Complementary realities: Public domain Internet measurements in the development of Canada’s universal access policies

Reza Rajabiun & Fenwick McKelvey

To cite this article: Reza Rajabiun & Fenwick McKelvey (2019): Complementary realities: Public domain Internet measurements in the development of Canada’s universal access policies, The Information Society

To link to this article: https://doi.org/10.1080/01972243.2019.1574533

Published online: 25 Mar 2019.
Complementary realities: Public domain Internet measurements in the development of Canada’s universal access policies

Reza Rajabiun and Fenwick McKelvey

ABSTRACT
As access to the Internet has become increasingly essential for social and economic participation, public domain Internet measurements have become indispensable for users to validate quality of service their network operator delivers and for policymakers to identify and address gaps in broadband infrastructure. This article evaluates public domain Internet performance measurements available for assessing the state of connectivity and developing universal access service quality standards in Canada. The analysis suggests that different approaches to Internet measurement represent complementary windows into a complex and fast evolving reality of broadband connectivity. Despite their potential shortcomings, large-scale crowdsourced open data network testing platforms have a central role to play in enabling broadband infrastructure policy coordination across different levels of government, empowerment of consumers, and achievement of universal service objectives for quality of service users experience when accessing the open Internet.

KEYWORDS
Canada; digital divide; internet connectivity; open data; universal access

Introduction
While there is significant agreement about the importance of broadband network infrastructure quality for the equitable development of the information society, there is little consensus about how it should be measured (Lehr, Bauer, and Clark 2013). Demand for information about actual speeds and quality of service has led to a variety of methodologies purporting to offer realistic estimates of speeds and quality (Hong and Morris 2016). These Internet measurements are routinely used (and abused) in telecommunications policy debates around the world. In this article, we analyze the application of public domain Internet performance in universal service policymaking in Canada (Rajabiun 2017a).

As in many other advanced economies, in Canada Internet measurement serves as an increasingly important means for evaluating broadband market outcomes, identifying capacity gaps, and developing policy responses to these gaps. In 2016, the Canadian Radio-television and Telecommunications Commission (CRTC) reclassified both fixed and mobile broadband Internet as a “basic service.”¹ It also set a goal: 90% of Canadians should have access to 50 Mbps download speed by 2021. Further, it planned in the near future to set standards for latency, jitter, and packet loss. These service quality standards have been adopted. They are to serve as evaluation criteria in the allocation of broadband infrastructure subsidies.

However, most data about Internet performance remain in the private domain, as network operators consider them to be sensitive information from competitiveness standpoint. The limited number of public domain Internet measurements restricts comparative analysis. We begin with an environmental scan and review of the Internet measurement methodologies used in Canada to flag the numerous obstacles to meaningful comparison of measures derived from different methodologies. Next, we analyze public data from the M-Lab Network Diagnostic Test (NDT) to better understand key sources of variation underlying aggregated network performance indicators and their implications for achieving universal service policy targets.² We conclude by highlighting the importance of consistent public domain Internet measurements for reducing information asymmetries in the retail market for Internet access services and also for monitoring connection quality users are actually experiencing.

CONTACT Fenwick McKelvey fenwick.mckelvey@concordia.ca Department of Communication Studies, CJ 3.230, 7141 Sherbrooke St W, Montreal, Quebec, Canada H4B 1R6.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/utis.

Published with license by Taylor & Francis © Reza Rajabiun and Fenwick McKelvey
Consistent public domain Internet measurements would enable evidence-based decision making by policymakers, consumers, and providers of content and application services.

Origins and applications of Internet performance measurements

Internet measurement is as old as the Internet itself. Computer science and electrical engineering researchers building prototypes of first digital networks had to devise methods to evaluate the performance of these new systems (McKelvey 2018). The field of Internet measurement emerged out of this research (Bauer, Clark, and Lehr 2010; Dischinger et al. 2007; Faris and Heacock Jones 2013; Lehr, Bauer, and Clark 2013; Prasad et al. 2003). Early on, researchers recognized the importance of Internet measurement data. As Vint Cerf, one of the key engineers behind the development of Internet’s protocols, highlighted more than 25 years ago, from the beginning public domain measurements were considered “vital to research and engineering planning activities, as well as to ensure the continued development of the operational infrastructure” (Cerf 1991, n.p.). A wide range of technical standards and indicators have been devised over the past decade that enable engineers managing large scale network infrastructure better understand and optimize their networks.

Internet measurement has largely been privatized along with the rest of the core Internet (Abbate 1999; Ceruzzi 2008; Rybczynski 2009). Early commercial networks did not standardize measurements; instead, network metrics largely became proprietary corporate data in the provision of public access to the Internet (Claffy 2007). With the convergence of different communications services onto Internet Protocol (IP)-enabled networks (Bar and Sandvig 2008; Winseck 2002), network performance data assumed strategic importance and is now considered highly confidential. Since private network operators are not usually willing to share their internal data, those who want to validate quality of service usually have little option but to do so themselves.

