On the Prevalence of Boomerang Routing in Africa: Analysis and Potential Solutions

Assane Gueye[1,2](✉) and Babacar Mbaye[1]

[1] Universite Alioune Diop de Bambey, Bambey, Senegal
{assane1.gueye, babacar.mbaye}@uadb.edu.sn
[2] University of Maryland, College Park, College Park, USA

Abstract. When an African Internet user sends a message to a friend in another country in the continent, the data travels around the world (mostly Europe and USA) before coming back to the continent. This phenomenon is called *boomerang routing* at the continent level. The implications of boomerang routing include: higher cost, increased delay and increased information exposure. In this paper, we use active measurement data (ICMP traceroute) from 2015 and 2016 to empirically study the prevalence of boomerang routing in Africa by focusing on the implications on cost and delay. We also discussed possible improvements of the current African connectivity map to reduce boomerang routing, which will eventually translate into reduced cost and lower delays for end-users.

Keywords: Africa · Boomerang · ICT · e-Trading · Internet access cost Delay

1 Introduction

The 21st century has been dubbed by many as the African century. In a recent printout, McKinsey reports that, following a decade of economic expansion, Africa is going digital. Although only 16% of the people in the continent are online today, that share is rising rapidly, thanks to recent growth in mobile networks and falling cost of Internet-capable devices. In 2016, there were more than 720 million mobile phone users in Africa, some 167 million African already used the Internet, and 52 million were on Facebook [1].

This follows a global trend of information and communication technologies (ICT) being an enabler for increased socio-economic growth. Not only ICTs help increase productivity, they also contribute to the overall GDP of an economy. ICTs can also help connect remote populations to markets, promote citizens' access to government and social services, expand educational opportunities, create platforms for innovation and increase people's freedoms [2].

This strong impact on socio-economic development is likely to carry-on in a much larger scale in the coming decade. This is significant for Africa because, according to the World Bank [3], mobile and broadband have more impact in developing economies than in developed economies. This has already been observed in the mobile communication sector. In 2014 the impact of mobile phones in term of GDP was 5.7% in Africa, which was more than any other continent [3].

C. M. F. Kebe et al. (Eds.): InterSol 2018, LNICST 249, pp. 3–12, 2018.
https://doi.org/10.1007/978-3-319-98878-8_1

Today, the Internet's contribution to Africa's GDP is relatively low, at 1.1%—just over half the levels seen in other emerging markets and well below the average of 3.7% in developed economies [4]. However, studies [1] have found that a 10% increase in broadband penetration in developing countries is correlated with a 1.35% increase in GDP, suggesting that Africa has even more to gain from the digital revolution. By 2025, the digital contribution to African GDP is expected to catchup with Sweden and Taiwan [1]. If the Internet matches or exceeds that level of impact, the result could be a leap forward in Africa's economic growth and development.

However, to meet those expectations, a certain number of barriers need to be removed. Today there is growing consensus that the primary barriers to connectivity in Africa are availability, affordability, relevance and readiness [5]. Relevance refers to information and service available to people's in their native language, while readiness reflects to people having the necessary basic skills, awareness, understanding, and cultural and social acceptance to the Internet.

In this paper, we are primarily interested in the availability of Internet infrastructures and the affordability to the end-users. More specifically, we are interested in understanding how the lack of a continent-wide connectivity infrastructures is impacting the cost to the end-users. For that, we analyze data (gathered between January 1, 2015 and December 31, 2016) to quantify the prevalence of *boomerang* routing in intra-Africa Internet communication.

Boomerang routing (or *tromboning* [7]) was originally used [6] to designate the situation whereby an Internet path starts and ends in a country, yet transits in an intermediate country. In this paper, we consider boomerang routing at the continent level and focus our attention to Africa. More precisely, we are interested in the implications of boomerang routing in cost and delay. This is important because today Africa is the region with the most expensive fixed-broadband prices, with an average price of more than 100% of GNI per capita [2]. Also, anyone who has used the Internet in Africa, might have experienced very long delays when trying to download a relatively large file.

For our study, we use publicly available data from the Center for Applied Internet Data Analysis (CAIDA) [8], which consists of router level communication traces for the entire 2015–2016 period. We then use the Max-Mind service [9] to geolocate each router within a country and we convert the router-level paths into country level paths. The dataset includes paths from the entire world. In this study, we only exploit paths for which both the origin and the destination are in Africa.

