



Measuring real broadband speeds using crowdsourcing data from the Internet Foundation

15 November 2017

Ivan Vallejo Vall

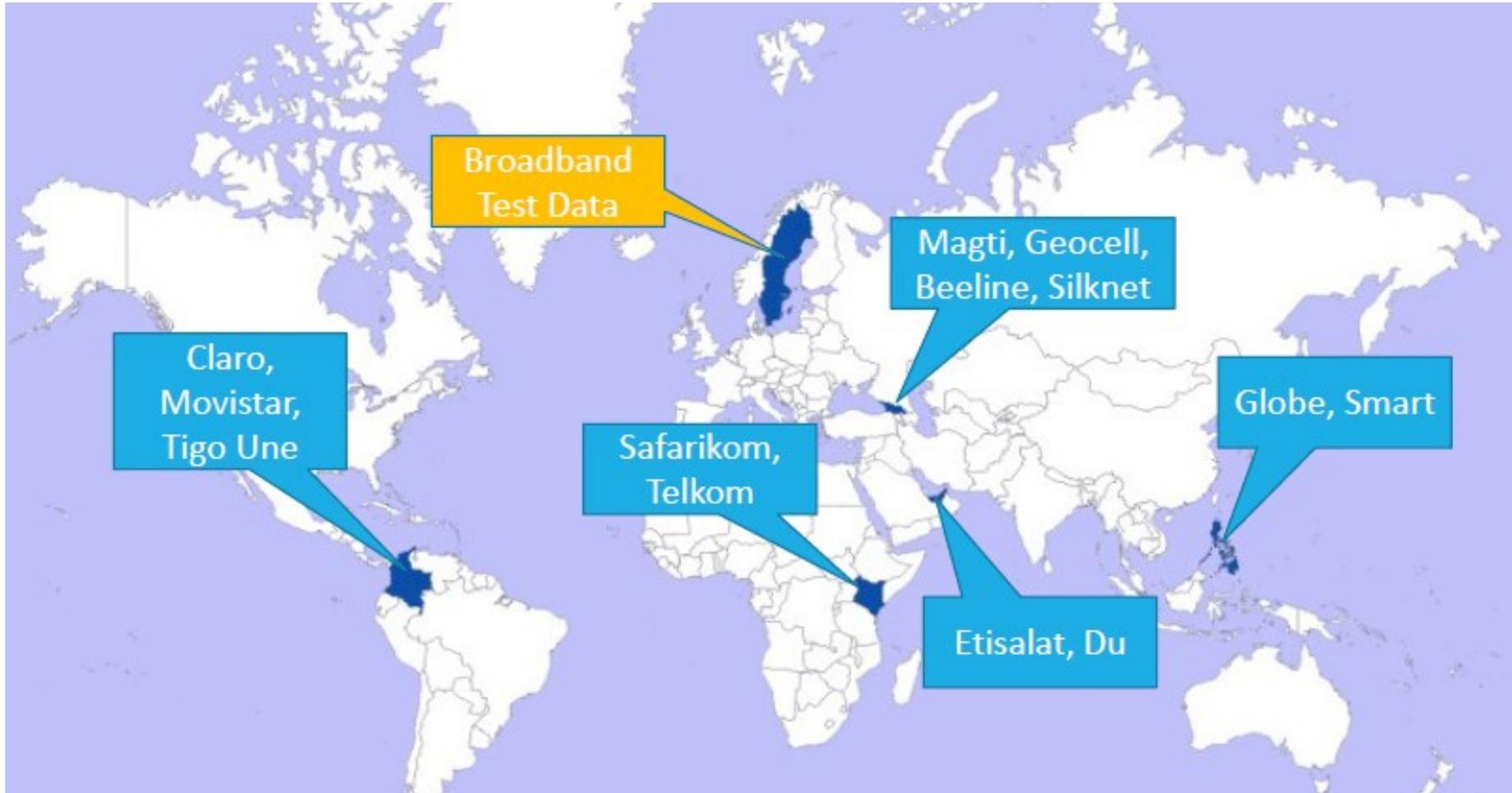
Market analyst, ITU

WTIS-17, Hammamet, Tunisia

Big Data for Measuring the Information Society



2



Bredbanskollen



3



Bredbanskollen

HOME SELF HELP SPEEDMAP REPORTS HELP

Bredbanskollen provides a simple way for you to measure your internet speed and get help on improving it.

START TEST!

