Transparent Estimation of Internet Penetration from Network Observations

Suso Benitez-Baleato[1,2(✉)], Nils B. Weidmann[1], Petros Gigis[3],
Xenofontas Dimitropoulos[3], Eduard Glatz[4], and Brian Trammell[4]

[1] Department of Politics and Public Administration,
University of Konstanz, Konstanz, Germany
`jesus.benitez-baleato@uni-konstanz.de`
[2] University of Santiago de Compostela, Santiago de Compostela, Spain
[3] Foundation of Research and Technology Hellas (FORTH), Heraklion, Greece
[4] Computer Engineering and Networks Laboratory (TIK),
ETH Zurich, Zurich, Switzerland

Abstract. The International Telecommunications Union (ITU) and the Organization for Economic Cooperation and Development (OECD) provide Internet penetration statistics, which are collected from official national sources worldwide, and they are widely used to inform policy-makers and researchers about the expansion of digital technologies. Nevertheless, these statistics are derived with methodologies, which are often opaque and inconsistent across countries. Even more, regimes may have incentives to misreport such statistics. In this work, we make a first attempt to evaluate the consistency of the ITU/OECD Internet penetration statistics with an alternative indicator of Internet penetration, which can be measured with a consistent methodology across countries and relies on public data. We compare, in particular, the ITU and OECD statistics with measurements of the used IPv4 address space across countries and find very high correlations ranging between 0.898 and 0.978 for all years between 2006 and 2010. We also observe that the level of consistency drops for less developed or less democratic countries. Besides, we show that measurements of the used IPv4 address space can serve as a more timely Internet penetration indicator with sub-national granularity, using two large developing countries as case studies.

1 Introduction

How has the usage of the Internet technologies increased in different countries? How has that growth affected economic and societal changes? A main source of empirical evidence to address these questions are the Internet penetration statistics provided by the ITU and the OECD. Those statistics are influential in debates about global technological development, Internet governance and its societal effects. Moreover, social scientists rely on these datasets to understand

the impact of technology on social and political systems [22]: Is the Internet really a catalyst of popular protest that can topple dictators? Or does it rather play into the hands of autocrats, increasing opportunities for surveillance and censoring?

While important both for policy-making and scientific research, these statistics exhibit some key shortcomings. The ITU and the OECD do not measure Internet penetration directly, but they rather collect and standardize information provided by different governments and their regulatory agencies. Each of these agencies has its own protocol for collecting these numbers at the national level. Thus, the final statistics that are ultimately included in the main datasets may be subject to error due to poor data collection standards, or even systematic inflation due to some countries' incentives to exaggerate economic progress because of aid conditionality. Similarly, differences in data collection across countries may significantly limit comparability and thus impede the main purpose of the data. Also, these statistics become available with significant delay (oftentimes a year or more). Last, only national-level statistics are provided, which makes analysis of variation in Internet penetration *within* countries impossible.

In this work, we make a first attempt to use Internet measurement techniques to independently verify and supplement existing penetration statistics. We introduce a reproducible methodology that uses publicly available data and circumvents the limitations of transparency, comparability availability and resolution. We derive Internet penetration estimates from geolocated network measurements of the globally used IPv4 address space using two different approaches, and then compare our estimates with the official ITU/OECD statistics between 2006 and 2010. We find that our estimates exhibit very high correlation (ranging between 0.898 and 0.978) with the official data for all studied years, which however drops for less developed or democratic countries. In addition, we show that our estimates are consistent with official statistics at the subnational level for two large developing countries. These observations are encouraging, because they suggest that readily available data (e.g. from RouteViews [29]) can be used to cross-check official statistics and derive Internet penetration estimates more timely and with finer geographical resolution than the ITU/OECD statistics.

Our paper is structured as follows. First, we discuss the importance of the Internet penetration statistics for debates on technological development, Internet governance and its societal effects. Then we describe the methods and the datasets that we use to map and geo-localize the used IPv4 address space. After that, we compare them with the official statistics both at the country and the regional level. Finally, we discuss the results, the limitations, and the potential uses of our estimates.

2 Data and Research on Internet Penetration

The ITU is the United Nations telecommunications agency in charge of the global radio spectrum and satellite orbits allocation, the development of technical standards and the fostering of ICT deployment in developing countries [9].