We first count the number of paths for which the traffic leaves the continent before coming back (i.e., there is a boomerang). Our analysis shows that in 2015, 70% of intra-Africa communication exited the continent before reaching the destination. In 2016, the fraction of boomerang traffic was 60%. When it exits, most of the traffic passes through the EU. US comes in second position. We also compute the number of hops a boomerang traffic goes through outside the continent. We found that in average, when exited, the traffic traverses 3.5 hops before returning to the continent.

We then focus on the communication delay (mainly round-trip-time) of the traffic. Our analysis suggests that, despite the fact that boomerang traffic roams in other continents before coming back to Africa, their RRT delay is smaller than the RTT delay of none-boomerang traffic that entirely stays within the continent.

The time that the traffic spends during the boomerang period is a key parameter for the cost of communication seen by end-users. In fact, when using international Internet connection, the African Internet Service Provider leases connectivity infrastructures and pays a fee known as transit cost. This fee is eventually transferred to the end-user who is then paying an additional (international) cost even when communicating with a friend in the same or neighboring country. A similar phenomenon is happening in the airline industry, where to go from one African country to another, travelers go through Europe, leading to additional cost.

A solution to this situation that has gained a large consensus is to build new communication infrastructures. This is certainly the reason why it is well considered in the African Internet Exchange System (AXIS) Project [10]. The main recommendation in this project is to build new Internet Exchange Points (IXPs) in the continent. Our study shows that, based on the data we gathered and the current connectivity between the existing IXPs, although it is important to build new IXPs, it is more pressing to connect the existing ones and to improve the bandwidth within the continent.

The rest of this paper is organized as follows. In Sect. 2, we give an overview of the current international Internet Peering principle and we talk about the boomerang routing phenomenon and its consequences in cost, delay and privacy. We then describe the dataset used in our study in Sect. 3 and perform our analysis in Sect. 4. In Sect. 5 we explore ways to improve the current African connectivity map to reduce boomerang routing. Concluding remarks and future work are presented in Sect. 6.

2 International Internet Connectivity: Peering, Transit and Boomerang

For international telephone routing, the provider in the country that originates the call makes a compensation payment to the provider in the destination country. If traffic is balanced in both directions, these payments usually cancel each other out. In the 1990s, traffic was higher in the direction of developed countries to developing countries and the ITU estimated that the net flow of settlement payments from developed countries to developing ones amounted to some USD 40 billion between 1993 and 1998 [11]. By the late 1990s, these payments have declined and even reversed as more traffic has shifted to the Internet.

Interconnectivity in the global Internet is essentially done through *peering* and *transit*. Peering refers to a relationship between two or more ISPs of similar size, in which the ISPs create direct links between each other and agree to forward each other's packets directly across this link. Transit refers to a bilateral interconnection where the customer pays the provider for connectivity to the global Internet. It is mainly the service provided by larger ISPs to smaller ISPs (such as those in Africa).

With the current international Internet charging principles (called full-circuit model), ISPs based in countries remote from Internet backbones, particularly in the developing countries, must pay the full (transit) cost of the international circuits. For example when a user in Kenya sends an email to a user in the USA, it is the Kenyan ISP who is bearing the full cost of the international connectivity. Conversely, when a user in the USA sends an email to a user in Kenya, it is still the Kenyan ISP who bears the international cost (different from the telephone network where payments goes in both directions) [12]. This cost is eventually transfer to the end user in Kenya.

Because most Africa ISPs rely on transit via ISPs in Europe on in the US, this very phenomenon is still experienced when an African Internet user sends a message to a friend in another country in the continent. In this case, the data travels around the world (mostly Europe and USA) before coming back to the continent. This phenomenon is dubbed *boomerang routing* at the continent level. In this case, ISPs also must pay the international transit cost. This represents an additional cost for intra-Africa communication that is due to the boomerang routing.