Bredbanskollen on your phone
Fast internet on your phone? Get our app for [IOS](#) or [Android](#)

Follow Bredbanskollen
If you also think Bredbanskollen is great, you can follow us on [Facebook](#) and [Twitter](#).

The screenshot shows the homepage of the Bredbanskollen website. It features a blue header with the title 'Bredbanskollen' and a navigation menu with icons for Home, Self Help, Speedmap, Reports, and Help. The main content area has a blue background with a large clock icon and a 'START TEST!' button. Below this, there are two promotional sections: one for a mobile app and another for social media follow-up.

Relevance of Internet speed



4

□ Consumers

- 2nd most important factor
- advertised \neq real speed + congestion [1]

□ Regulators

- Quality of service (QoS) [2]

□ Policy makers

- Universal Service Funds, broadband plans [3]

Research question



5

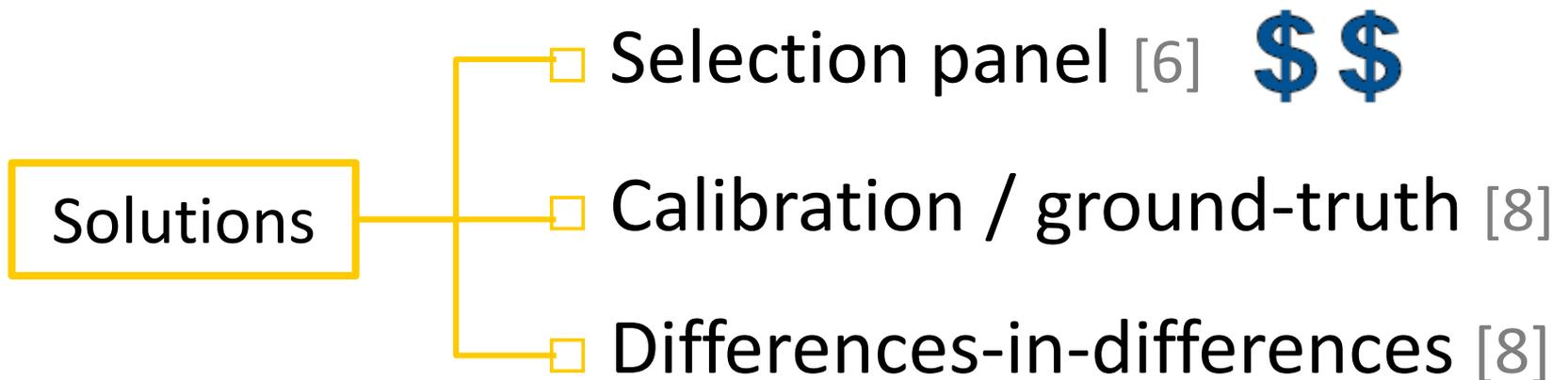
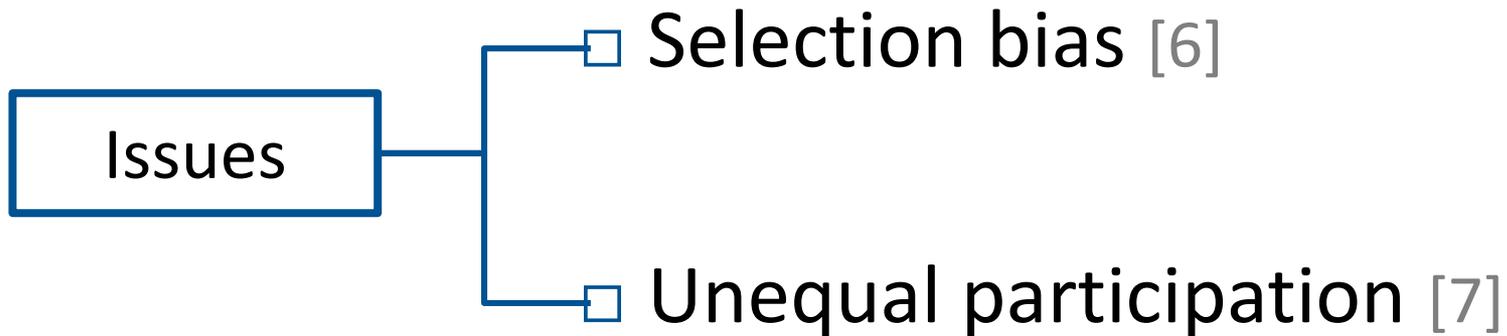
Can crowdsourcing Internet data be used to measure real broadband speeds?