As part of its role in technological development, the ITU collects, verifies and harmonizes ICT statistics. The outcome of this work is disseminated through the World Telecommunications/ICT Indicators Database (WTID), a chronological time series for over 200 countries regularly updated from 1960 on [12]. The WTID is made of more than 150 indicators describing aspects like coverage, traffic, price or quality of several communication technologies, including access and use of the Internet. The Internet penetration indicators are available for 192 countries, starting in 1990. The ITU retrieves this data from questionnaires submitted to the official country contacts. There are two types of national contact points in charge of providing the information to the ITU. The first one is the national telecommunication ministries and regulatory authorities, which provide Internet penetration estimates based on data from fixed and mobile Internet providers. The second source is the national statistical offices, which typically obtain data on access and use of the Internet through surveys. The collected data is then harmonized by the statistical division of the ITU, consistently with a set of guidelines intended to ensure the comparability of the data measurement and collection efforts performed by the respective countries [10,11].

Other organisations providing Internet penetration statistics, such as the Organisation for the Economic Cooperation and Development (OECD), follow a similar procedure. The OECD indicators also rely on data provided by the administrative bodies of the member states and from the EU Community Survey on household use of ICT. They are available for 34 countries starting in 2006 [20]. However, despite the similar data collection method, this does not mean that the values correspond to those in the ITU dataset; the correlation between the two is only 0.705 during our period of analysis (2006–2010). Thus, we will treat the OECD estimates as separate datasets in the analysis below.

2.1 Existing Work

In the following paragraphs, we discuss existing work that relies primarily on the ITU WTID database. Due to the fact that it is the only global cross-national database on ICT penetration, the WTID has been widely used in policy and research. The WTID is the main reference for many other UN agencies, including the Department of Economic and Social Affairs (DESA) and the World Intellectual Property Organisation (WIPO), who use it for their e-Government survey and the Global Innovation Index [27,32]. Also, the ITU data is used to measure the progress of the Millenium Development Goals, a road map adopted by 189 countries to make available the benefits of the ICT for developing countries [26]. Moreover, the WTID is extensively used in the Global Internet Report by the Internet Society (ISOC) [14] and the Global Information Technology Report by the World Economic Forum (WEF) [30], which describe the state of the Internet.

ITU statistics have also been used in research. Economists, for example, have used the WTID to analyze the effect of ICT investment on economic growth [23,24]. In political science, one strand of research has focused on the role of political institutions and economic development for technology adoption [19,21]. Here, again, the methodological approach is cross-national statistical

comparison using ITU indicators. Another question political scientists have focused on is the impact of ICT on democratization. Earlier work using ITU data concludes that the Internet fosters democracy through less restrictive channels of communication [2,6]. However, more recent results provide a more cautious view, as they show that closed autocratic regimes are keen adopters of this technology and are no more likely to democratize as a result of ICT introduction [22].

2.2 Limitations of Existing Databases

As the previous section has shown, the ITU indicators are a useful resource both for policy-makers and researchers alike. However, these valuable datasets suffer from a number of shortcomings.

Transparency. The lack of a standardized methodology across countries makes it difficult to understand and verify how the data are generated. For example, many countries will not have systematic data collection routines in place, requiring rough "approximations". Hence, it is not inconceivable that data provided to the ITU is subject to errors and biases in reporting. This may not be a problem affecting ICT statistics alone, but has been shown to be a more general issue with statistics from less developed countries [13].

Comparability. Because of the different quality and accuracy of numbers across countries, comparability in cross-national analyses may be severely hampered. The reason is that differences in Internet penetration across countries as picked up by the WTID can be partly the result of different data collection methodologies. With little information about the procedures employed by each country, it is difficult to even assess the severity of the problem, rather than correct it.

Availability. While the ITU offers semestral updates of their database, these updates are applied only to a selected number of indicators. The final revised edition of the full indicators is only available with one year delay, and is subject of retrospective revision caused by changes either in external datasets (like the population statistics) or by amendments submitted by national agencies. Those delivery times do also affect the other datasets considered; for example, some of the OECD indicators are only delivered with the publication of the Communications Outlook, once every two years [20].

Resolution. The WTID and OECD databases provide national level data only. However Internet penetration does not need to be uniform in a given country; regions of high coverage can exist next to those with low coverage. For many research projects, it would be useful to have indicators at the subnational level (for example, provinces or districts), to study how Internet coverage is provided sub-nationally, and what effects it has. The available statistics are of no use for this.