The information asymmetry between network providers, on the one hand, and consumers and policymakers, on the other, has important implications for the governance of the broadband ecosystem and the quality of access to the open Internet. While larger organizations may have the resources to negotiate Service Level Agreements (SLAs) with Internet access providers and validate performance of their supplier, in the mass residential and small and medium size (SME) business markets, technical and resource constraints of users have traditionally limited the scope for this type of decentralized monitoring. Furthermore, standard form “best effort” contracts in which sellers promise only a maximum theoretical speed up to x Mbps are the norm in residential and SME markets. This makes it challenging for buyers to hold sellers accountable for the actual speeds they deliver relative to what the customer was expecting per their retail contract with the service provider (Rajabiun 2017b). The asymmetry of information between buyers and sellers in the retail market for Internet connectivity can generate perverse incentives as in other markets, where buyers cannot easily observe the quality of what they are paying for (i.e. The Lemons Problem a la Akerlof 1970). In addition to its negative impact on consumer welfare and protection of vulnerable consumers (Sluijs, Schuett, and Henze 2011), the information asymmetry between buyers and sellers has broader implications for the technological evolution of the networks: when potential investors in higher quality sunrise platforms cannot credibly signal their superiority compared to lower quality suppliers of legacy technologies ex post, this can reduce their ex ante incentives to invest in new technologies and deliver the superior services (Rajabiun and Middleton 2015).

Public domain Internet measurements have emerged as the only available tool for independent monitoring of the quality of physical broadband access networks. These public tools – often relying on crowdsourced user participation – provide some of the few indicators of local and global broadband performance. Although competition in the market for Internet testing platforms has expanded their range, the proliferation of data sources has exacerbated confusion as many of these tests are sponsored by the sellers to validate their speed claims to customers and there is limited standardization. There is little consensus among operators, policymakers, and experts about the relevance of particular sources of Internet measurements and outcome indicators, leading national regulators and other public agencies to adopt different measurement approaches (Bajpai and Schönewälder 2015; Lehr, Bauer, and Clark 2013). Furthermore, data when available tend to be designed for the needs of the computer science community, complicating its application to public policy (Asghari, van Eeten, and Mueller 2013). Despite these challenges and potential limitations of particular testing methodologies, public domain Internet measurement from large-scale testing
platforms represent a critical source of knowledge for market participants and policymakers.

**Internet measurements in Canada**

There are five major global testing platforms commonly cited in telecom policy debates around the world: Ookla’s Speedtest, Measurement Lab’s NDT, Réseaux IP Européens’s (RIPE) Atlas Program, Samknows, and indicators from Akamai Technologies State of Internet Report. Each has a distinct methodology. Ookla/Speedtest and M-Lab both rely on a crowdsourcing methodology that involves users actively initiating a test, while Akamai measures connection quality passively while optimizing the delivery of content and applications for its clients. RIPE Atlas and SamKnows represent testing platforms where the measurements are collected by a specialized white-box placed in a particular vantage point within the network (e.g. “probes” on edge of the cloud in nearby large city, local switching facilities, home and business routers).

All these types of initiatives are underway in Canada to varying degrees. Ookla and Akamai have both been present in Canada for some time. By contrast, SamKnows and M-Lab have only recently begun testing. RIPE Atlas is just beginning to expand into Canada. The expansion of Internet measurement programs began in 2015 with both CRTC and the national domain registrar the Canadian Internet Registration Authority (CIRA) launching their own programs. CIRA deployed a set of test servers based on the standardized M-Lab/Network Diagnostic Tool (NDT) methodology in a number of urban centers across the country (Calgary, Montreal, and Toronto). CRTC commissioned the network performance testing company SamKnows (in a no-bid contract) to measure broadband service in Canada.5 Table 1 provides an overview of key sources of public domain fixed broadband performance indicators available to users and utilized in network mapping and broadband policy development in Canada.6 It summarizes the widespread differences in the architecture, testing method, and data license of these programs, which are discussed at depth below.

**Architecture**

Internet measurements estimate connection performance by measuring transmission rates between two points. Usually, these points are described as server side and client side. Most initiatives, including M-Lab NDT, use off-net servers outside the ISP’s infrastructure, which reflects widespread consensus around off-net testing as a preferred vantage point to measure the end user experience when accessing content and applications on the open Internet (i.e. outside of the network of their operators; as opposed to content applications cached within that network). By contrast, Ookla uses an on-net approach albeit without strict guidelines on the location of the server leading to inconsistent results. This lack of standardization in this testing methodology leads to substantive inconsistencies in measured speeds from the Ookla/Speedtest platform across jurisdictions and service providers (Sandvine Inc. 2013). In general, client locations are either through a user device, such as a test running in a browser of a home computer, or on a white-box (probes). The SamKnows test bed the CRTC has commissioned utilizes around 5000 white-boxes on the customer’s premise, which helps control for potential bottlenecks in home networks and other variables that can impact access link capacity to off-net servers. While increasing the stability of data, using specialized equipment adds significant barriers to participation.