- Advertised speeds 
- Hardware-based measurements  
- Software-based crowdsourcing?

Non-probabilistic samples

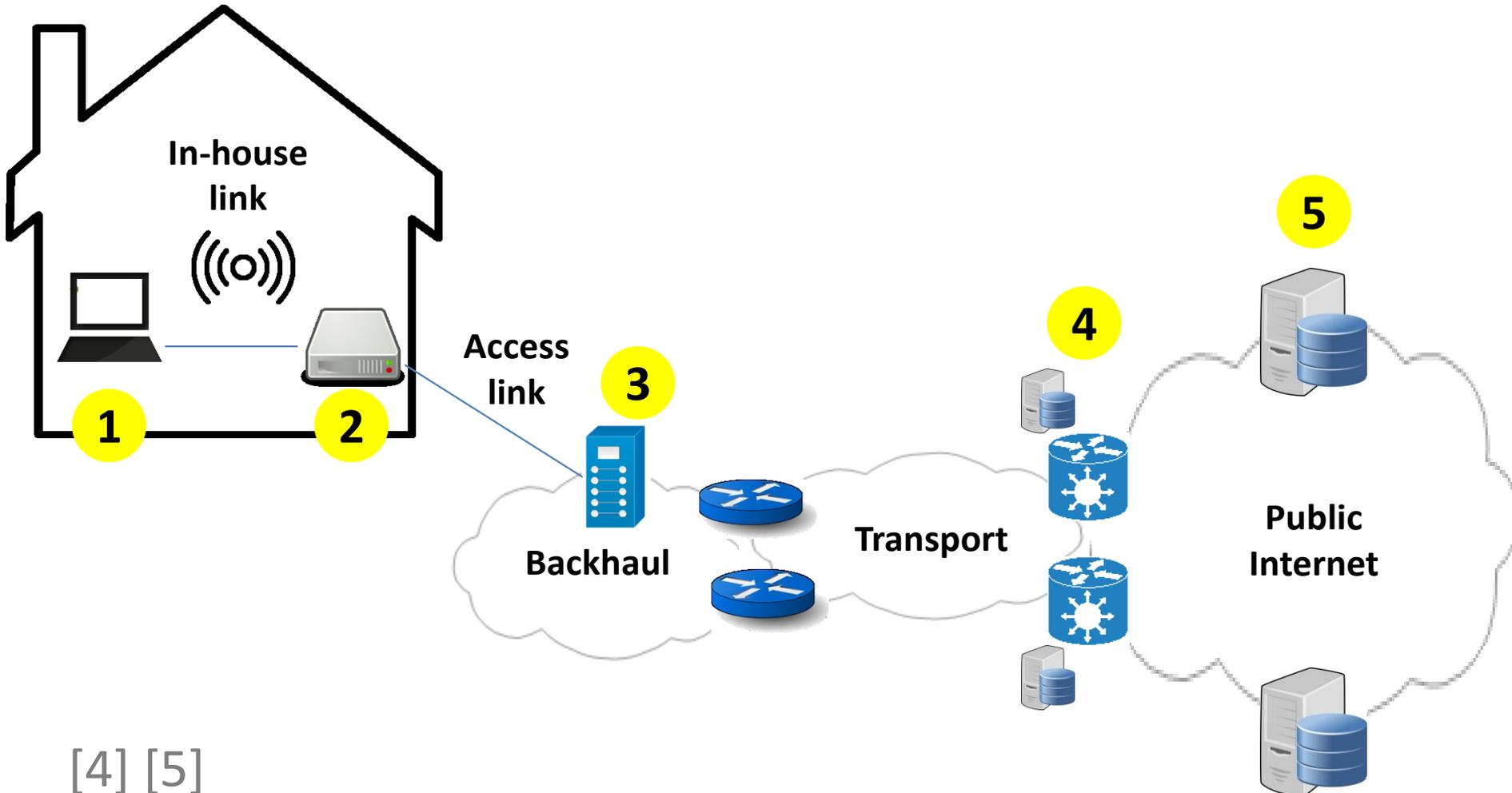


6



From where to where?

