



On the African Peering Connectivity Revealable via BGP Route Collectors

Enrico Gregori, Alessandro Improta^(✉), and Luca Sani

Institute of Informatics and Telematics,
National Research Council of Italy, Pisa, Italy
{enrico.gregori,alessandro.improta,luca.sani}@iit.cnr.it

Abstract. Internet pervasiveness in Africa has been slowly but steadily increasing since the beginning of this millennium. Thanks to several organisations which donated time and resources, it is nowadays possible to claim that the AS ecosystems of several countries in Africa are now experiencing an early stage of the peering era. In this paper, we investigate the capability of the BGP route collectors publicly available to reveal the newborn peering connectivity in African countries. By analysing BGP data available with existing techniques we found that a lot of this connectivity is missing from the dataset, mainly due to the lack of data sources in the region. In most countries, this could theoretically be solved by introducing no more than ten new ASes sharing their full routing information to route collectors.

Keywords: Internet · AS-level · BGP · Route collectors · Set coverage

1 Introduction

The Internet is composed of as a set of heterogeneous and independent networks, each of which compete and cooperate with each other by means of the Border Gateway Protocol (BGP) to build the routes that will carry the actual traffic. This ecosystem can be analysed at different levels of abstraction depending on the type of analysis that has to be carried on [1, 16], e.g. Autonomous Systems (ASes) [15, 22], IP/router [6, 10, 13], Points of Presence (PoPs) [4]. The AS-level in particular is helpful to analyse how the different players composing the Internet (i.e. ASes) interact with each other – in terms of BGP routing – without focusing too much on details concerning the inner structure of each player. The complete knowledge of the AS-level of a given region would also help network administrators in that very same region to plan their inter-domain routing in advance, therefore introducing the proper amount of redundancies in their provider choices so that possible problems of the regional Internet would have the minimum impact on the performances of their own networks.

Thanks to several heuristics and inference techniques developed in the past years, it is now possible to infer the economic agreements existing between pairs of ASes (e.g. [5, 8]) and geolocate each AS to perform economic and regional

analyses [7, 20]. In this paper we exploit some of these latter heuristics to analyse the Internet AS-level of the African Internet ecosystem as inferred from BGP data made publicly available by BGP route collecting projects. In particular, we focus on quantifying the completeness of the available BGP data in terms of data sources and their distribution in Africa, relying on the *p2c-distance* metric introduced in [8]. We found that the Internet ecosystem reflects the heterogeneity already shown by Africa in terms of culture and development and highlighted by recent work [2]. In the very same continent coexist countries where the ecosystem is rich in terms of ASes and Internet eXchange Points (IXPs), and countries where the Internet is at its very early stage of development [18] due to poor infrastructures and expensive transport services [24]. Most of the peering ecosystem in the area is however invisible to current route collecting infrastructure due to the overall absence of African ASes sharing their routing information, with exception of South Africa. Nevertheless, we also found that it could be theoretically possible to reveal the full peering connectivity of most countries just by adding less than ten new ASes sharing their full routing information to the current route collector infrastructure.

The rest of the paper is organised as follows. Section 2 describes the data sources exploited in our analysis and the methodologies applied to infer the AS-level. Section 3 gives a general overview of the African Internet AS-level ecosystem and Sect. 4 analyses the completeness of available BGP data. Finally, Sect. 5 concludes the paper.

2 Data Sources and Methodology

The Internet AS-level is typically represented as a graph where nodes are ASes and connections are BGP sessions established between ASes. Two ASes that decide to establish a BGP session essentially exchange a set of network reachability information which is used to route part of their Internet traffic between them. The amount and the nature of the reachability information exchanged totally depend on the economic agreements signed between the two ASes, which can typically be classified as provider-to-customer (p2c) or peer-to-peer (p2p) [5]. In the former, the provider announces to the customer the routes to reach all the Internet networks, whereas the customer announces to the provider the routes to reach its networks and its client networks (if any). In the latter, the two ASes exchange routes to reach their respective clients, typically to *keep local traffic local* and to reduce transit costs [17].