3 Data Sources and Methodology

In this section we describe the datasets and the processing methodology. First
we describe how we use routing data from Border Gateway Protocol (BGP)
collectors, namely from RouteViews [29], to estimate Internet penetration. Since
not all routed addresses are actually used, we use in addition passive traffic
measurements based on NetFlow from an academic Internet Service Provider
(ISP) in Switzerland (SWITCH [25]) to estimate the globally active IPv4 address
space based on the methodology of [3]. We collapse IPv4 address blocks to /24s
which is the longest unfiltered IPv4 prefix; and then geolocalize /24 prefixes in
national or large subnational administrative units.

Globally Routed IPv4 Addresses. We first extract routed IPv4 addresses from
publicly available BGP tables. In particular, we used a daily routing table snap-
shot from route-views2.oregon-ix.net for the first 16 days of each February and
August between 2004 and 2012. We selected these days to align with the times-
pan of the NetFlow data that were available to us. We collapsed all prefixes to
unique /24s per day. To mitigate the effect of misconfigurations and route leaks,
which could pollute our data, we filtered out reserved IPv4 addresses and prefixes
larger than /8. In addition, for each 16-day interval we kept the /24 prefixes that
were observed in all 16 days. In Fig. 1 we show the number of unique /24 prefixes
in each 16-day interval over time with and without our filtering. We observe two
large spikes, which were the result of route leaks, which our filtering effectively
mitigated. Apart from these two spikes, the effect of filtering is negligible. From
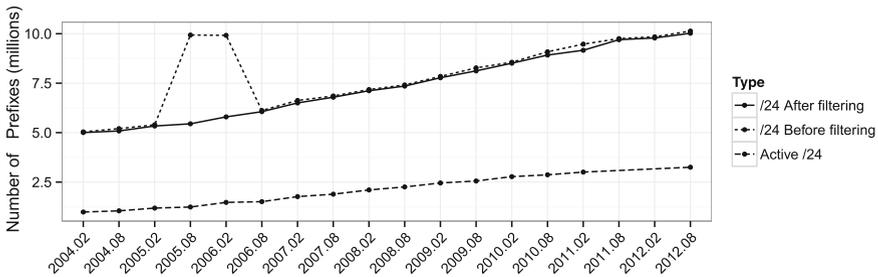2004 to 2012 the number of routed /24 prefixes doubled.

Fig. 1. Routed and active /24 IPv4 address blocks over time.

Active IPv4 Addresses. We also infer active /24 IPv4 address blocks using pri-
vate network traffic data from an ISP. We use the inferred active addresses
as a sanity check on the methodology based on the publicly available routed
addresses. Specifically, we used unsampled NetFlow records collected from the
border routers of SWITCH the first 16 days of each February and August
between 2004 and 2010. For the years 2011 and 2012 we do not use any
further the August and February samples due to anomalous or missing data.

We then extracted two-way TCP flows (to eliminate the effect of spoofing) and /24 blocks seen from SWITCH based on the methodology of [3]. Our previous work showed that this approach provides rich visibility (although not complete) into the globally used IPv4 address space [3]. In this paper, we extend our analysis to span a period of 9 years, for which we processed 218 billion flows (corresponding to 8.05 petabytes of traffic) in total. In Fig. 1 we compare how the active address space compares with the routed address space. We observe that on average 27.7 % of the routed address space is seen in the collected netflow data.

Geolocation. We then geo-reference each /24 block using the *Maxmind* GeoIP2 City database [18] and assign it to a country. The GeoIP2 City database is the most accurate geo-database provided by Maxmind, which claims 99.8 % accuracy at the country level and also high levels of accuracy within several different countries worldwide (for more details see [17]). We note though that geolocalization at finer granularity (e.g. the city level) or in cellular networks is still an open research problem and can be inaccurate, which is an issue that may affect our subnational results (cf. Sect. 4.4). We assign the /24 blocks to countries based on their spatial coordinates after removing 13,431 blocks georeferenced to 'EU' with coordinates in Switzerland, which account for approximately 0.05 % of the total active prefixes. We assign coordinates to countries using the CShapes dataset, a Geographical Information Systems (GIS) dataset on international borders that also incorporates border changes over time [31]. Our final indicator of Internet penetration is the number of routed or active /24 blocks in each country.