7



Speed test comparison

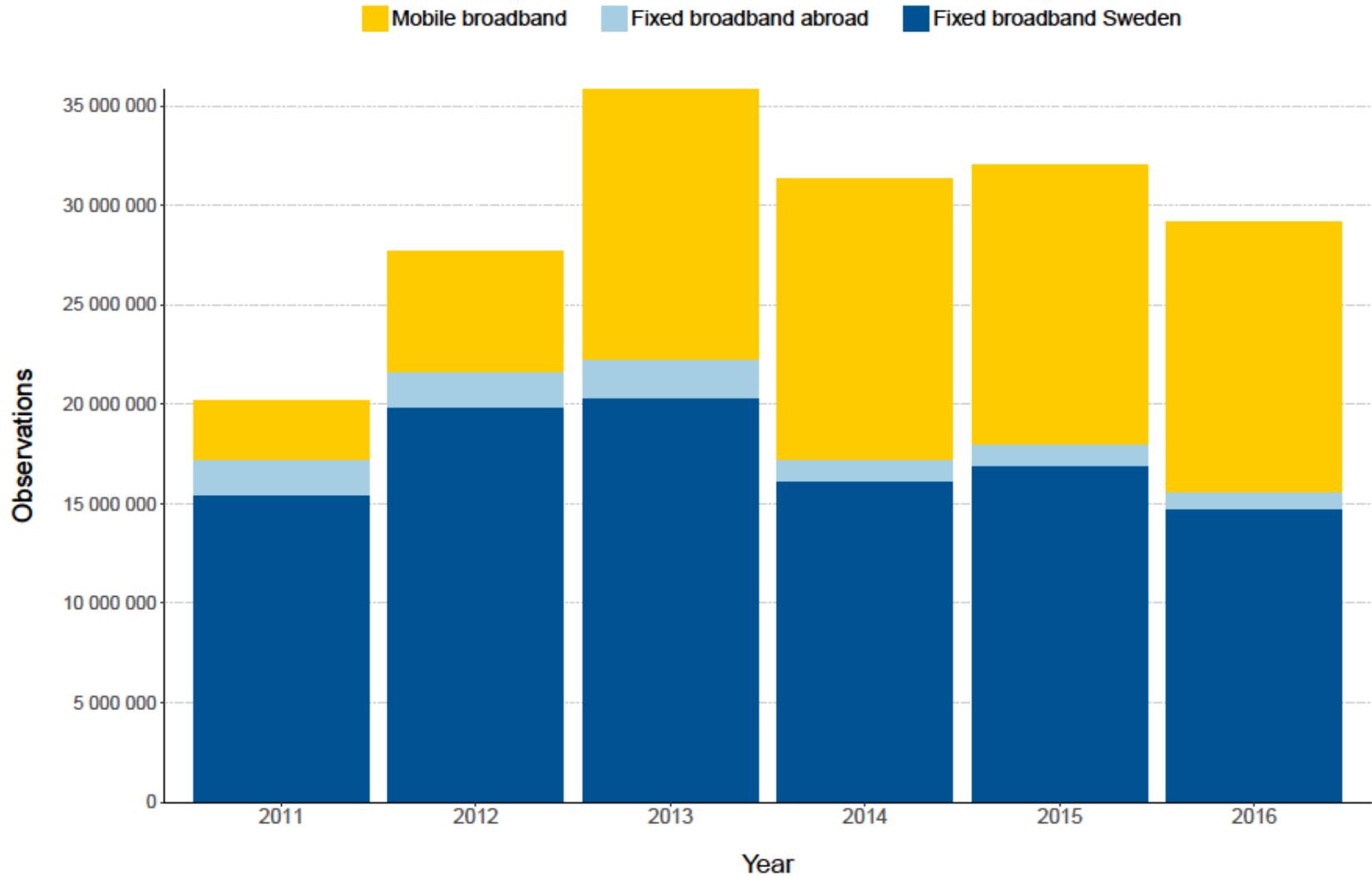


	Test path	Active Passive	Voluntary Automatic	Statistical aggregation
Akamai	1-4 or 1-5	Passive	Real traffic	Unknown
Ookla	1-4 or 1-5	Active	Voluntary	Avg. first 2s or avg. 10s
Bredbands kollen	1-4 or 1-5	Active	Voluntary	Avg. after ramp-up. Excl. top 10% & bottom 30%
SamKnows	2-5	Active	Automatic	Avg. after ramp-up. Excl. top & bottom percentiles

