Consolidated Review of

Peeking behind the NAT: An Empirical Study of Home Networks

1. Strengths

Home networks are an important component of Internet ecosystem. However, due to the difficulty of measurement instrument deployment and data collection inside home networks, there are few studies in this space. This paper leverages the actual deployment of BISmark to hundreds of home networks, and sheds light on the availability, infrastructure and usage of these home networks. The authors have put a lot of effort collecting this data compared to similar prior attempts that usually rely on remote measurements. The duration of the measurements is also long enough to make measurements meaningful. Some of the results are interesting and show potential of the measurement approach.

This paper was a joy to read. Very well written, and interesting measurement results.

2. Weaknesses

Some statements can be more accurate. For instance, a claim that "we observed that in the United States, most users leave their routers powered on all of the time", can become more clear by rigorously defining what is meant by "most users".

Also, more in depth information about data could've been provided.

The deployment is really small (and potentially biased, see below). The dataset is fairly small, especially when subsetted in several ways, and this limitation goes nearly without mention. It makes the dataset more useful for comparison among different regions rather than representing general behavior.

Outage analysis conflates user and ISP outages. Talking about reliability and characterization of outages when they often really measure user behavior is technically incorrect.

Most of the findings are not really surprising. One of the most interesting results is about the non-saturation of access link capacity. However, it is difficult to understand how such findings are affected by the bias (bias highlighted by the authors) in the population of measurement routers deployed.

The impact of the results is not clear. The goals/claims are a bit weak for a long paper: "preliminary view into many home networks" (preliminary); "show how measurements from a home router can yield significant information about home networks" (not surprising).

3. Comments

This paper is a first look at the important problem of what happens inside people's homes. The data is not perfect, but as a look in a new area, this work seems important.

Overall, this is a nicely written paper. It is very nice to see the authors are releasing measurement data collected from the experiment to the networking research community. Due to the difficulty of deploying customized routers in large scale home networks, this data will be valuable to many researchers who want to study home networks.

Parts of the paper that are generic descriptions could be shortened in favor of providing more detailed information and data analysis.

This paper considers an outage as a gap in the heartbeat logs lasting longer than 10 minutes. What is the distribution of the temporal gaps of the heartbeat logs? Are there a lot of gaps less than 10 minutes? Rather than using Tables 3-5 to summarize the highlights of findings, this paper could use a short summary at the end of the sections.

It is very difficult to understand the highlights without first reading the details of the results. The highlights provided in the paper are great, but given the strict page limits, they could've been traded for more data analysis. I think italicized text in the paper can have the same effect, without the need for extra space.

This paper is unable to distinguish access network failures and the routers being turned down, both of which cause persistent heartbeat losses. Why not use a simple script (or a self-monitoring tool) that continuously writes time-stamps or other information to a log file? The log file will have a gap if the router is turned down. In this way, one can distinguish network failures and router failures.

I have serious problems in calling an "outage" when the owner of the home router simply switches it off before going to sleep! I think the word outage should be removed.

This paper lists the geographical locations of 126 routers. It would be interesting to show the breakdown number of home routers in each country. For instance, per-country charts (at least for countries that there are enough data points available) would've been very interesting. Explicitly mentioning the limitations of the data set is a very good idea that can help limit the potential future generalizations and conclusions reached based on this data set by others.

A concern about the work is that it ends up making very sweeping conclusions about developing countries from a very, very small sample size. It starts with 36 routers in developing countries, but when one subsets it, outages in developing countries is based on 27 users. A pretty small sample for 4 billion people. I recognize the challenge to deploy infrastructure, but what about the huge variation across both the developed and developing world, where outage rates may vary hugely between rural, suburban, and urban locations? Should I assume your 27 developing users are all in cities? Given the small sample size, averages in figures 8 and 9 could easily be skewed by an individual with a single device, or 10 devices. Why not show distributions? (The bars do have whiskers, which are good, but I don't see their definition. Standard deviation? 95%ile? Please share. But CDFs or characterization of the distribution would be better.) For devices (page 7-8): is your analysis over the whole dataset? If so, do visitors skew the data? (If I have a LAN party and 20 friends, does that mean my home seem to have 21 devices forever?)

Same for "reliable", see title of section 4.1 "How reliable is home broadband access". This is not only about wording. The whole section 4.1 refers to reliability and outages whereas there simply

seem to be different behaviors/habits in different regions of the world (cost of electricity is one of my first guesses about the causes). The authors present the paper as the "the first empirical study of home network availability, infrastructure, and usage, using data collected from home networks around the world." However, reference [17] does something very similar (and with a deployment 1 order of magnitude larger) except that they don't measure from the home router but from the end hosts. The authors should better relate to reference [17], especially when they discuss findings on topics on which there is overlap, such as the wireless spectrum. First paragraph of Section 6.1 is a bit vague. Section 6 seems the most interesting. The paper is well written and very readable.

4. Summary from PC Discussion

Discussion summary from PC meeting:
- ❖ Good paper, interesting results (maybe not surprising).
- ❖ Need to distinguish between availability and user behavior.

5. Authors' Response

We are grateful to the anonymous reviewers for their feedback and made changes in response to their comments.

We added a breakdown of the number of routers in each country.

We also elaborated on several unclear statements by adding numbers or distributions to support our claims. One reviewer suggested we substantiate our claim that most users in the United States keep their routers on all of the time. We found that the median time that a router is on across the 63 routers in the United States is 98.25% of the collection period, which quantitatively supports this claim.

Some reviewers felt the word *outage* was inappropriate to describe situations when users voluntarily powered off their router, so we replaced *outage* with *downtime* and provided a clearer description of the nature of the downtimes we observed. Although our dataset makes it difficult to distinguish between network outages and users who power off their routers, we did our best to identify causes of downtime.

We clarified how we count the number of devices in a home by distinguishing between devices we see briefly (*guest devices*) and devices that we see for a longer time.

We added an explanation of the error bars in Figures 8 and 9. Although we considered replacing these figures with box plots to highlight the nature of the distributions more clearly, the distributions actually have fairly low variance; thus, bar plots conveyed the intended information in a simpler format, so we left the figures unchanged.

One reviewer suggested that our work overlaps with HomeNet Profiler [17], yet HomeNet Profiler does not perform longitudinal measurements of devices on the network, comparisons of 2.4 vs. 5 GHz, usage characterization, and so forth. In fact, the only real overlap is an analysis of device counts in home networks, but our work even contributes on that front with longitudinal measurements. We clarified these differences and more thoroughly compared the two studies.