<table>
<thead>
<tr>
<th>Metric/provider</th>
<th>Akamai (avg.)</th>
<th>Speedtest (avg.)</th>
<th>SamKnows (unknown)</th>
<th>M-Lab/CIRA (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation method</td>
<td>Passive</td>
<td>Active</td>
<td>Active</td>
<td>Active</td>
</tr>
<tr>
<td>Architecture</td>
<td>Home user to Content Delivery Network servers</td>
<td>Home user to various point (off-net or on-net) located across Canada</td>
<td>White-box (probes) to off-net measurement server located in Halifax, Montreal, Toronto, Vancouver, and Winnipeg</td>
<td>Home user to off-net measurement server located in Calgary, Montreal, and Toronto (with nodes planned in Moncton, Vancouver, and Winnipeg)</td>
</tr>
<tr>
<td>Testing method</td>
<td>Measured while delivering apps</td>
<td>User-initiated, crowdsourced</td>
<td>Hardware-initiated, fixed testing program, curated participation</td>
<td>User-initiated, crowdsourced</td>
</tr>
<tr>
<td>Data license</td>
<td>No</td>
<td>Commercial</td>
<td>Commercial</td>
<td>Public domain, value-added commercial services</td>
</tr>
<tr>
<td>Sample size</td>
<td>Large</td>
<td>Large</td>
<td>Small</td>
<td>Large</td>
</tr>
</tbody>
</table>
including cost and finding volunteers, complicating large-scale deployments.

**Testing method**

Internet measurement is either passive (i.e. made during the normal course of operations) or active (i.e. made when test packets are injected into the network to observe their behavior) (Murray and Claffy 2001). Active testing further varies in its testing schedule, either initiated by users on demand or set up on a predetermined schedule. Akamai estimates connection performance while delivering content and application services from its global Content Delivery Network (CDN) to users. As such, indicators from Akamai can be viewed as an example of passive tests where the measurement is taken while the user is trying to access content and application services from Akamai’s CDN. SamKnows hardware placed in customer premises runs on a fixed schedule, typically when the connection is not under a heavy load from other devices on the home network. Ookla and M-Lab rely on user-initiated testing whenever a client elects to run a test. This is likely to lead to tests taking place more frequently when users want to deploy applications they require, but the quality of service they are able to achieve falls short of their expectations. User-initiated testing approaches therefore tend to capture network quality under normal use conditions.

There is no clear consensus about the preferred methodology or testing schedule. Passive testing offers more “real-world” results, though only as relevant for the delivery of a narrow range of services and applications (e.g. Akami’s CDN, Google/YouTube Video Quality Report, Netflix ISP Speed Index), whereas active testing can help evaluate wider infrastructures and develop a broader perspective (e.g. Ookla, M-Lab). SamKnows white-boxes enable a high degree of scheduling control, but its testing schedule might omit periods of high home use that might be more indicative of performance users experience under normal use conditions.

Tests also differ in how they simulate the experience of transferring data. Ookla tests use multiple threads to increase speeds and discard a proportion of slower measurements during a test, while M-Lab’s NDT tool uses a single-threaded methodology (Bauer, Clark, and Lehr 2010). Both represent Internet traffic, but there is significant debate over which is more representative. Bennet (2016) suggests that the Ookla/Speedtest approach is superior because its accounts for the fact that modern web browsers utilize multiple TCP streams. The difference may be relevant for evaluating multi-threaded web applications, but as documented in the subsequent section there is little evidence to suggest that single threaded tests underreport speed.

**Data license**

Only M-Lab releases raw data for public analysis. Notably, even the regulator does not release performance data from the SamKnows test bed it has commissioned (unlike the similar Measuring Broadband America program). In its three-year plan, the public agency has committed to releasing this data, but there is currently no schedule for it. The commercial nature of many of the initiatives also raises issues related to the transparency of the testing and validation of their results.

**Problem of comparability**

Different testing methodologies capture different perspectives that can complement each other in evaluating Internet connectivity. Table 2 provides a perspective on the different measurement of broadband speeds provided by the different providers discussed above. While there are no public estimates of speeds from the CRTC SamKnows test bed, they do publish high-level estimates of some quality of service metrics. The results are consistent with those from Ookla/Speedtest in that they tend to represent the upper bound of observed speeds. The Ookla/Speedtest and SamKnows indicators may therefore represent a good proxy for assessing maximum theoretical capacity of links (respectively from computer to on-net server and from white-box to off-net server). Broadband performance outcomes captured by Akamai and M-Lab are broadly consistent as indicators of a complete path (from computer to off-net server), despite their methodologies differing significantly.

<table>
<thead>
<tr>
<th>Metric/provider</th>
<th>Akamai (avg.)</th>
<th>Speedtest (avg.)</th>
<th>SamKnows (unknown)</th>
<th>M-Lab/NDT (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Download Mbps</td>
<td>14</td>
<td>44</td>
<td>+104% of advertised</td>
<td>19</td>
</tr>
<tr>
<td>Upload Mbps</td>
<td>11</td>
<td>11</td>
<td>+100% of advertised</td>
<td>7</td>
</tr>
<tr>
<td>Other indicators</td>
<td>Peak speed: 60 Mbps</td>
<td>Latency: 10-30ms Packet loss: 0.04-0.23%</td>
<td>Latency: 96ms Jitter: 300ms</td>
<td></td>
</tr>
</tbody>
</table>
Estimate of average download speeds from Ookla/Speedtest is in fact closer to that of average peak measured speeds as detected by Akamai, suggesting it represents more of an upper bound for possible speeds rather than an average of real speeds.10

The differences in the results from these different testing methodologies are consistent with findings in previous research (Bauer, Clark, and Lehr 2010; Lehr, Bauer, and Clark 2013; Rajabiun and Middleton 2015) on their design, what they measure, and potential biases. These differences in the results of the tests conducted with different methodologies are noticeable internationally as well. As Figure 1 shows, the range of broadband speeds across the tests in Canada (and the U.S.) has been generally smaller than those observed in leading countries from Asia and Europe (on the left-hand side of the figure) where operators have stronger incentives to decommission legacy copper/DSL (digital subscriber line networks) and deploy fiber-to-the-premises (FTTP) broadband access technologies. In absolute terms, estimate of the upper bound of average speeds from Ookla/Speedtest in 2014 for Canada was about 20 Mbps, which was over two times higher than measures of download bandwidth capacity from Akamai and M-Lab. Given that these tests are measuring different things and offer different windows into the world of Internet connectivity, this is not surprising. Nevertheless, differences across results of these tests, which are publicly available, feeds confusion and abuse in policy design and implementation. Development of policy relevant indicators from standardized public domain measurements can mitigate this problem.