Observations

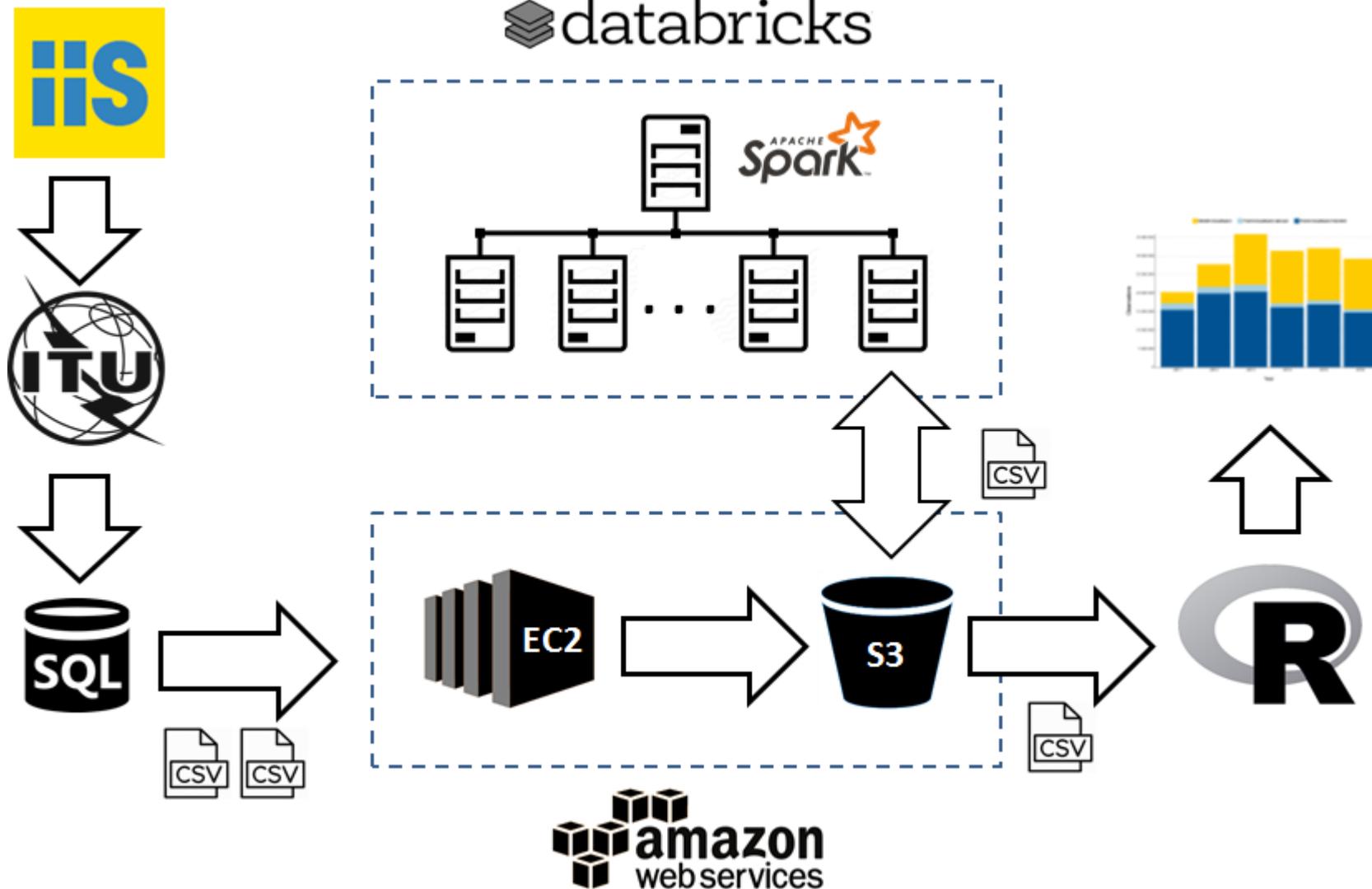


9



Data processing