The primary data source [25] to analyse the Internet at the AS-level of abstraction is the BGP data collected and provided by organisations running route collectors, such as the Réseaux IP Européens Network Coordination Centre (RIPE NCC) with the Routing Information Service¹ (RIS), the University of Oregon with the Route Views project² and the Institute of Informatics and Telematics of the Italian National Research Council (IIT-CNR) with the Isolario

¹ <http://ris.ripe.net>.

² <http://www.routeviews.org/>.

project³. A route collector is a server running a BGP routing daemon which collects and stores routing information in Multi-Threaded Routing Toolkit (MRT) format, and does not announce any reachability information back to its BGP neighbors. Every AS is free to join and share its routing information with the public, contributing to improving the amount – and thus the quality – of BGP data available for research purposes. The fundamental piece of information found in collected BGP data to analyse the AS-level ecosystem is the `AS_PATH` attribute, which can be used to infer both the nodes (ASes) and the connections (AS adjacencies) of the AS-level topology.

As described in [7], it is possible to infer the AS geolocation by geolocating each of the subnets announced by the AS thanks to the availability of databases mapping IP addresses in countries. This technique allows to infer regional topologies just considering that each AS adjacency can be geolocated if both ASes are announcing (at least) one subnet in the very same country. In the end, the set of ASes geolocated in a given country/continent will contain both ASes owned by local organisations and strictly linked to the territory where they operate – hereafter *local ASes* – and ASes owned by international organisations that are operating in the very same territory for marketing purposes – hereafter *international ASes*. In this paper we will consider a given AS to be local if there is an entry in the AFRINIC registry related to that AS.

The analyses performed in the following sections are computed on BGP data collected by every route collector made available by Isolario, RIS and Route Views on August 8th, 2017. Geolocation is performed using the Maxmind GeoLite2 Country⁴ database, while economic relationships are inferred using the algorithm described in [8].

3 A First Glance at the African AS-Level Ecosystem

Africa is an extremely heterogeneous continent in terms of language, culture, and economics, and this heterogeneity can also be recognized also in its AS-level ecosystem. In the very same continent coexist a quite large set of countries which are clearly behind the digital divide – and where the Internet is the least of their peoples’ worries – with a set of progressive countries where Internet is starting to be a consolidated part of their economies. Table 1 shows details about the distribution of ASes in each country. Out of 1084 local ASes, South Africa receives the lion’s share with 322 ASes, followed by Nigeria (145 ASes), Kenya (79 ASes), Tanzania (63 ASes) and Ghana (56 ASes). One of the most impressive feature that can be noticed at glance is the poor pervasiveness of IPv6 despite the efforts spent by several organisations in training sessions and IPv6 focused conferences. Every local AS announces on the Internet at least one IPv4 network while only 203 of them announce (at least) one IPv6 network. The latter set of ASes is mainly distributed among South Africa (47%), Tanzania (13%), Kenya (10%), Mauritius (9%) and Nigeria (6%). Another interesting

³ <https://isolario.it>.

⁴ <https://dev.maxmind.com/geoip/geoip2/geolite2/>.

Table 1. African AS-level ecosystem data (IXP list provided by AF-IX website).