Other consequences of boomerang (that are not detailed in this paper) are increased delay and privacy or information exposure. Indeed, since communication is bound by the speed of light, when the traffic roams around the world before coming back to Africa, this creates additional delays, which are commonly experienced by users in Africa. Also, data sent over the Internet can be monitored and manipulated by the entities that transmit that data from the original source to the destination. For unencrypted communications (and some encrypted communications with known weaknesses), eavesdropping and man-in-the-middle attacks are possible. For encrypted communication, the identification of the communicating endpoints is still revealed. In addition, encrypted communications may be stored until newly discovered weaknesses in the encryption algorithm or advances in computer hardware render them readable by attackers. Thus, when two African countries are communicating, the information is exposed to the rest of the world because of the boomerang routing phenomenon. In this paper, we study the prevalence of boomerang routing in intra-Africa communication. For that, we use real-life communication data gathered during the period of 2015–2016. In the next section we describe the data utilized in this study.

3 Data Acquisition and Description

3.1 Data Acquisition

We obtained our data from the public datasets provided by the Center for Applied Internet Data Analysis (CAIDA) [8] covering the years 2015 and 2016. We collected traceroute type data from a worldwide set of monitors. Every 2–3 min, each monitor probes a random prefix using ICMP traceroute and stores the router-level path information. We convert these paths into country paths by using the MaxMind [9] geolocation database. While geolocation data of routers can be inaccurate, previous work has found that it is more accurate at a country level of abstraction [13]. We then filter out the traces for which both the source and the destination are within Africa and keep them for our analysis. Overall, we collected a total of or 140K traces for 2015 and 350K traces for 2016.

3.2 Data Limitation

The main limitation of the data is the low number of sources countries of the collected traces. In fact, the Ark infrastructure only has a limited number of monitors within

Africa (only in five African countries in 2015 and in 13 countries in 2016). Since the source of a given trace is the country where it is geo-located, this limits the number of source countries in our analysis. However, previous studies have shown that these countries are very representative when it comes to Internet communication in the continent. Therefore, we recognize that our study might not be complete but believe that it generates well to the whole continent.

Another limitation of the data is that the Ark infrastructure is only capable of discovering preferred paths to and from subnets containing a monitor. Routes between subnets that do not lie on a preferred route from a monitor to a target subnet will not be discovered. However, given that each subnet will be probed from a random monitor every two to three days, the traceroutes to each subnet will come from many different directions. Thus, we have confidence that we are discovering the major pathways from each source countries to the rest of the continent.

Despite these limitations, we believe that our study can be used to have a very good insight about Internet communication in Africa. Furthermore, since data is in general very rarely available in Africa, we consider this study as one step towards the general effort to understanding the evolution and implications of Internet communication within the continent.

4 Data Analysis

With the data collected for each year, we perform a certain number of experiments. First, we count the number of boomerang traces. Those are traces that contain routers that were geo-located outside the continent. Then, for each boomerang trace, we compute the number of countries that are traversed during the "*exit period*". Finally, we record the round-trip-time (RTT) for each trace. For non-boomerang traces, we only record the RTT to the destination. For boomerang traces, we record the time for the three portions of the communication: the first portion before exiting the continent, the exit portion (when the traffic roams outside the continent), and the return portion when the traffic comes back to the continent. It is worth to notice that we did not find any trace that boomerang more than once.

4.1 Number of Boomerang Routes

Table 1 summarizes the results of our study. Among the 138505 traces collected in 2015, 97157 exited the continent before arriving to the destination country. This represents 70% of the traffic. This percentage improved slightly in 2016 where 60% of the traffic was found to boomerang.

These numbers are consistent with previous studies [14, 15]. In [14], Fanou *et al.* collected and analyzed traceroutes data from November 2013 to April 2014, while Chavula *et al.* analyzed data collected during a two-week period between April 6th and April 20th 2014. Both studies revealed a percentage of more than 70% of boomerang routes. Our study (which can be considered as a complement and update to these previous studies) shows that although there is a slight improvement in 2016, most of the intra-Africa communication traffic still leaves the continent before coming back.

This translates into high costs of intra-Africa communication that are mainly paid to ISPs in Europe and the US. According a study of the African Union [16], Africa spends between US $400 millions and $600 million per year in transit fees for intra-African traffic [14].

