Need, Want, or Can Afford Broadband Markets and the Behavior of Users

Z. Bischof, F. Bustamante and Rade Stanojevic
Northwestern U. | °Telefonica Research
IMC 2014
The importance of broadband

- One of the fastest growing sides of the Internet
  - Over 70 countries with 25% of the population online

- Key for social and economic development
  - Increasing capacity may increase GDP
  - National broadband plan in many countries
  - A basic human right in several countries
Several efforts on broadband characterization

- Web-based, from routers, gateways and end hosts

Target characterizing coverage & network-level performance, but miss broader context

- Service may be available, but is it affordable?
- How does link capacity affect user behavior?
- How do performance and reliability impact usage?
- Does the cost of upgrading services affect demand?
Study’s methodology

**Goals**
- Explore impact of capacity, price, cost of upgrading and connection quality on ..
- Broadband user’s behavior – usage as a stand-in
- Check for longitudinal usage/capacity trends

**Challenges**
- Dataset covering a range of broadband markets
  - Performance, usage, relative costs …
- Experimentation at scale
  - Different ISPs, technologies, regions, contexts …
- Natural settings
- Reproducibility
Control and natural experiments

- Classical controlled experiments
  - Control and treatment groups randomly selected
    - If large enough, roughly equivalent groups
  - Difference in outcome likely due to treatment

- Reproducibility, but
  - Poor scalability
  - No natural settings
  - Ethical and practical issues

- Instead …
Natural experiments and related study designs
- Common in epidemiology and economics
- Assignments to treatment is as-if random, controlling for co-founding factors
  - E.g., impact of military service on lifetime earnings, pump location and cholera
  - Cholera’s method of transmission
  - Land titling’s impact on property investment and child education

London’s cholera epidemic, 1854
Datasets

- Broadband performance and usage
  - AquaLab’s Dasu and FCC/SamKnows
    - 53,000 users in 160 countries (6,000 from FCC/SamKnows)
  - Capacity, loss, latency, network usage
  - World-wide (Dasu) and US (FCC)

- Retail broadband connectivity plans
  - ~1,800 broadband plans, ~100 countries
  - Monthly cost adjusted based on purchasing power parity (PPP)
How are service capacity and usage related?
– Do users with higher service capacities generate more traffic than those with lower capacities?
– Does user demand increase on faster networks?
Service capacity and usage

95th percentile usage

At ~12 Mbps, about 10% of capacity
Confidence intervals begin overlapping at ~12 Mbps

At ~1 Mbps, only using about 30% of capacity

Relationship follows a law of diminishing returns

Error bars represent 95% confidence interval

$r = 0.885$
Does the relationship changes over time?
  – Considering growth on data avail and Internet traffic
  …
Usage consistent over 3 years for a given capacity
Does this mean users don’t need higher capacity than what they have?

– Or, would their demand increase if upgraded?
– A natural experiment to determine impact of upgrades
  • Same user before and after a “service upgrade”
95\textsuperscript{th} percentile usage before and after upgrade

For 70\% of users, peak usage increased after upgrading.

And impact of upgrade is higher for lower initial capacities.

$p < 10^{-24}$
Affordability – Cost of access

How does the cost of access impact usage?
– Are similar services in expensive markets used more?
The impact of cost of access

- Cost is a relative term
  - Adjust for purchasing power (PPP)

- Cost of broadband services varies by market
  - $50/mo (PPP) for 1 Gbps in Japan
  - $100/mo (PPP) for 512 Kbps in Botswana

- Use monthly price of cheapest 1+ Mbps service as an estimate of the country’s “cost of broadband”
The impact of cost of access

- Nearest neighbor matching with caliper

In markets where access cost was >$60/mo usage was higher 72% of the time

\[ p < 10^{-9} \]
## Market features – Case study

<table>
<thead>
<tr>
<th>Country</th>
<th>Common service (Mbps)*</th>
<th>Price in PPP USD (relative cost)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Botswana</td>
<td>0.512</td>
<td>$100 (7.1%)</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>4</td>
<td>$78 (3.8%)</td>
</tr>
<tr>
<td>US</td>
<td>20</td>
<td>$53 (1.3%)</td>
</tr>
<tr>
<td>Japan</td>
<td>30</td>
<td>$38 (1.3%)</td>
</tr>
</tbody>
</table>

*Defined as the plan closest to median user’s capacity
+As % of monthly GDP per capita

1/60th the capacity of Japan at 2.5x the cost!
Cost of access – Case study

More expensive services consistently see higher utilization
Cost of access – Case study

Peak utilization in the US

Fraction of the link utilized decreases as capacity increases

The graph shows the cumulative distribution function (CDF) of users across different capacity tiers in the US. The fraction of the link utilized at the 95th percentile is plotted against the CDF of users for different capacity ranges: <1 Mbps, 1-8 Mbps, 8-16 Mbps, 16-32 Mbps, and >32 Mbps. The trend indicates that as capacity increases, the fraction of the link utilized by users decreases, suggesting more efficient use of bandwidth.

[Graph showing CDF against 95th percentile link utilization for different capacity tiers]
Utilization for same tier is higher in Botswana
Is the cost of upgrade holding people back?
  – Is utilization higher in markets with expensive upgrades?

Cost of upgrade also varies by market
  – Approximate upgrade cost using linear regression
Cost of upgrading

Cost of access similar in US and Japan (~$25), but upgrading is cheaper in Japan.
US demand and utilization are higher than Japan
• In general, more expensive services see higher utilization …
  – But India’s cost of service >3x that of the US, and utilization is lower?!?
  – Many with high latency (>500 ms) and loss (>10%) …

• Does service quality impact service usage?
  – Comparing usage of “similar” (nearest neighbor) users
Connection quality – Latency

Control group (512, 2048]

With lower latencies, higher utilization (and vise versa)

50% - random

\[ p = \{0.033, 0.00766, 0.0062, 0.00825\} \]
Summary

- Context and network measurement up-the-stack
  - Up-the-stack – engagement, demand, management complexity …
  - For broadband – economic factors, service dependability, alternative services, …

- New experimental and study designs models
  - From social science, psychology, economics …
  - Natural experiments, quasi-experiments, …
Today

- Internet design principles and measurements
- Standards and strategies
- Experimental approaches

**Ethical considerations**