4 Correlation Analysis of Internet Penetration Estimates

In order to evaluate the consistency between our estimates of Internet penetration and those provided by the ITU and the OECD, we first conduct a bivariate analysis of the correlation of both estimates and how it changes over time. In particular, we correlate the number of routed or active /24 blocks with the number of Internet users according to the ITU or OECD data. For both estimates, we use the logarithm, since the numbers span several scales. We also analyze the agreement between the ITU numbers and our estimates when distinguishing between countries with different levels of economic development and democracy. Lastly, we take our analysis to the subnational level, evaluating the agreement of official and estimated Internet penetration estimates within two large developing countries: India and Turkey. Although we have verified that our findings hold for the entire duration of the studied datasets, we present results primarily for the period between 2006 and 2010 which is covered by all datasets.

4.1 ITU/OECD Statistics vs. Internet Measurements at the Country Level

The scatter plots in Fig. 2 illustrate and quantify the correlation between the proposed and the official Internet penetration indicators for a single time slice
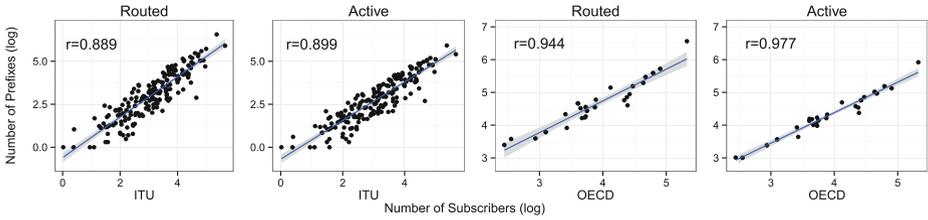
Fig. 2. Scatterplots of the proposed and the ITU/OECD country-level Internet penetration estimates for February 2010. For each scatterplot, we mark the correlation coefficient.

(February 2010). Both for routed and active prefixes, the correlation with the ITU data is very high, i.e. 0.889 or higher, as shown in the first two panels of Fig. 2. This suggests that both the number of routed or active prefixes can serve as good alternative indicators of Internet penetration. This picture improves for OECD statistics, which include more developed countries, as shown in the last two panels of Fig. 2.

The high correlation is also stable over time. The average correlation with ITU across all semi-annual samples between 2006 and 2010 is 0.898 and 0.911 for routed and active prefixes, respectively, and 0.944 (routed) and 0.978 (active) when comparing with the OECD data. Figure 3 shows the evolution over time, both for different validation datasets (ITU/OECD) and different measurement approaches (routed/active).
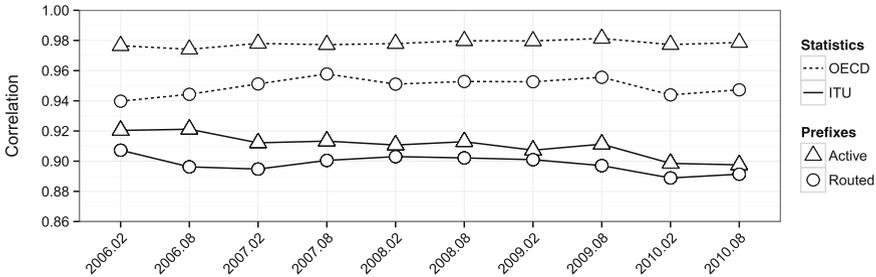

Fig. 3. Correlation over time between official statistics (ITU and OECD) and the number of routed/active prefixes.

4.2 Internet Penetration by Level of Economic Development and Democracy

We have seen that Internet penetration estimates based on network measurements achieve high correlation with the official Internet penetration statistics provided by the ITU and the OECD. However, so far we have treated all countries equally. In order to evaluate how our methods fares in different contexts,

we analyze correlations in different types of countries. We conduct this analysis on a global level using the ITU data only. As discussed above, the statistics provided by the ITU may be particularly problematic (i) in less developed countries with poor bureaucracies and (ii) in non-democratic countries where governments are not required or not willing to collect and share data. For these reasons, our analysis aims to establish how the agreement between ITU Internet penetration statistics and our network-based ones varies across different levels of development and different regime types (non-democratic to democratic ones).