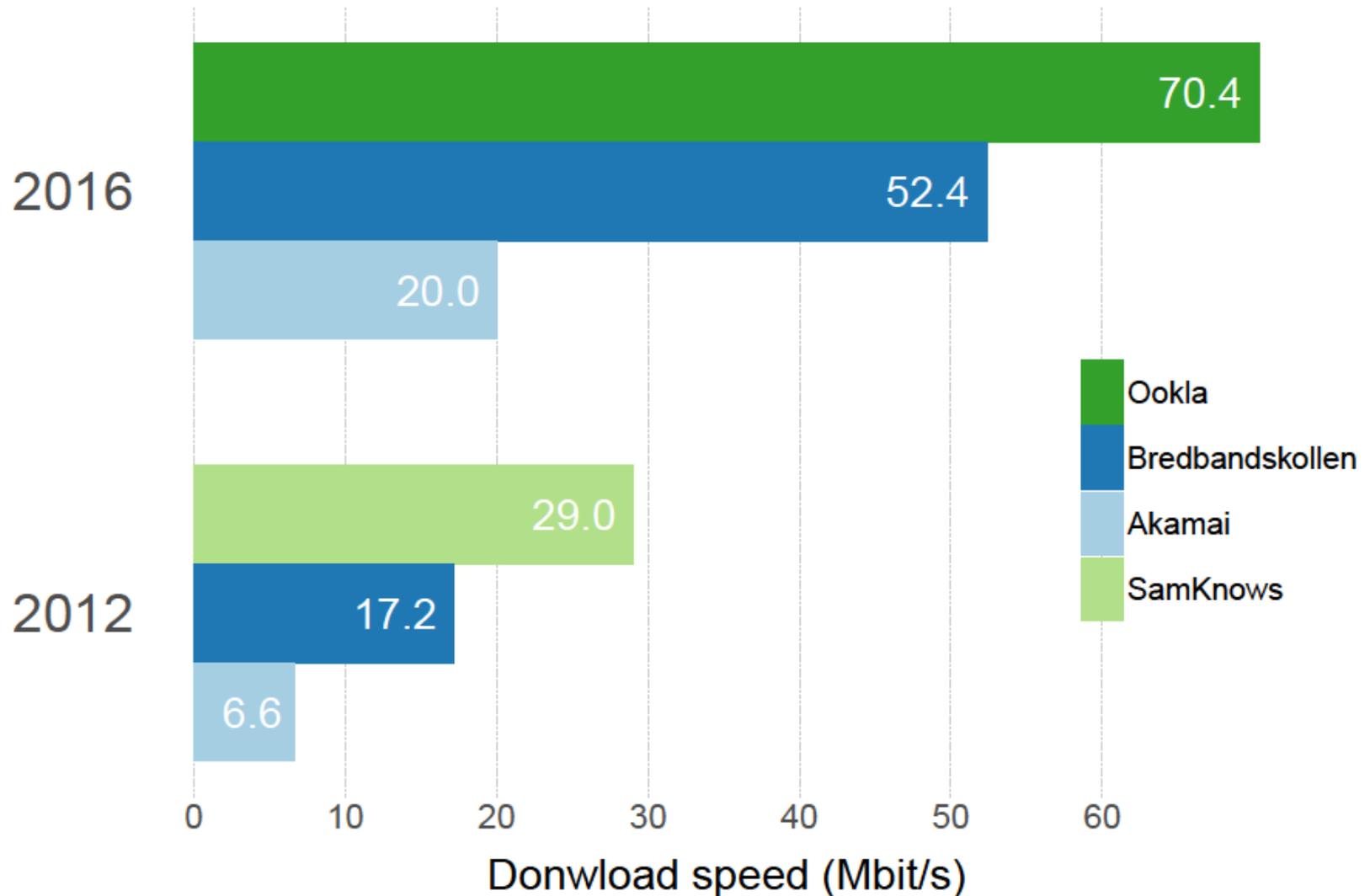
10



Comparison speed tests