**Application to universal service policy**

Broadband measurements have an immediate application in Canada in the context of the 2016 decision by the CRTC to redefine “high-speed” connectivity as a “basic service” under the *Telecommunications Act*.11 While the CRTC did not adopt legally binding universal access obligations to deliver high-speed connectivity or minimum standards of service, it increased aspirational speed targets for all Canadians from 5 Mbps download/1 Mbps upload speed to 50 Mbps and 10 Mbps, respectively. Based on data on “best effort” providers advertise, the CRTC “estimates that 82% of Canadians currently have access to fixed broadband Internet access services at speeds of at least 50 Mbps download and 10 Mbps upload.”12 Since more than 80% of Canada’s population live in a small number of urban centers, CRTC’s estimate based on advertised speeds that are theoretically available is consistent with the common assumption that gaps in universal access are only a concern for users in rural and remote communities. The agency nevertheless determined that “these speeds are to be the actual speeds delivered, not merely those advertised.”13

Internet measurement tools can be used to evaluate these claims and the possibility of reaching these goals as per the Commission’s timeline. Given that M-Lab is the only consistent public domain source of data, here we will use M-Lab data to provide a granular picture of the state of connectivity. Specifically, we explore key sources of variation in a sample of approximately four million individual M-Lab NDT connection performance tests between 2015 and
This analysis helps us identify challenges and opportunities in utilizing Internet measurements as a basis for benchmarking and monitoring the pace of progress in the development of Internet connectivity. We move from a high-level national perspective to increasingly lower level perspectives, which are useful for validating and interpreting measurements.

Figure 2 provides an overview of the evolution of download speeds and the size of the M-Lab database of tests in Canada since 2015. Although access to services with theoretical and actual speeds close to or higher than the CRTC’s 50 Mbps targets might be available in some communities where very high-capacity fiber access networks have been deployed, average and median measured speeds are substantially below the regulators’ aspirational benchmarks. Lack of growth in measured speeds in the presence of substantive private and public-sector capital expenditures on network infrastructure suggests that capacity supply is barely keeping up with growth in demand for network intensive content and application services.15

The rapid rise in the number of tests in 2016, we suspect, results primarily from Google starting to promote the M-Lab NDT platform on its search results.

Figure 2. Download speeds in Canada, January 1, 2015–March 21, 2017. Source: Measurement Lab.

Figure 3. Google promoting NDT Measurement Lab in search returns. Source: Screenshot from Google search for speed test via Chrome browser, 2018.
for “speed test,” which then allows people to easily run an NDT connection diagnostic test (Figure 3). Indicators of effective downstream bandwidth documented in Figure 2 suggest that volatility of the measurements declined as CIRA test servers were deployed in major Canadian cities and Google started to include a link to the M-Lab/NDT testing platform in its search results in the second half of 2016. The M-Lab/NDT sample sizes increased rapidly as more people started to leverage this measurement platform to ascertain effective connection speeds their providers are delivering to them. In addition to documenting the power of Google as a search platform, this highlights strong demand from subscribers for independent information about the quality of Internet access they receive from their service providers.16

M-Lab/NDT measurements for Canada suggest average measured speeds tend to be around two times the median rates. This reflects the “long tail” in the distribution of speed measurements (i.e. a large number of slower observations combined with a small number of very fast connections). For this reason, using average rates in universal access gap analysis can be misleading because it will overestimate basic service quality level due to the presence of a small number of very high-speed connections in sample. For high-level estimates required for gap identification and prioritization, median speeds are likely to offer a more realistic metric of basic service quality that is available to users. On the other hand, aggregating tests in terms of averages provides a more optimistic picture by capturing the impact of the small number of higher speed connections in the long tail of the data distribution.

Table 3 shows the distribution of tests below and above the CRTC aspirational target of 50 Mbps. In addition to the lack of growth in average, median, and fastest quintile measurements, the proportion of tests on the two sides of the CRTC universal service target remained relatively stable over this period. The fact that connections with measured speeds above 50 Mbps make up less than 10% of the total may appear to contradict estimate by the CRTC that 82% of Canadians currently can access services that meet this target. However, the two are perfectly consistent as the CRTC estimate is based on indicators constructed from maximum advertised speeds that are theoretically available in particular areas (defined in terms of hexagons with a diameter of approximately 5 km), while we are looking at measured speeds for accessing the open Internet (i.e. off-net servers). Services with speeds that meet CRTC targets may be available in some areas and parts of the hexagonal units but might be hard to actually achieve by users in the rest of that area; for example, due to distance from local fiber node. What is particularly notable from a technological perspective is that the share of connections above the 50 Mbps rate is consistent with the penetration of high-capacity fiber-to-the-premises (FTTP) connections in Canada, which also remains just below 10% of total broadband connections.17

These high-level aggregated estimates suggest gaps relative to CRTC’s aspirational targets tend to be pervasive, and achieving its universal service objectives in terms of actual speeds will remain a challenge. More generally, our findings show that measurement of Internet speeds over time can help evaluate accomplishment of universal service policy objectives. Nevertheless, aggregation hides significant variation in the data associated with the dynamic and differentiated nature of the market for Internet connectivity. We will now explore four different sources of variation in the data that are particularly relevant for using Internet measurements in universal access policy development.