Economic development is typically measured using GDP per capita values, which we obtain from [5]. The political regime type is measured using the Polity IV dataset [16]. Polity IV is the most frequently used quantitative database on political regimes, and encodes regime type using a numeric indicator ranging from −10 to 10. Low values correspond to strongly autocratic countries such as North Korea, and high values to established democracies such as the US. We bin the countries in our sample into low-, medium- and high GDP countries (Fig. 4a) and those with low-, medium- and high democracy scores (Fig. 4b). The split is done such that each category includes exactly a third of the countries in our sample. In each category, we compute the correlation between our network-based Internet penetration estimates and the ones from the ITU.
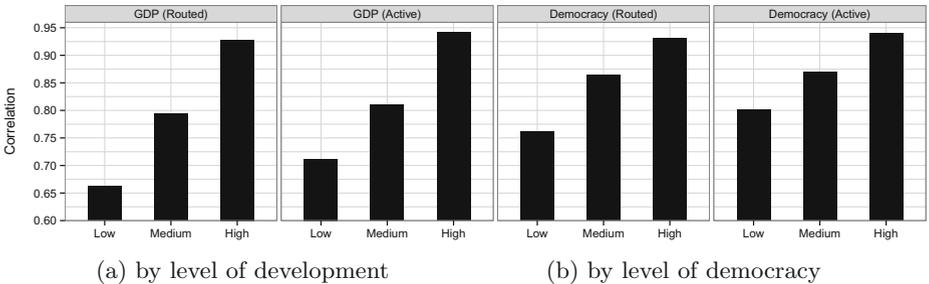

(a) by level of development   (b) by level of democracy

Fig. 4. Correlation between the ITU Internet penetration statistics and the number of routed/active IPv4 /24 prefixes of 192 countries grouped by their GDP and democracy index (Polity IV) into three equal-size groups.

The results show small but distinct trends in the correlations. Figure 4a reveals that the agreement between our estimates and the ITU figures increases for more developed countries, regardless of whether we use estimates based on routed prefixes or active networks. A similar trend can be identified in Fig. 4b for more democratic countries. These trends could have different causes: First, they could be due to limitations of data collection and biases in reporting that affect primarily less developed or less democratic countries. Second, it is possible that our method suffers from a lower accuracy of geo-localization in these countries, which could render it less precise in these contexts. However, due to fairly high accuracy of geo-localization at the country level (see below), we believe that the

second reason is probably less influential. This would mean that the trends in the correlations could indeed be due to differences in the quality of the ITU estimates across the different groups of countries.

4.3   Internet Penetration within Countries

Our analysis above compared country-level Internet penetration statistics to those inferred from network observations. In principle, however, the proposed approach can also be applied to the subnational level, by estimating penetration in sub-national units (such as provinces or districts) from the number of routed or active /24 blocks in that unit. Although not the main contribution of this paper, we provide a first analysis here. We focus on two countries for which subnational Internet penetration statistics are publicly available and which have an interesting political profile (India and Turkey). For both countries, the number of routed/active /24 blocks was computed using the geo-localized prefixes as described above, which were assigned to first-tier subnational units (states in India, and provinces in Turkey). The boundaries of these units were taken from the Global Administrative Areas database, a spatial dataset of internal administrative units [28].

India. For India, we used data from 2006 to 2010 for 29 states collected from the official statistics of Internet and Broadband Subscribers released by the Lower House of the Parliament [4]. As above, we reported correlations between the (logged) number of Internet users according to the official sources, and the (logged) number of prefixes in the respective state and time period. We match the statistics from India to the closest in time measurements from the February/August snapshot of the respective year. Figure 5 shows the results, where the official statistics show an average correlation coefficient of 0.956 with active and 0.942 with routed prefixes.
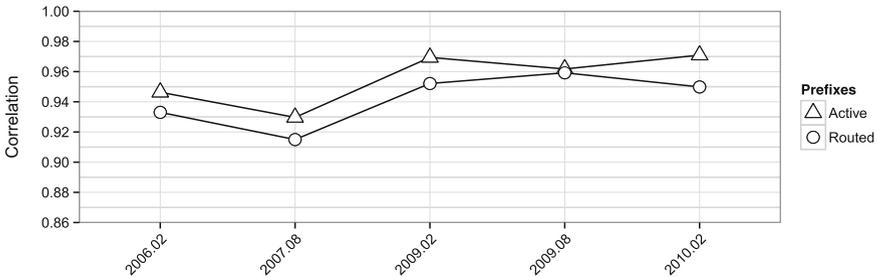

Fig. 5. Correlation over time between official Internet penetration statistics and the number of active/routed prefixes for 29 states within India.