Table 1. Results summary

	Year 2015	Year 2016
Number of traces	138505	354575
Number of boomerang	97157	215259
% boomerang	70.14	60.07

In our analysis, we were also interested in two other questions: (1) where does the "exit traffic" go? And (2) how many countries are traversed outside Africa?

The data shows that most of the boomerang traffic goes to Europe (58%). We could not perform a more granular study because there are many routers for which the geo-location database just returns "EU" as country of residence, without further information about which European country it is. The US comes in second position receiving 36% of intra-Africa traffic. In very rare occasion the traffic roams through Asia and Latin America. When the traffic gets outside of Africa, we found several patterns: some traces remain entirely within Europe, some within the US and some traces go both through Europe and the US. Finally, the data shows that a boomerang traffic will traverse in average 3.5 countries. This indicates a high degree of information exposure when African countries are communicating over the Internet.

4.2 Delay Analysis

In this section, we want to understand the causes of the large delays experienced in Africa-to-Africa communication. For that, we record the round-trip-time (RTT) for each trace. For non-boomerang traces, we only record the RTT to the final destination. For boomerang traces, we record the RTT delay for the three portions of the communication: (2) the first portion before exiting the continent. We approximate this delay as the RTT between the source of the trace and the last route of the trace that is still within Africa. (b) the exit portion (when the traffic roams outside the continent). We estimate this delay as the RTT between the last router just before the traffic exits the continent and the first router just after coming back. (c) the return portion when the traffic comes back to the continent. We approximate this delay as the RTT between the first router when the trace comes back to the continent and the destination. We then derive the average delay of each of these portions (computed using all traces with boomerang). The result is shown in the three bottom columns of Table 2. Note that our 2015 data and 2016 data produced similar results (for all portions) and thus we present only the results for 2016 in this paper.

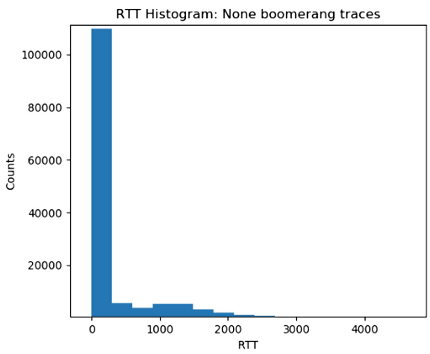

Fig. 1. Histogram of RTT (milli-second) for none boomerang traces

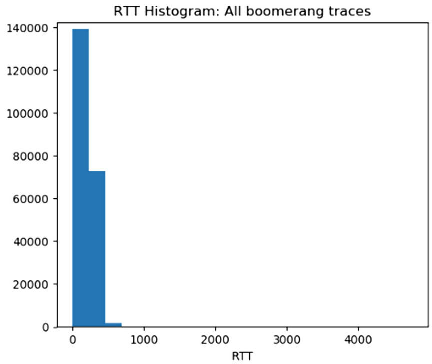

Fig. 2. Histogram of RTT (milli-second) for boomerang traces

Figure 1 shows the histogram of the end-to-end RTT for traces that stayed within Africa, while Fig. 2 shows a similar histogram for traces that exited the continent. We have observed (not shown here) that both types of traces experience larges delays (up to 30 s). However, the figures show that, despite the additional distance travelled by boomerang traffic, none-boomerang traces (i.e., traces that stays within the continent), tend to have larger delays in general. Table 2 shows the average delay for each type of traffic: the average round trip time delay for none-boomerang traces is 407 ms (top column) while the average RTT delay for boomerang traces is 240 ms.

We then zoom into the boomerang traces to analyze the delays in the three different portions: (a) Portion 1: when the traffic is still within Africa, (b) Portion 2: when the traffic exists the continent, and (c) Portion 3: when the traffic returns back to the continent. The average RTT delay for these portions are given in the three bottom columns of the table.

Table 2. Average delay analysis

Traffic type		Average RTT delay (ms)
None-boomerang traffic (entire trace within Africa)		407
Boomerang traffic	End-2-End trace	240
	Portion 1: within Africa	146
	Portion 2: out of Africa	31
	Portion 3: within Africa	63

As can be seen in the table, the delay experienced by the traffic that stays within the continent is much larger. The first portion (from the source to the last router within the continent that is traversed before the boomerang) has the largest delay. It has an average RTT delay of 146 ms, while the last portion (from the first router within the continent when the traffic comes back, to the destination) has an average RTT delay of 63 ms. The boomerang delay is just 30 ms.