Country	ASes (v4)		ASes (v6)		IXPs	Country	ASes (v4)		ASes (v6)		IXPs
	<i>Loc</i>	<i>Int</i>	<i>Loc</i>	<i>Int</i>			<i>Loc</i>	<i>Int</i>	<i>Loc</i>	<i>Int</i>	
Algeria	10	17	2	1	0	Mali	3	7	1	1	0
Angola	40	15	4	2	1	Mauritania	3	13	0	1	0
Benin	10	4	2	1	1	Mauritius	29	22	11	3	1
Botswana	18	8	2	1	1	Mayotte	1	5	0	0	0
Burkina Faso	8	8	0	1	1	Morocco	10	19	4	1	0
Burundi	9	6	0	1	0	Mozambique	19	13	1	1	1
Cameroon	17	9	1	1	0	Namibia	12	6	3	1	1
Cape Verde	3	8	1	1	0	Niger	10	10	0	1	0
Central African Rep	3	9	0	0	0	Nigeria	144	30	7	1	3
Chad	8	10	0	1	0	Rep. of Congo	11	9	1	1	1
Comore	2	7	0	1	0	Réunion	3	19	1	2	0
Côte d'Ivoire	15	5	2	1	1	Rwanda	14	6	1	1	1
Djibouti	6	6	1	2	1	São Tomé and Príncipe	2	7	0	1	0
DR Congo	23	20	0	1	1	Senegal	4	9	0	2	0
Egypt	56	30	4	2	1	Seychelles	15	36	1	4	0
Equatorial Guinea	6	8	2	1	0	Sierra Leone	10	7	0	1	0
Eritrea	1	7	0	0	0	Somalia	13	8	0	1	0
Ethiopia	1	8	0	1	0	South Africa	315	87	87	24	6
Gabon	11	19	2	1	1	South Sudan	8	5	0	1	0
Gambia	7	8	1	1	1	St. Helena	0	3	0	0	0
Ghana	56	6	4	3	1	Sudan	6	8	4	1	1
Guinea	9	8	0	1	0	Swaziland	10	8	0	1	0
Guinea-Bissau	1	6	0	1	0	Tanzania	62	19	23	1	2
Kenya	78	31	16	6	2	Togo	4	7	1	1	0
Lesotho	9	7	0	1	1	Tunisia	13	16	4	3	1
Liberia	7	11	1	1	1	Uganda	32	17	3	1	1
Libya	5	13	0	1	0	Western Sahara	2	5	0	0	0
Madagascar	5	12	1	1	1	Zambia	23	13	1	1	1
Malawi	11	9	2	1	1	Zimbabwe	17	15	4	1	0

aspect is that just 90 local ASes (about 8%) are located in more than a single country, highlighting how traffic transiting between neighboring countries is still demanded to international providers.

Similarly to the rest of the Internet ecosystem, the peering ecosystem in Africa is at a very early stage of development. Not many years ago most of local traffic was routed via Europe and North America [11], causing issues in performance due to high latencies. Things started changing during last decade, when initiatives like the African Internet Exchange System (AXIS) Project⁵

⁵ <https://au.int/en/axis>.

led to a dramatic increase in the number of IXPs in the region. Nowadays in Africa, there can be found 37 active IXPs located in 34 cities in 28 countries⁶ (Table 1). Scraping the websites of each IXP, it is easy to see that most of them have currently less than 20 ASes connected, with the notable exception of NAPAfrica in South Africa (273 ASes among Johannesburg, Cape Town and Durban), JINX in South Africa (82 ASes in Johannesburg), IXPN in Nigeria (54 ASes among Lagos, Abuja and Port Harcourt), KIXP in Kenya (36 ASes between Nairobi and Mombasa), TIX in Tanzania (36 ASes between Dar es Salaam and Arusha) and UIXP in Uganda (26 ASes in Kampala). The presence of an IXP as crowded as NAPAfrica in South Africa stress even more how South Africa's Internet ecosystem is totally different from the rest of Africa, resembling the ecosystem of a European country. Finally, it must be stressed out that most IXPs interconnect ASes only at country-level [3] and it has been shown that ASes often do not peer with one another at local IXPs [11].

4 On the Completeness of African AS-Level Graph

It is well known that BGP data is far from being completely representative of the Internet AS-level ecosystem [8, 19, 21]. First, the number of ASes participating in any route collecting project is extremely low if compared to the whole size of ASes composing the Internet. During our analysis, only 525 ASes were sharing their routing information with Isolario, RIS and/or Route Views, while the total number of ASes routed on the Internet was 59,005. Second, route collectors are not receiving complete routing information from all of their feeders. Several collectors are placed on IXPs across the world and many feeders apply to them the very same export policies applied to other IXP participants. In other words, they announce to the route collectors only their customer cone [14], which provides an extremely limited view of the Internet. During our analysis, about half of the feeders were showing this kind of behavior, with only 257 ASes sharing an IPv4 space and 200 ASes sharing an IPv6 space close to a full routing table. The feeders announcing their full routing table to route collectors will be hereafter referred to as *full feeders*. Finally, BGP data is known to miss a large part of p2p connectivity established at IXPs and via private peering [12]. This is mostly caused by the location of full feeders in the AS graph and the presence of BGP export policies and economic relationships between ASes. Given the standard economic relationships established between ASes [5], an AS announces to the other AS its full routing information – containing routes learned from its peers, providers and customers – but only if it is a provider of the other party. As a consequence, a route collector is able to see routes established via IXPs and private peering of a given AS X only if exists a chain of transit relationships from the route collector towards X. This concept has been formalised as *p2c-distance* in [8], and it has been used to quantify the number of ASes for which it is possible to discover the full connectivity given a set of full feeders. The resulting graph incompleteness must be taken into serious consideration when