**Diurnal variation**

Varying congestion on local aggregation links and routers affects performance. When service providers have not made sufficient investments in these local facilities, localized congestion can substantially degrade actual service quality compared to “best effort” rates the seller advertise. As Table 4 shows, measured speeds tend to vary around 20–30% across low and peak traffic periods of the day. Even during times of the day with relatively low traffic levels, average measured speeds were less than half of the 50 Mbps CRTC’s basic service target.

**Smart pricing/speed tiering**

Using flow control mechanisms and data caps on basic service plans, operators can try to maximize

---

Table 3. Sample size and speed distribution of M-Lab tests in Canada. Source: M-Lab.

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50 Mbps</td>
<td>356,246</td>
<td>1,324,309</td>
<td>1,778,865</td>
</tr>
<tr>
<td>50 Mbps and higher</td>
<td>25,058</td>
<td>93,576</td>
<td>108,482</td>
</tr>
<tr>
<td>Total</td>
<td>381,304</td>
<td>1,417,885</td>
<td>1,887,347</td>
</tr>
</tbody>
</table>
revenues from their fixed network assets with these “smart pricing” strategies (Odlyzko 2014; McKelvey 2018). Asymmetric information between sellers and buyers about the quality of service further disadvantages the buyers, as they are unable to accurately differentiate low from high quality offers in the market. To provide a sense of the extent of smart pricing through speed tiering, Figure 4 shows the distribution of downstream bandwidth M-Lab tests in Canada for a sample of 99,999 unique IP addresses. The sample distribution provides substantive empirical validation of the standard-based M-Lab NDT testing methodology by clearly capturing effective speed tiering practices of operators in the Canadian market for fixed broadband access services (i.e., tall peaks around 5, 10, 15 Mbps, and smaller ones around 30 and 40 Mbps). Despite potential limitations of the M-Lab/NDT in terms of both over and underestimating connection quality noted earlier, in practice it appears to offer a reasonably accurate window into the differentiated market for Internet access services.18

Table 4. Daily variation in broadband speeds (Mbps, 2016).

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Median</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midnight to Morning</td>
<td>6.9</td>
<td>16.5</td>
</tr>
<tr>
<td>9–5 h</td>
<td>8.4</td>
<td>20.3</td>
</tr>
<tr>
<td>Supper (17–19 h)</td>
<td>8.8</td>
<td>18.9</td>
</tr>
<tr>
<td>Prime Time (20–22 h)</td>
<td>8.2</td>
<td>18.2</td>
</tr>
<tr>
<td>Night (23–24 h)</td>
<td>7.7</td>
<td>17.1</td>
</tr>
</tbody>
</table>

Figure 4. Distribution of measured broadband speeds in Canada. Source: Measurement Lab/RIPEstat; Jan.-July 2016, n = 99,999 unique IP addresses.

Technological endowments and strategic choice

Previous research on the increasing unevenness of broadband network quality in Canada and the U.S. suggests that technological, organizational, and financial strategies of service providers are the key determinant of the quality of service they deliver to their customers (Rajabiun and Middleton 2018). Figure 5 shows the relationship between median measured download speeds and quality of service in terms of latency provided by large Internet service providers that dominate regional markets in Canada, as well as a number of smaller ones. Lowest performing providers tend to be smaller ones serving higher cost rural areas with long loop DSL, fixed wireless, and satellite “high speed” connectivity (e.g. Xplornet, EastLink, Saskatchewan Telecom, MTS). Median downstream speeds and latency rates for the largest two incumbent copper/DSL providers (Bell and Telus) were broadly similar at the beginning of 2016 (~10 Mbps and 40 ms, respectively). Cable companies and incumbents with accelerated FTTP deployment strategies (on the right hand side of the figure; e.g. Rogers, Cogeco, Bell Aliant) delivered the highest measured median speeds and lowest latency (~20 Mbps and 20 ms, respectively).

Notably, according to this data the highest performing service provider was Cogeco, a relatively small cable broadband provider that primarily serves
rural communities in Ontario and Quebec. Another one of higher performing providers, Bell Aliant, represents an incumbent operator of legacy copper networks that also serves relatively less densely populated areas. However, unlike its larger counterparts Bell and Telus, a few years ago Bell Aliant decided to accelerate FTTP deployment and decommission some of its century old copper plants in Atlantic Canada. On the other hand, the quality of service delivered by the largest cable broadband provider in Western Canada, Shaw, substantially lagged its counterparts in Central and Eastern Canada (Rogers, Cogeco, and Videotron). Firm level variation in the data highlight the importance of strategic choices of providers that tend to dominate regional markets in shaping broadband outcomes consumers experience over exogenous factors such as their endowments, geography, or market regulations.