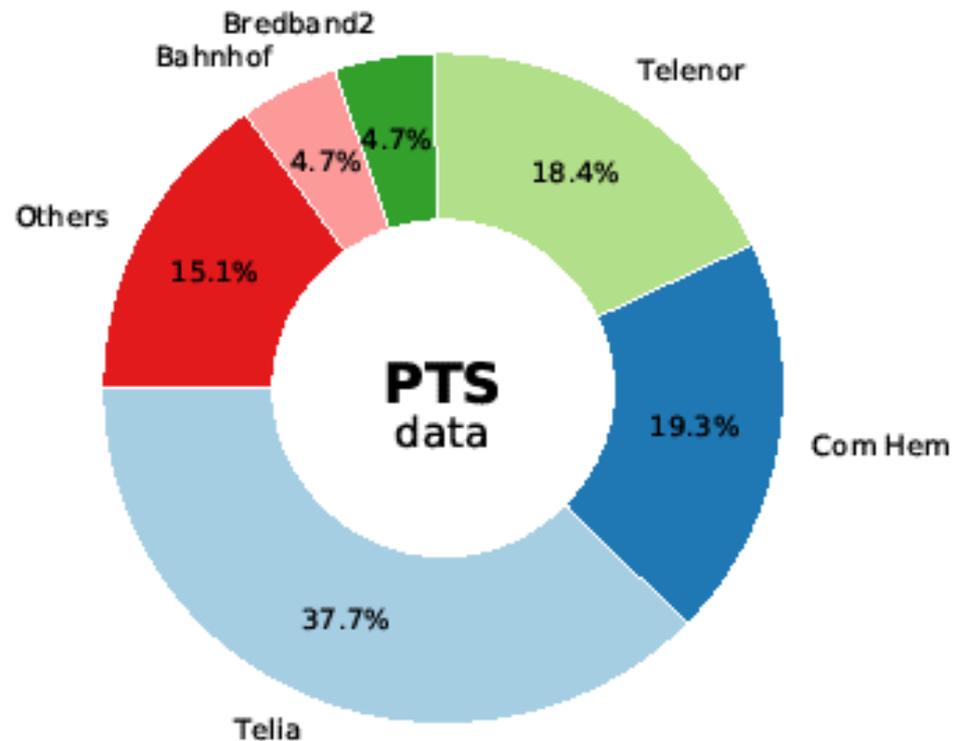
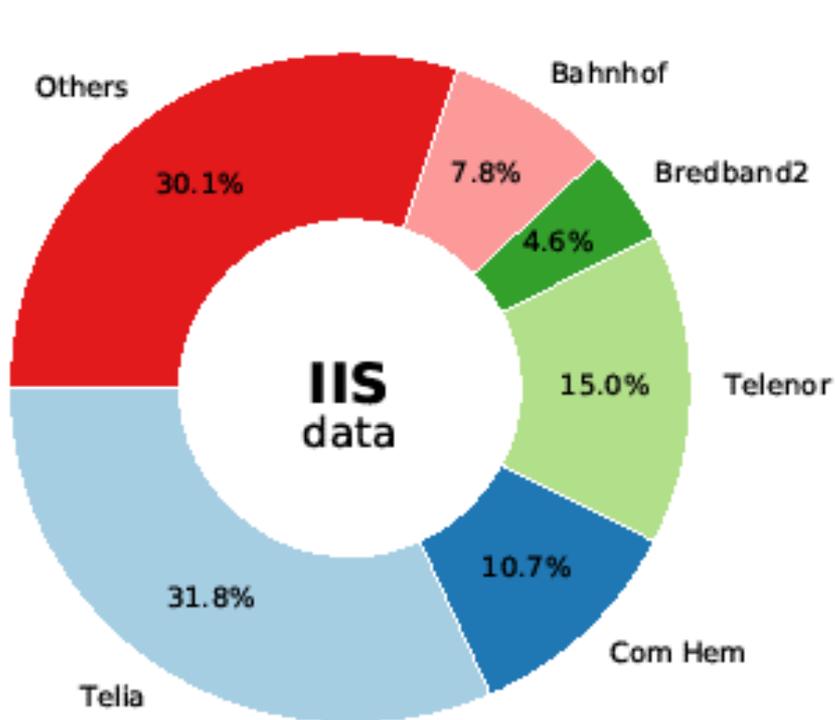
11