⁶ <http://af-ix.net/>.

analysing the Internet at the AS-level of abstraction since it can easily lead to wrong conclusions, especially when analysing the graph properties [23]. To the best of our knowledge, this is the first work analysing the coverage of route collectors in the African region. Only Fanou et al. [2] indirectly tackled the vantage point placement issue in Africa by introducing several RIPE Atlas probes in the region.

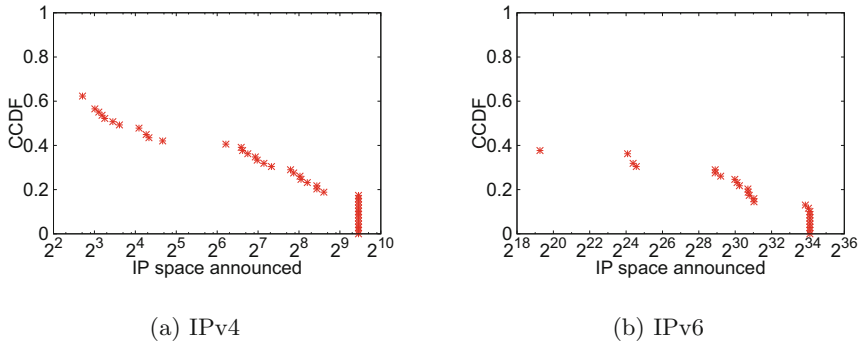


Fig. 1. CCDF of the amount of IPvX space announced by each feeder

The coverage situation in Africa is not very different from the rest of the world. Currently there are three route collectors in Africa physically deployed at KIXP in Kenya (Route Views), JINX in South Africa (Route Views) and NAPAfrica in South Africa (RIS). Those collectors receive data from 69 feeders, 63 located in South Africa, 4 in Kenya and 2 in Mauritius. An additional feeder from South Africa is connected to Isolario via multihop BGP. Figure 1 shows the Complementary Cumulative Distribution Function (CCDF) of the IPvX space announced to the route collectors by African feeders. As can be seen, the vast majority of feeders announce only a small portion of IP space, much smaller than the respective full routing table space which nowadays is composed of around 600k (v4) and 40k (v6) routes. Out of 69 feeders, only 13 can be considered v4 full feeders and only 9 v6 full feeders. All of them are located in South Africa, with the exception of one v4 and one v6 full feeders located in the island of Mauritius. Thus, it is almost straightforward to understand that the peering connectivity established at the 30 IXPs in Africa located neither in South Africa nor in Mauritius is currently completely hidden to BGP route collectors, while the small number of full feeders available in South Africa and Mauritius do not allow them to reveal much of the peering connectivity in their countries. Taking into account the p2c-distance metric, it is possible to claim that the current full feeders allow to reveal the full connectivity of 29 transit IPv4 ASes out of 129 (22.5%) and 5 out of 28 transit IPv6 ASes (17.9%) in South Africa whereas it is possible to discover the full connectivity of 6 transit IPv4 ASes out of 31 (19.4%) and no IPv6 transit ASes over 9 in Mauritius. In the rest of the African