Geospatial variation

Differences in speeds/service quality across jurisdictions and within communities represent some of the key source of variation in network quality indicators relevant in broadband infrastructure benchmarking, gap analysis, and policy development. To explore this source of variation, below we present and analyze the implications of increasingly disaggregated perspectives on the M-Lab data analyzed at the national level above.

Provincial variation

As broadband has become more essential, provincial governments in Canada, which are responsible for delivering social and business infrastructure, have become increasingly interested in benchmarking and monitoring the state of connectivity. Ontario has been particularly active in using broadband speed measurements to benchmark and map the state of broadband connectivity and in policy development. Table 5 details indicators of measured download speeds across the provinces in 2016 and percentage change relative to this baseline in 2017.

There appears to be some clustering in median and average speed levels within the provinces and territories. Not surprisingly, measured speeds in the three sparsely populated Northern territories, where there is significant reliance on slower fixed wireless and satellite services, form the slowest cluster. Starting from a lower base, median speeds in Nunavut and Yukon Territories experienced the fastest improvement over this short period. The Prairie Provinces that tend to be more rural and DSL incumbents face relatively limited competition from faster cable broadband providers represent the second slowest cluster; with measured speeds in Saskatchewan falling significantly behind Manitoba and Alberta. At the other extreme, operators in British Columbia and New Brunswick delivered fastest speeds in the country in 2016. With the exception of British Columbia and Yukon

Table 5. Provincial variation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>5.9</td>
<td>14.2</td>
<td>–9.0</td>
<td>–0.9</td>
</tr>
<tr>
<td>BC</td>
<td>10.9</td>
<td>22.5</td>
<td>6.1</td>
<td>16.2</td>
</tr>
<tr>
<td>MB</td>
<td>5.7</td>
<td>13.5</td>
<td>–2.3</td>
<td>4.2</td>
</tr>
<tr>
<td>NB</td>
<td>10.0</td>
<td>28.0</td>
<td>–16.3</td>
<td>–9.9</td>
</tr>
<tr>
<td>NL</td>
<td>6.5</td>
<td>21.9</td>
<td>–16.2</td>
<td>–16.3</td>
</tr>
<tr>
<td>NS</td>
<td>5.4</td>
<td>20.1</td>
<td>–24.7</td>
<td>–12.7</td>
</tr>
<tr>
<td>NT</td>
<td>2.8</td>
<td>7.9</td>
<td>–4.4</td>
<td>7.3</td>
</tr>
<tr>
<td>NU</td>
<td>1.6</td>
<td>3.2</td>
<td>46.2</td>
<td>17.4</td>
</tr>
<tr>
<td>ON</td>
<td>8.1</td>
<td>17.5</td>
<td>–6.7</td>
<td>–10.7</td>
</tr>
<tr>
<td>PE</td>
<td>4.0</td>
<td>15.5</td>
<td>–18.8</td>
<td>–13.8</td>
</tr>
<tr>
<td>QC</td>
<td>9.2</td>
<td>17.7</td>
<td>–1.9</td>
<td>19.0</td>
</tr>
<tr>
<td>SK</td>
<td>3.8</td>
<td>8.9</td>
<td>–2.2</td>
<td>1.9</td>
</tr>
<tr>
<td>YT</td>
<td>2.9</td>
<td>7.7</td>
<td>35.9</td>
<td>9.2</td>
</tr>
</tbody>
</table>
where both median and average measures experienced substantive growth, median speeds declined or stayed broadly stable in all other provinces. In some provinces and territories that experienced a measure of growth in average speeds, median speeds appear to have plateaued (e.g. Quebec, Manitoba, Saskatchewan). This suggests that an increase in average connection speeds does not necessarily lead to improvements in basic service quality available, as indicated by the median rate. Overall, provincial variations in the data offer little evidence to indicate broadband speeds at either the basic service end or higher speed end of the scale are on a positive trajectory: Growth in network capacity appears to be barely keeping up with increasing demand by users for network resources in most of the country.

**Municipal divergence**

Aggregating Internet measurements to assess performance of large areas can obfuscate local differences in the development of network infrastructure and service quality. Conventional wisdom suggests that in relatively low cost/higher income urban areas infrastructure competition between operators can create stronger incentives to increase network capacity and deploy new technologies than in higher cost/lower income rural communities and small towns. To better understand the extent and nature of local variation in service quality across Canada’s diverse municipalities, we aggregate the nearly two million tests conducted in 2016 into around 2500 local jurisdictional entities (cities, municipalities, towns, villages, etc.) and estimate their average download speed and latency rates. We use the number of tests from a particular jurisdiction as a proxy for the “size” of the community (i.e. assume places from which users conducted more/less tests tend to be larger/smaller). We find little statistical association between our proxy for community size and measured download speeds or latency.

This lack of empirical association between community size and broadband infrastructure quality indicators may seem surprising. However, some of the providers with highest measured speeds in the country tend to serve rural communities and smaller cities in Central and Atlantic Canada (Cogeco and Bell Aliant; see Figure 5). While measured speed and latency do not appear to be correlated with community size, as Figure 6 shows there is a negative nonlinear relationship between the two indicators of Internet access infrastructure quality: In communities where providers have provisioned more effective bandwidth, their customers tend to experience relatively higher quality of service (lower latency as measured by Round Trip Time).