Observations per operator



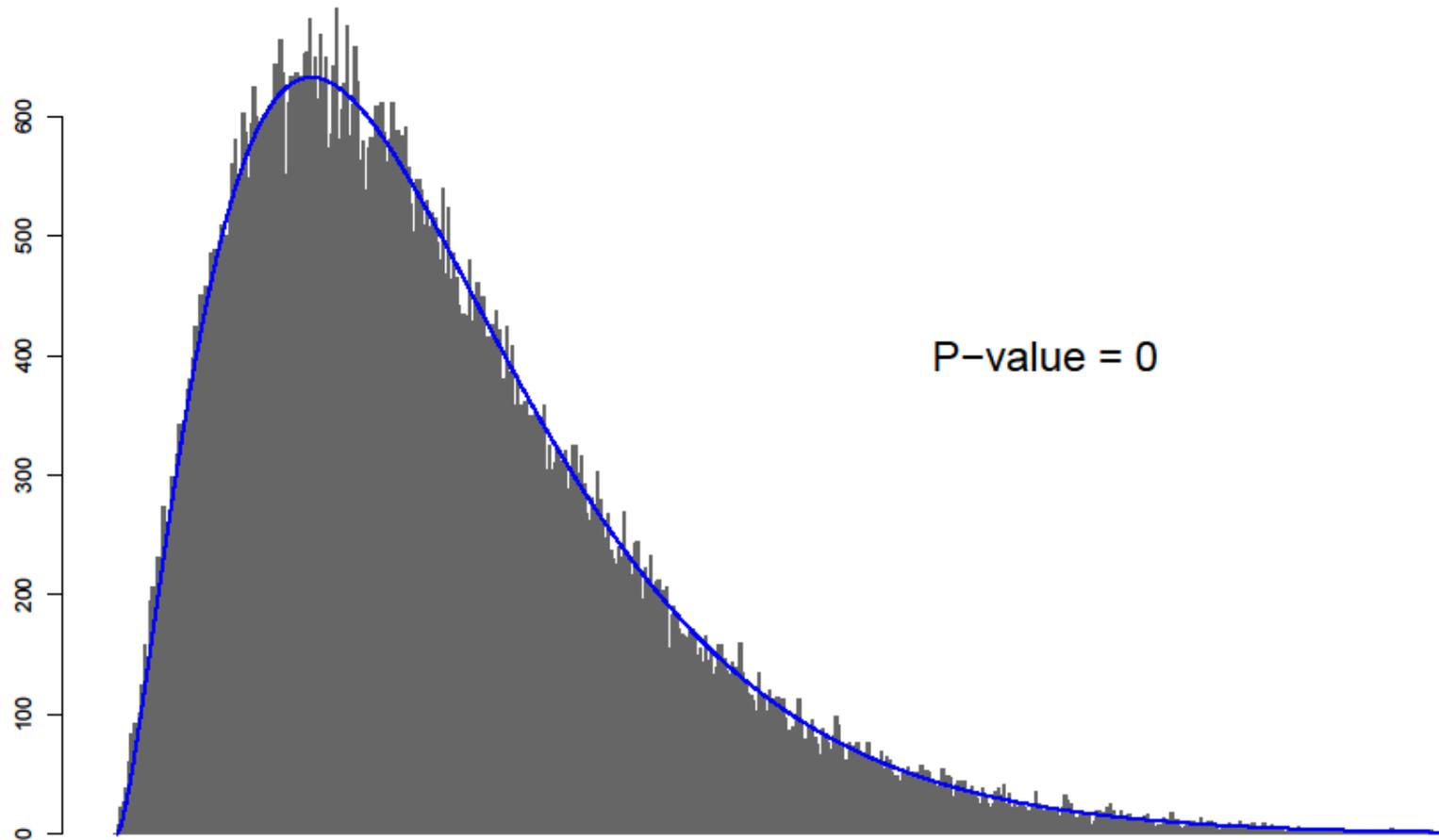
13



Selection bias



14



Statistic: $-2 \ln(\text{LR}) = 2.446 \times 10^6$ Range: (0 - 20.52)

Conclusions



16

Findings

- Measurement error
- In-house congestion (hypothesis)
- Selection bias
- Importance of ground truth data
(incl. robust panel/survey data)



List of references

17

1. TNS Opinion & Social (2014). Special Eurobarometer 414 “E-Communications and Telecom Single Market Household Survey”. Technical report, European Commission.
2. Telecom Regulatory Authority of India (2006). Regulation quality of service standards for broadband services.
3. Davies, R. (2016). Broadband as a universal service. Technical report, European Parliamentary Research Service. PE 581.977.
4. Sundaresan, S., Feamster, N., and Teixeira, R. (2016). Home Network or Access Link? Locating Last-Mile Downstream Throughput Bottlenecks. In PAM 2016 - Passive and Active Measurement Conference, pages 111–123, Heraklion, Greece.
5. M-Lab Research Team and others (2014). ISP interconnection and its impact on consumer Internet performance – a measurement lab consortium technical report.
6. Rood, H., Yoshikawa, D., Wevers, R., Post, R., Tetteroo, A.-J., Kuipers, A., and Schurmann, R. (2012). Usability of broadband performance measurements for statistical surveys. In 2012 TRPC.
7. Couper, M. (2013). Is the sky falling? new technology, changing media, and the future of surveys. *Survey Research Methods*, 7(3):145–156.
8. Zagheni, E. and Weber, I. (2015). Demographic research with non-representative internet data. *International Journal of Manpower*, 36(1):13–25.

Thank you