Turkey. In the case of Turkey, we use the statistics of Internet usage for 2010 released for the 81 provinces by the largest ISP (TTNET) and the Information and Communication Technologies Authority [8]. After removing missing cases

and matching the administrative units, our final sample includes 65 provinces. Again, we find high correlations. The number of routed /24 blocks correlates with the official number of Internet subscribers at 0.89, which is slightly higher for active networks (0.907).

Thus, our method works well also at the subnational level. A key issue here, however, is the resolution and quality of IP geo-localization. The subnational units we use in this analysis are still fairly large; once we increase resolution down to the level of municipalities or even cities, low geo-localization accuracy becomes a key limitation as discussed next.

4.4 Discussion and Shortcomings

Address Space Over-/Underpopulation. One complication in comparing Internet penetration based on used address space is that an IP address may be used by a different number of subscribers in different regions of the world. Network address translation (NAT) has long broken any assumption of an 1:1 mapping between addresses and users. Further, the causes are not only political and economic, but related to Internet governance as well. IP addresses are allocated by five Regional Internet Registries (RIRs): ARIN for North America, LACNIC for Latin America and the Caribbean, RIPE for Europe and West Asia, APNIC for Asia and the Pacific, and AFRINIC for Africa. Each of these RIRs has a member base made up of Internet service providers and enterprises, a mission to allocate IP address space based on need, and its own framework for deciding policies for allocation of addresses to the members. As global IPv4 space has been exhausted, the different approaches within the different regions [15] have led to regionally linked amounts of pressure to conserve addresses by sharing them more broadly. A complete analysis of this phenomenon is outside the scope of this work, but this should be kept in mind when comparing Internet penetration numbers based on address counting across different RIR regions. Despite these differences, our analysis shows very high correlation coefficients across regions.

IP Address Geolocation. The accuracy of the MaxMind GeoIP database we use for geolocation, and of IP geolocation databases in general, is difficult to evaluate, and generally lacking in good sources of ground truth. Nonetheless, previous research has evaluated the accuracy of a set of these databases, including Max-Mind GeoIP, in 2011 [7]. For national-level data, the MaxMind GeoIP database we use agreed with the majority of other databases 99.1 % of the time, which was the best agreement ratio of any of the evaluated geolocation databases. For subnational data, the authors found that 78 % of the geolocated IP addresses globally were within 40 km of the centroid of the region most probably containing the IP address, with a great deal of regional variation: 75th percentile distances range from about 10 km in the ARIN region, to about 40 km in the APNIC region (containing India), to about 400 km in the LACNIC region. Any analysis of subnational-level IP geolocation data must therefore take the probable error into account, as well as the size of the regions in question. Given the comparison

to other databases, however, we have confidence in our selection of MaxMind GeoIP, and in our broad conclusions at both the national and subnational levels.

5 Conclusions

Official statistics about Internet penetration in different countries provided by the ITU and the OECD are widely-used in research studies and policy debates. However, due to the reliance on governments as the source of information, these statistics are derived from opaque methodologies, which may not be comparable. In addition, they are provided with significant delay and only at the national level. In this work, we propose an alternative Internet penetration indicator based on readily available measurements of the routed IP address space per country and show that this approach provides largely consistent results with the official ITU/OECD statistics. This helps both to increase confidence in the ITU/OECD data and to provide an alternative methodology with better data transparency, comparability, resolution, and availability. Furthermore, we showed that the high level of consistency drops for less developed or democratic countries. Finally, we also found that our approach is able to pick up variation in Internet penetration within two large developing countries.

To support our analysis and make our data more broadly accessible to the community, we provide visualisations of the growth of the Internet between 2004 and 2012, measured in terms of globally routed IPv4 addresses, versus the Gross Domestic Product (GDP), the income per capita, the population, and the polity index of 92 large countries in [1].