This empirical relationship indicates small improvements in effective bandwidth in communities that are further behind can lead to significant improvements in service quality users experience (left-hand side of Figure). However, as effective bandwidth increases it becomes exceedingly harder to reduce latency. This observation is important because, while some Internet applications might be tolerant to latency peaks when traffic starts to reach maximum available bandwidth (i.e. congestive collapse), reliable utilization of many others requires consistently low latency levels (e.g. voice and video communications, cloud-based applications and control, multimedia). Overall, the evidence suggests broadband networks in only a small number of communities within Canada are sufficiently developed to make available actual average speeds approaching CRTC’s 50 Mbps aspirational universal service target. Further exploring factors that help
explain variations in the pace of network development at regional and local levels might offer useful insights and help identify innovative solutions for improving connectivity in lagging communities.

**Local disparities**

One of the key assumptions in universal access policy development in mature markets such as Canada has been that broadband connectivity problems are primarily a rural concern as “market forces” should be sufficient to deliver services of sufficient quality in urban areas. However, even in large cities with low costs and high income, incentives of providers to invest in capacity enhancements and scalable fiber networks can be limited, for example in relatively low-income neighborhoods and/or older ones where upgrading legacy assets might be expensive. To help explore the relevance of local variations in connectivity, Figure 7 maps measured download speeds in Canada’s most populous region, the Greater Toronto and Hamilton Area (GTHA). GTHA represents a diverse geographic region, which includes a number of large modern cities, newer suburbs, older towns and villages now part of suburban sprawl, and relatively sparsely populated rural communities on the edges (with a population approaching 8 million). We aggregate speed tests in 2016 into submunicipal hexagonal areas the federal authorities use to estimate service availability based on information about advertised speeds operators provide. Speed mapping of the region challenges a number of common assumptions. While measured speeds in some rural communities on the GTHA periphery were particularly poor (in red), it is notable that some had average speeds comparable to densely populated urban cores of the region in Toronto and Hamilton (in dark and light blue).

Furthermore, with the exception of a few areas in downtown business cores of larger cities in the region and some higher income/newer suburban areas (e.g. North and East of Toronto), average speeds in most areas/neighborhoods in the country’s largest urban center were less than half of CRTC’s 50 Mbps (in white and light blue). Particularly notable is the cluster of lower income/older neighborhoods in the northwest of Toronto which had average measured speeds below 10 Mbps. This picture from GTHA suggests there are local conditions that promote private sector incentives to deploy relatively high-quality networks in some rural communities, while service quality in some urban neighborhoods remains subpar. Understanding why this is the case is particularly relevant in a country where more than 80% of the population lives in a small number of urban metropolitan regions such as GTHA. Capacity gaps relative to CRTC’s 50 Mbps universal service targets might be particularly wide in some rural and remote communities. However, public domain measurements illustrate private sector incentives to provision sufficient network capacity to enable higher speeds can also be relatively weak in some urban neighborhoods. Variations in broadband infrastructure quality available within urban areas partly explains increasing calls by municipal stakeholders and the business community on higher levels of government to evaluate the state of broadband infrastructure and adopt policies aimed at making very high speed/low latency fiber access technologies more widely available in both urban and rural areas.

Increasing test sample sizes within communities/neighborhoods and improving geolocation accuracy represent two key methodological challenges in utilizing public domain Internet measurements to build more fine-grained perspectives on the state and evolution of localized digital divides in terms of access to “high-speed” Internet connectivity. A pilot project that helps identify challenges and develop a methodology for community-based testing represents the next step in this research.

**Conclusions**

Emergence of broadband as an essential input for social and economic participation necessitates users, researchers, and policymakers to measure broadband
infrastructure quality. Our article aids in bridging the gap between computer science and public policy by highlighting challenges in translating technical measures of connection performance into indicators relevant for broadband gap identification and policy development. Using Canada as a case, we have compared the strengths and weaknesses of different Internet testing methodologies and indicators utilized by users and policymakers to assess the quality of service network providers deliver (i.e., versus "best effort" rates specified in retail contracts). We further analyzed key sources of variation underlying a large-scale public domain Internet measurement dataset and drew a number of insights from the results that challenge commonplace conceptualizations of the digital divide in terms of broadband service quality.

The analysis shows that users in only a handful of communities in Canada may be able to get the basic service speed targets established by the CRTC in 2016. This is not surprising since traditionally policymakers in Canada and other countries have used “best effort” speeds Internet service providers advertise (up to x Mbps) to set Internet access service quality targets and identify gaps relative to them. More importantly, with regard to public policy and business strategy development, findings show that inequities in broadband connection speeds and service quality tend to be driven by local conditions that transcend the rural–urban digital divide. This finding challenges common conceptualizations in Canada and in other countries that universal access is a concern only in high cost rural communities.

