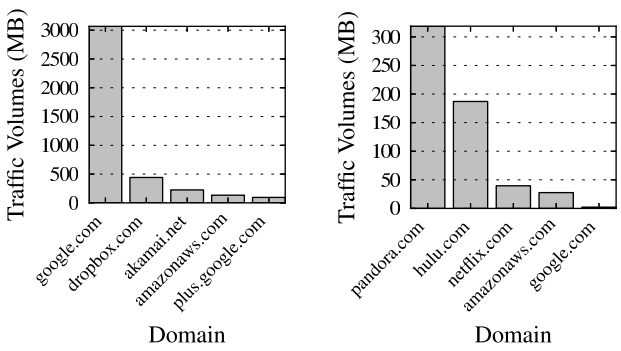


Figure 18: *The number of home networks in the Traffic data set for which a particular domain appeared in the top-five or top-ten domains, ranked by total traffic volume.*

long, with many domains being popular for just one or two homes (*e.g.*, streaming sites, news sites). This distribution confirms on a smaller scale other reports concerning the rise of "super peers" who are responsible for sending much of the content into access networks [29].

How much traffic are the most popular domains responsible for? Figure 19 shows the distribution of domains visited in terms of both total traffic volume and the number of connections made to the domain (which indicates the frequency of visits). Figure 19a shows the average volume of traffic from each home to the most popular domains in the whitelist. (We use rank indexes for domains rather than actual names, since the domain ranking is not exactly the same across all homes.)

The results show that the most popular whitelisted domain by volume on average accounts for about 38% of the total traffic volume, even though these sites are responsible for less than 14% of connections (Figure 19c); the next most popular domain accounts for about 11% of all traffic and only about 7% of all connections. This disproportionality reflects the fact that these popular domains are most likely serving streaming media content over long-running TCP connections. Hence, it is fairly safe to assume that these top two or three most popular domains by traffic volume are likely to represent streaming content, which would corroborate other reports that more than 40% of traffic into home networks is streaming video traffic. Figure 19b shows that the domain with the most number of



(a) *iMac (Desktop).* **(b)** *Roku box.*

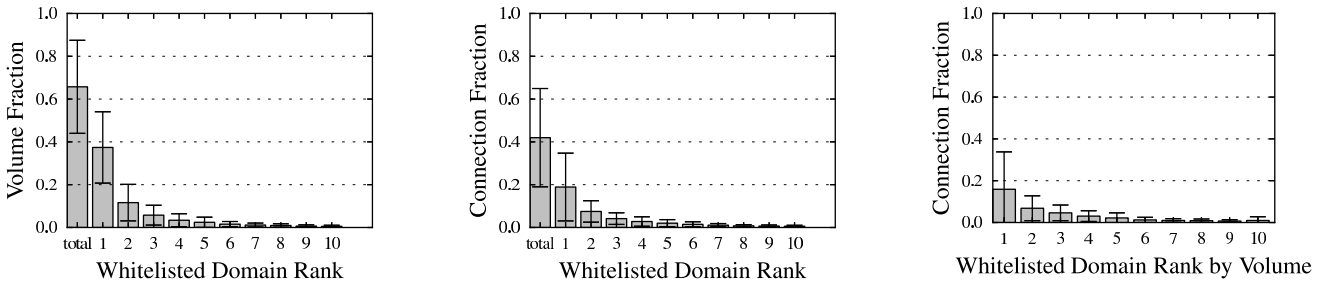
Figure 20: *Traffic distribution from an Apple iMac (desktop) and a Roku streaming player. For the desktop, Dropbox creates significant traffic to* `dropbox.com` *in the process of syncing large files. The Roku is used almost exclusively for streaming, as evidenced by the large fraction of traffic to* `pandora.com`*,* `hulu.com`*, and* `netflix.com`*.*

TCP connections is responsible for about 19% of all connections on average; again, this distribution has a very long tail.

It is worth cautioning that our anonymization of some of the domains in the Traffic data set could bias some of these results. In particular, if many homes in our data set for some reason sent a significant amount of traffic to domains that were not in our whitelist (*e.g.*, domains for pornographic content, which we explicitly removed from the whitelist), our results would not reflect this phenomenon. Nevertheless, because traffic to whitelisted domains represents about 65% of all traffic volume to and from our home networks on average, we believe that we have captured a representative sample of usage. Additionally, the Traffic data set currently only represents homes in the United States; as we begin to collect this data from home networks in other countries, we will be able to compare differences in domain popularity from different home networks.

Do different devices look up different sets of domains? We also examined domain popularity for different devices in home networks, to see whether the distribution of traffic volumes to domains differed by device. Our hypothesis was that certain devices might look up considerably different sets of domains than others. For example, an Apple device might exchange more traffic with `apple.com`, a streaming set top box might exchange more traffic with `hulu.com`, and so forth. If true, such a finding could prove extremely valuable for applications and utilities inside the home that want to automatically "fingerprint" devices—typically, the manufacturer ID of a device's MAC address may narrow down the device to a manufacturer, but it is not fine-grained enough to distinguish, say, a laptop from a smart phone.