References

1. Internet growth versus economic and political indicators, October 2014. http://www.ics.forth.gr/tnl/ipen/index.html
2. Best, M.L., Wade, K.W.: The internet and democracy: Global catalyst or democratic dud? Bull. Sci. Technol. Soc. **29**(4), 255–271 (2009)
3. Dainotti, A., Benson, K., King, A., Claffy, K., Kallitsis, M., Glatz, E., Dimitropoulos, X.: Estimating Internet address space usage through passive measurements. ACM SIGCOMM Comput. Commun. Rev. (CCR) **44**(1), 42–49 (2014)
4. Datanet India Pvt. Ltd.: Indiastat.com (2014). http://www.indiastat.com/
5. Gleditsch, K.S.: Expanded trade and GDP data. J. Confl. Resolut. **46**, 712–724 (2002)
6. Groshek, J.: The democratic effects of the Internet, 1994–2003. Int. Commun. Gaz. **71**(3), 115–136 (2009)
7. Huffaker, B., Fomenkov, M., Claffy, K.: Geocompare: A comparison of public and commercial geolocation databases. CAIDA Technical report, May 2011. http://www.caida.org/publications/papers/2011/geocompare-tr/geocompare-tr.pdf
8. ICAT: Turkish electronic communications sector quarterly market reports (2013). http://www.btk.gov.tr/kutuphane_ve_veribankasi/yil_istatistikleri/ehsyib.pdf
9. ITU: Telecommunications development sector. http://www.itu.int/en/ITU-D/

10. ITU: Handbook for the collection of administrative data on telecommunications/ICT, 2011 (2011). http://www.itu.int/en/ITU-D/Statistics/Pages/publications/handbook.aspx
11. ITU: Manual for measuring ICT access and use by households and individuals (2011). http://www.itu.int/en/ITU-D/Statistics/Pages/publications/manual2014.aspx
12. ITU: World telecommunication/ICT indicators database (2013). http://www.itu.int/en/ITU-D/Statistics/Pages/publications/wtid.aspx
13. Jerven, M.: Poor numbers: How We Are misled by African Development Statistics and What to Do About It. Cornell University Press, Ithaca (2013)
14. Kende, M.: Global Internet report. Internet Society (2014)
15. Lehr, M., Lear, E., Vest, T.: Running on empty: The challenge of managing Internet addresses. In: Proceedings of the 36th Annual Telecommunications Policy Research Conference (TPRC), Arlington, VA, USA, September 2008
16. Marshall, M.G., Jaggers, K.: Polity IV project: Political regime characteristics and transitions, 1800–2012 (2013). http://www.systemicpeace.org/polity/polity4.htm
17. Maxmind GeoIP2 City Accuracy. https://www.maxmind.com/en/geoip2-city-accuracy
18. Maxmind: GeoIP2 Databases. http://www.maxmind.com/en/geoip2-databases
19. Milner, H.V.: The digital divide: The role of political institutions in technology diffusion. Comp. Polit. Stud. **39**(2), 176–199 (2006)
20. OECD: Key ICT indicators (2013). http://www.oecd.org/internet/broadband/oecdkeyictindicators.htm
21. Oyelaran-Oyeyinka, B., Lal, K.: Internet diffusion in Sub-Saharan Africa: A cross-country analysis. Telecommun. Policy **29**(7), 507–527 (2005)
22. Rød, E.G., Weidmann, N.B.: Empowering activists or autocrats? The Internet in authoritarian regimes. J. Peace Res. **52**(3), (2015, forthcoming)
23. Roeller, L.H., Waverman, L.: Telecommunications infrastructure and economic development: A simultaneous approach. Am. Econ. Rev. **91**(4), 909–923 (2001). http://www.jstor.org/stable/2677818
24. Sridhar, K.S., Sridhar, V.: Telecommunications and growth: Causal model, quantitative and qualitative evidence. Econ. Polit. Wkly. **41**(25), 2611–2619 (2006). http://www.jstor.org/stable/4418381
25. SWITCH: Swiss National Research and Education Network (NREN). http://www.switch.ch/
26. United Nations: Millenium development goals (2014). http://www.un.org/millenniumgoals/
27. United Nations Department of Economic & Social Affairs: United Nations e-government survey (2014). http://www.un.org/en/development/desa/publications/e-government-survey-2014.html
28. University of California, Berkeley Museum of Vertebrate Zoology and the International Rice Research Institute: Global Administrative Areas Dataset (2012). http://www.gadm.org/
29. University of Oregon: Route Views Project. http://www.routeviews.org/
30. WEF: Global information technology report (2014). http://www.weforum.org/issues/global-information-technology
31. Weidmann, N.B., Kuse, D., Gleditsch, K.S.: The geography of the international system: the cshapes dataset. Int. Interact. **36**(1), 86–106 (2010). http://dx.doi.org/10.1080/03050620903554614
32. World Intellectual Property Organization: Global Innovation Index (2014). http://www.wipo.int/econ_stat/en/economics/gii/