In addition, the analysis highlights the value of utilizing multiple sources of data and testing methodologies for developing indicators of quality of Internet access. “On-net” testing methods or others that assess the quality of well-provisioned links might be relevant for capturing the experience of users when accessing content and applications cached within the operators’ infrastructure (or otherwise prioritized for delivery). “Off-net” measurements are more relevant for assessing the quality of connectivity to content and applications on the edge of clouds in nearby cities or in far away countries. New testing initiatives with distinct methodologies that offer their data in the public domain will be needed to enable a “big data” analytic approach to characterizing Internet access in the future. While small scale test beds designed to measure particular aspects of connectivity under controlled conditions can generate valuable information, large scale user and community-based testing initiatives are likely required for moving from a world of “best effort” retail contracts to one with differentiated and guaranteed standards of access to an open Internet that enables competing over-the-top (OTT) content and application providers. The capacity of users to validate counterparty performance in retail contracts they have with service providers will depend on the availability of public domain Internet measurement tools and indicators such as the ones we have analyzed in this article.

Our study contributes to a needed synthesis of big data computational approaches and Internet performance data in the service of broadband policy evaluation. Without such research, Internet service providers that manipulate Internet measurement methodologies and act as if they are delivering on their contractual or regulatory performance targets can easily get away. For example in the United States, the New York Attorney General has brought a lawsuit against Charter/Spectrum for “overprovisioning” connections to test servers “to conceal the fact that most subscribers, particularly those on congested service groups, received far less than their promised speed”. Without reliable measurements of quality of Internet connectivity, universal service policies may not be able to deliver on the promise to provide universal access in our information society.

Notes

1. Telecom Regulatory Policy CRTC 2016-496.
2. See https://www.measurementlab.net/tools/ndt/
3. For example, service providers, system management vendors, and policy makers typically consider increasing latency at particular nodes as an indicator of congestion, which could call for investment for provision more capacity (see Sandvine Inc. 2015).
4. In theory, Internet performance tests can be designed and configured in a way that helps shape the results.
6. Please note there are various other types of tests in use and multiple vendors in the international market offer different testing platforms to operators, users, and public authorities.
7. For an overview of some policy relevant measurement issues see Hong and Morris (2016).
9. Confidential estimates of throughput levels using probe-based tests from the edge of clouds in large Canadian cities to users are also broadly consistent with measured speeds from Akamai and M-Lab. This class of tests also shows significant regional variation in available bandwidth for connecting to “off-net” services on the edge of clouds in nearby metropolitan areas.
10. This observation is particularly important for broadband gap identification as using the Ookla/Speedtest data can lead to significant underestimation of their magnitude.


14. All Measurement Lab data is accessible through Google BigQuery. SQL queries used our data analysis can be found at: https://github.com/Zeitsperre/AMO-OMA. All queries selected tests occurring in Canada to M-Lab servers inside and outside of Canada.

15. The lack of sustained growth in measured speeds we observe over this period may be partly a function of the "off-net" nature of the M-Lab test that captures connection quality to a third-party server in a large city nearby (i.e., to the "open Internet"). Speeds/service quality levels for accessing services that are cached "on-net" or otherwise prioritized may have increased as a result of network investments. Due to a lack of consistent on and off net measurements, it is not really feasible to evaluate how they co-evolve. Furthermore, the rapid rise in the number of tests in the 2nd half of 2016, which can be attributed to: (a) launch of the CIRA IPT, and (b) Google search engine started to direct users looking for a "speed test" to the M-Lab/NDT platform. There is a possibility that increasing demand from users could have tested the resources of the links and servers to the M-Lab/CIRA testing infrastructure. On the other hand, the growth in the number of servers in large Canadian cities with the launch of the CIRA program could have had the opposite effect. It is difficult to control for the impact of such changes in the testing infrastructure. Given the popularity of M-Lab’s standardized approach to network performance measurement among provincial and municipal governments, it would be prudent to explore the relevance of this and other potential causes of observed variation in more detail to ensure that exogenous changes to the testing platform do not have an adverse impact on the quality of data.

16. It also provides further evidence of the power of the Google search platform as an arbiter in the market for information people value from the Internet.


18. Please note that in theory it is also feasible for operators to identify M-Lab test servers and overprovision connections from their customers to them. Although we have not seen any evidence of this in Canada, increasing reliance on the CIRA/IPT by users and lower levels of government can increase the incentives of some operators to do so and care should be taken in validation of results from M-Lab/CIRA tests in the future.

19. What is somewhat more surprising within this group is the extent to which Nunavut has fallen behind Yukon and Northwest Territories.

20. We use average rates here as it better captures the proportion of higher speed connections that might be available in a particular area in the long tail of the distribution. Consequently, it is more reflective of potential access to higher speed connections than actual basic service levels.

21. Unfortunately, the names of entities from which tests originate in the M-Lab database do not conform to available socio-economic data and there are also substantive geolocation problems in integrating Internet measurements from IP addresses into locally refined maps. Small sample sizes in rural communities also limits the usability of the data and highlights the relevance of community driven testing initiatives or probe-based tests to enable accurate rural broadband mapping.

22. For example, some service providers may be able and willing to identify test servers and overprovision links to them. For a detailed analysis of this problem see: Complaint by The State of New York vs. Charter/ Spectrum/Time Warner. 31 January 2017. File no: 450318/2017. Available at: https://ag.ny.gov/sites/default/files/summons_and_complaint.pdf

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This research was supported, in part, by the Social Sciences and Humanities Research Council of Canada and the Ministry of Infrastructure, Government of Ontario, Canada.

ORCID
Fenwick McKelvey http://orcid.org/0000-0002-7584-1133

References


R. RAJABIUN AND F. MCKELVEY

14