The reasons of the differences in these delays are worth investigating. One plausible explanation for the smaller delay experienced by the boomerang traffic (portion 2) might the larger bandwidth that is available in Europe and in the US. We believe that more data is needed in order to understand why the first portion of the traffic has a considerably larger delay than the last portion (despite the fact that they both stay within the continent).

5 Network Improvements

To correct intra-Africa boomerang and eventually improve cost and delay for the African end-users, many have suggested building new infrastructures and mainly Internet Exchange Points (IXPs). This has led the African Union to adopt in 2012 the Program for Infrastructure Development In Africa (PIDA) [17]. In its priority action plan the PIDA has highlighted the importance of establishing Internet Exchange Points. It is in this context that the African Union Commission initiated the African Internet Exchange System (AXIS) project to promote keeping of intra-Africa's internet traffic within the continent by supporting the establishment of National and Regional IXPs in Africa [10]. Through the support of the African Internet Exchange System (AXIS) project, AU Member States with internet exchange points (IXPs) have increased from eighteen in 2013 to thirty-eight in 2017, in 29 countries [18].

We collected the names/urls of all the thirty-eight African IXPs and used reversed DNS lookup to compute their IP addresses. We then use their IP addressed to check the appearance of the IXPs in our router-level traces. We also checked the appearance of IXP links in the traces, where *a link is considered to exist between two IXPs if their network address (subnet mask/16) appeared together in one trace and are adjacent*. From these links, we constructed an IXP-level connectivity graph for Africa. Since most of the IXPs were built after 2015, we only consider 2016 traces.

Although many of the IXPs were observed in the traces (thirty-two among the thirty-eight), occurrences of links were very rare. We only found 40 links that were mostly between IXPs in coastal countries and specially in the south side of the continent. As a consequence, the IXP-level connectivity graph was disconnected with many isolated nodes. This suggests that the additional IXPs built in the continent are still not connected. This is mainly due to the lack of cross-border links, specially for landlocked countries.

Since the IXPs are not connected among themselves, there is still the need to go through transit ISPs for communication between most of African countries. In other terms, despite the construction of new IXPs inside the continent, boomerang routing is still present. In our opinion, the optimal solution for the AXIS project is to combine the building of IXPs with that of links between them. To illustrate that, we use the IXP connectivity graph built earlier and consider the following optimization problem: given that n (cross-border) links are to be added between neighboring countries in Africa, where should we add them to maximally reduce the fraction of boomerang routes? How much this should reduce the fraction of boomerang? Notice that n is a parameter that take different values. In this paper, we consider $n = 1, 2, 3, 4$.

As a first step, we use a naïve approach in which we consider that all links have the same cost (which is certainly not the case). Our analysis shows that by adding only one links, we reduce 0.7% of the boomerang. With two links, we are able to reduce 5% of the boomerang, while with 3 links, 15% of the boomerang are removed. With four additional links, 27% of the boomerang are corrected.

The analysis above shows that adding only IXPs without links between them will not solve the boomerang problem in the continent. We believe that the effort of constructing news IXPs should be backed with an effort to building high speed links between them.

6 Conclusion

In this paper, we use publicly available traceroute data to study the prevalence of boomerang routing in intra-African Internet communication. The analysis has shown that in 2016, still 60% of traffic between two African countries has to go through Europe or the US before reaching their final destination. This raises issues with regards to cost, delay, as well as information exposure. We have also analyzed the delay profile of the communication paths and have discovered that, despite the boomerang, the traffic that stays within the continent experiences larger delays. Finally, we have used our traces to build a connectivity graph between the existing Internet Exchange Points (IXPs) in the continent. Our study has shown that the African IXPs are not connected. This indicates that despite the effort spent in building these IXPs, the boomerang phenomenon remains current in the continent. Consequently, we believe that efforts by the African Union (AU) to support individual member states to build IXPs should be backed with parallel effort to build high speed cross-border links between the countries. Related to this, we plan to follow up this paper with a study on how to optimally determine where to build those links by considering cost, regulation and geographical constraints.