To explore this hypothesis, we surveyed users from six homes in the Traffic data set and asked them to manually identify the devices corresponding to each of the MAC addresses in their home. These labels provide ground truth identification for these devices. Here, we show an example of how different devices send different distributions of traffic volumes to various domains. Figures 20a and 20b show the distributions for an Apple iMac Desktop and a Roku Streaming Player, respectively. Whereas a device's MAC address only reveals the manufacturer, further examination of traffic behavior suggests that usage patterns may differ significantly enough across *types* of devices to serve as fingerprints for device



(a) *Distribution of the most popular domains by mean traffic volume.*

(b) *Distribution of the most popular domains by mean number of connections.*

(c) *Fraction of connections for the most popular domains by volume.*

Figure 19: *Breakdown of data usage by domain. The most popular domain by volume consumes about 38% of total traffic. "Total" refers to the portion of traffic to whitelisted domains (Alexa top 200, plus any domains that the user manually whitelists) and accounts for about 65% of the total traffic on average.*

identification (either automatically, or with some help from the user). Exploring how these traffic patterns can assist with device fingerprinting is an area of future work.

7. DISCUSSION

This study offers a glimpse into the characteristics of a variety of home networks; the findings in the paper suggest many avenues for future work.

Combining network measurement with qualitative studies about Internet use. Previous work in home networking has explored various characteristics of home networks via detailed user studies and interviews [13, 14]. Our results can complement these studies, which have been primarily qualitative to date. Future work might entail doing a study that jointly performs user studies in conjunction with network traffic monitoring, to determine whether users' perceptions about their network use are consistent with the reality (*e.g.*, whether people spend more or less time online than they claim).

Device fingerprinting for security alerts. Various Internet service providers offer services that alert users about possible infected devices in the home network; unfortunately, because ISPs typically cannot map offending traffic to a particular MAC address, it is difficult for them to attribute traffic to a particular device. Future work could follow up on device fingerprinting using traffic patterns to develop a system that provides more fine-grained alerts to Internet service providers about the suspicious activities of an individual device within a home.

Expanding the study of usage to more countries. Our study of home network usage focused on the United States alone. Future work could expand this part of the study to determine how usage patterns and other traffic characteristics (*e.g.*, device usage, popular domains) differ by country.

8. CONCLUSION

Despite the proliferation of home networks, very little is known about the properties of these networks, in terms of their availability, infrastructure, or usage. Although continual, longitudinal measurements of these networks can provide insight into how people build, configure, and use these networks, there has been little attempt to instrument these networks to gather such data. This paper represents the first attempt to instrument a significant number of home networks to learn about their properties. We presented the first large-scale, longitudinal measurement study of home networks, based on data from 126 homes and 19 countries. Our set of passive and active measurements allows us to characterize properties of these networks such as network availability, infrastructure, and usage. We have publicly released all data sets that do not involve personally identifying information, and we plan to continue expanding the deployment into a broader, more diverse set of environments.

Our study yields many interesting findings with respect to the availability, infrastructure, and usage of home networks that could have broader implications for ISPs, users, and policymakers. With respect to availability, we found that developing countries experience far more frequent connectivity interruptions, some of which are due to poor connectivity, but others that are due to behavioral patterns (*e.g.*, turning the router on only during times when a user wants to access the Internet). More insights into how behavioral patterns differ across countries may help both ISPs and application designers. We also found that the 2.4 GHz spectrum is significantly more crowded than the 5 GHz spectrum, both in terms of number of devices and in terms of the number of visible access points; more widespread statistics about the usage of wireless spectrum (as will hopefully be possible as we continue to expand the BISmark deployment) can ultimately help ISPs debug connectivity problems in home networks and provide policymakers important data about spectrum usage. Finally, we see that most of the traffic from homes is destined for only a few domains, and that, on average, most of the traffic originates from just a small handful of devices. These usage statistics may ultimately help ISPs with provisioning and peering decisions, and they also offer a rare picture into how people use and interact with their home networks. Although this study has offered a first glimpse into many aspects home networks around the world, we expect that more lessons will come with more experience and a broader deployment.

Acknowledgments

This research was funded in part by NSF award CNS-1059350 and a Google Focused Research Award. We thank the hundreds of BISmark users, without whom this research would not have been possible. We also thank our shepherd, Yashar Ganjali, and the anonymous reviewers for their valuable feedback.

REFERENCES

[1] Internet world stats. `http://www.internetworldstats.com/`.

[2] Netgear WNDR3800. http://www.netgear.com/home/products/wirelessrouters/high-performance/wndr3800.aspx.

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

[4] Reports from Federal Communications Commission. http://www.fcc.gov/reports. FCC.

[5] SamKnows & Ofcom Working together for better broadband. http://www.samknows.com/broadband/signup/ofcom.

[6] Sign up with us today to accurately measure your broadband performance. https://www.samknows.eu/.

[7] The World in 2011: ICT Facts and Figures. http://www.itu.int/ITU-D/ict/facts/2011/material/ICTFactsFigures2011.pdf. International Telecommunication Union.