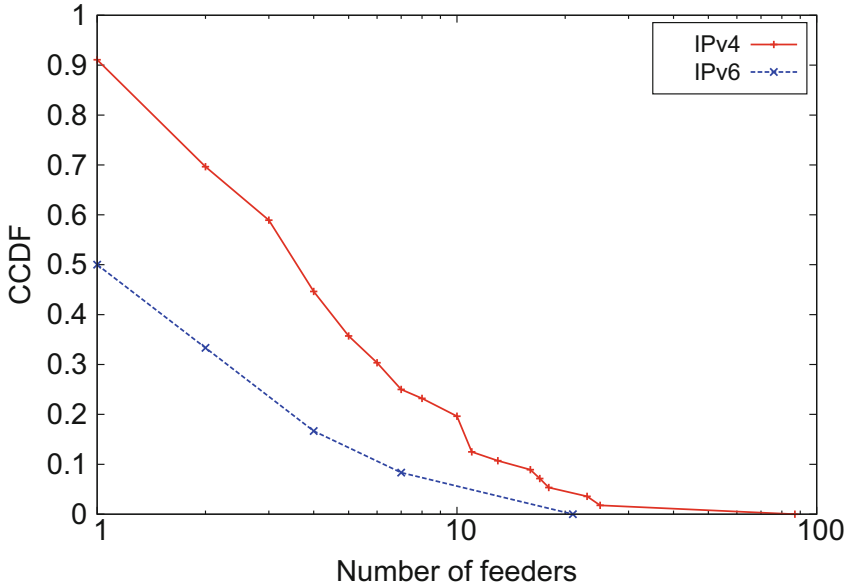


Fig. 2. CCDF of the solution cardinality of MSC problem in each African country

countries, the only ASes covered are the international ASes, which are covered from feeders outside of Africa.

To better understand how far the current BGP measurement system is from the ideal condition, where the entire p2p connectivity of each country is revealable and potentially visible, we applied the Minimum Set Cover (MSC) problem described in [9] to each regional topology gathered from BGP data. In each regional scenario every AS available is considered to be a potential feeder with its own *covering set* – i.e. the set of transit ASes having a finite p2c-distance from the AS – and the goal of the problem is to find the minimum number of ASes whose covering sets cover the whole set of transit ASes in that region. Figure 2 shows the CCDFs of the number of feeders required in each African country. Note that the v6 scenario is computed based only on 12 countries⁷ where ASes were connected to each other. The scenario of South Africa is the most distinguishable in both pictures. Given the large amount of ASes in both v4 and v6 scenarios, a rather large number of feeders is required to obtain the full coverage of transit ASes. In all the other cases though, the number of feeders required is quite low, often smaller than 10 either in v4 or v6. This means that with a considerably small effort – 10 full BGP sessions to be established – it could be possible to reveal the full peering connectivity of 90% of countries in Africa.

⁷ Angola, Botswana, Djibouti, Equatorial Guinea, Kenya, Mauritius, Morocco, Nigeria, South Africa, Sudan, Tanzania and Tunisia.

5 Conclusion

Africa shows in its AS-level ecosystem the same heterogeneity it shows terms of culture, economics and development. The most developed AS-level ecosystem can be found in South Africa, where the peering ecosystem is extremely similar to most European countries, as proved by the number of IXP available. Then, there is a small set of progressive countries (e.g. Egypt, Kenya, Nigeria and Tanzania) where Internet pervasiveness is steadily increasing and being more and more an important part of their economy. Finally, there is a large set of countries where Internet is at the very early stage of development. In this ecosystem, we found that BGP route collectors almost completely fail to reveal the peering connectivity established among ASes, thus affecting any possible graph analysis concerning the African ecosystem. Despite that, we found that theoretically it could be possible to solve this situation in most African countries by introducing just ten new full feeders.

References

1. Donnet, B., Friedman, T.: Internet topology discovery: a survey. *IEEE Commun. Surv. Tutor.* **9**(4), 56–69 (2007). <https://hal.archives-ouvertes.fr/hal-01151820>
2. Fanou, R., Francois, P., Aben, E., Mwangi, M., Goburdhan, N., Valera, F.: Four years tracking unrevealed topological changes in the african interdomain. *Comput. Commun.* **106**, 117–135 (2017)
3. Fanou, R., Valera, F., Francois, P., Dhamdhere, A.: Reshaping the African Internet: from scattered islands to a connected continent. *Comput. Commun.* **113**, 25–42 (2017)
4. Feldman, D., Shavitt, Y.: Automatic large scale generation of internet pop level maps. In: *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pp. 1–6 (2008)
5. Gao, L.: On inferring autonomous system relationships in the internet. *IEEE/ACM Trans. Netw.* **9**(6), 733–745 (2001)
6. Govindan, R., Tangmunarunkit, H.: Heuristics for internet map discovery. *Proc. IEEE INFOCOM 2000*, 1371–1380 (2000)
7. Gregori, E., Improta, A., Lenzini, L., Rossi, L., Sani, L.: Discovering the geographic properties of the internet AS-level topology. *Netw. Sci.* **3**(1–4), 34–42 (2013)
8. Gregori, E., Improta, A., Lenzini, L., Rossi, L., Sani, L.: Improving the reliability of inter-AS economic inferences through a hygiene phase on BGP data. *Comput. Netw.* **62**, 197–207 (2014)
9. Gregori, E., Improta, A., Lenzini, L., Rossi, L., Sani, L.: A novel methodology to address the internet AS-level data incompleteness. *IEEE/ACM Trans. Netw.* **23**(4), 1314–1327 (2015)
10. Gunes, M.H., Saraç, K.: Analytical IP alias resolution. In: *2006 IEEE International Conference on Communications*, vol. 1, pp. 459–464 (2006)
11. Gupta, A., Calder, M., Feamster, N., Chetty, M., Calandro, E., Katz-Bassett, E.: Peering at the internet’s frontier: a first look at ISP interconnectivity in Africa. In: Faloutsos, M., Kuzmanovic, A. (eds.) *PAM 2014. LNCS*, vol. 8362, pp. 204–213. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-04918-2_20