Acknowledgements. This work was partially accomplished under NIST Cooperative Agreement No. 70NANB16H024 with the University of Maryland. The authors thank Dr. Peter Mell (CSD/NIST) for his support (time and advice). The authors would like to thank Christopher Schanzle (ACMD/NIST) for his help in gathering the data.

References

1. McKinsey Global Institute: Lions Go Digital. The Internet's Transformative Potential for Africa, 2013 Annual Report. https://www.mckinsey.com/industries/high-tech/our-insights/lions-go-digital-the-internets-transformative-potential-in-africa. Accessed 03 Dec 2017
2. Guerrioro, M.: The impact of internet connectivity on economic development in Sud-Saharan Africa. EPS Speaks, January 2015
3. World Bank (WB): Information and communications for development 2009: extending reach and increasing impact. https://issuu.com/world.bank.publications/docs/9780821376058. 15 Dec 2017

4. ITU: White Paper on Broadband Regulation and Policy in Asia-Pacific Region: Facilitating faster Broadband Deployment, November 2016
5. Internet.org by Facebook: The State of Connectivity 2015–2016: A Report on Global Internet Access https://fbnewsroomus.files.wordpress.com/2016/02/state-of-connectivity-2015-2016-02-21-final.pdf. Accessed 03 Dec 2017
6. Edmundson, A., Ensafi, R., Feamster, N., Rexford, J.: Characterizing and avoiding routing detours through surveillance states. ArXiv. http://arxiv.org/abs/1605.07685. Accessed 31 Dec 17
7. Obar, J.A., Clement, A.: Internet surveillance and boomerang routing: a call for canadian network sovereignty. In: Ross, P., Shtern, J. (eds.) Proceedings of the Technology and Emerging Media, Track–Annual Conference of the Canadian Communication Association, TEM2013, Victoria, 5–7 June 2012 (2013). http://www.tem.fl.ulaval.ca/fr/victoriaP2013
8. CAIDA: Macroscopic Internet Topology Data Kit. http://www.caida.org/data/internet-topology-data-kit/. Accessed 13 Dec 2017
9. Maxmind geolocation service, December 2017. https://www.maxmind.com/en/home
10. The African Internet Exchange System (AXIS) Project. https://au.int/en/african-internet-exchange-system-axis-project-overview. Accessed 24 Dec 2017
11. ITU: Are poor countries subsidizing the rich? The Issues of International Internet Connectivity. https://www.itu.int/itunews/manager/display.asp?lang=en&year=2005&issue=03&ipage=interconnectiv&ext=html. Accessed 03 Dec 2017
12. The Halfway Proposition, Background Paper on Reverse Subsidy of G8 Countries by African ISPs. In: Conference of African Ministers of Finance, Planning and Economic Development, Johannesburg, South Africa, 19 October 2002. www.afrispa.org/HalfwayDocs/HalfwayProposition_Draft4.pdf
13. Huffaker, B., Fomenkov, M., Claffy, K.: Geocompare: a comparison of public and commercial geolocation databases. In: Proceedings of NMMC, pp. 1–12 (2011)
14. Fanou, R., Francois, P., Aben, E.: On the diversity of interdomain routing in Africa. In: Mirkovic, J., Liu, Y. (eds.) PAM 2015. LNCS, vol. 8995, pp. 41–54. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-15509-8_4
15. Chavula, J., Feamster, N., Bagula, A., Suleman, H.: Quantifying the effects of circuitous routes on the latency of intra-Africa internet traffic: a study of research and education networks. In: Nungu, A., Pehrson, B., Sansa-Otim, J. (eds.) AFRICOMM 2014. LNICST, vol. 147, pp. 64–73. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16886-9_7
16. African Union: Study on Harmonisation of Telecommunication, Information and Communication Technologies Policies and Regulation in Africa, March 2008. http://www.itu.int/ITU-D/projects/ITU_EC_ACP/hipssa/docs/2_Draft_Report_Study_on_Telecom_ICT_Policy_31_March_08.pdf
17. Program for Infrastructure Development in Africa. http://www.au-pida.org/. Accessed 02 Jan 2018
18. The African Peering & Interconnection Forum (AFPIF) 2017 Conference Report. https://www.afpif.org/afpif2017/. Accessed 02 Jan 2018