[8] Top Sites in United States. http://www.alexa.com/topsites/countries/US. Alexa.

[9] The World Bank GDP per capita. http://data.worldbank.org/indicator/NY.GDP.PCAP.CD, 2011. The World Bank.

[10] Y. Benkler. Next generation connectivity: A review of broadband internet transitions and policy from around the world. Working paper 8, Regulation2point0, Oct. 2009.

[11] K. L. Calvert, W. K. Edwards, and R. E. Grinter. Moving toward the middle: The case against the end-to-end argument in home networking. In *Proc. 6th ACM Workshop on Hot Topics in Networks (Hotnets-VI), Atlanta, GA*, 2007.

[12] J. Chen, S. Amershi, A. Dhananjay, and L. Subramanian. Comparing web interaction models in developing regions. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10, pages 6:1–6:9, New York, NY, USA, 2010. ACM.

[13] M. Chetty, R. Banks, R. Harper, T. Regan, A. Sellen, C. Gkantsidis, T. Karagiannis, and P. Key. Who's hogging the bandwidth: the consequences of revealing the invisible in the home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 659–668, New York, NY, USA, 2010. ACM.

[14] M. Chetty, D. Haslem, A. Baird, U. Ofoha, B. Sumner, and R. Grinter. Why is my internet slow?: making network speeds visible. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1889–1898, New York, NY, USA, 2011. ACM.

[15] M. Chetty, J.-Y. Sung, and R. E. Grinter. How smart homes learn: the evolution of the networked home and household. In *Proceedings of the 9th international conference on Ubiquitous computing*, UbiComp '07, pages 127–144, Berlin, Heidelberg, 2007. Springer-Verlag.

[16] L. DiCioccio, R. Teixeira, M. May, and C. Kreibich. Probe and pray: using upnp for home network measurements. In *Proceedings of the 13th international conference on Passive and Active Measurement*, PAM'12, pages 96–105, Berlin, Heidelberg, 2012. Springer-Verlag.

[17] L. DiCioccio, R. Teixeira, and C. Rosenberg. Measuring home networks with HomeNet Profiler. In *Passive & Active Measurement (PAM)*, Hong Kong, Mar. 2013.

[18] M. Dischinger, A. Haeberlen, K. P. Gummadi, and S. Saroiu. Characterizing residential broadband networks. In *Proc. ACM SIGCOMM Internet Measurement Conference*, San Diego, CA, USA, Oct. 2007.

[19] J. Gettys. Bufferbloat. http://www.bufferbloat.net/.

[20] R. Grinter, W. Edwards, M. Chetty, E. Poole, J. Sung, J. Yang, A. Crabtree, P. Tolmie, T. Rodden, C. Greenhalgh, et al. The ins and outs of home networking: The case for useful and usable domestic networking. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(2):1–28, 2009.

[21] R. Grinter, W. Edwards, M. Newman, and N. Ducheneaut. The work to make a home network work. In *ECSCW 2005*, pages 469–488. Springer, 2005.

[22] International Monetary Fund. World economic outlook database. http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx, Apr. 2013.

[23] W. Keith Edwards, R. Mahajan, and D. Wetherall. Advancing the state of home networking. *Communications of the ACM*, 54(6), 2011.

[24] H. Kim, S. Sundaresan, M. Chetty, N. Feamster, and W. K. Edwards. Communicating with caps: Managing usage caps in home networks. In *ACM SIGCOMM Computer Communication Review*, volume 41, pages 470–471. ACM, 2011.

[25] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: Illuminating the edge network. In *Proc. Internet Measurement Conference*, Melbourne, Australia, Nov. 2010.

[26] Netalyzr. http://netalyzr.icsi.berkeley.edu/.

[27] E. S. Poole, M. Chetty, R. E. Grinter, and W. K. Edwards. More than meets the eye: transforming the user experience of home network management. In *Proceedings of the 7th ACM conference on Designing interactive systems*, pages 455–464. ACM, 2008.

[28] M. A. Sánchez, J. S. Otto, Z. S. Bischof, and F. E. Bustamante. Trying broadband characterization at home. In *Passive & Active Measurement (PAM)*, Hong Kong, Mar. 2013.

[29] Sandvine. Global Internet Phenomena Report. http://www.sandvine.com/news/global_broadband_trends.asp, 2012.

[30] Shaperprobe. http://www.cc.gatech.edu/~partha/diffprobe/shaperprobe.html.

[31] T. N. Smyth, S. Kumar, I. Medhi, and K. Toyama. Where there's a will there's a way: mobile media sharing in urban india. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 753–762, New York, NY, USA, 2010. ACM.

[32] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè. Broadband internet performance: A view from the gateway. In *Proc. ACM SIGCOMM*, Toronto, Ontario, Canada, Aug. 2011.

[33] S. P. Wyche, T. N. Smyth, M. Chetty, P. M. Aoki, and R. E. Grinter. Deliberate interactions: characterizing technology use in nairobi, kenya. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2593–2602, New York, NY, USA, 2010. ACM.