12. He, Y., Siganos, G., Faloutsos, M., Krishnamurthy, S.V.: Lord of the links: a framework for discovering missing links in the internet topology. *IEEE/ACM Trans. Netw.* **17**(2), 391–404 (2009)
13. Li, L., Alderson, D., Willinger, W., Doyle, J.: A first-principles approach to understanding the internet’s router-level topology. *ACM SIGCOMM Comput. Commun. Rev.* **34**(4), 3–14 (2004)
14. Luckie, M., Huffaker, B., Dhamdhere, A., Giotsas, V., Claffy, K.C.: AS relationships, customer cones, and validation. In: *Proceedings of the 13th ACM SIGCOMM Conference on Internet Measurement, IMC 2013*, pp. 243–256 (2013)
15. Mahadevan, P., Krioukov, D., Fomenkov, M., Dimitropoulos, X., Claffy, K.C., Vahdat, A.: The Internet AS-level topology: three data sources and one definitive metric. *ACM SIGCOMM Comput. Commun. Rev.* **36**(1), 17–26 (2006)
16. Motamedi, R., Rejaie, R., Willinger, W.: A survey of techniques for internet topology discovery. *IEEE Comm. Surv. Tutor.* **17**(2), 1044–1065 (2015)
17. Norton, W.B.: *The Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press, Palo Alto (2011)
18. Nyirenda-Jere T., Biru T.: Internet development and Internet governance in Africa. <https://www.internetsociety.org/doc/internet-development-and-internet-governance-africa>. Accessed 06 June 2018
19. Oliveira, R., Pei, D., Willinger, W., Zhang, B., Zhang, L.: The (in)completeness of the observed internet AS-level structure. *IEEE/ACM Trans. Netw.* **18**(1), 109–122 (2010)
20. Rasti, A.H., Magharei, N., Rejaie, R., Willinger, W.: Eyeball ASes: from geography to connectivity. In: *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, IMC 2010*, pp. 192–198 (2010)
21. Roughan, M., Tuke, S.J., Maennel, O.: Bigfoot, Sasquatch, the Yeti and other missing links: what we don’t know about the AS graph. In: *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, IMC 2008*, pp. 325–330 (2008)
22. Roughan, M., Willinger, W., Maennel, O., Perouli, D., Bush, R.: 10 lessons from 10 years of measuring and modeling the internet’s autonomous systems. *IEEE J. Sel. Areas Commun.* **29**(9), 1810–1821 (2011)
23. Willinger, W., Alderson, D., Doyle, J.C.: Mathematics and the internet: a source of enormous confusion and great potential. *Not. Am. Math. Soc.* **56**(5), 586–599 (2009)
24. World Bank: World development report 2016: digital dividends. <http://documents.worldbank.org/curated/en/896971468194972881/World-development-report-2016-digital-dividends>. Accessed 06 June 2018
25. Zhang, B., Liu, R., Massey, D., Zhang, L.: Collecting the internet AS-level topology. *ACM SIGCOMM Comput. Commun. Rev.* **35**(1), 53–61 (2005)